Diode calibration. Figure 1 shows a plot of the diode voltage versus temperature for three diodes with a wide variation in their room temperature voltages. The difference in voltage between these units was discovered to be almost linear in temperature. (See figure 2.) This allows for a single point calibration of the diodes. The room temperature voltage of 'your' diode is compared to the standard diode, (diode C in the figures). The voltage at any other temperature is then calculated from the known voltage of the standard diode at that temperature and recorded in the table below. A linear extrapolation between the recorded temperatures will give the voltage at any temperature. For the best accuracy you should also record the DC offset of the monitor output (typically +/-1 mV)

NF163

Temperature (K)	Voltage (mV)
77.320	994.419
90.000	967.032
100.000	944.477
110.000	921.305
120.000	897.616
130.000	873.417
140.000	848.864
150.000	823.964
160.000	798.691
170.000	773.213
180.000	747.485
190.000	721.533
200.000	695.363
210.000	669.002
220.000	642.448
230.000	615.748
240.000	588.907
250.000	561.837
260.000	534.655
270.000	507.311
280.000	479.881
290.000	452.258
300.000	424.465
310.000	396.645
320.000	368.731
330.000	340.625
340.000	312.542
350.000	284.206
360.000	255.855
370.000	227.428
380.000	198.940
390.000	170.463
400.000	142.124

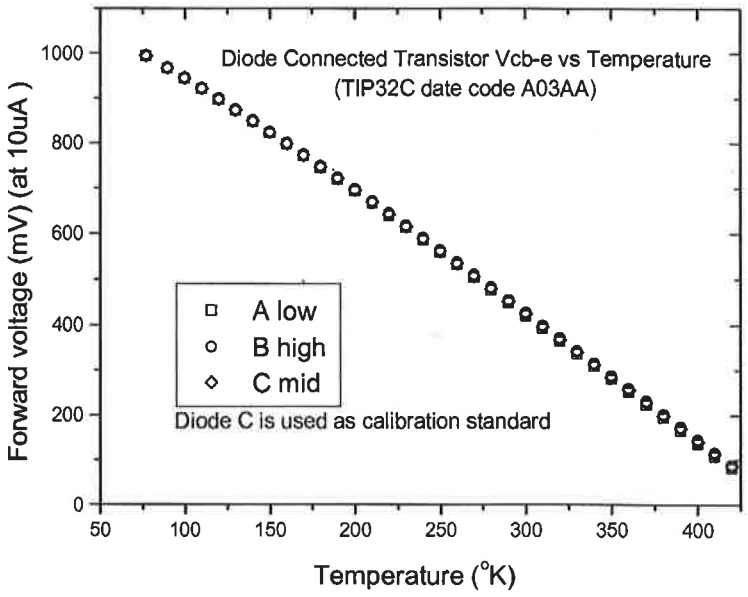


Figure 1. Diode voltage vs. Temperature

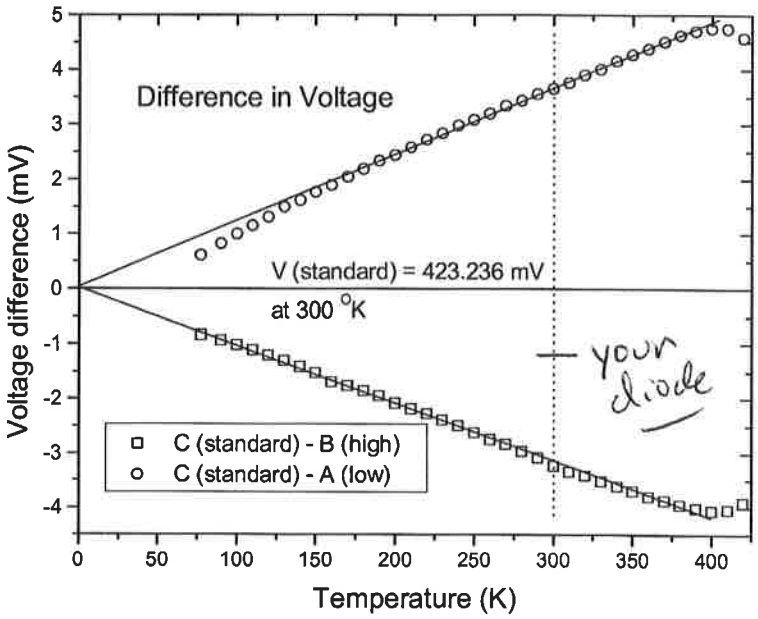
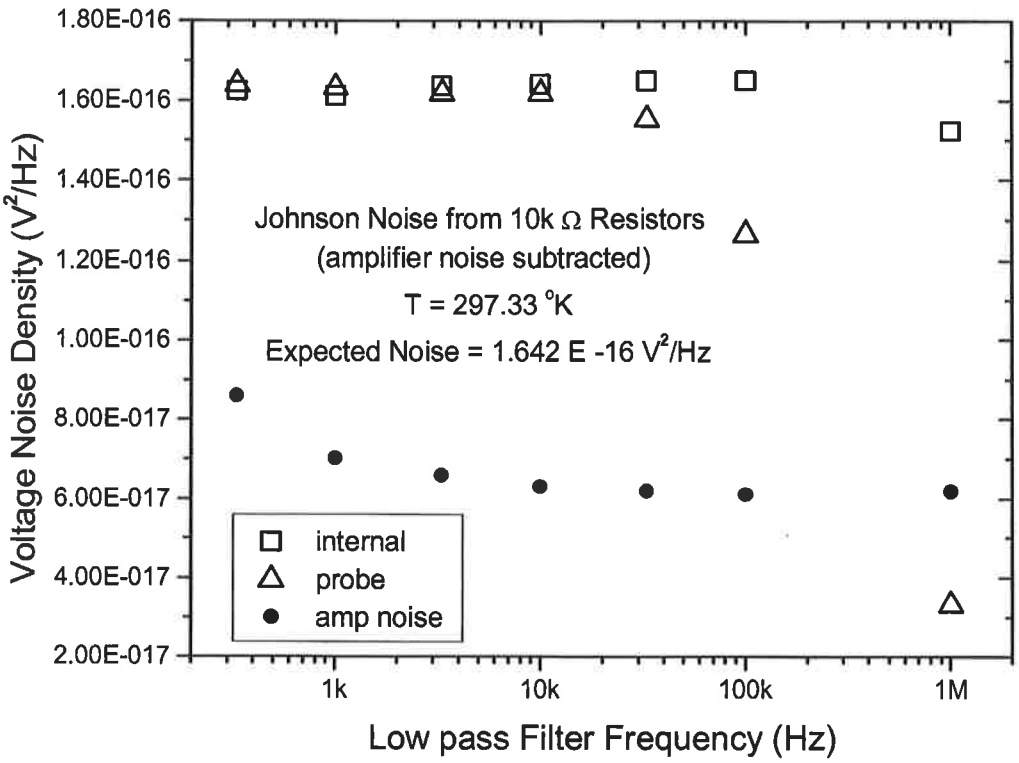
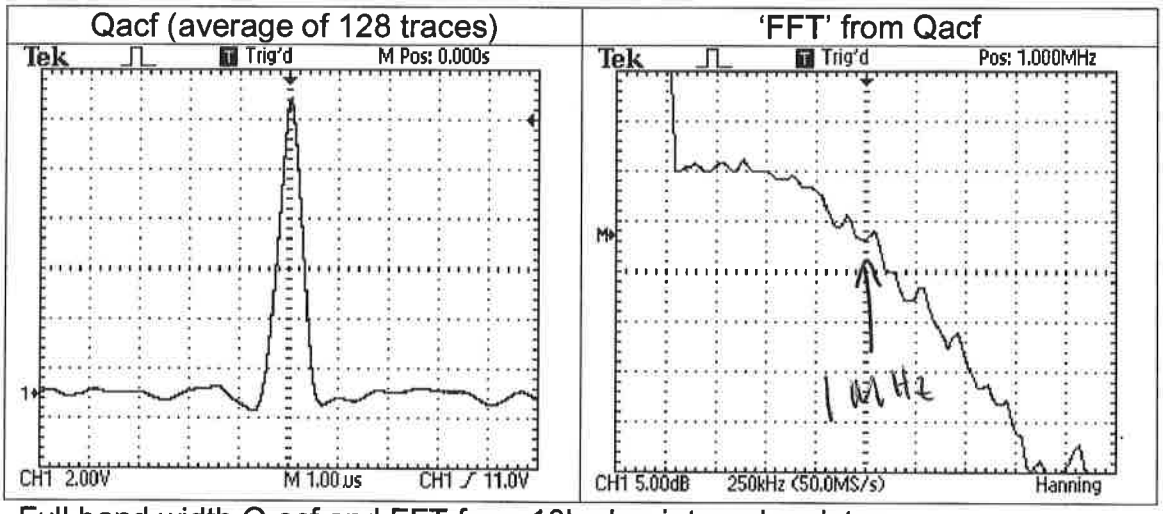


Figure 2. Difference in diode voltage

May 28, 2014





Full band width Q acf and FFT from 10k ohm internal resistor.
Low level gain = 600, High level gain = 300.

High frequency filter values					Low frequency filter values			
Nominal freq. (Hz)	Measured freq. (Hz)	Q	damping factor γ=1/2Q		Nominal freq. (Hz)	Measured freq. (Hz)	Q	damping factor γ=1/2Q
330	336.5	0.706	0.708		10	10.02	0.707	0.707
1 k	1.010k	0.706	0.708		30	30.06	0.707	0.707
3.3 k	3.365k	0.707	0.707		100	100.2	0.707	0.707
10 k	10.09k	0.7088	0.7055		300	300.6	0.707	0.707
33 k	33.55k	0.7159	0.6984		1.00k	1.002k	0.707	0.707
100 k	99.95k	0.7388	0.6768		3.00k	3.006k	0.708	0.706

High level electronics filter measurements

Diode voltage T =297.33 K

Current	Voltage (mV)
10nA	248.72
100nA	311.47
1uA	371.96
10uA	431.75
100uA	491.13
1mA	550.60
offset	-0.14

60

Temperature sensing diode installed in probe.

IMPORTANT INFORMATION

for

INSTRUCTORS TEACHING WITH THIS APPARATUS

To: All Instructors using TeachSpin Apparatus
From: Your TeachSpin Support Staff

Please make sure you are on our TeachSpin Contact List so that you will hear about any revisions or additions to your instrument manual as well as any new experiments you can do with this apparatus.

Send Current Contact Information to info@teachspin.com.

All of our TeachSpin apparatus is designed and built by physicists who have taught (or are still teaching) in the undergraduate lab. In addition to being available for both technical and instructional support, our TeachSpin physicists are always finding new ways to use our apparatus. We want you to know about them.

We look forward to welcoming you into the TeachSpin community

TeachSpin Apparatus

Reliable, Robust, and Always Supported

Warranty: 2 years - Support: Lifetime

Questions: 716-885-4701 or info@teachspin.com

TEACH SPIN

Instruments Designed for Teaching

Noise Fundamentals

NF1-A

INSTRUCTOR'S MANUAL

A PRODUCT OF TEACHSPIN, INC.

Written by David Van Baak
Edited by George Herold
Jonathan Reichert

TeachSpin, Inc.
2495 Main Street Suite 409 Buffalo, NY 14214-2153
Phone: (716) 885-4701
Fax: (716) 836-1077
www.TeachSpin.com

Table of Contents

Table of Contents (Annotated)

Appendices

0. Introduction

0.0. Parts of this apparatus

Your instrument has been carefully packed in two boxes with an enclosed packing list. It is important that you check to be sure all the parts have been received and are in good condition. Please inform TeachSpin immediately if any parts are missing or have been damaged in transit.

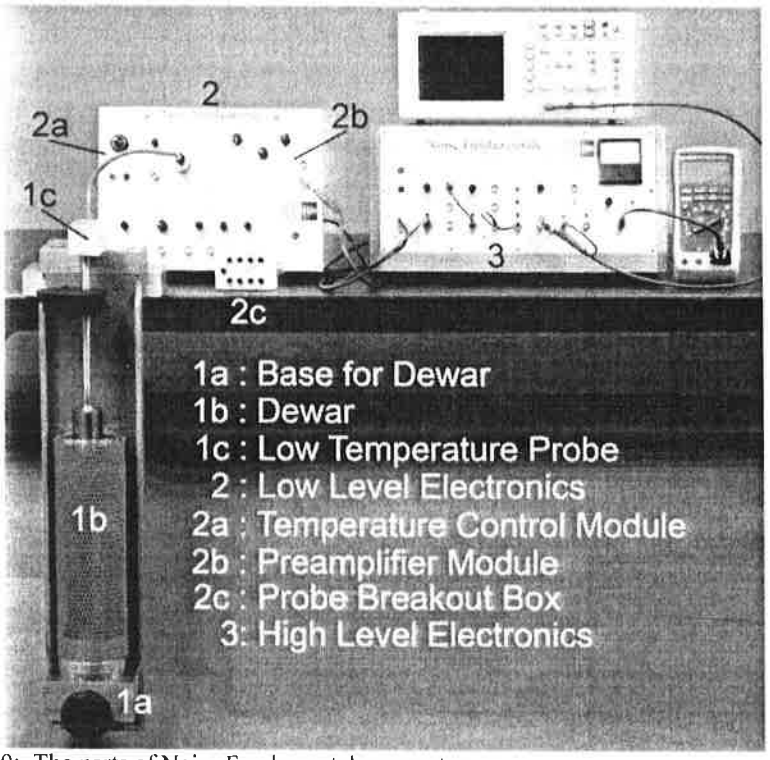


Figure 0.0: The parts of Noise Fundamentals apparatus

The major components of this apparatus are shown in Fig 0.0. Box 1 has the high level electronics in a wooden case, the low level electronics in a steel and aluminum case, the universal power supply (not shown), plastic parts boxes 1 and 2 containing various tools and spare parts (not shown), this instruction manual, three 36-inch BNC cables, and a a.c. power cable (for US – Canada).

Box 2 contains the apparatus for the variable-temperature measurements. These include the variable-temperature sample probe, the Dewar vessel, the Dewar support, and the breakout box. All these are shown in Fig 0.0. Please note that TeachSpin does not supply the oscilloscope or the digital voltmeter.

This instruction manual should contain two pages of actual data taken on your instrument. This data was taken at TeachSpin not only to test your unit, but also to provide you with benchmark data that you and your students should be able to reproduce.

0.1. Definition, kinds, uses of noise

Just as a weed is an unwanted plant, a noise, *ordinarily speaking*, is an unwanted sound. In the fields of physics, electrical engineering, and many other places, we extend the definition of 'noise' beyond acoustics to the general field of information. Since almost any signal that's a function of time can be translated into a voltage, we will often use the concept of a voltage signal. We'll call it a 'noisy signal' if, in addition to the voltage we expect or wish to see, there is unwanted, typically (but not always) a randomly-fluctuating, voltage. Surprisingly, the noise signal is sometimes not only wanted, but is the essence of the measurement.

There are several kinds of noise. One of them is 'interference', which is the presence of an unwanted signal, added to the desired signal. It's easy to imagine that your neighbor's electronic apparatus is polluting your TV or radio signal with some sort of interference. The kind of interference students are likely to encounter in these experiments probably comes from three sources: electrostatic coupling to the apparatus from fluorescent lights in the laboratory, electromagnetic coupling due to nearby transformers or motors, and vibrational coupling due to microphonic components within the unit.

Another source of noise we will call 'technical noise' since it is the noise generated by the technique of the investigation, or that gets into the circuits due to *faulty* experimental techniques. For example, a student's failure to tighten the cover on the preamplifier section, or a poor electrical connection to the first-stage op-amp, can add extraneous noise to the signal path.

Of greatest interest to us is 'fundamental noise', noise that is intrinsic and inevitable because of the physical nature of an apparatus. We'll observe noise sources that arise from the Second Law of Thermodynamics, and from the quantization of electrical charge. Physicists and electrical engineers know these as Johnson and shot noise respectively. Noise sources like this display the characteristics of non-periodic, unpredictable, random waveforms, but nevertheless conforming, in their statistical properties, to universal laws.

Fundamental noise is especially worthy of study, for at least two reasons. The first reason is that fundamental noise presents us with a physics-based limit on the degree to which we can measure in a given experiment. In many cases in research and technology, it often defines what is possible within the limits of physical law. In particular, fundamental noise can and does set limits to the rate of data-transfer in a host of contexts in communication.

The second reason we care about noise is that it becomes possible to use noise to measure the values of some fundamental constants. Boltzmann's constant k_B can be determined from the voltage or Johnson noise of resistors; and the magnitude of the charge on the electron, e, can be determined from the current or shot noise of a photocurrent.

But measurement of 'fundamental noise' has its experimental challenges. There is a saying about noise measurements: 'you're either measuring too much or too little signal'. You will understand this quip better after you have had some experience with these measurements. Our advice here is to read both the manual and some of the references and do your measurements carefully. **But most of all, have fun!**

0.2. Tactics for using this apparatus

'Noise Fundamentals' is an unusually versatile apparatus which can be used to explore noise in a wide variety of ways. It is incumbent **on the instructor** to devise a plan for the use of NF-1 that is appropriate for the level of students in the class, the time available, and the pedagogical goals of the program. This manual is designed to help the instructor create this plan.

To begin the process of creating your plan, we would like to point out some of the features of the unit which you should consider. The low-level electronics (LLE) is designed so that part of it can be rewired and reconfigured by the user. The basic motivation for this design is that *no single pre-amplifier configuration* can be suitable for all types of noise measurements. The question of who does this rewiring and reconfiguration is up to the instructor. For example, you might have your students make a few simple changes in the wiring, or select different feedback resistors in the circuit. If the time is available, and the skills that can be acquired within the goals of the lab, you might have students configure the first stage pre-amplifier from a schematic diagram alone.

The manual has electronic schematic diagrams, as well as the point-to-point wiring diagrams for the various configurations. The unit has both terminal blocks and IC sockets so that almost all this rewiring and reconfiguration can be done without soldering. Also, we have current-limited all of the power supplies so that the inevitable wiring mistakes and shorts will not cause permanent damage to the electronics. Students often learn the most from their mistakes, so let them make them.

The unit has the capability of measuring noise as a function of temperature from 77 - 400 K, but this requires liquid nitrogen and a fair amount of time. Should you desire to examine noise in one three-hour class period, you might consider configuring the electronics yourself and examining Johnson noise as a function of resistance at a constant temperature and bandwidth.

This apparatus could also be used in an electronics course to teach analog electronics. Not only can students learn how to configure low-noise electronics, they can learn the uses of low-noise, high-gain, variable-bandwidth electronics. This electronics is capable of measuring very small signals. For example, George Herold (TeachSpin Senior Scientist and the principal designer of NF-1) used it in another project, to compare the reverse-bias leakage current of a 1N4148 diode (≈5 nA) and of the base-collector 'diode' of a 2N4401 NPN transistor (≈1 pA).

The unit comes with a wide variety of components, all of which will be explained in this manual. Some components are hard-wired into the unit, while others need to be inserted. We would encourage faculty and students to make noise measurements on other components that intrigue them. If a particular set of measurements turns out to be very interesting, possibly with unexpected results, TeachSpin would love to hear about it. We have long had a standing 'reward' available for students who have come up with new experiments to do with our equipment.

0.3. How to use this Manual

Noise Fundamentals (NF1-A) can be used in a wide variety of modes in both the undergraduate and graduate educational experience. Both the level of experiences and the amount of time available must be considered in designing the appropriate use of the equipment. We have created some 'routes' by which an instructor might reach the different goals they have set for their students. These routes have been organized in a kind of 'matrix', where the horizontal direction indicates student level, and the vertical direction is categorized by laboratory time available.

Although the matrix refers to specific sections in the manual, we do not intend that these sections should necessarily be copied for student use. ***This manual is written for the instructors***, but you should feel free to use any part that you deem appropriate for your students. Since there are so many different levels of students, and so many different objectives that can be achieved, ***instructors*** will have to create manuals which will fit their own particular needs. Please remember, this matrix is only an approximate guide.

Remember too that different users will want to follow distinct *themes*. These might include physics explorations, electronics proficiency, and metrological calibrations. To follow the theme you judge appropriate for your students, it is essential to become familiar with the instrument and its capabilities. The best way we know is to read (or at the least, to browse) the entire Manual, and to get hands-on familiarity by doing most of the experiments.

The following notes may be of help in finding a particular topic.

- If you want to learn how to quantify noise, do Johnson noise, via Sections 1.0 - 1.5.
- If you want to measure Boltzmann's constant k_B, and are content with a single temperature, add Section 1.6.
- If you want to see the temperature variation of Johnson noise, you need Sections 4.1 - 4.2; to do temperature variation systematically, add sections 4.3 - 4.4.
- If you want to use shot noise to measure the electronic charge '*e*', you'll need Sections 1.1 - 1.5 to learn how to quantify noise, and then Sections 3.0 - 3.3 to produce and quantify shot noise.
- When you want to review how to measure noise, and noise density, see Sections 2.1 - 2.2.
- When you want to understand details of calibrations and uncertainties in noise measurement, see Sections 5.0 - 5.3.
- If you want to know what a 'Volt per root Hertz' or V/√Hz is, see Section 5.4.
- There is a special project in 'noise thermometry' found in Section 6.4.
- If you want to learn the definitions of 'noise temperature' and 'noise figure', see the end of Appendix A.3.
- If you want to learn about computer-based Fourier methods for extracting the frequency spectrum of noise, you'll want to use Appendix A.10.

Tactical Matrix

This matrix assumes all students have read and carried out Section 7, Getting Started

(**Instructor CE** or **Student CE** indicates whether the Instructor or Student Configures the Electronics.)

Total Lab Time (hours)	INTRODUCTORY	INTERMEDIATE	ADVANCED
3 - 4	**Instructor CE** Johnson Noise vs. *R* with fixed *T* and Δf *Do 1.0-1.4* **or** Johnson Noise vs. Δf with fixed *R*, *T* *Do 1.0-1.5* **or** Shot Noise vs. *i* with fixed Δf *Do 3.0-3.3*	**Instructor CE** Johnson Noise vs. *R* with fixed *T* and Δf *Do 1.0-1.4* **or** Johnson Noise vs. Δf with fixed *R*, *T* *Do 1.0-1.5* **or** Shot Noise vs. *i* with fixed Δf, *T* *Do 3.0-3.3*	NOT APPROPRIATE
6 - 8	**Instructor CE** Johnson Noise vs. *R*, Δf with fixed *T* *Do 1.0-1.6,* *Read 2.0-2.2* **or** Shot Noise vs. *i*, Δf *Do 3.0-3.3* *Read 3.6, 3.7*	**Instructor CE** Johnson Noise vs. *R*, Δf with fixed *T* **and** Shot Noise vs. *i*, Δf *Do 1.0-1.6, 2.0-2.2, 3.0-3.3* **or** Johnson Noise vs. *R*, Δf and varying *T* *Read 1.0-1.6, 2.0-2.2, 4.1-4.5*	**Instructor CE** Johnson Noise vs. *R*, Δf with fixed *T* **and** Shot Noise vs. *i*, Δf *Do 1.0-1.6, 2.0-2.2, 3.0-3.3* **or** Johnson Noise vs. *R*, Δf and varying *T* *Read 1.0-1.6, 2.0-2.2, 4.1-4.5*
12 to 20	NOT APPROPRIATE	**Student CE** Johnson Noise vs. *R*, Δf with fixed *T* **and** Shot Noise vs. *i*, Δf *Do 1.0-1.6, 2.0-2.2, 3.0-3.4* **or** Drop one experiment and add Calibrations, *Do 5.0-5.3*	**Student CE** Johnson Noise vs. *R*, Δf with fixed *T* **and** Shot Noise vs. *i*, Δf *Read 1.0-1.6, 2.0-2.3, 3.0-3.3* **or** Johnson Noise vs. *R*, Δf and *T* *Read 1.0-1.6, 2.0-2.2, 4.1-4.5* Advanced students can do calibrations (*Ch. 5*), projects (*Ch. 6*), or explore other parts of the manual.

0.4 Getting Started

After reading several sections, particularly Chapters 1 and 2, you may tire of reading theory and want to fire up the unit. **Section 7, Practical Guide** describes how to begin making measurements. It will hold your hand through your first measurements of Johnson noise, and will demonstrate how to analyze the raw data. Feel free to start there.

Self-configured electronic instruments present unique problems to students who are new to electronics, and who have little or no built-in instinct for recognizing problems with a circuit they have built or configured. In Newtonian Mechanics, by contrast, if a glider fails to slide on an air track, even the most naïve student recognizes that something is very wrong with the apparatus. If a student hears a scraping sound when operating a torsional oscillator, he or she immediately recognizes there is a problem to be cured. But if a beginning student dutifully builds a required circuit as shown in a manual, and has a lab partner check it, they will probably assume that the circuit will perform as expected.

But that may not happen – why? Experienced practitioners can think of many reasons:
1. One of the components may be defective, or far out of specifications.
2. There may be one or more poor connections or bad contacts.
3. The power supply may have problems – for example, low voltage, noisy voltage, oscillating voltage, limited current, etc.

It is also possible, or even likely, that the student has made one or more mistakes. For example, a student might have:
1. Put in an additional wire.
2. Made the wrong connection
3. Used the wrong components
4. Reversed the polarity on a component for which it matters.
5. Left out a part or a connection.
6. Left *in* a part or wire from a previous investigation.

Clearly, students need an independent way to test a circuit set-up before it is used to take data. Among other things, students need to know if they are measuring external interference, rather than the intended internal noise, and they also need to know if the bandwidth of their system is correct. We have put a whole collection of diagnostic techniques in Appendices A.5 and A.6, and in Chapter 8, so that they can be used any time that students suspect there may be a problem, or that they wish to check out an experimental configuration.

0.5 Care and Maintenance of the Apparatus

The front panel of the low-level electronics has different properties than you might expect. In order for shielding (see Appendix A.5) to be most effective, the panel must be electrically conductive where it touches both the screw-in modules' panels and the black steel enclosure. So these aluminum panels are *not* anodized. Instead, their surface treatment leaves them conductive, but also more susceptible to scratches than anodized aluminum.

If your panels get discolored, you are free to use denatured alcohol and paper wipes, along the direction of their brushed finish, to clean them of any oil or grease they may have accumulated. This will leave their electrical conduction unimpaired.

The wooden box of the high-level electronics is best maintained by occasionally wiping it with a damp (not wet) cloth, perhaps using a small amount of liquid dishwashing soap.

If you want to put the apparatus in the 'default condition' in which it's shipped, consult Section 7.2 to see, in particular, how to restore the as-shipped state of the wiring inside the pre-amplifier of the low-level electronics.

1. Johnson noise at room temperature

1.0 The reasons for Johnson noise, and its predicted size

Every student knows $V = i\,R$, which really says that there's a potential difference ΔV across any resistor R which has a current i passing through it. This of course predicts a ΔV of *zero* for a resistor with no current. But for deep reasons, any actual resistor at any temperature above absolute zero, will display a 'noise voltage' $V_J(t)$ across its terminals, a potential difference that has all the character of an internal (a.c.) *emf* built into the resistor. The emf which the resistor generates is called 'Johnson noise', and it arises because of the deep thermodynamic connection between dissipation (which any resistor surely has) and fluctuations (which here show up as a fluctuating emf). The size of this emf is also predicted by fundamental theory, and it should not surprise you to learn that $V_J(t)$ is, *on average*, zero. But $V_J(t)$ exhibits fluctuations, positive and negative, about that average value of zero. To quantify these, we form the (always-positive) *square* of $V_J(t)$, and time-average that, giving a 'mean square' voltage which we denote as $\langle V_J^2(t)\rangle$. The predicted value for $\langle V_J^2(t)\rangle$ was first deduced by Nyquist, following Johnson's empirical discovery of the noise, and it's given by the expression

$$\langle V_J^2(t)\rangle = 4\,k_B\,R\,T\,\Delta f\ .$$

Here k_B is Boltzmann's constant, T is the (absolute) temperature of the resistor, and Δf is the novel factor -- it is the 'bandwidth' used in the measurement electronics.

The involvement of bandwidth Δf is a first hint that 'noise' is quite distinct from 'signal'. Everyone starts with 'd.c. signals', which have nothing but a sign and a value, in Volts. Then there are 'a.c. signals', which have a magnitude (perhaps specified by amplitude, or rms value, or peak-to-peak excursion) but also a *frequency*, or a mixture of frequencies. But it is the essence of fundamental noise that it contains, or is composed of, *all frequencies*. In fact, the amount of energy we can get out of a 'noise source' depends on the *range* of frequencies to which we arrange to be sensitive, and this is the reason for the inclusion of the bandwidth-factor Δf in the expression above.

How large a Johnson-noise voltage should we expect from a typical resistor? Let's calculate this mean-square voltage for a 100-kΩ resistor at room temperature. Suppose that our electronics for detecting and measuring $V_J(t)$ are fully sensitive to all frequencies from 0 to 100 kHz, but entirely insensitive to higher frequencies. Then:

$T = 22\ ^\circ\text{C} = 295\ \text{K}$
$k_B = 1.38 \times 10^{-23}\ \text{J/K}$ (textbook value)
$\Delta f = 100\ \text{kHz} = 10^5\ \text{Hz}$

$$\begin{aligned}\langle V_J^2(t)\rangle &= 4\,(1.38 \times 10^{-23}\ \text{J/K})\,(295\ \text{K})\,(10^5\ \Omega)\,(10^5\ \text{Hz}) \\ &= (1.63 \times 10^{-20}\ \text{J})\,(10^5\ \text{V/A})\,(10^5\ \text{/s}) \\ &= 1.63 \times 10^{-10}\ \text{V}^2.\end{aligned}$$

Not everyone is familiar with the curious unit of the square-of-a-Volt, so we often take the square root of this mean-square noise voltage, to give a 'root-mean-square' or 'rms' measure of the noise voltage,

$$V_J(\text{rms}) \equiv \langle V_J^2(t) \rangle^{1/2} = 1.28 \times 10^{-5} \text{ V} = 12.8 \ \mu\text{V}.$$

So if we have a room-temperature 100-kΩ resistor simply hooked up to an ideal voltmeter, and if that voltmeter responds to all (but only) frequencies under 100 kHz, then the voltmeter's instantaneous reading will *not* be zero volts, but instead will fluctuate (rapidly: in this case, on a microsecond time scale) around zero, with typical excursions of order ±10 μV. We further assert that this is an actual emf intrinsic to the resistor, and it will still be present, though typically unwanted, in addition to any *iR*-drop that the resistor may exhibit. It follows that measurement of any *iR*-drop to microVolt precision in such a case would require thinking about this effect.

There are many textbook derivations of Nyquist's prediction, and the best of them emphasize the connection to thermodynamics and to blackbody radiation. Here's a 'thought experiment' to help you see that some sort of Johnson noise must exist. First imagine a cubic meter of iron at room temperature and another cubic meter of cold iron (say, at temperature $T = 4$ K), spaced 10 meters apart in empty space. (If you like, think of them as located at the two focal points of a large evacuated ellipsoidal reflecting cavity which surrounds them both, and isolates them from the external universe.) It should be clear to you that each iron block is giving off blackbody radiation, with a range of frequencies and in all directions -- but that the warm block is giving off lots more. Since the blackbody radiation of each block will run into the other block, there will be a net flow of (radiant) energy from the warmer block to the colder one, and their temperatures will therefore start to equilibrate.

Now imagine a 50-Ω resistor at room temperature, connected to nothing but a lossless coaxial cable of 50-Ω impedance; and imagine there's another 50-Ω resistor, but down in a Dewar at $T = 4$ K, connected to the far end of this cable. Even if there is no thermal conductivity in the cable, there is still electrical conductivity. It's the 'Johnson emf' in each resistor which still acts like a black-body source, here generating travelling waves of (confined) radiation along the one-dimensional cable structure, and that 'radiation' is caught and dissipated in the far end's resistor. This is the mechanism by which the two resistors will tend toward thermal equilibrium, as the hotter resistor will experience a net outflow, and the colder a net inflow, of electrical energy.

1.1 'Seeing' Johnson noise

This exercise will let you see, directly on an oscilloscope, a time-dependent waveform which can be traced all the way back to the Johnson noise generated in a resistor. You'll need to ensure that you've restored the 'default condition' of the system for this to work -- see Section 7.2, Getting Started.

You need to plug, into your 100-to-240-V outlet, the line cord of the the universal power supply which supplies power to the high-level electronics (HLE). You should see a green LED on the transformer unit light up. Now connect the output of this supply to the receptacle on the back of the HLE. You should see a green LED on the front panel of the HLE light up. (Note there is no power switch in the HLE box; instead, it gets powered up as soon as you establish the power-supply connections.) Now find the power cable emerging from the LLE box, and plug it into the connector on the front panel of the HLE box. You should see a green LED light up on the front panel of the LLE. Once you have three green LEDs lit, everything in your system is being powered.

Set the switch to select a 'source resistor' of R_{in} = 100 kΩ in the pre-amplifier module installed in the LLE box. This resistor is connected only to the high-impedance input of the first stage of amplification in the pre-amp. That first stage is wired to give a 'gain', or amplification factor, of 6.00, *provided* you set the feedback resistor, R_f, to its 1-kΩ setting. (The feedback capacitance C_f is not connected in the default mode, so its setting is irrelevant.) Read the graphics on the panel of the pre-amp to see that there is an additional amplification stage, with gain 100., following this first stage. Now you can connect the pre-amp's output, by a coaxial cable, to an oscilloscope, to see if there is any signal present. Use a rather sensitive vertical scale on your 'scope (of perhaps 10 mV/division sensitivity), a sweep speed of 5 μs/div on the horizontal axis, and trigger near zero volts.

Below are the schematic, and the wiring, diagrams of the circuit you're using.

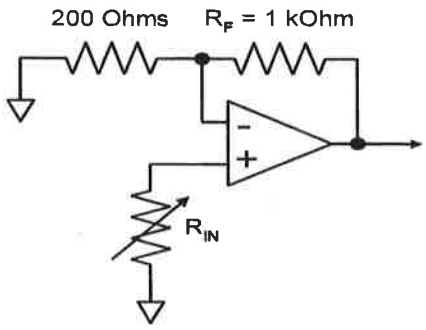


Fig 1.1a: Johnson noise preamplifier schematic

The wiring diagram for this configuration is shown in Figure 1.1b. The connections indicated in grey-scale printing are those you need to check, or establish. (By contrast, connections shown in thin solid lines are already established for you on the printed-circuit boards.)

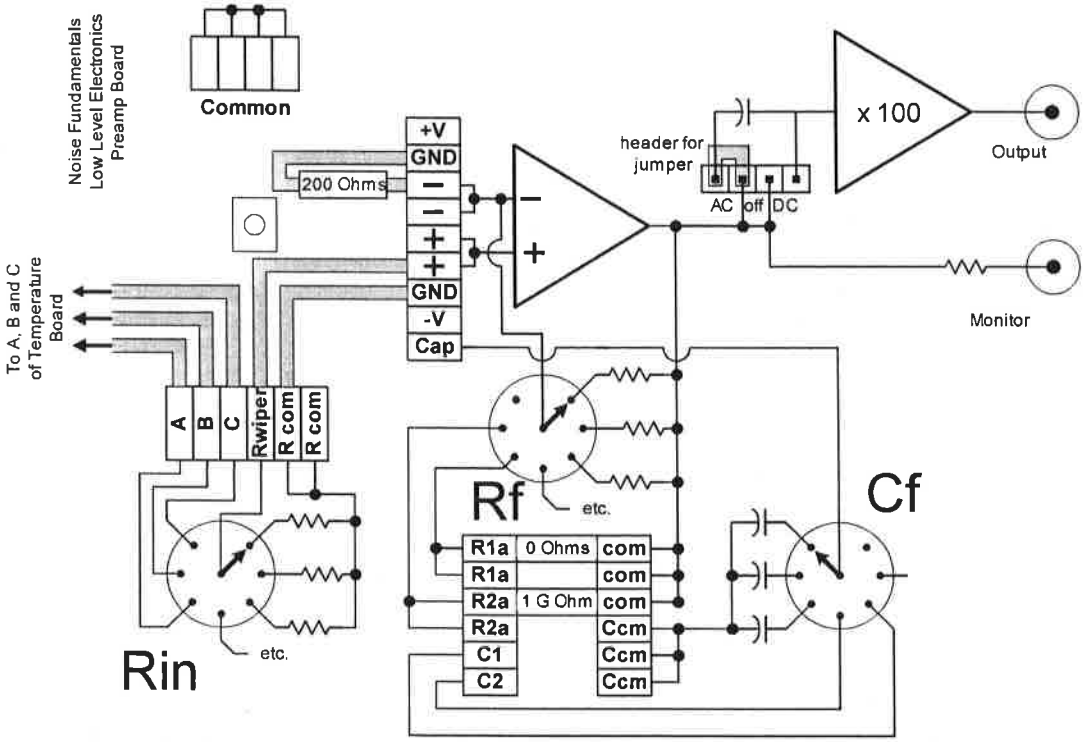


Fig. 1.1b: Wiring diagram of the default condition of the interior of the low-level electronics.

The signals you've seen emerging from the pre-amp are rather small. So next use a BNC cable to convey the pre-amp output to the HLE box instead, where you can filter and amplify the still-small noise signals. If you use the settings and the cabling shown in Fig.1.1c, you will be selecting a frequency band, extending from about 100 Hz to about 100 kHz, to pass along to the main amplification stages. The first filter shown has its high-pass output in use; you may think of this as passing frequencies on the high side of 100 Hz, or equivalently as blocking frequencies below 100 Hz. The second stage is used as a low-pass filter, here passing all frequencies on the low side a chosen 100 kHz. So after the output of the two filters, you have Johnson noise, pre-amplified by factor 600., and then filtered to pass only the 0.1 - 100 kHz frequency band.

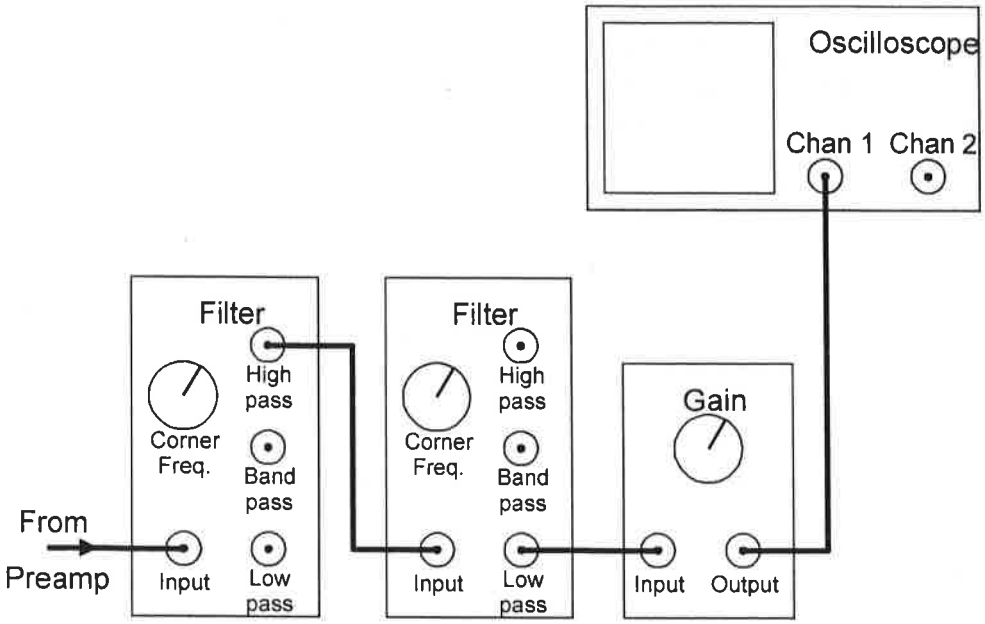


Fig. 1.1c: Cabling diagram for first use of the high-level electronics.
(left) Filter: selector to .1k (for 0.1 kHz), switch to AC (for a.c. coupling)
(right) Filter: selector to 100k (for 100 kHz), switch to AC (for a.c. coupling)
Gain Fine Adjust 30, toggle x1, toggle x10

Notice the figure shows more cabling, now to amplify this signal by a further factor of 300. You achieve this by a setting of gain x1 and x10 at two toggle-switch settings, and a further gain of x30 on the rotary switch setting. (Here too you can switch to AC for a.c. coupling at the input.) Finally, at the output of this main amplifier, you'll have a signal large enough to see easily on a 'scope. A view of it, using a 2 V/div vertical sensitivity, and a 10 μs/div horizontal scale, is shown in Figure 1.1d.

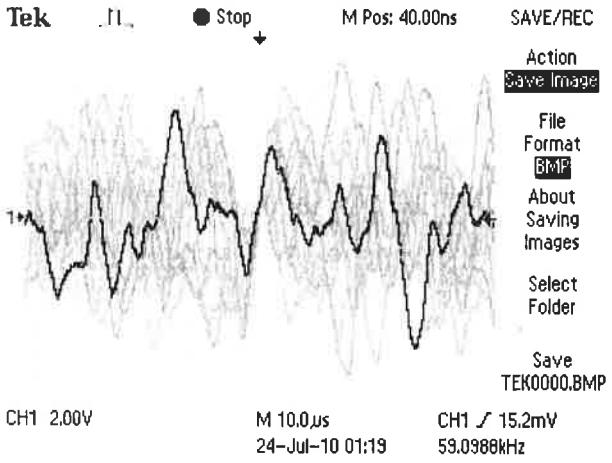


Fig. 1.1d: Samples of amplified Johnson noise from a 100-kΩ resistor, using pre-amp gain 600, filtering to 0.1 - 100 kHz bandwidth, and main-amp gain 300. Vertical scale 2 V/div, horizontal scale 10 μs/div, triggering on positive-going zero-crossings.

To get a first, qualitative, indication that this 'noise signal' has something to do with the original source resistor at the front end of this pre-amp/filter/main-amp chain go back to the pre-amp, and change the source resistor from 100 kΩ to 10 kΩ. You should see the size of the noise signal on your 'scope *change* -- it should decrease, and by a factor of about three.

For a first rough understanding of the *size* of these 'scope signals, consider our claimed 13-μV (rms measure, in the 0-100 kHz band) Johnson-noise signal emerging from a 100-kΩ source resistor. The pre-amp gain of 600 ought to raise this to about 8 mV (rms), and further main-amp gain of 300 ought to raise this to about 2.5 V (rms). (The intervening filter stages enforce the limitation to the 0.1 – 100 kHz band, and they provide a gain very near 1.00 within that band.) We'll see later a good way to measure the rms value of signals such as shown in Fig. 1.1d, but you can now see why those voltage excursions fall (mostly) in the ±5-V range.

If your signals differ dramatically from those shown here, something is amiss. (See Appendix A.6 for some suggestions about 'troubleshooting'.) It's certainly possible for the signal you see to be smaller, say if you've made wrong connections or wrong settings. It's also possible for the signal to be 'too large', particularly if there are unwanted (interference) signals present. (Appendix A.5 discusses interference, its possible sources, and cures.) But the apparatus you're using, in the configuration you've set up, ought to be displaying a noise almost wholly due to nothing else than the Johnson noise of your source resistors. It is the universality of Johnson noise that lets us be sure that your signals should match, in rms measure, those shown here, certainly to within a factor smaller than two!

1.2 Quantifying Johnson noise

If you've done section 1.1, you've seen a rapidly-fluctuating signal on an oscilloscope which we claim is due mostly to Johnson noise, and which you now want to quantify. The method we'll describe here executes quite directly, in analog electronics, the very operation built into the mean-square definition of noise. You need one more cable to convey the filtered-and-amplified noise signal to the Multiplier module, configured as a 'squarer' as shown in the Fig.1.2a. Conduct the noise signal to the 'A' input, and choose the AxA on the toggle switch. The multiplier circuit delivers at the MONITOR point, a real-time output voltage

$$V_{out}(t) = [V_{in}(t)]^2 / (10\ V)\ ,$$

which still has dimensions Volts (due to the fixed 'scale factor' of 10 Volts in the denominator above). Take a look at $V_{out}(t)$ on your 'scope, and notice that it is always positive, *un*like your input noise signal $V_{in}(t)$, which is as often negative as positive.

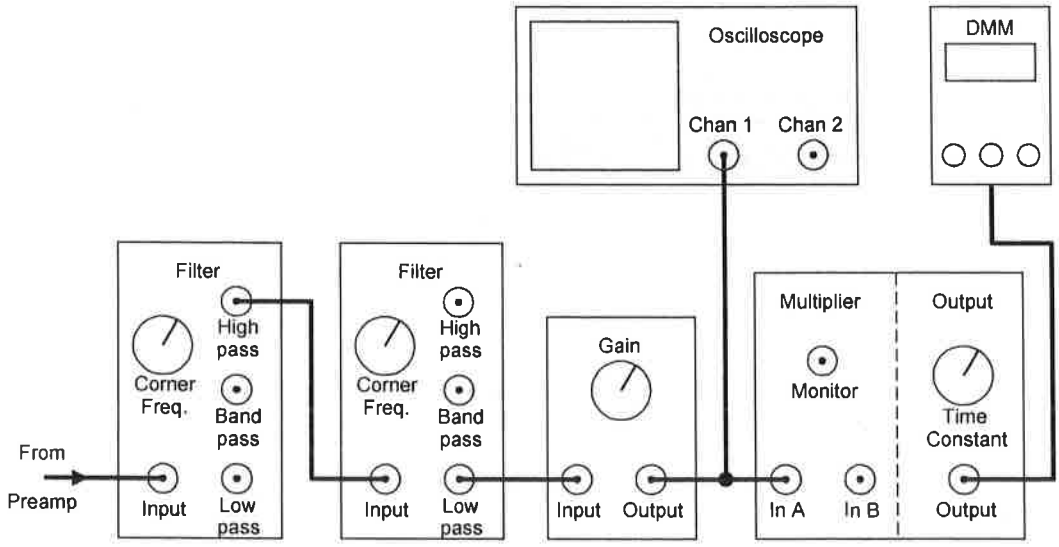


Fig. 1.2a: Cabling diagram for using the multiplier as squarer. High-pass filter 0.1 kHz, a.c. coupling; Low-pass filter 100kHz, a.c .coupling; Gain 400, a.c. coupling ; multiplier AxA, a.c. coupling

In fact, to persuade yourself that the squarer is working, use the XY-display capability on your 'scope. Convey the squarer's input $V_{in}(t)$, both to the squarer and to the X-channel of your 'scope, and convey $V_{out}(t)$ to the Y-channel, and have a look at a real-time XY-display. You should see a parabola emerge. See to it that you understand the origin of your XY-coordinate system, and then try changing some things: What are the right sensitivities to choose on the two axes? What would happen to your parabola if you raised the gain in the main-amplifier module of the HLE? Why does your data lie on a parabola, after all?

Now with*out* the need for a further cable, the output of the squarer is already being sent internally to the Meter module of your HLE. What this module does is to take the time-average of $V_{out}(t)$, averaged over a time interval you can select (by switch) to 1.0 second. This time average will *not* be zero, since $V_{out}(t)$, though fluctuating, is always and only on the *positive* side of zero. (Recall that the multiplier's squaring function ensures that $V_{out}(t)$ is proportional to the *square* of $V_{in}(t)$.) The meter will display that time-average, either on its 0-10 V or its 0-2 V scale. We suggest the use of the 0-2 V scale, and also suggest you go back and change the main-amp gain until the meter reaches a value near mid-scale, about 1 Volt on the 0-2 V scale.

What can you infer from this? Start with $V_J(t)$, the actual instantaneous Johnson-noise voltage generated by the source resistor. At the output of the pre-amp, you have a signal:

$$(6.00)(100.) \cdot V_J(t).$$

After the filter stages, you have the 0.1-100 kHz bandwidth-selected, or filtered, part of this signal. After the main amp, you have a signal

$$G_2 \cdot (600) \cdot V_J(t),$$

where G_2 is the main-amp gain, perhaps 300. Then after the squarer, you have a signal

$$[(300) \cdot (600) \cdot V_J(t)]^2 / (10\ \text{V}) .$$

Finally, using the <...> brackets to indicate a time average, what you have displayed on your meter is the signal

$$V_{meter} = \langle V_J^2(t) \rangle \cdot (600 \cdot 300)^2/(10\ \text{V}) .$$

From this result and the meter reading, you can work all the way backwards to find $\langle V_J^2(t) \rangle$, the mean-square voltage present (within your chosen bandwidth) across the source resistor.

Now use a cable to carry this time-averaged positive voltage to a digital multimeter. You should see a number consistent with your analog-meter indication, and you should see it fluctuate. (The expected size, and speed, of the fluctuations are treated in Appendix A.12.) Note that with the use of a 1-second time constant, you'll have to wait rather *longer* than one second for results to stabilize to any new value, especially if you're waiting for the 3rd or 4th digit of a multimeter display to settle down. Once the reading *has* settled, you'll notice the residual fluctuations, but go ahead and write down multiple readings from the multimeter, taking a new reading every second or so. See if you can persuade yourself that the readings display fluctuations about a mean value, and compute that mean value. It is connected, by a known chain of amplification and filtering, to the mean-square Johnson-noise voltage at the source.

1.3 Observing and Correcting for Amplifier Noise

You've now seen how all-analog electronics can take you all the way from a Johnson-noise source voltage $V_J(t)$ to a time-averaged d.c. voltage which is a traceable measure of $\langle V_J^2(t)\rangle$. This section teaches you how to

a) make that measurement optimally, and
b) correct that measurement for amplifier noise.

a) The noise measurements you perform all depend on the linear operation of the amplifiers, and they (like all analog electronics) have only a ***finite range of output voltages over which they remain linear***. For the high level electronic amplifiers, that range is (-10 V, +10 V). If you were to put a simple sinusoid through the amplifiers, you could use the full ±10-V excursions . But since you are amplifying *noise*, you have to ensure that even the rare large fluctuations of the noise stay within the ±10-V 'span' of the amplifier. In practice, a maximum ***average noise signal*** of 3 Volts (rms) is a safe choice. This should avoid serious distortion of the signal, called 'clipping', like that shown in Figure 1.3a. For an average noise signal of 3 Volts rms, an excursion beyond ±10-V is so rare as not to spoil the accuracy of your measurement.

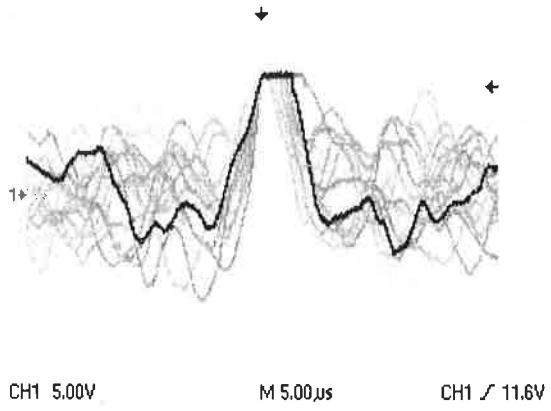


Figure 1.3a: Clipped signal from HLE – notice the clipping level is near +12 Volts.

Now if the rms measure of the signal at the A-input of the squarer, $V_A(t)$, is 3 V, then (by definition) its mean-square value is

$$\langle V_A^2(t)\rangle = (3\text{ V})^2 = 9\text{ V}^2\,,$$

and under these circumstances, the squarer's MONITOR output will give

$$V_{sq}(t) = [V_A(t)]^2/(10\text{ V})$$

so that the time-average at the OUTPUT will be

$$\langle V_{sq}(t)\rangle = \langle V_A^2(t)\rangle/(10\text{ V}) = (9\text{ V}^2)/(10\text{ V}) = 0.9\text{ V}\,.$$

You could use a smaller rms size for the input $V_A(t)$, but you'd be getting an even smaller output from the squarer, and your readings might be affected by zero-offsets in the squarer's output. (See Section 5.3 for details.)

So from here onwards, whenever you measure a noise voltage, you should check the main-amp output to see that it fits easily into the ±10-V range. If it exceeds these limits, reduce the gain. And you should look at the squarer's output on the panel meter, to see a time-averaged output near, or a bit below, 1 Volt. Again, if it's much larger, you want to reduce the gain, or if much smaller, raise the gain. Whenever you do take a reading of the time-average of the squarer's output, be sure to record also the net gain you've used to attain that reading, since this is your ticket to tracing the meter reading back to the desired mean-square noise $\langle V_J^2(t)\rangle$.

b) Now back to Johnson noise. The problem you're now going to address is tracing noise back to a source, because here you have to consider the possibility that some of the noise you're seeing is *not* due to the Johnson noise of the of source resistor, but instead due to the amplifier chain which follows it. Since this 'amplifier noise' is just as featureless and random as the resistor's Johnson noise, there's apparently no way to separate the two waveforms *once they're added.* But there *is* a way to separate their effects, if we can assume that the amplifier noise does not depend on the source resistor's value. Here's the demonstration: let $V_J(t)$ be the instantaneous noise voltage from the source resistor, and let $V_N(t)$ be the instantaneous noise voltage apparently present at the input of the amplifier. That is to say, $V_N(t)$ is a model for a noise emf which, applied to the input of an ideal *noiseless* amplifier, would match the noise actually observed at the output of the real amplifier, driven only by its internal noise. If the gain of the amplifier is G, its output will be

$$V_{out}(t) = G\,[V_J(t) + V_N(t)]\,,$$

and the mean-square of this output will be

$$\langle V_{out}^2(t)\rangle = G^2 \langle [V_J(t) + V_N(t)]^2$$

$$= G^2\{\langle V_J^2(t)\rangle + 2\,\langle V_J(t)\cdot V_N(t)\rangle + \langle V_N^2(t)\rangle\}\,.$$

There's a 'cross term' in this expression, the time average of the product $V_J(t)\cdot V_N(t)$, but this time average is <u>zero</u>. The reason is that $V_J(t)$ and $V_N(t)$ can be safely assumed to be **<u>un</u>correlated**, arising as they do from distinct physical mechanisms in two different objects. So when $V_J(t)$ happens to be positive, the amplifier noise $V_N(t)$ is just as likely to be negative as it is positive; thus the product of the two factors is also as likely to be negative as positive. That's why the absence of correlation enforces a zero for the time-average of the product. But that fact leaves

$$\langle V_{out}^2(t)\rangle = G^2\{\langle V_J^2(t)\rangle + 0 + \langle V_N^2(t)\rangle\}\,,$$

which says that ***mean-square voltages from uncorrelated sources are simply additive.*** In particular, it gives us a way to measure the amplifier noise -- we just change temporarily to a configuration in which the Johnson-noise term in this sum is negligible. Theory says that a choice of $R = 0$ for source resistance would give $\langle V_J^2(t)\rangle = 0$, but in practice, it suffices to use the R = 1-Ω or 10-Ω settings for giving a $\langle V_J^2(t)\rangle$ which is small enough that the result is a good measure of the amplifier noise, $\langle V_N^2(t)\rangle$.

And once that latter value is measured, ***it can be assumed to be present, and unchanged, in any use of (the same configuration of) the amplifier.***[1] So for any source resistor R_{in} > 10 Ω, the amplifier noise contribution previously established can be subtracted off, leaving $\langle V_J^2(t)\rangle$ isolated by itself.

Here's a concrete illustration: we have the values R_{in} = 1 Ω, 10 Ω, etc. We pick the 0.1 - 100 kHz bandwidth as before, and we pick gains to give good results at the squarer. In a particular example, the time-averaged outputs of the squarer we find are the $\langle V_{sq}\rangle$ values below:

R_{in} chosen	gain G_2 (HLE)	$\langle V_{sq}\rangle$ read	$\langle V_J^2 + V_N^2\rangle$ inferred	$\langle V_J^2\rangle$ derived
1 Ω	1500	0.6353 V	7.843×10^{-12} V^2	$\approx 0.002 \times 10^{-12}$ V^2
10 Ω	1500	0.6372	7.867	0.026
100 Ω	1500	0.6516	8.044	0.203
1 kΩ	1500	0.7911	9.767	1.926
10 kΩ	1000	0.9801	27.225	19.384

Now we expect, for the time-averaged output of the squarer,

$$\langle V_{sq}(t)\rangle = \langle V_{in}^2(t)\rangle / (10\text{ V})$$
$$= \{(G_1 G_2)^2 / (10\text{ V})\}\ \langle V_J^2 + V_N^2\rangle ,$$

so we can use the G_1 = 600 and G_2-as-listed values to compute the column with $\langle V_J^2 + V_N^2\rangle$ values. We can eyeball-extrapolate to the $R_{in} \to 0$ limit, and deduce a contribution of 7.841×10^{-12} V^2 for $\langle V_N^2\rangle$ alone, the amplifier noise contribution (for this particular amplifier chip, at this particular bandwidth -- your number will vary!). Subtracting this contribution from all the entries gives the rightmost column for $\langle V_J^2\rangle$, our estimate of the mean-square Johnson noise of the source resistor, corrected for the effects of amplifier noise. Notice that the amplifier-noise corrections are large, even dominant, for small values of source resistance! You'll find (for the present choice of pre-amp input stage) that Johnson noise surpasses amplifier noise only when the source resistance has risen to about 3 kΩ.

[1] Under the assumption of negligible op-amp current noise, and no noise from external interference, both of which may depend on R_{in}

1.4 Johnson noise dependence on resistance

The previous sections have taught you how to configure the pre-amp/filter/main-amp combination, and how to select a gain for optimal use of the squarer. The results can also be corrected for amplifier noise, and traced back to an inferred mean-square measure of Johnson noise, $\langle V_J^2(t) \rangle$, for any source resistor from $R = 10\ \Omega$ upwards.

You should now investigate systematically the dependence of $\langle V_J^2(t) \rangle$ upon source resistance R. To do so, you can use the $R = 10\ \Omega$ through 10 MΩ choices built into the pre-amp module. (These internal source resistors have tolerances of 0.1% to 1 MΩ, and 1% thereafter.) But the selector switch also gives you access to three more test positions, A_{ext}, B_{ext}, and C_{ext}, which you are free to 'populate' with devices of your choice behind the pre-amp panel.

Here's how to do so: You can 'flip' the pre-amp panel, as illustrated in Figure 7.2a, to expose the back (component) side of the pre-amp's circuit board. You can also find the pre-amp power switch (near the internal power-on red LED inside the low-level electronics), and **turn OFF the pre-amp power** before making any changes to the board. Now use the diagram below to find the screw-connect terminal strips, and find also the location of the two endpoints for the components you're putting into the A, B, and C positions.

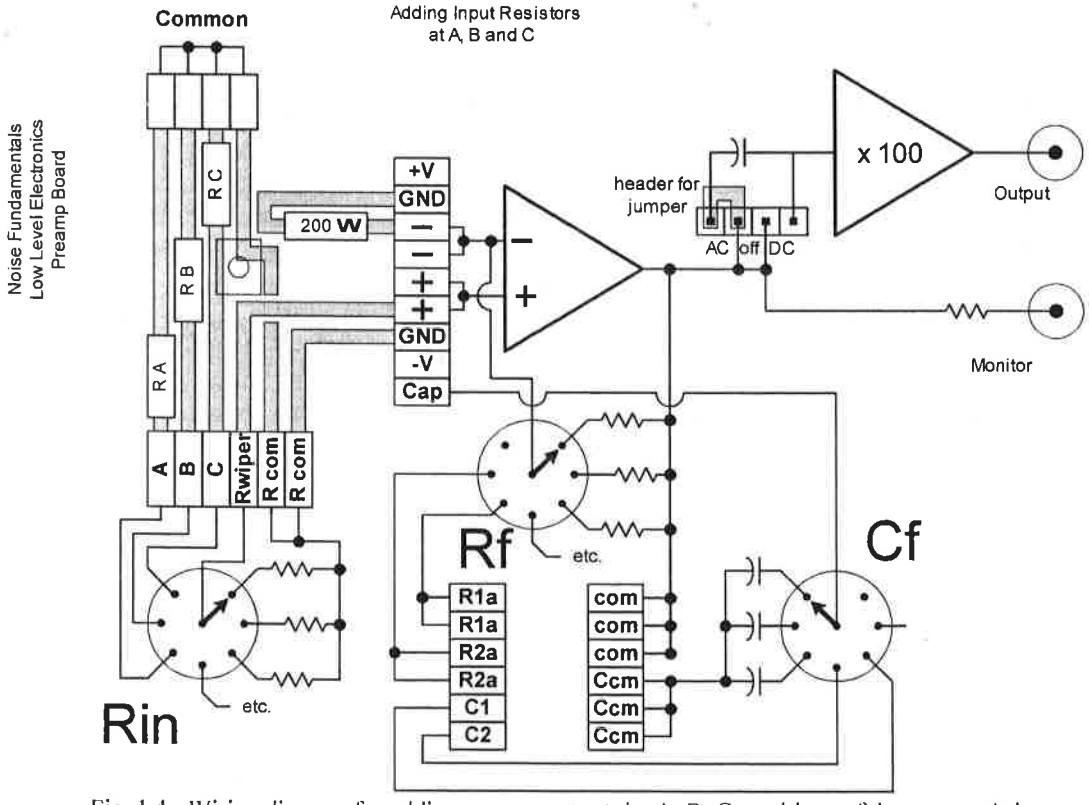


Fig. 1.4: Wiring diagram for adding components at the A, B, C, positions of the pre-amp's input. Note all input resistors have a common ground.

You can choose resistors of any value in the 20 Ω to 5 MΩ range; you can even choose different kinds of resistors. (Most resistors sold nowadays are of metal-film construction, but ask around for some carbon-composition or wirewound resistors -- and look up what kinds of resistors Johnson himself used.) You can clearly use resistors of any power capability you like -- their internal Johnson-noise emf is *not* going to overheat them! If you wish, you can have a comrade *hide* from you the resistance values, so you'll be measuring some actual unknowns. Don't forget to turn the pre-amp power back ON before you re-flip the front panel and close up the box.

Now you can take noise data for your own resistors, as well as for the built-in source resistors. Once you have values for $\langle V_J^2(t) \rangle$, each corrected for amplifier noise, you can plot those values as a function of R. Since both axes will vary over many orders of magnitude, a log-log plot is appropriate. The vertical axis has units of Volts-squared, the horizontal axis has units of Ohms. Nyquist's theory predicts a first-power power-law dependence on resistance R, namely

$$\langle V_J^2(t) \rangle = (4\, k_B\, T\, \Delta f) \cdot R^1 ,$$

and you might see this confirmed. There will be deviations from this behavior at the high-R end of the plot, for reasons to be discussed in sections 1.5 and 2.2, and Appendix A.8.

At the *low*-resistance end of the plot, you'll see the amplifier-noise-corrected values enable you to follow Johnson noise to a regime *well* below the apparent limit set by amplifier noise. You'll be able to establish values of $\langle V_J^2(t) \rangle$ which are less than 1% of the amplifier noise $\langle V_N^2(t) \rangle$ that overlays them. Of course, the corrected value of Johnson noise will be the difference between two nearly equal quantities, so the results will be subject to larger uncertainties than other data points.

1.5 Johnson noise dependence on bandwidth

Thus far you've learned how to observe and quantify Johnson noise, and you've seen how to isolate its mean-square value from amplifier noise. You've also seen its dependence on source resistance R. But Nyquist's formula claims that $\langle V_J^2(t)\rangle$ also depends on the bandwidth Δf; ie. on the range of frequencies to which your system is sensitive. [2]

So for now you should stay at room temperature, and stay at a fixed R-value; we suggest a starting value of R_{in} = 10 kΩ. The goal is to see how the choice of bandwidth matters. The method is to imagine a 'white noise spectrum', ie. noise power uniformly spread in frequency at its origin, but subsequently modified by the high-pass and low-pass filter sections as depicted below.[3]

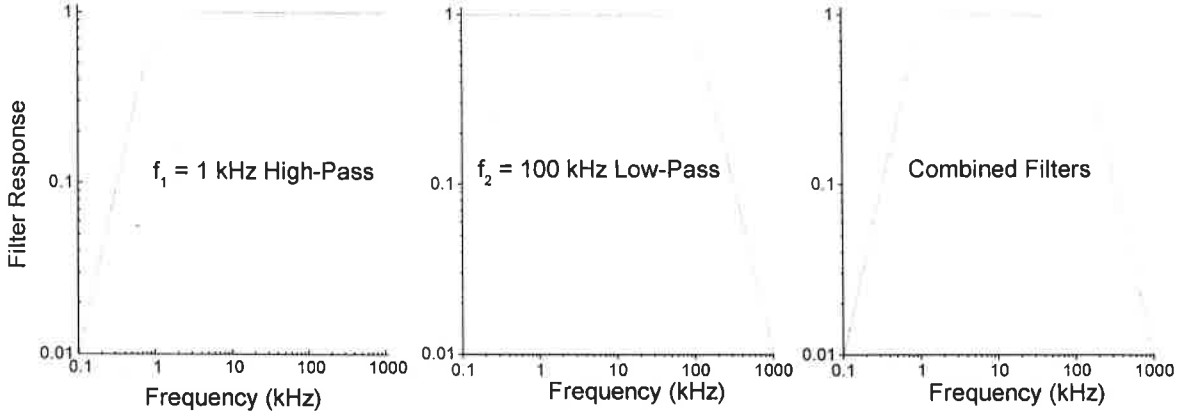


Fig. 1.5a: Representation (left) of the transmission of a high-pass filter, of corner frequency f_1; (center) of a low-pass filter, of corner frequency f_2; (right) the combined effect of both filters. The graph's scales, horizontal and vertical, are all logarithmic.

You have a range of choices for the 'lower corner' frequency f_1 or high-pass filter setting, and a separate range of choices for the 'upper corner' frequency f_2 or low-pass filter setting. You might first think that the bandwidth Δf should be given by $|f_2 - f_1|$, which is a decent approximation, but subject to significant corrections. These so-called 'corrections' are discussed in great detail in Sections 2 and 5. But for now we present you with the generic corrections which are the result of a model calculation. The model of Section 2.2 predicts the effective bandwidth Δf for each combination of f_1 and f_2, and gives the results shown in Table 1.5.

[2] The further prediction that Johnson noise depends on the resistor's *temperature* is tested in Chapter 4.

[3] Section 2.0 teaches you how to get data of this form.

Table 1.5 Effective noise bandwidths, Δf, given in Hertz, computed for model filter responses

	f_2 = 0.33 kHz	1 kHz	3.3 kHz	10 kHz	33 kHz	100 kHz
f_1 = 10 Hz	355	1,100	3,654	11,096	36,643	111,061
30 Hz	333	1,077	3,632	11,074	36,620	111,039
100 Hz	258	1,000	3,554	10,996	36,543	110,961
300 Hz	105	784	3,332	10,774	36,321	110,739
1000 Hz	9	278	2,576	9,997	35,543	109,961
3000 Hz	0.4	28	1,051	7,839	33,324	107,740

These computed values are all subject to uncertainties of order 4%; (see Section 5.2 for details on how any of them can be more carefully calibrated). They are all computed (by the methods of Section 2.2) for ideal filter responses, ignoring systematic effects. Inclusion of those effects may raise values in the rightmost column by (3±1)%, and may raise values in the next-to-rightmost column by (1±1)%. There are further corrections to effective noise bandwidths for large f_2-values, in the case of large source resistance, due to capacitive effects -- see Appendix A.8.

Your goal is to measure the mean-square Johnson noise of the resistor, $\langle V_J^2(t)\rangle$, for as many (f_1, f_2) combinations as you wish. Recall that for each choice of filter settings, you'll want to adjust the gain so as to use the squarer optimally. Recall that each mean-square value you measure needs to be corrected for amplifier noise (measured at *that* bandwidth setting: the amplifier-noise contribution to the mean-square depends, as does the Johnson-noise contribution, on the bandwidth you use.)

You can plot your data for $\langle V_J^2(t)\rangle$ in various ways:
as a function of the f_1-value used to obtain it;
as a function of the f_2-value used to obtain it;
as a function of the difference $|f_2 - f_1|$ of the f_1- and f_2-values used to obtain it; or
as a function of the equivalent noise bandwidth, from the table above.
Which plot is the most nearly linear? Try again using log-log scales, to be able to see all you data points, spread as they are over many orders of magnitude.

If your plot is consistent with $\langle V_J^2(t)\rangle \propto \Delta f$, then the coefficient of this proportionality tells you a 'noise power spectral density', as you'll see in the next Section. Its units are V^2/Hz, and it's usually denoted by S.

1.6 Johnson noise density, and Boltzmann's constant

Previous sections have shown you how to measure noise, and have tested its dependence on source resistance R and on measurement bandwidth Δf. This section introduces you to noise *density*, and then relates your measured values, via Nyquist's formula, to Boltzmann's constant.

If you have shown that measured mean-square noise $\langle V_J^2(t)\rangle$ has a linear dependence on the bandwidth Δf used, you are entitled to infer the existence of a 'noise density' that's uniform in frequency. Here's an analogy to mass density that should make this clear -- we'll use a one-dimensional example. Suppose you have a string, of unknown composition, laid out on an x-axis, and that you can make clean cuts at arbitrary locations x_1 and x_2, and then weigh the piece of string you've extracted. If (and only if) you find that the observed mass M is always proportional to $|x_2 - x_1|$, you may conclude the string is of uniform density. You can also see that the quotient

$$(\text{mass } M) \,/\, |x_2 - x_1|$$

gives the value for this density, given in units of mass per unit length.

Similarly, if you've shown that mean-square noise $\langle V_J^2(t)\rangle$ is always proportional to the bandwidth Δf you used to obtain it, then you can define the 'noise power density'

$$\langle V_J^2(t)\rangle \;/\; \Delta f\,,$$

in this case with units of Volts-squared per Hertz, or V^2/Hz. [Strictly speaking, this is not a power density -- but if a voltage $V(t)$ is applied across a resistance R then the quotient $V^2(t)/R$ is a power. So the quotient above is just a factor-of-R away from being an actual power density, with units Watts per Hertz.]

Your data for a single source resistance R = 10 kΩ has given you a noise power density; you can go back to your data of Section 1.4 and convert that data to noise power density as well, to check the dependence-on-R of this density. The motivation for all of this is that Nyquist's formula can be written as

$$\text{noise density } S = \langle V_J^2(t)\rangle \,/\, \Delta f = 4\, k_B\, T\, R\,.$$

So you should plot all of your data thus far, and perhaps more data that you now take for various R- and Δf-values, to see if you can further establish the linear-in-R claim of the prediction above. (In practice, you'll see deviations in the regime where R and/or Δf is large, for reasons discussed in Appendix A.8.)

If you establish a regime of linear dependence on R, your plot, or fit, will give you a value for a slope, $(4\, k_B\, T)$. What *units* should it have? (Answer: rise over run, so V^2/Hz per Ohm -- and what unit is *that*?) What *value* does it have? Hardest: what *uncertainty* can you assign to your value? (Do so before you look up any 'book values', because the uncertainty intrinsic to your experiment is conceptually a matter quite separate from any discrepancy between your value and anyone else's.)

Finally, if you know you room's temperature T (and express it in absolute, ie. Kelvin, units), you can now conclude by finding a value (and uncertainty!) for Boltzmann's constant k_B.

What's the nature of k_B? At one level, it connects historical choices of temperature units to a 'common language' of energy units, via $E = k_B T$. At another level, k_B is a 'microscopic' version of the macroscopic gas constant R (here, not a resistance), as you can see by writing the ideal-gas law in two ways,

$$p V = n R T \qquad \text{and} \qquad p V = N k_B T .$$

The first form has n = (number of moles of gas), and that gives to R the units of Joules per (mole·Kelvin). The second form has N = (number of molecules of gas), and it gives to k_B the units of Joules per (molecule·Kelvin), or just J/K. This double form of the law also makes it clear how R and k_B have to be related: since $(n R)$ and $(N k_B)$ both give pV/T, we have

$$n R = N k_B , \qquad \text{or} \qquad R = (N / n) k_B .$$

But (N/n) is Avogadro's number N_A, the number of molecules per mole. Hence you expect the numbers to obey the relation

$$R \approx 8.31 \text{ J/mole·K} = N_A \cdot k_B \approx (6.02 \times 10^{23} \text{ /mole}) (1.38 \times 10^{-23} \text{ J/K}) .$$

Check that claim. Does this mean that electrons inside a resistor are acting like molecules in a gas, bouncing around between the resistor's two ends? And is the Johnson-noise emf akin to the pressure fluctuations which kinetic theory predicts for a gas? What's the connection to Brownian motion? See if you can find any guidance on these conceptual points.

2. Noise Density

2.0 Setting up to see a bandwidth

One of the new features of noise measurements (compared to other measurements you've previously made) is that the signal measured depends on the bandwidth. You've seen in Section 1.5 that the amount of Johnson-noise signal (in the 'mean square' sense) depends on the choices made for the difference between f_1 and f_2, the high-pass and low-pass corner frequencies chosen in the electronic filtering applied to the raw noise waveform. Here are some exercises that temporarily set aside noise, and concentrate on the depiction of bandwidth. These require a signal generator for sinusoids, capable of about 1-Volt output, covering the 1 Hz to 100 kHz range, to serve as a signal source, and a 2-channel digital oscilloscope. The goal is to set up the filter system and drive it with sinusoids, to get and to graph data showing you the gain-vs.-frequency profile $G(f)$.

A cabling diagram for the system is shown in Figure 2.0a. It requires only the filter sections of the high-level electronics.

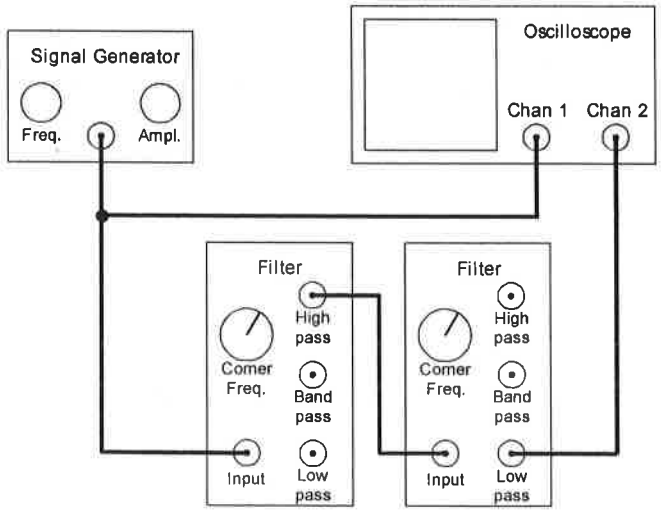


Fig. 2.0a: A cabling diagram for testing the action of combined filters on sinusoids.

Note that the generator's output signal, with some amplitude A and frequency f, is applied both to the 'scope's ch. 1 input and to the input of a first filter stage. Note that the left-hand filter section is used as a high-pass filter (with corner frequency set to some low frequency f_1), and that the right-hand filter section is used as a low-pass filter (with its corner frequency set to some higher value f_2). For a first experiment, we suggest setting f_1 = 1 kHz and f_2 = 10 kHz. Note that the output of the cascaded filters is sent to the 'scope's ch. 2.

If possible, you should configure your 'scope to give the rms measure of the filter's input signal via its ch. 1, and also to give the rms measure of the filter's output via its ch. 2.

Finally, for any chosen value f of generator frequency, you can define the empirical measure of the gain of the filter assembly by

$$G(f) = [\text{rms measure of output signal}] / [\text{rms measure of input signal}] .$$

In practice, for the TeachSpin filter design, you'll find the 'gain' might actually be a loss – you'll find $G(f)$-values less than or equal to one for most frequencies. But it's the frequency dependence of $G(f)$ that you want to study.

Notice that for each choice of f, you need to pick a sensible time base for your 'scope – for best results, choose a horizontal-axis scale which permits several cycles of the sinusoid to be displayed. You will also need to pick a sensible triggering option – why is triggering on ch. 1 the better choice?

Notice that for each choice of frequency and amplitude, you should pick sensible vertical sensitivities for your 'scope as well – for best results, choose scales such that both sinusoids you observe will fill more than half, but less than the whole, of the vertical-axis display range.

Notice that for a generic choice of frequency, the output signal will not be in phase with the input. That is to say, the times of occurrence of peaks (or of valleys, or of zero-crossings) of input and output need not coincide. Any real-time filtering system will have such phase shifts of output relative to input. But these phase shifts are happily *not* relevant to noise measurements, and not part of the $G(f)$ definition above.

It is also possible to measure the gain using a good digital multimeter to perform these rms measurements. Typically the specifications of true-rms a.c. voltmeters extend to 300 kHz, but you'll need to check the limits of your own meter.

Exercise 1. Measure $G(f)$ as a function of f for filter settings of f_1 = 1 kHz and f_2 = 10 kHz. You will want to cover at least the range 0.1 to 100 kHz. It is not necessary to take points at equal spacing in frequency, as in the list (0.1, 0.2, 0.3, . . . 99.8, 99.9, 100.0) kHz! Instead, you want a coverage that'll be roughly uniform on a logarithmic scale, and you should aim for 2, 3, or perhaps 5 points per 'decade' in frequency space. Whatever list of frequencies you pick, measure $G(f)$ at each frequency in the list. Now plot $G(f)$ as a function of f, using logarithmic scales for both axes.

For at least one frequency, try changing the amplitude of the input signal, to see if the output amplitude changes in proportion. This is (one) test of the linearity property of the filter.

In your plot, identify the 'low-block' or high-pass corner near f_1, above which frequency, sinusoids are passed through the first section of the filter. Also identify the 'high-block' or low-pass corner near f_2, below which frequency, sinusoids also pass the second section of the filter. Also identify the 'pass band' as the region in which the combined filter assembly gives $G(f) \approx 1$. Compare with Fig. 1.5a.

Exercise 2. Measure the $G(f)$ for the single band-pass output from one of the filter sections. The cabling diagram is shown in Figure 2.0b. You should observe that the gain at the center

frequency is less than 1.0 (closer to 0.707), and that the wings of the filter response drop off more slowly than for the low-pass or high-pass filters.

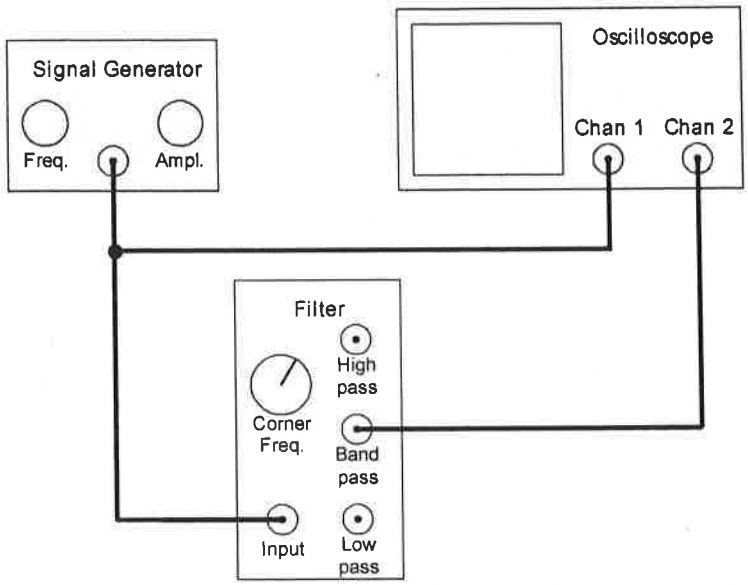


Figure 2.0b Cabling diagram for a single (band-pass) filter measurement

Exercise 3. For one of your (f_1, f_2) choices, plot the squared function $G^2(f)$ as a function of f, this time keeping both scales linear. You'll see the motivation in Section 2.2. Since the 'area' under this $G^2(f)$ curve is vitally important, you'll see that to draw the curve through the data points you've gotten will require either more data points to 'fill it in', or else a best-fit to a model that can provide the equivalent interpolation of the points you have taken.

All these exercises have used (only) sinusoids as test signals to be injected into your filter assembly, so what's their relevance to noise waveforms which are nothing like sinusoidal? The answer comes from the linearity property, which you've already tested (for single-frequency sinusoids) in Exercise 1 above. But systems which are linear (as your filters are, within limits) display another linearity property, in that:

their response to a sum-of-inputs
is equal to
the sum of their responses to the inputs taken individually.

Now it's time for you to exercise a Fourier imagination, and to think of a noise waveform as being made up out of (or, as analyzable back into) a whole collection of sinusoids, of all frequencies ranging from $f << 1$ Hz to $f >> 1$ MHz. The deep part of the linearity property is now to understand the operation of the filter as

filter's output (when driven by noise)
= filter's output (when driven by a sum of sinusoids)
= sum-of-(filter's output when driven by individual sinusoids).[1]

[1] In Appendix A.11 we discuss the technique by which you can observe the filters' response to noise signals using a digital oscilloscope which has FFT capabilities.

That is to say, the filter's effect on a sum-of-inputs is the same as the sum of the (filter's effect on individual inputs). And the filter's response to individual sinusoids is just what's described by the $G(f)$ function you've already measured. Now you understand better that raw noise is composed of all frequencies, but that filtered noise has had its frequency content well below f_1, and frequency content well above f_2, ***suppressed***. That is to say, filtered noised has had its frequency content in the range $f_1 < f < f_2$ passed along with gain about 1, but frequencies outside the 'pass band' suppressed.

If the 'edges' of the filter's response curve were perfectly sharp-edged corners, you can see that your filter would have bitten out, from the entire frequency spectrum of noise, that portion lying in the range from f_1 to f_2. In the sharp-cornered-filter limit, that would define a pass-band of width $f_2 - f_1$. We'd call that a 'bandwidth' $\Delta f = f_2 - f_1$. (Notice that a filter's pass-band *width* is a different matter than the pass-band's *center location*.) In Section 2.2 you'll see how the 'equivalent noise bandwidth', still labeled as Δf, can be defined when the actual $G(f)$-function has rounded, rather than sharp, corners.

2.1 Summary of how to measure a noise density

Recall the procedure we've found which takes us from a tiny, fluctuating, zero-on-average noise signal $V_n(t)$ to a nearly-steady d.c. value which is traceably related to $\langle V_n^2(t) \rangle$:

- we pre-amplify $V_n(t)$ by 'gain' or amplification factor G_1, to give $G_1\ V_n(t)$;
- we filter-in-frequency, to isolate a band of frequencies (which get gain factor 1) from frequencies we reject (these get gain factor 0);
- we further amplify, by gain G_2, to a level convenient to use in the next stage; within our band of frequency width Δf, we have amplified signal $G_2 \cdot 1 \cdot G_1 \cdot V_n(t)$;
- then we square (with a scale factor of 10 V) to give the squarer's output

$$V_{sq}(t) = [G_2 \cdot 1 \cdot G_1 \cdot V_n(t)]^2 / (10\ \text{V}) ;$$

- finally, we time-average, over an chosen interval of time, to give a result

$$V_{meter} = \langle V_{sq}(t) \rangle = \langle V_n^2(t) \rangle \cdot [G_1\ G_2]^2 / (10\ \text{V}) .$$

It follows that the mean-square noise at the pre-amp's input can be recovered from known quantities:

$$\langle V_n^2(t) \rangle = (10\ \text{V}) \cdot V_{meter} / (G_1\ G_2)^2 ,$$

and this has the proper units of Volts-squared.

We have also shown, in cases where the noise is 'white noise', ie. has a uniform distribution in frequency, that the quotient $\langle V_n^2(t) \rangle / \Delta f$ gives the value for the 'power spectral density' S of the noise, with units of V^2/Hz. Here Δf is the bandwidth, the effective range in frequency space to which the noise is restricted before it's quantified.

There are several fine points worthy of attention for this method to succeed.

1. Clearly, it's crucial that every amplifier in the whole chain of amplification stay within its linear range. This can be checked at the MONITOR and OUTPUT test points of the pre-amp, and at every stage of the high-level electronics as well. Typical use of the amplifier chain will use a.c.-coupling at each opportunity, to insure that an initially-small d.c. offset is not amplified so much as to drive some subsequent signal up against the ±12-V 'rails' of the modules' outputs. Even with the use of a.c.-coupling, clipping could still occur if too much gain is used.

2. It's also important that the squarer be used in the 'good part' of its range -- so as to produce a time-averaged output lying in the 0.6 - 1.2 V range. The reasons are laid out in Section 1.3.

3. The accuracy of the inferred noise density depends on knowing the bandwidth Δf. Some claimed values of Δf are listed in a Table in Section 1.5, and the method for computing those values is in the next Section. The precision of the Δf-values is of order 5%, and can be improved via the calibration procedures of Section 5.2. These corrections are most important with the use of filter bandwidths extending to 100 kHz.

4. The accuracy of noise density measurements further depends on knowing the pre- and main-amp gains factors G_1 and G_2. These ratios are ultimately established by resistor values and by operational-amplifier (op-amp) performance. All the gain-critical resistors are of 0.1% precision, so the overall gain can be trusted at the 1% level. At higher frequencies, the limitations of op-amp performance start to create larger uncertainties, which are addressed in Section 5.1.

5. The noise-density measurements depend on the precision of the squarer. The scale factor (of 10 V) is laser-trimmed by the manufacturer to 0.4% worst-case, 0.1% typical tolerance, and can be checked by the methods of Section 5.3. That section also addresses the equally important issues of d.c. offsets in the squarer's input and output. To get a hint of the existence and importance of these offsets, you can use the GND/AC/DC-coupling switch at the squarer's input, and see what output emerges when the input is forced to the 'ground' or 0-V condition.

Finally, there is the implicit claim that the quotient we've called 'noise density' is the physically-important result worthy of measurement. This is not *a priori* obvious, but it is true that Johnson noise (in Chapter 1 and 4) and shot noise (in Chapter 3) both are described theoretically in terms of a noise power density, $< V^2 > / \Delta f$. And clearly, there's no expectation that instantaneous noise values $V(t)$ could be predicted; nor is there any hope of quantifying a 'maximum value' to the noise. Neither is there any chance of predicting how much 'noise signal' there is right 'at' any given frequency, since (nearly by definition) noise is a waveform lacking any concentration or location in frequency space. It is true that (thus far) the noise at its source has been assumed to be of uniform distribution in frequency, but this assumption can be checked, by methods including those of Section 2.3.

2.2 How the 'equivalent noise bandwidth' is defined

The previous section followed noise signals $V_n(t)$ to the squarer by way of a 'brick wall' model of the filtering process, in which the filter sections together gave

$$\text{gain factor} = 1 \text{ for } f_1 < f < f_2 \text{ , but gain} = 0 \text{ elsewhere.}$$

Under this assumption, the filter bandwidth Δf is clearly given by $\Delta f = |f_2 - f_1|$. ***But real filters do not have such sharp-edged characteristics. TeachSpin filters are optimized for predictability of performance, rather than sharpness of edge.***

Let's examine what happens in case of a *general* filter gain function $G(f)$. For any Fourier component of a noise signal at frequency f_a, the signal reaching the squarer's input will then be

$$V(t) = G_2\, G(f_a)\, G_1\, A_a \cos(2\pi f_a\, t - \phi_a) ,$$

where A_a is the amplitude, and ϕ_a is the phase, of the original noise-signal's component. If two or more components are present, and if the amplifiers and filters are all linear (in the mathematical sense, of obeying the superposition principle), then a typical input to the squarer is

$$V(t) = G_2\, G(f_a)\, G_1\, A_a \cos(2\pi f_a\, t - \phi_a) + G_2\, G(f_b)\, G_1\, A_b \cos(2\pi f_b\, t - \phi_b) + \ldots$$

The squarer will then produce an output containing two kinds of terms. One kind is the squares of individual terms in the sum above, such as

$$[G_2\, G(f_a)\, G_1\, A_a]^2 \cos^2(2\pi f_a\, t - \phi_a) ,$$

and in all such terms, the cosine-squared function will time-average to a non-zero value (of 1/2, in fact). (Notice the emergence of the quantity $G^2(f_a)$. This squared-function will end up in the equivalent noise bandwidth.) By contrast, the other sort of terms that emerge are cross terms, of character

$$G_2^2\, G(f_a)\, G(f_b)\, G_1^2\, A_a\, A_b \cos(2\pi f_a\, t - \phi_a) \cos(2\pi f_b\, t - \phi_b) .$$

By a trigonometric identity, the product of two cosines can be replaced by a sum of two new cosine functions, having frequencies $(f_a + f_b)$ and $|f_a - f_b|$. But both of these cosine terms *vanish* upon taking the time average, for any $f_a \neq f_b$, so we can *drop* them from the result.

Hence it is correct to assume that 'each frequency component acts independently', not only in all the linear stages up to the squarer, but also in the time average of the output of the squarer. So we can now see that the previous result

$$\langle V_{sq}(t) \rangle = \langle [G_2\, G_1\, V_n(t)\,]^2 \rangle / (10\ \text{V}) = \langle V_n^2(t) \rangle\, (G_2\, G_1)^2 / (10\ \text{V})$$

can be replaced, in case of a filter-function of frequency-dependent gain $G(f)$, with

$$\langle V_{sq}(t) \rangle = \langle V_n^2(t) \rangle\, [G(f)]^2\, (G_2\, G_1)^2 / (10\ \text{V})$$

for any single frequency component's contribution.

Now suppose that the noise is distributed by a power-spectral-density function $S(f)$, with units V^2/Hz, such that

$$\int_{f_1}^{f_2} S(f)\, df$$

gives the mean-square value $\langle V_n^2(t)\rangle$ measured for frequency components lying in the range $f_1 < f < f_2$. Hence $S(f)\, df$ gives that part of $\langle V_n^2(t)\rangle$ attributable to the range in frequency space $(f, f+df)$. (The absence of any time-averaged effect of the cross terms generated by the squarer is essential for this argument to make any sense.)

But with that assumption, each such frequency interval gives a partial contribution to the squarer's total result, in amount

$$S(f)\, df\, [G(f)]^2\, (G_2\, G_1)^2/(10\ \mathrm{V})\ ,$$

and summing over all source frequencies gives us the total of the time-averaged output of the squarer:

$$\langle V_{sq} \rangle = \frac{(G_2\, G_1)^2}{10\ V} \int_0^\infty G^2(f)\, S(f)\, df \quad .$$

The previous cases have assumed white noise: $S(f)$ = a constant S, giving the uniform spectral density. We'll continue to make that assumption, which allows us to write

$$\langle V_{sq} \rangle = \frac{(G_2\, G_1)^2}{10\ V}\, S \int_0^\infty G^2(f)\, df \quad .$$

The previous assumption of a brick-wall filter-function made the integration trivial: in such a case, $G(f) = 1$ only within a range of width Δf, so the integral gave $1^2 \cdot \Delta f$ for the noise band width. But now we see that even in cases where $G(f)$ varies continuously in frequency, the 'equivalent noise bandwidth' is given by

$$ENBW = \int_0^\infty G^2(f)\, df \quad ,$$

and we'll continue to denote this result as Δf. (Notice that this result only is 'equivalent' if the noise can be assumed to be white. Notice too that the gain $G(f)$, and its square, are dimensionless, but df is not, so the integration gives the result Δf a value whose units are Hertz.) So with this definition of noise bandwidth Δf, and the assumption of noise that's white at its source, we have as time-averaged output of the squarer the result

$$\langle V_{sq} \rangle = S\, \Delta f\, (G_2\, G_1)^2\, /(10\ \mathrm{V})\ ,$$

which of course allows us to infer the noise power spectral density S in terms of measured quantities,

$$S = (10\ \mathrm{V})\, \langle V_{sq} \rangle\, /\, [(G_2\, G_1)^2\, \Delta f]\ .$$

What's new here is that we now have a way to compute the bandwidth Δf as soon as we have the filter-transmission function $G(f)$.

It's worth tracking back through this whole argument to see why the gain, but *not* the phase shift, of any of the amplifier or filter section of the measurement system, is what matters for purposes of noise-density measurement.

Finally, we're ready to compute one example of an equivalent noise bandwidth. Suppose we use only one filter section, a single low-pass filter, set to a 'corner frequency' f_c. The filters built into the TeachSpin modules are two-pole state-variable filters, whose low-pass response is given to a very good approximation by the 'Butterworth response'

$$G(f) = [1 + (f/f_c)^4]^{-1/2} .$$

Insofar as this model is correct, we get by an analytic or numerical integration

$$\Delta f = [\pi/(2\sqrt{2})]\, f_c = 1.1108\, f_c$$

The following figure shows this example's $G^2(f)$ function on a linear-in-frequency scale; the area under the curve shows where the contributions to the integral defining Δf arise. On the same plot, in dotted lines, is the curve for the 'equivalent brick-wall filter', which achieves the same equivalent bandwidth by having its hypothetical sharp corner lying 11.08% higher in frequency than the rounded corner of the actual filter response.

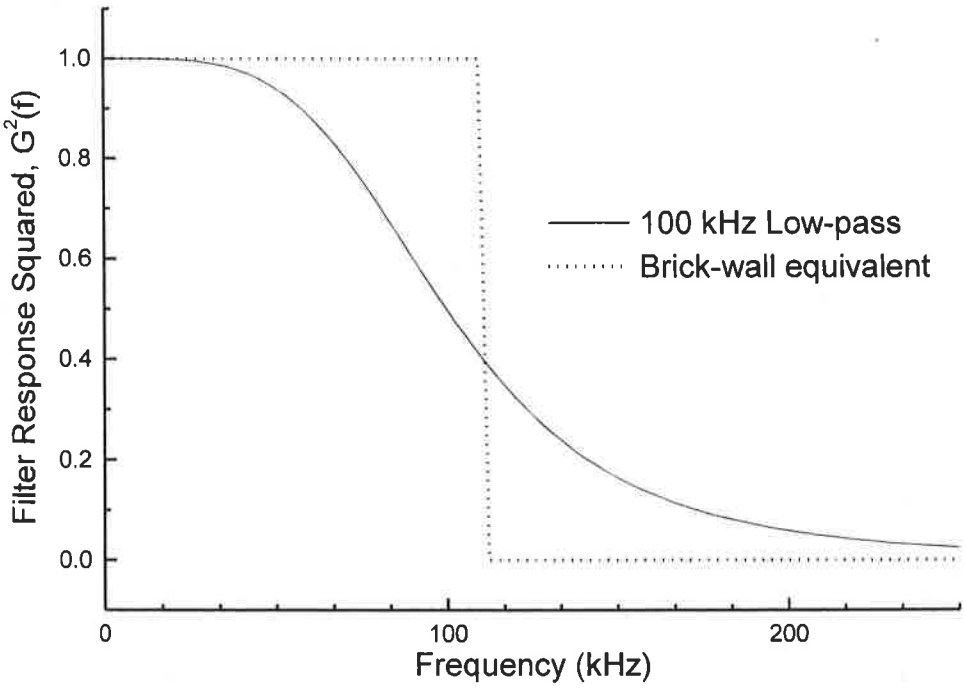


Fig. 2.2a: The square of the gain, or $G^2(f)$, function for an ideal Butterworth-response 2-pole low-pass filter (solid line), and the brick-wall filter response of the same equivalent bandwidth (dashed line).

Note that the Butterworth response curve of the filter contains a long 'tail' extending far above the nominal corner frequency. This can also be seen in a log-log display of the same $G^2(f)$ function (but note that such a plot mis-represents the idea of equal areas in the plot standing for equal contributions to the integral):

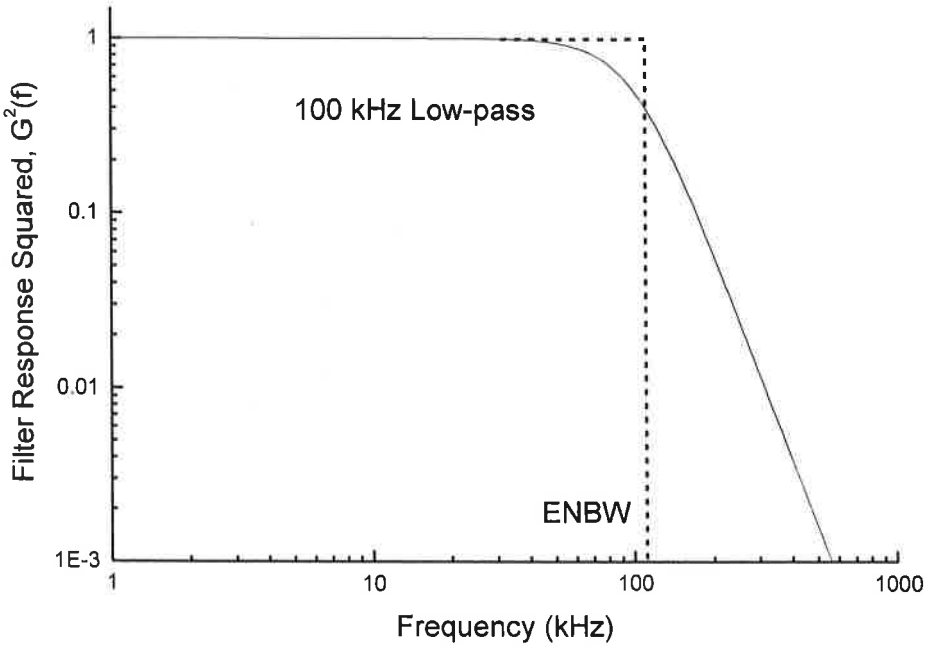


Fig. 2.2b: A log-log display of $G^2(f)$ for the same filter.

Now let's be concrete about the implications of this 'tail' for the case of a low-pass filter chosen with **corner frequency f_c = 100 kHz.** The equivalent noise bandwidth is 111.08 kHz. This might create the illusion that the performance of the pre-amp and main-amplifier at (say) 200 kHz is irrelevant, but now that's seen to be an error. Instead, we can do the integrals

$$\int_0^{f_c} G^2(f)\,df \quad , \int_{f_c}^{2f_c} G^2(f)\,df \quad , \int_{2f_c}^{4f_c} G^2(f)\,df \quad , \quad etc.,$$

to learn the relative contributions:

white noise in band dc to 100 kHz gives 86.697 kHz, which is 78.06% of the total ,
white noise in band 100 to 200 kHz gives 20.315 kHz, which is 18.29% of the total,
white noise in band 200 to 400 kHz gives 3.539 kHz, which is 3.19% of the total,
white noise in band 400 to 800 kHz gives 0.454 kHz, which is 0.41% of the total;

and these together account for 99.94% of the total. But if such a filter were to be driven by a pre-amp whose gain was the proper G_1 all the way from dc to 400 kHz, but fell to 0 above 400 kHz, we'd have lost 0.47% of the total that should appear after the filter. The lesson is that with this sort of filter, and with the wish to achieve errors under 1% in accounting for the total noise power, the spectrum of the noise has to be faithfully amplified to frequencies lying up to 4, or better 8, times that of the highest nominal corner frequency. Since the highest corner frequency selectable is 100 kHz, this is the reason that the pre-amp, main amp, and squarer have been designed to offer good performance all the way to 1 MHz.

The methods above also allow the computation of the equivalent noise bandwidth, or *ENBW*, of *combinations* of filters. The state-variable filters used in the TeachSpin design, set to corner frequency f_c, offer responses quite accurately modeled by

$$\begin{aligned} &\text{low-pass:} && G_{LP}(f) = [1 + (f/f_c)^4\,]^{-1/2} \\ &\text{band-pass:} && G_{BP}(f) = (f/f_c)\,[1 + (f/f_c)^4\,]^{-1/2} \\ &\text{high-pass:} && G_{HP}(f) = (f/f_c)^2\,[1 + (f/f_c)^4\,]^{-1/2}\ . \end{aligned}$$

So the combined use of a high-pass filter set to (smallish) frequency f_1, and a low-pass filter set to (largish) frequency f_2, give a net response close to

$$G(f) = \{\,(f/f_1)^2\,[1 + (f/f_1)^4]^{-1/2}\,\} \cdot \{\,[1 + (f/f_2)^4]^{-1/2}\,\}\ ,$$

and the (numerical) computation of the integral of $G^2(f)$ for this function, for two given f_1, f_2 values, is used to generate the entries of the Table in Section 1.5.

Naturally, the same methods can be used to compute the *ENBW* Δf, for any combination of filters.

Note that all of these computations have assumed ideal Butterworth high- and low-pass filters, with corners at f_1 and f_2, as the only frequency-selective effects in the entire amplifier chain. In practice, there are additional effects:

a) The amplifier gains, modeled heretofore as constants G_1 and G_2, in fact will drop off with frequency, starting in the vicinity of 1 MHz;

b) Capacitive effects at the input of the pre-amp create an additional one-pole low-pass effect of the form

$$G_{\text{one-pole}}(f) = [1 + (f/f_c)^2\,]^{-1/2}\ ,$$

where $f_c = (2\pi\, R_{in}\, C)^{-1}$. This roll-off of response has the biggest effects for large source resistance R_{in}, and for large input capacitance. For 'local' resistors the C-value is of order 10 pF, but $C \approx 100$ pF for the 'remote' resistors in the Temperature Probe -- see Chapter 4. For a source resistance $R_{in} = 100$ kΩ, the product $R_{in}\, C$ is then $(10^5\ \Omega)(10^{-11}\ \text{F}) = 10^{-6}$ s, and the resulting corner is located near $f_c = 160$ kHz. This is *not* large enough, compared to a choice $f_2 = 100$ kHz or even 33 kHz, to preserve full 1% accuracy. See Appendix A.8 for more details on modeling this effect.

c) Capacitance in parallel with the pre-amp's feedback resistor may also reduce the bandwidth. This can be observed when measuring shot noise at low currents (when using a feedback resistor greater than or equal to 1 MΩ).

d) The pre-amp configuration can also change the pre-amp bandwidth. Raising the gain of the preamp (perhaps by increasing the feedback resistor R_f) will in general reduce the bandwidth of the pre-amp.

e) The filter functions $G(f)$ suffer from finite gain-bandwidth product of the op-amps used to build them, with effects (especially at the largest bandwidths) that are mentioned at Table 1.5.

2.3 Bandpass filters and 'spot' noise-density measurements

Every result thus far has assumed that the noise source has an original spectral distribution which is uniform: the noise is assumed to be white, as 'white light' would deliver equal amounts of power into any two frequency intervals of equal width in frequency. (Note that actual sunlight is far from 'white' by this criterion!) In the previous section we've written $S(f)$ = a constant S, allowing the relevant integral to be simplified via

$$\int_0^\infty G^2(f)\, S(f)\, df \rightarrow S \int_0^\infty G^2(f)\, df = S\, \Delta f \quad .$$

There are ways of testing whether $S(f)$ is in fact a constant. The method of Section 1.5 is to use bandwidths Δf which are computed from f_2 and f_1, with f_2 and f_1 well separated, and then to see if a variety of choices of f_1 and f_2 entail the same value of S. This section offers an optional route to a more 'localized' test of $S(f)$.

The ultimate test would be to pick some small value for Δf ($<< f_0$), and to take as filter-function $G(f)$ the narrow and sharp-edged function having

$$G(f) = 1, \qquad \text{for } f_0 - \Delta f/2 < f < f_0 + \Delta f/2\,, \qquad \text{but } G(f) = 0 \text{ elsewhere}\,.$$

This would give $\int G^2(f)\, df = 1^2 \cdot \Delta f$ as required, and it would clearly isolate a narrow band of frequencies around f_0, and hence it would only be sensitive to values of $S(f)$ with $f \approx f_0$. For small values of Δf, it would give an adequate approximation to measuring the 'spot value' $S(f_0)$.

An approximation to this idea of narrow-band filtering can be achieved by using the band-pass filters in one (or two) sections of the high-level electronics. These offer a gain function very near to

$$G(f) = (f/f_c)\, [1 + (f/f_c)^4]^{-1/2}\,,$$

where you may think of f_c as the nominal 'corner frequency' or 'center frequency' -- plot this function to see why that's a good name. You'll see that $G(f)$ is peaked near $f = f_c$, though you'll also note the filter-function's peak value is not 1 but near $1/\sqrt{2} \approx 0.707$.

Note that this $G(f)$ function defines an equivalent noise bandwidth, given by the integral

$$\Delta f = \int_0^\infty G^2(f)\, df = \int_0^\infty \frac{(f/f_c)^2}{1+(f/f_c)^4}\, df = f_c \frac{\pi}{2\sqrt{2}} \approx 1.1108\, f_c \quad .$$

A log-log plot of the integrand $G^2(f)$ is traditional, and reveals the way $S(f)$ is going to be 'sampled' chiefly in the (rather broad) region around $f \approx f_c$. But since you can choose f_c-values ranging from 10 Hz to 100 kHz, you can make spot-checks of $S(f)$'s regional values in frequency.

There are two modules for filtering, and you could use either one, to get center frequencies 10 Hz, 30 Hz, . . . 3 kHz, or to get center frequencies 330 Hz, 1 kHz, . . . 100 kHz. Note that these distinct choices also give distinct Δf-values -- there is no pretence that you are sliding a filter of fixed bandwidth across the frequency spectrum. If you use each of these band-centers to

measure a 'local noise density', you'll have found a good way to test, or even to falsify, the hypothesis that a given noise source is white.

As an example of this technique, you might look at the 'amplifier noise' -- practically speaking, the noise which emerges from the R_{in} = 1-Ω or 10-Ω choice in Johnson noise. In section 1.3 you'll have seen that with so small a source resistor, the noise power (in the sense of voltage-squared) is >99.9% attributable to the noise generated inside the pre-amplifier. There is no necessary theoretical reason for *this* sort of noise to be 'white'. The empirical approach is to use the band-pass technique above to get a whole series of 'spot checks' on $S(f)$. You'll find the measurements are easiest for the choice of large values of f_c; as you go to smaller values of f_c, you are also getting smaller values of Δf, so less noise comes through, and you'll need higher main-amp gain.

Hint #1: At bandwidths below about 1 to 3 kHz, you will find that there is not enough gain in the high level electronics to get your signal up to the volt level. You can still use the squarer at these low levels, but you need to a bit more carful and measure the small d.c. offset voltage from, the squarer. (See Section 5.3.) Another option would be to increase the gain of the preamp. What would be the gain of the preamp in Figure 1.1a if the feedback resistor were changed to 10 kΩ?

Hint #2: When looking at center frequencies below 100 Hz you will want to use d.c.-coupling in the pre-amp and all the filter and amplifier sections. (See Appendix A.2 for the methods for doing so.) The reason is that a.c.-coupling is a way to kill response at d.c., but it also suppresses response below 'corners' at 16 Hz . (The different modules have different a.c. coupling high-pass corner frequencies. You can look at the specifications listed in Appendix A.1.) You *will* probably want to retain a.c.-coupling at the input of the squarer; at the cost of losing some noise power below a 1-Hz corner, you will be free of any d.c. component of the signal getting squared and giving a non-zero error in your mean-square result.

Hint #3: When using the lowest values of center frequency f_c, your *ENBW* gets very small. In addition to requiring the use of higher main-amp gain to compensate, you'll see that this results in much larger fluctuations at the time-averaged output of your squarer. (Appendix A.12 shows why.) The cure is first to use the longest (τ = 3 s) choice of averaging time, and then to take a series of fresh meter readings (say, every 3 or 5 s) and average those readings after-the-fact, until you get a mean value of adequately small statistical uncertainty.

Hint #4: (optional) For a few values of f_c, you can 'stack' or cascade two band-pass filters in series. Make some $G(f)$-plots to show that this can give you better selectivity in frequency. You'll need to do the necessary integrals to compute the new *ENBW*. If you can create a narrower, as well as your original wider, version of a band-pass filter, and still get consistent values of S at the same center-location in frequency space, you'll have more reason to be confident in your methods of measurement.

The motivation for pursuing this check of pre-amp 'noise power spectral density' to low frequencies is that you can thereby enter the glamorous world of 'excess low-frequency noise' or '1/*f* noise'. This is a physical phenomenon amazingly widespread in all sorts of time series emerging from a host of different phenomena, but no fundamental physical reason for its generality is yet known. It is also of considerable technical interest, providing as it does a motivation to use lock-in or other signal-averaging techniques that evade the excess low-frequency noise which so many systems display. You might even subtract off the constant white noise contribution from your low frequency data and see if the excess noise does behave as f^{-1}. (See the technical data included with your instrument for some evidence of such excess low-frequency noise.)

Now back to spot checks of $S(f)$, but for another kind of noise: you might take the same kind of data for *Johnson* noise from a 10-kΩ resistor. For each filter-function chosen, you can measure the result of R_{in} = 10 kΩ (and you already have the result for R_{in} = 1 Ω), and correct it for amplifier noise. What's left is just the Johnson noise, and theory *does* predict that it will be flat-in-frequency. What do the *data* say?

It's interesting that the ENBW for the band-pass filter is the same as for the low-pass filter with the same corner frequency. You might also check this prediction.

3. Shot noise

3.0 The reason for shot noise, and its predicted size

As Johnson noise is due to thermal fluctuations, so *shot noise* is due to charge quantization. It's quite remarkable that certain macroscopic electric currents will display noise, ie. random fluctuations about their average d.c. value, and that this noise is directly attributable to the microscopic 'graininess' of electric current. While shot noise might be a nuisance, or a limiting factor in certain measurements, its existence makes possible the determination of the magnitude of the quantum of charge, just by measuring noise. So shot noise provides a route to a tabletop measurement of the value of '*e*', the fundamental charge.

Here's some reasoning about why there *might be* fluctuations about the mean value of a current. Let's think about a d.c. current of average value i_{dc}, created by a flow of electrons in a wire. In a time interval τ, the (average) amount of charge arriving is

$$Q = i_{dc}\,\tau \ ,$$

and because that charge arrives in charge packets, each of size (-)*e*, the number of electrons arriving is

$$n = Q/e = i_{dc}\,\tau/e \ .$$

The next step in the argument depends entirely on physical characteristics of the source of the electron current. If the electrons were arriving with perfect regularity, then in any time interval of duration τ, you'd expect only a ± 1-count uncertainty in the number n.[1] At the opposite extreme from regular, periodic, and predictable arrivals, think of an electron current consisting of the statistically-independent arrivals of entirely *uncorrelated* electrons. In such a case, the average number of electrons arriving in time τ would be n, but the actual number on any particular trial would be subject to the same $n \pm \sqrt{n}$ 'counting statistics' that you'd get in radioactive decay or any other Poisson process.

It is *not* obvious which limiting case you'll achieve with actual electric currents! In fact, there are simple room-temperature ways to generate electric currents that show 'full shot noise', but there are also ways to produce currents which are well below the 'shot-noise limit'. (In fact, if you deliver electrons in bursts, you can get *above*-shot-noise fluctuations -- think about the ns-long, 10^6-electron pulses arriving at the anode of a photomultiplier tube.)

Here are the consequences of those statistical fluctuations: Suppose we have in a time τ the arrival of n electrons on average, delivering an average charge of $Q = i_{dc}\,\tau$. Now there will be fluctuations in charge about this mean, with standard deviation

$$\sigma_Q = e\sqrt{n} = e\,(i_{dc}\,\tau/e)^{1/2} \ .$$

Since current is charge per unit time, the instantaneous current $i(t)$ will also show statistical fluctuations about its mean i_{dc}, with standard deviation $\sigma_i = \sigma_Q/\tau = (i_{dc}\,e/\tau)^{1/2}$.

---

[1] Look up the topic 'electron turnstile', an example of state-of-the-art cryogenic electronics, to encounter a device that actually produces such a current.

Since that standard deviation is defined as $< [i(t) - i_{dc}]^2 >^{1/2}$, we see that the deviation from the average, $\delta i(t) \equiv i(t) - i_{dc}$, has a mean-square value given by

$$< [\delta i(t)]^2 > = i_{dc}\, e/\tau \ ,$$

which is, in fact, a result for the mean-square noise in the current $i(t)$. That we can measure! The only complication in this simple derivation is to relate the effective measurement time τ to the bandwidth Δf of the measurement system, as previously defined. A low-pass filter of corner frequency f_c, and a bandwidth approximately given by $\Delta f = f_c - 0$, defines a time-scale of $(2\pi f_c)^{-1}$, which is approximately the time scale on which statistically-independent readings can be made. So we might expect a result near

$$<\delta i^2(t)> \approx i_{dc}\, e/\tau = i_{dc}\, e\, (2\pi f_c)^{+1} \ .$$

In fact, the correct result has exactly of this character, but with different numerical constants; Schottky's prediction for the mean-square noise in a current of uncorrelated electrons is

$$<\delta i^2(t)> = i_{dc}\, e\, (2\, \Delta f) = 2\, e\, i_{dc}\, \Delta f \, ,$$

where Δf is the same 'equivalent noise bandwidth' defined in Section 2.2.

A numerical example is worthwhile. Suppose we have a bandwidth of 100 kHz, and think of the output of such a filter as providing a 'fresh answer' every 10 μs. And suppose we have a 10 μA current, and have arranged for it to display full statistical fluctuations. That current delivers 10^{-5} Coulombs per second, so it delivers 10^{-10} C (on average) in each fresh 10-μs time interval. That amount of charge consists of a number of electrons,

$$n = (10^{-10}\ \text{C})/(1.6 \times 10^{-19}\ \text{C}) = 6.3 \times 10^{8}\ \text{electrons.}$$

The expected statistical fluctuations in that number are given by the square root of the count, or 2.5×10^4. Thus the fluctuations are 1 part in 25,000 of the total. So the average current of 10 μA = 10,000 nA is subject, under the conditions of this experiment, to fluctuations of order

$$(10{,}000\ \text{nA})/25{,}000 = 0.4\ \text{nA} \quad \text{in every 10-μs time interval.}$$

In fact, Schottky's formula predicts the fluctuations will have rms measure

$$\delta i_{rms} \equiv < [i(t) - i_{dc}]^2 >^{1/2} = (2\, e\, i_{dc}\, \Delta f)^{1/2}$$

$$= (2 \cdot 1.6 \times 10^{-19}\ \text{C} \cdot 10^{-5}\ \text{A} \cdot 10^{5}\ \text{Hz})^{1/2} = 5.7 \times 10^{-10}\ \text{A} = 0.57\ \text{nA.}$$

This answer depends on the value assumed for e.(!) Turning the calculation around, we can see that carefully measured results for average current i_{dc}, bandwidth Δf, and mean-square current fluctuations $< [i(t) - i_{dc}]^2 >$ can produce a value for the fundamental quantum of charge, e.

Notice in this example that the fluctuations in charge, which give the noise in the current, are small compared to the charge itself (rms fluctuations $\pm 2.5 \times 10^4\ e$ << charge $6.3 \times 10^8\ e$), but that the fluctuations are nevertheless large compared to the charge $1e$. So even though we use room-temperature electronics *lacking* the ability to register the electrons' arrival on a one-by-one basis, we still gain the possibility of quantifying fluctuations which depend on the size of e.

3.1 Operating a photodiode

In order to see 'shot noise', you need a current consisting of uncorrelated electrons, and we get this current from an illuminated **photodiode**. This section is not yet about noise, but it will show you a first way for mounting and illuminating a photodiode, and for reading out the current which the light generates. The successful assembly of this circuit, and the confirmation of the presence of an average photocurrent i_{dc}, is a pre-requisite for the subsequent measurement of shot noise. If you would like more contextual orientation to photodiodes, consult Sections 3.6 and 7.

A photodiode is a two-terminal device, and it has (like any other diode) connections to cathode and anode. ***For all the experiments you do, you'll need to identify and distinguish the leads connected to these electrodes of the photodiode.***

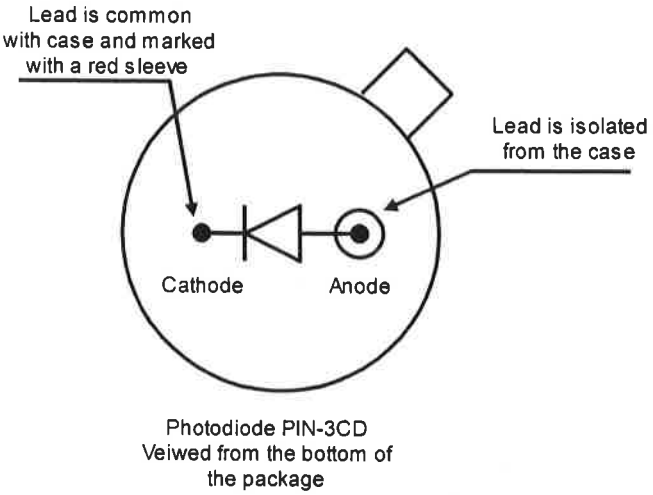


Fig. 3.1a: Identifying the leads of the PIN-3CD photodiode.

We'll adopt the circuit shown in Fig. 3.1b to illuminate a photodiode, and get a first look at the photocurrent it produces.

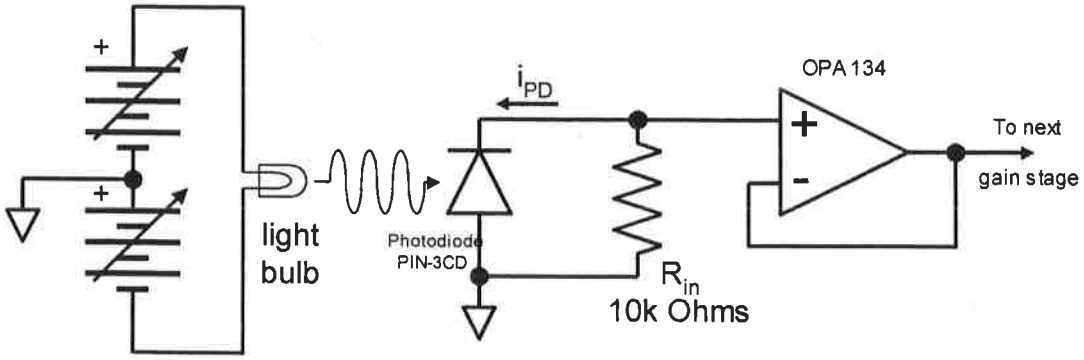


Fig. 3.1b: Schematic diagram for the connection of a photodiode and a load resistor to a pre-amp input stage configured as a voltage follower.

Here we indicate schematically the arrival of light onto the photodiode, and show the direction of the conventional photocurrent which results. (Note that the photocurrent flows *oppositely* to the 'forward direction' of conventional current through the diode.) In this circuit, the anode is held at ground potential, and the photocurrent passes through a 10-kΩ load resistor. The amplifier shown schematically is a voltage-follower configuration of the first stage of the pre-amplifier section of the low-level electronics (LLE) of your apparatus. The method for creating this circuit is shown in Fig. 3.2c. Note that the load resistor is created by using the R_{in} switch set to the 10-kΩ position, and the voltage-follower circuit by setting the R_f switch to the R_1 position (which, as shipped, is populated by a 0-Ω wire) or to the 10-Ω position (since that's low enough for these applications).

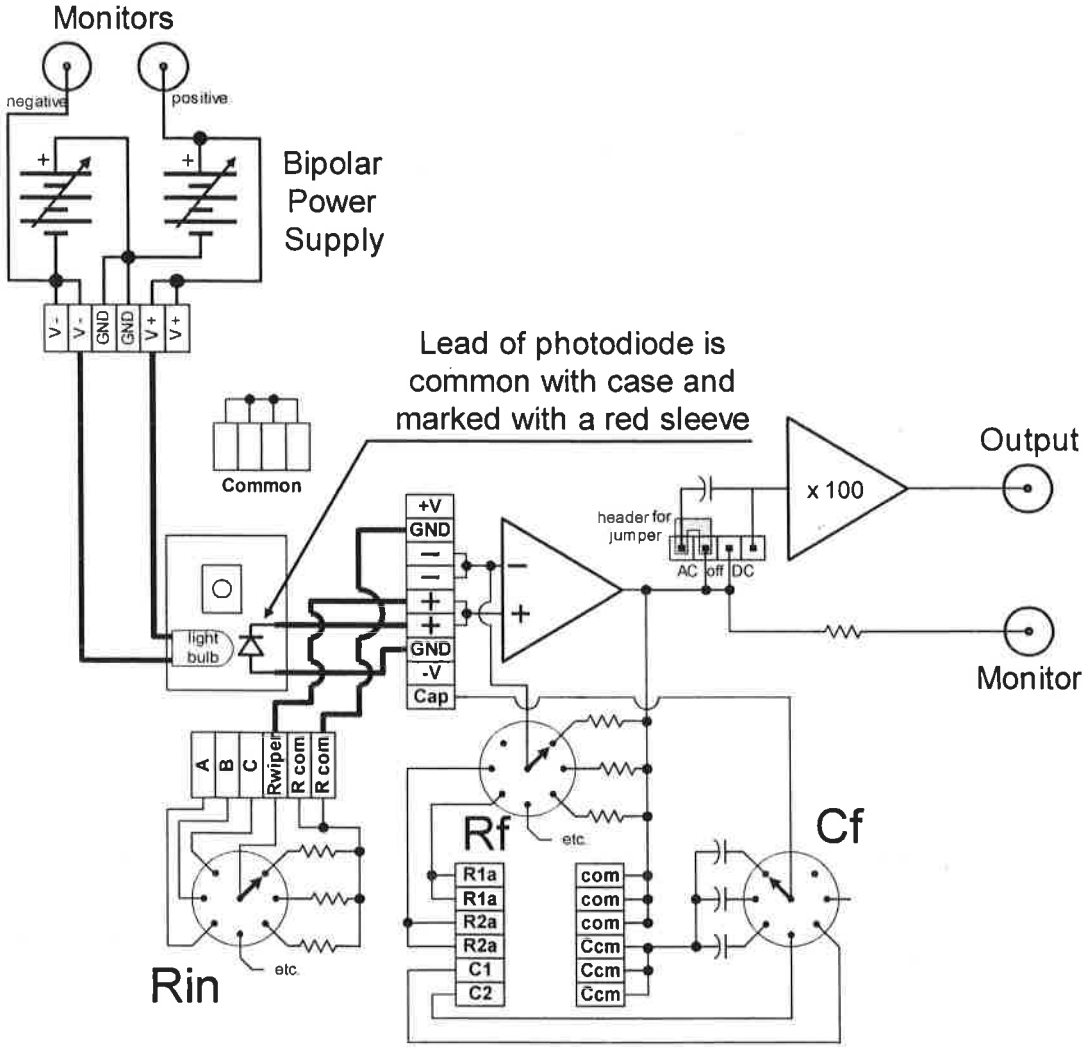


Fig. 3.1c: Wiring diagram for the lamp and photodiode combination, and the voltage-follower topology of the input stage of the pre-amp.

The photo in Fig. 3.1d shows how the photodiode is mounted in a black plastic block. This holds it stably, and ready to be illuminated by special light sources. This arrangement also holds the photodiode very close to the first-stage amplifier, which is important for minimizing the effects of wiring capacitance.

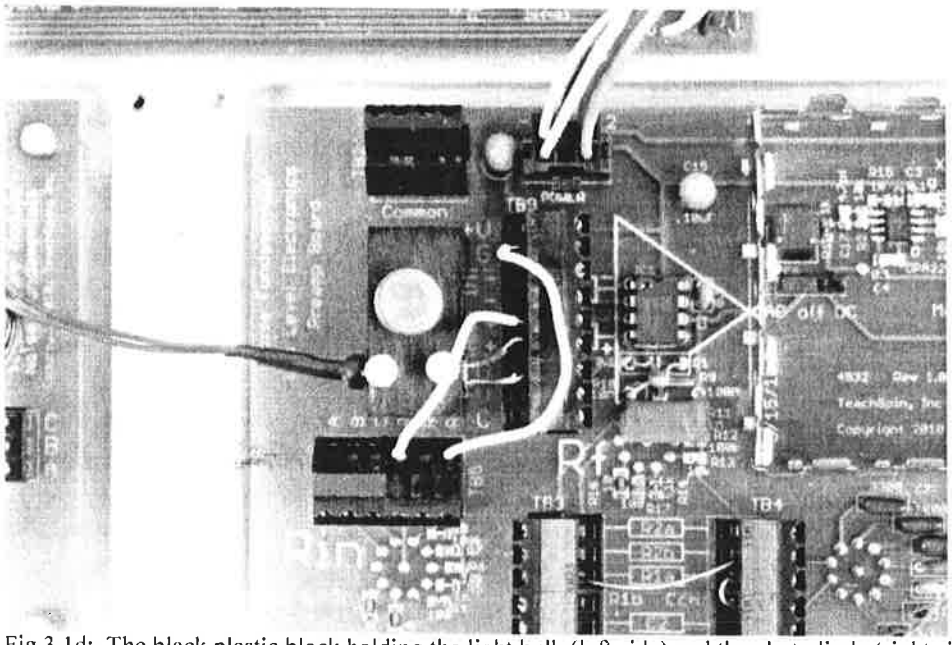


Fig 3.1d: The black plastic block holding the light bulb (left side) and the photodiode (right side), mounted near the first stage of the pre-amp.

Now suppose that we arrange enough illumination to create a photocurrent of 1 μA – it turns out that this would require about 2 μW of red light to reach the photocathode. If that photodiode current $i_{dc} = 10^{-6}$ A flows through the 10-kΩ load resistor, we expect a potential drop of $(10^{-6}\ \text{A})\ (10^{4}\ \Omega) = 10^{-2}\ \text{V} = 10\ \text{mV}$ across the resistor. We also expect the ungrounded end of this resistor to be at a negative 10-mV potential. Since a non-inverting amplifier's input is attached there, we expect illumination to create a negative, -10-mV, output to be measured at the pre-amp's MONITOR output.

To confirm the operation of this circuit, we suggest a first use of the tiny incandescent bulb supplied, which can also be installed in the black plastic block of Fig. 3.1d, to illuminate the photodiode.

Caution: Only finger-tighten, and do not over-tighten, the plastic screw holding the light bulb in place. It is easy to crack the glass of the bulb and destroy it.

To light the bulb, you can use the positive and negative outputs of the 0 - 11 V bipolar power supply built into the LLE, which will give you a variable potential difference of 0 - 22 V across the bulb. Wire that lamp as shown in Fig. 3.1c, and before installing it in the plastic block confirm that you can get it to light up, and that you can vary its brightness over a wide range. Now arrange for the bulb to be glowing dimly, and mount it into the black plastic block. Reassemble your LLE by flipping over and thumb-screwing down its front panel, and use a multimeter, at the MONITOR output of the pre-amp, to look for an output voltage. Check to see if you get a negative voltage, which grows *more* negative as you dial up the voltage you're supplying to the incandescent bulb.

There are other ways to illuminate the photodiode. As alternatives to the bulb, we have provided some light-emitting diodes (LEDs) in the parts bin. The LEDs are equipped with pre-soldered leads, and they too will fit into the black plastic block which holds the photodiode.

In using the bulb, polarity of connections didn't matter, and a 0 - 22-V potential difference as a voltage source would drive the bulb. By contrast, in using the LEDs,

- polarity *does* matter -- attach the red lead (of either LED) to a positive polarity.
- a *current-limited* source is recommended for driving the LED -- see Fig. 3.1e for one method of using the positive variable-voltage supply in your LLE, together with a series current-limiting resistor and the LED.

The circuit shown also provides, at a BNC output, an iR-drop which is a surrogate for the LED's current. For the series resistor, we suggest the use of $R = 1$ kΩ when using the red-emitting LED, and $R = 100$ Ω when using the infrared-emitting LED supplied.

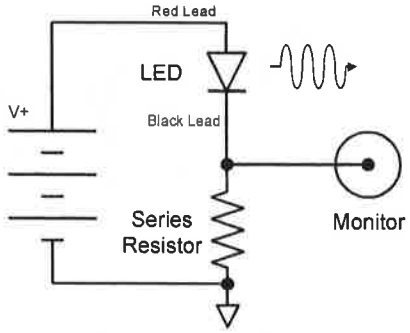


Fig. 3.1e: Schematic diagram for an LED, allowing voltage-drop monitoring of the LED current.

The electrical properties of the three light sources are summarized in a table:

Wavelength	Package	typical bias	Absolute max.	Pulsed	Photo-current	Mfg part #
Red LED ≈ 650 nm	clear plastic	20 mA	30 mA 50 mA with possible damage	100mA	0.25 mA	Avago- HLMP-EG08-Y200 Digikey-516-1377-ND
IR LED ≈ 930 nm	metal can with lens	100mA	100mA	2A for 0.1 us	1.2 mA	Optek OP233 Newark-08F2954
Light Bulb	glass bulb	24 Volts	28 Volts		0.4 mA	SPC Tech. -2185 Newark – 50N8119

Some notes on the light sources:

1) The IR (infrared) LED can tolerate the largest drive current.

2) A 100-mA current may ***NOT*** be safely passed through the 100-Ω, 1/4-Watt resistor provided by the Series-Resistor switch! To measure currents as large as this, mount a suitable resistor into the terminal blocks provided, and access it via the A_{ext} or B_{ext} position of the selector switch.[1]

3) The red LED, and the light bulb, can both be made bright enough to create >0.2 mA of photocurrent in the photodiode, while the IR LED may give you >1.0 mA of photocurrent. Any of these devices can be dimmed to reduce the photocurrent to 1 μA or even smaller values.

4) One further advantage of the LED sources is that they (*un*like the light bulb) can be turned on and off very rapidly (<< 1 μs). This is not directly needed for shot-noise studies, but one use of this on-off modulation capability is illustrated in Sec. 3.4.

[1] Early units of NF-1 (serial numbers 101 to ≈115) were shipped using ¼-W resistors at the Series-Resistor Switch, and owners of those units have been sent replacement 10-Ω and 100-Ω resistors of 3-Watt rating to be used in the A_{ext} and B_{ext} positions. Later units have had the 3-W resistors built in at the 10-Ω and 100-Ω positions of the Series-Resistor switch. If you look at the interior of your low-level electronics, you should be able to see the situation with your unit – by eyeball distinction of a resistor of 3-W rating from one of ¼-W rating.

3.2 First views of noise on a photocurrent

The previous section has explained how to confirm that a d.c. photocurrent from an illuminated photodiode has been generated. Now it's time to look for the fluctuations, ie. the noise, in such a photocurrent. Those fluctuations will display 'full shot noise' if the electrons flowing to constitute the photocurrent are uncorrelated, statistically-independent arrivals of charge.

The most persuasively independent electrons are photo-electrons, the more so if the light producing them is 'thermal' light. In your first shot-noise experiment, the light involved will be produced by an incandescent bulb, which has negligible spatial and temporal coherence. So it's appropriate to think of the bulb as shedding a rain of independent photons down onto any surface. (The word 'shot' in shot-noise is meant to remind you of the sound of pellets of birdshot, or perhaps raindrops, falling onto a sheet-metal roof.) When those photons fall onto a photo-diode, it is a fair picture to think of each photon absorbed in the p-n junction as producing an electron-hole pair. The internal electric field of the junction separates that pair, and drives the electron through an external circuit as part of an electric current. It's easy to think of the photons as statistically-independent (since they're independently produced, and thereafter non-interacting); what's not quite so obvious is that the photo-electrons thereby produced will create a current of still-statistically-independent electrons (given that electrons in a wire certainly *can* interact through their electric field with other electrons).

Use the same circuit as in Sec. 3.1, with the use of the incandescent bulb to illuminate the photodiode, adjusted so as to produce 10 μA (but not more, to avoid complications) of photocurrent. The evidence for this will be a d.c. voltage of $i_{dc}\, R_{in}$ = (-)(10 μA)(10 kΩ) = (-)10^{-1} V = (-)100 mV at the MONITOR point in the pre-amplifier of the LLE.

Now here's a calculation of the expected size of the shot-noise fluctuations in that current, and their detectable consequences. Recall that for i_{dc} = 10 μA, and for an equivalent noise bandwidth of Δf = 100 kHz in the processing chain downstream, Schottky's formula predicts the rms measure of current fluctuations will be

$$\delta i_{rms} \equiv < [i(t) - i_{dc}]^2 >^{1/2} = (2\, e\, i_{dc}\, \Delta f)^{1/2}$$

$$= (2 \cdot 1.6 \times 10^{-19}\ \text{C} \cdot 10^{-5}\ \text{A} \cdot 10^{5}\ \text{Hz})^{1/2} = 5.7 \times 10^{-10}\ \text{A} = 0.57\ \text{nA}.$$

So the 10 μA, or 10,000 nA, photocurrent is predicted to exhibit fluctuations of only 0.57 nA (rms).

How will these fluctuations be made visible? The load resistor R_{in} = 10 kΩ not only 'maps' i_{dc} to a voltage $V_{mon} = i_{dc}\, R_{in}$, it also maps a fluctuation δi to a fluctuation $\delta V = \delta i\, R_{in}$. In rms measure, we map 0.57 nA to (0.57 nA)(10 kΩ) = 5.7 μV. Such voltage fluctuations are still too small to see directly. But as suggested in Fig. 3.1b, this signal is sent, by a.c. coupling, to the further gain stage (G_I = 100) in the pre-amplifier. So the output of the pre-amp ought to exhibit fluctuations of rms measure 570 μV (in a 100-kHz bandwidth). These sub-mV fluctuations might still be too small to be shown directly on an oscilloscope.

So to see if these amplified fluctuations are present, set up the high-level electronics (HLE) to include a 100-kHz low-pass filter, and a gain G_2 of 10 x 10 x 60 = 6000. This ought to give a noise signal of rms measure 3.4 Volts (in a 0-100 kHz bandwidth). Applied in the usual way to the squarer, you can expect an output with non-zero d.c. average value

$$< V_{sq} > = < V_{in}(t)^2 > / 10 \text{ V} = (3.42 \text{ V})^2 / 10 \text{ V} = 1.17 \text{ Volts} \ .$$

Your value will *differ* from this, because of complications and deficiencies in this circuit. But you can now reduce the illumination on your photodiode, to show that part of this $< V_{sq} >$ is attributable to the light. (Another part of it represents Johnson noise, and amplifier noise.) The part of the 'noise signal' that *is* connected to the illumination level is connected, by a quantifiable series of transformations, to the value of the electron charge *e*.

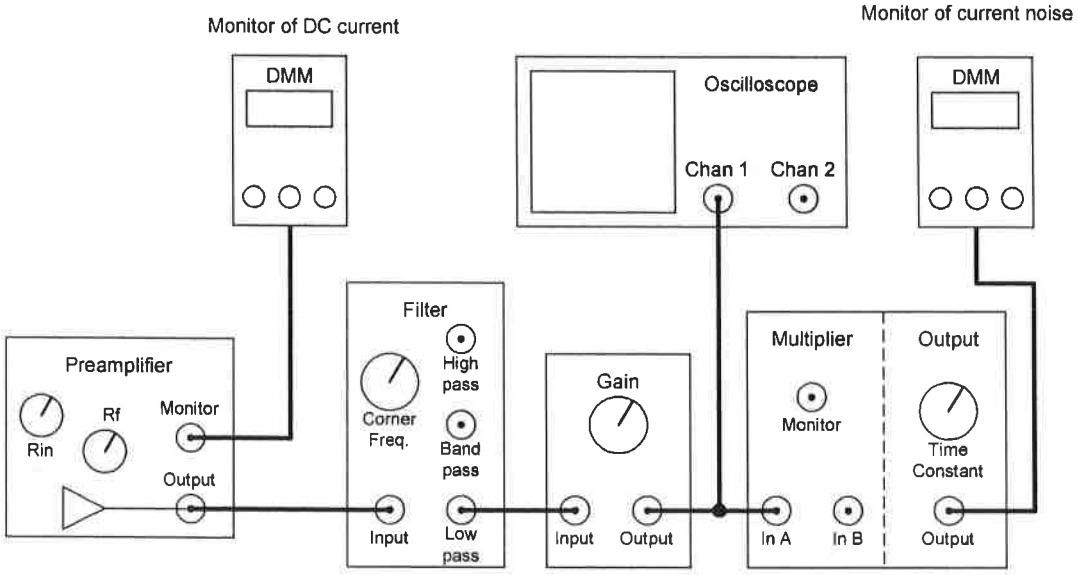


Fig. 3.2: Cabling diagram for LLE and HLE interconnections for first studies of shot noise

Postponing detailed quantitative questions until the next section, you ought here to confirm that you can get noise signals of this approximate size from this configuration of the system. Next, turn down the supply to the incandescent bulb to its minimum value, which ought to reduce the photocurrent i_{dc} to zero as well. Check a surrogate for i_{dc} at the MONITOR point on the pre-amp. But the noise is predicted to go to zero as well. In practice, it will go down, but not to zero, since there remains noise from the pre-amplifier itself.

What are the deficiencies of this circuit? There are at least two problems:

a) Changing the bulb's intensity will change the photodiode's current i_{dc}. But you will *not* be operating the photodiode at the special case of short-circuit current. Already at i_{dc} = 10 μA, there is a voltage drop of 100 mV across the load resistor -- and reference to Fig. 3.1b shows that there is also a (forward) bias of +100 mV across the photodiode. Because of this, the total diode current is not simply proportional to the illumination, because forward conduction is beginning to occur. The problem gets much worse as the photocurrent is raised to 100 μA or even 1 mA.

b) The photodiode, operating with zero or positive potential difference across it, exhibits maximal capacitance. While this does not affect i_{dc}, it *does* affect the noise, because it reduces the bandwidth over which the pre-amp system exhibits its full gain. Because of this reduction in bandwidth, the rms measure of the noise will fall short of the value predicted by the Schottky prediction.

Both these difficulties can be dealt with by reconfiguring the 'front end' of the pre-amplifier, as is described in the next section.

3.3 Shot-noise measurement using a transimpedance amplifier

The previous section showed you the existence, and approximate size, of shot-noise fluctuations in a photocurrent. In this section, we describe a circuit that has been optimized for measuring shot noise, and thus accurately determines the electronic charge, e. The novelty of this section is a method for operating an illuminated photodiode in a reverse-biased configuration, and converting its photocurrent to a voltage by a precisely known coefficient.

The circuit we use in illustrated in Fig. 3.3a. The cathode of the photodiode is connected to a point at potential +12 V. Meanwhile, the anode is actively held at potential zero, by being attached to the input of a current-to-voltage converter. (Appendix A.4 describes this among other op-amp topologies.)

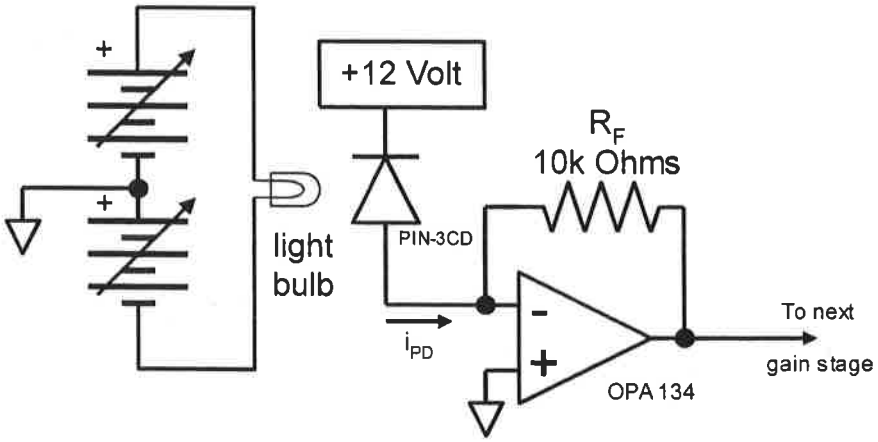


Fig. 3.3a: Schematic diagram for the connection of a reverse-biased photodiode to a pre-amp input stage configured as a current-to-voltage converter, also called a **transimpedance amplifier (TIA)**.

You will create this circuit by changing the default wiring of the pre-amp's first stage, and reconfiguring it to act as a current-to-voltage converter (see Fig. 3.3b, and also the photo in Fig. 3.3c.). Since the non-inverting input is grounded, feedback through R_f will actively hold the inverting input at near-zero potential as well. This ensures that the voltage drop across the photodiode always has the full value that is set on the biasing supply. Now with all the connections made, turn the pre-amp power back on, and flip the low-level electronics panel back into its ordinary position. Use the thumbscrews to close up the box.

Notice that in this circuit the photodiode current i_{dc} passes entirely through R_f (since the inverting input of the op-amp draws negligible current). This ensures that

$$0 - i_{dc} R_f = V_{out}, \text{ ie. } V_{out} = - R_f i_{dc} .$$

So the photocurrent has been mapped to an output voltage, which is measurable by a d.c.-coupled path at the pre-amp's MONITOR output. An a.c.-coupled path passes the a.c. components of this signal (including all the noise components with $f \geq 16$ Hz) to the subsequent gain stages of the pre-amp.

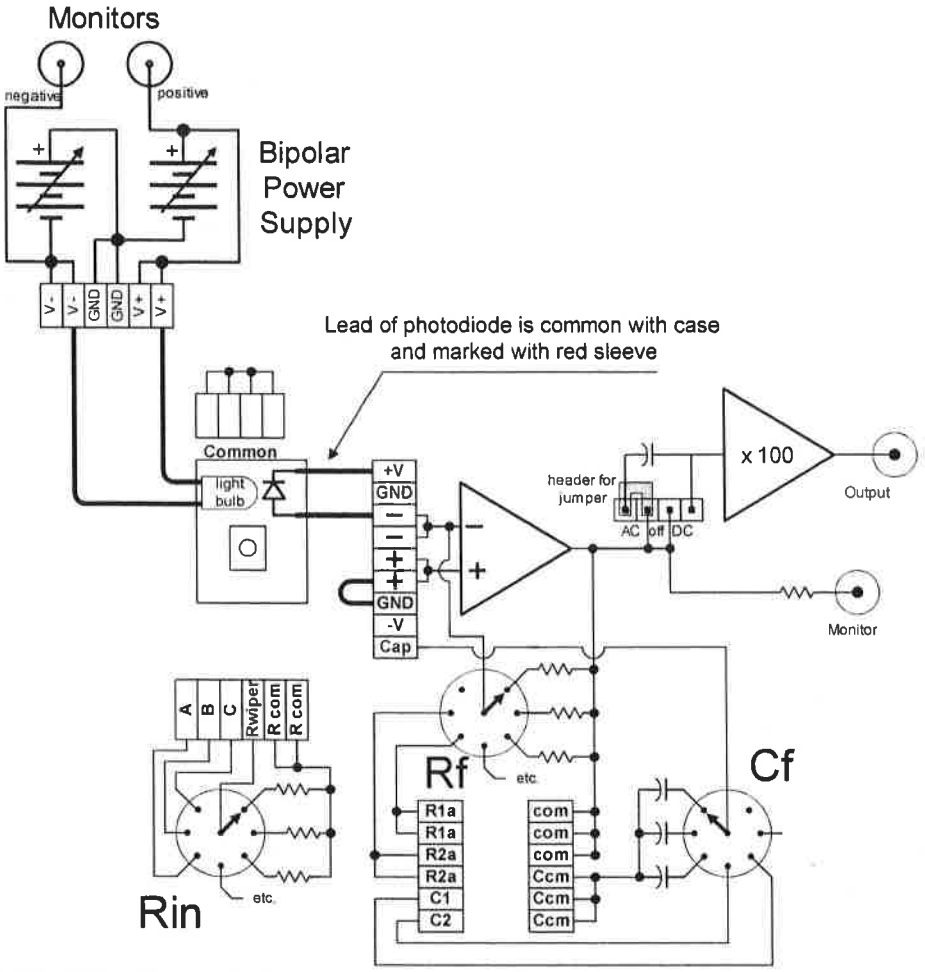


Fig. 3.3b: Wiring diagram for the input stage configured as an i-to-V converter.

Now if the photodiode current $i(t) = i_{dc} + \delta i(t)$ shows the sum of a d.c. average current plus the current fluctuations representing shot noise, then the i-to-V converter will give an output

$$V_{out}(t) = (-1)\, i_{dc}\, R_f + (-1)\, \delta i(t)\, R_f\,.$$

In the TeachSpin pre-amp, what follows in the a.c.-coupled path is 100-fold voltage amplification, so the cable connecting the pre-amp output to the high-level electronics will be conveying the voltage signal

$$V_{pre}(t) = 100 \text{ x } (-1)\, R_f\, \delta i(t)\,.$$

Inside the high-level box, use high- and low-pass filters just as before to create some chosen bandwidth Δf, and then use the main amp to provide further gain G_2 to bring the fluctuating signal up to the size suitable for the squarer. The output of the squarer will thus be

$$V_{sq}(t) = [-\, G_2\, 100\, R_f\, \delta i(t)]^2 / (10\text{ V})\,,$$

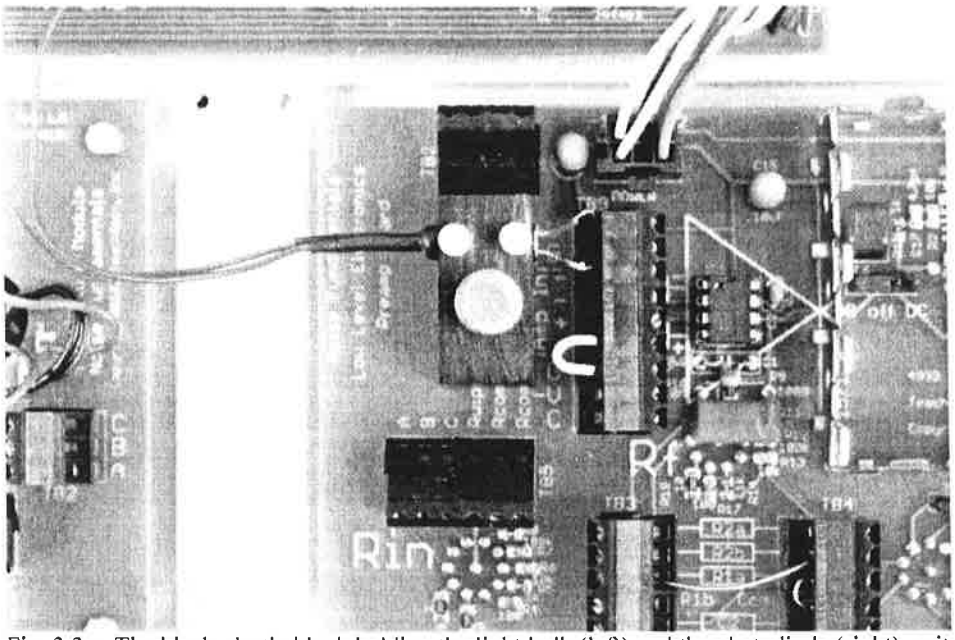


Fig. 3.3c: The black plastic block holding the light bulb (left) and the photodiode (right), suitably mounted for use with the transimpedance amplifier (TIA). (Notice that the plastic block has been rotated by 180°, relative to its orientation in Fig. 3.1c.)

and the familiar time-average of the squarer's output will give

$$< V_{sq}(t) > = < \delta i^2(t) > (100\ G_2\ R_f)^2 / (10\ \text{V}) .$$

If Schottky's formula is correct, then this gives

$$< V_{sq}(t) > = 2\ e\ i_{dc}\ \Delta f (100\ G_2\ R_f)^2 / (10\ \text{V}) \quad ,$$

which relates the quantum of charge e to quantities measurable in your experiment

To confirm that the bulb-photodiode combination is working, connect a multimeter (set to d.c. Volts) to the MONITOR point on the panel of the pre-amp. This will display a voltage $V_{mon} = (-)\ i_{dc}\ R_f$ (where R_f is the value of the feedback resistor you choose). This will be your way to measure the value of i_{dc}.

For the first attempt at measuring the electronic charge, we recommend you use $R_f = 10\ \text{k}\Omega$. Now dial up the bulb supply until you see evidence of photocurrent, via V_{mon}. For example, if you dial up the bulb until $V_{mon} = (-)\ 1$ V, you'll know

$$i_{dc} = (-)\ (-1\ \text{V})/(\ 10\ \text{k}\Omega) = \ 100\ \mu\text{A} .$$

By this means, you can explore photocurrents from <1 μA to >100 μA.

Now connect the rest of the cables just as was shown in Figure 3.2. Before connecting the output of the pre-amp to the input of the filter, you might want to observe the small noise signal on a 'scope, and confirm that the d.c. average value is near zero. (Notice we're not making use of the leftmost filter section.) In the adjacent filter section, use the low-pass output and set the corner

frequency to 100 kHz. Set all the AC/DC toggles switches to AC, for a.c. coupling between the stages. Adjust the HLE gain G_2 to keep the multiplier output in its 'good' range (0.6 to 1.2 Volts). Now turn the voltage to the light bulb to its minimum value. You should observe a decrease in the noise signal, and also notice, at the separate monitor point, that the d.c. current goes to zero.

So after noting the squarer's output in the presence of, and then in the absence of, the average photocurrent i_{dc}, and correcting for 'amplifier noise' by subtraction, you can infer a value for the mean-square fluctuation of the photocurrent:

$$< \delta i^2(t) > = [< V_{sq}(t) > \cdot 10 \text{ V}] / (100 G_2 R_f)^2 \ .$$

From these quantified mean-square fluctuations in the photocurrent, you can solve for the electron charge e; you will of course also need to know the d.c. photocurrent i_{dc} you measured indirectly, and the equivalent bandwidth Δf you used. (Recall that the ENBW of the 100-kHz low-pass filter is about 114 kHz.)

There are a few fine points:

1. The first stage in the pre-amp will have some small d.c. offset, in the range ± 2 mV. With the light bulb off, you can read and record this d.c. offset. To get the most accurate values of e, you should subtract this offset in forming your d.c. current measurement, i_{dc}. Make sure that you get the signs correct in your subtraction.

2. You've always thought of a digital multimeter as a passive device, merely reading the voltage presented to it. But some DMMs actively generate interference that's sent out *from* their input terminals. (It's also possible that the DMM generates no noise, but that its test leads are acting as antennas coupling interference into the pre-amp.) In either case, your DMM might be injecting some noise into the gain-100 stage of your pre-amp. This signal will make it all the way to the squarer, and add to the noise there. So when you're using a DMM at the MONITOR point to measure a surrogate for i_{dc}, check the squarer's output to test if you're subject to any of this interference. With the power to the light bulb set to zero, observe the noise voltage. This is a measure of the amplifier noise. Now remove the DMM's test leads from the d.c. current-monitor point, and see if there is any change in the noise voltage.

3. What we are calling the amplifier noise is now a bit more than just the voltage noise of the OPA134 op-amp. In this circuit, it also includes the Johnson noise of the feedback resistor[2]. A 10 kΩ resistor at room temperature has a noise of about 13 nV/√Hz. We add these two (uncorrelated) noise sources in quadrature to get an expected amplifier noise density of

$$S_{amp} = (8 \text{ nV}/\sqrt{\text{Hz}})^2 + (13 \text{ nV}/\sqrt{\text{Hz}})^2 \approx (15 \text{ nV}/\sqrt{\text{Hz}})^2 = 2.25 \times 10^{-16} \text{ V}^2/\text{Hz} \ .$$

For example, we measured a (bulb-off) value of 0.755 V from the squarer using a HLE gain of 5000.

[2] In our first measurements of Johnson noise, the gain setting resistors were chosen such that their Johnson noise was much smaller than the op-amp voltage noise.

$$S_{meas} = [< V_{sq} > \cdot 10\ V] / [G_1^2\, G_2^2\, \Delta f]$$
$$= [0.755\ V \cdot 10\ V] / [100^2 \cdot 5000^2 \cdot 114\ kHz] = 2.65 \times 10^{-16}\ V^2/Hz\ .$$

The slightly larger-than-expected amplifier noise is due to 'gain peaking' in the preamp. We will look at this in more detail in Section 3.4, but for the moment we note that the amplifier noise remains constant as the current is varied. So record your own number for the amplifier noise, and use it as a quantity to be subtracted from values obtained with non-zero photocurrent present.

You can now investigate shot noise systematically. With a fixed bandwidth Δf, you might first check the dependence of current noise $<\delta i^2(t)>$ on the average photocurrent i_{dc}. As was the case with Johnson noise, you can see changes in the noise which are much smaller than the amplifier noise. You should be able to see the noise increase slightly with only 0.1 μA of photocurrent. (That's 1 mV/10 kΩ measured at the MONITOR on the preamp module; to see a 1-mV level here, you'll need to note the d.c. offset at this MONITOR point.)

You will find it profitable to compute the values of mean-square current fluctuation per unit bandwidth, or 'current noise power density', $<\delta i^2(t)>/\Delta f$, with units of A^2/Hz. That's because the Schottky formula predicts this quotient ought to have a simple dependence on i_{dc}.

Thus far we have assumed that the shot noise is 'white', ie. spectrally uniform. To test this, temporarily fix i_{dc} and R_f at some suitable values, and test the effect of changing the choice of filter bandwidth in the HLE. As you lower the bandwidth, the amount of noise emerging should drop, and as usual, you'll increase the main-amplifier gain to keep the squarer in its optimum regime. But you should test to see if the quotient $<\delta i^2(t)> / \Delta f$ stays fixed -- this is a test of the claim that the shot noise is 'white'. (Because of the effects of capacitance, this uniformity may fail at choices of largest bandwidth -- the spectral uniformity of the response of the electronics chain is hardest to maintain at the high-frequency end of the large-bandwidth coverage.)

If you can confirm that the current noise power spectral density' $<\delta i^2(t)>/\Delta f$ is in fact independent of your bandwidth choices, but that it does depend on the average photocurrent i_{dc}, then you can try a log-log plot of $<\delta i^2(t)>/\Delta f$ as a function of i_{dc}. Your plot will have x-axis values in Amperes, and y-axis values in A^2/Hz. If you get a linear variation, that line will have slope with units (rise over run) of $(A^2/Hz)/A = A/Hz = A \cdot s = C$, Coulombs. In fact, Schottky's theory predicts a power-law fit, with power-law exponent 1 and coefficient $(2\ e)$, since

$$<\delta i^2(t)> / \Delta f = (2\ e)\ i_{dc}^{\ 1}$$

expresses the theory. Is your plot consistent with a power-law exponent of 1.00? If so, you can read off $(2\ e)$, in Coulombs, as the coefficient of your fit to the data! As is often the case, estimating the uncertainty in your value for e might be the hardest part of your experiment. (Remember this is *not* the same as the discrepancy, if any, between your value and the 'book value'.)

Despite its apparent simplicity, there can be experimental and conceptual pitfalls to this experiment. The most annoying is that some light bulbs, under some conditions, give unstable light output, despite stable voltage input. The effect, in changing the photocurrent in the photodiode, looks just like excess noise, which can falsify the derived value of e. The cause is (apparently) the intermittent shorting of adjacent turns of the finely-coiled tungsten wire in the

filament of these bulbs. It may help to isolate the low-level electronics box from mechanical vibration.

Another complication of the same character is the apparent presence of excess low-frequency fluctuations in the light output of the red LED when it is operated at its highest currents. (By contrast, we have not seen these fluctuations in the even larger output of the infrared LED.)

One cure for either of these effects is to change the bandwidth over which the experiment is sensitive. Previously you've been using coverage from about 16 Hz all the way to 100 kHz, but if there is excess noise at low frequencies, a good strategy is give up the contribution of the lowest frequencies. To do this, add in a previously-unused filter section as a high-pass filter (set to perhaps 1 or 3 kHz) at the input of the HLE. You will need to use the Table in Section 1.5 to give a revised value of Δf for your calculations. Now you've given up about 1 or 3% of the total white noise you have been detecting, but you've also blocked a disproportionately greater fraction of any excess low-frequency noise.

3.4 Diagnosing proper high-frequency behavior

Up to this point you have assumed that the current-to-voltage converter of Section 3.3 exhibits a conversion coefficient R_f which is frequency-independent. This is important, because while the d.c. or $f = 0$ value of this conversion coefficient relates i_{dc} to V_{mon}, it's the behavior of this coefficient at (and well beyond) frequencies of 100 kHz which applies to the various frequency components of the noise you're measuring. Now we'll take up a way to check the 'gain flatness' of your pre-amplifier section, and to compensate for the 'gain peaking' that tends to occur in the neighborhood of $f \approx 2$ MHz because of capacitive effects in the photodiode and the pre-amp's first stage (and the op-amp's finite gain-bandwidth product).

3.4a Step response as a diagnostic

To do these experiments, leave the input stage configured as in Section 3.3, but use the red LED as source of illumination. Wire that LED as suggested here, so that you can drive it with an external square-wave generator.

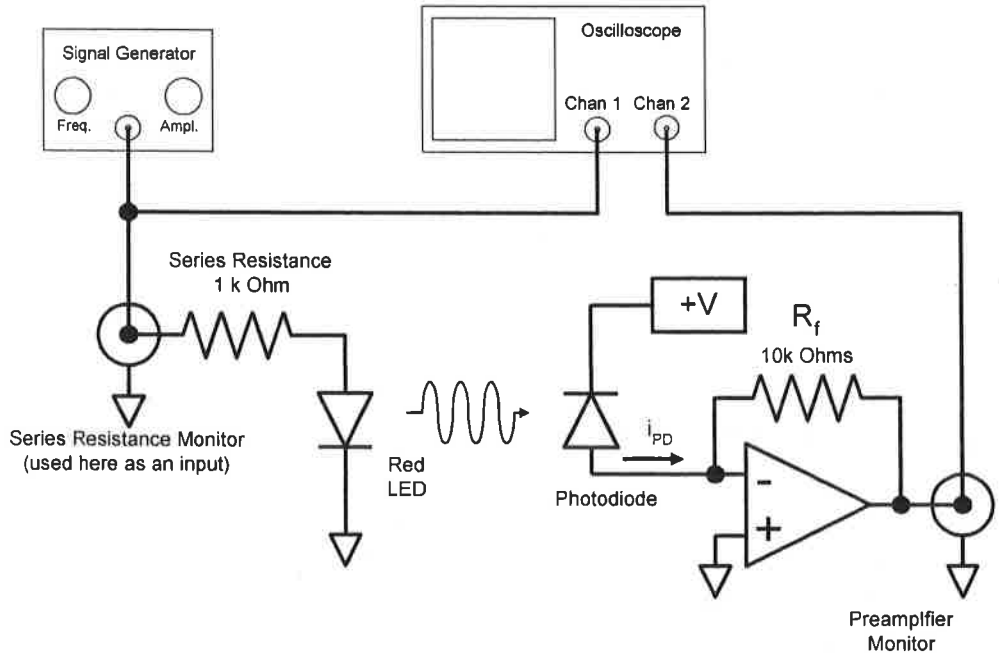


Fig. 3.4a: Schematic diagram for wiring an illuminating LED to permit drive by an external generator.

Here's the idea: a square-wave drive can cause the LED to alternate, at any desired rate, between off and on states. You could use a 1 kHz square wave, for example, to spend 500 μs in the 'off' mode, and 500 μs in the 'on' mode, in every period of 1 ms. An oscilloscope looking at your pre-amp's MONITOR point will show you the direct-coupled version of the photocurrent resulting from your modulated LED output. Since you're about to look at high-frequency behavior at the MONITOR output point, we suggest the use of a 10x 'scope probe, inserted into the monitor BNC connector in place of a BNC cable, so as to minimize capacitive effects.

Use a second channel of your 'scope to display the drive waveform you're sending to the LED, and trigger the 'scope on this waveform. Now you have a twin view of 'cause' and 'effect', LED drive and photodiode/pre-amp response.

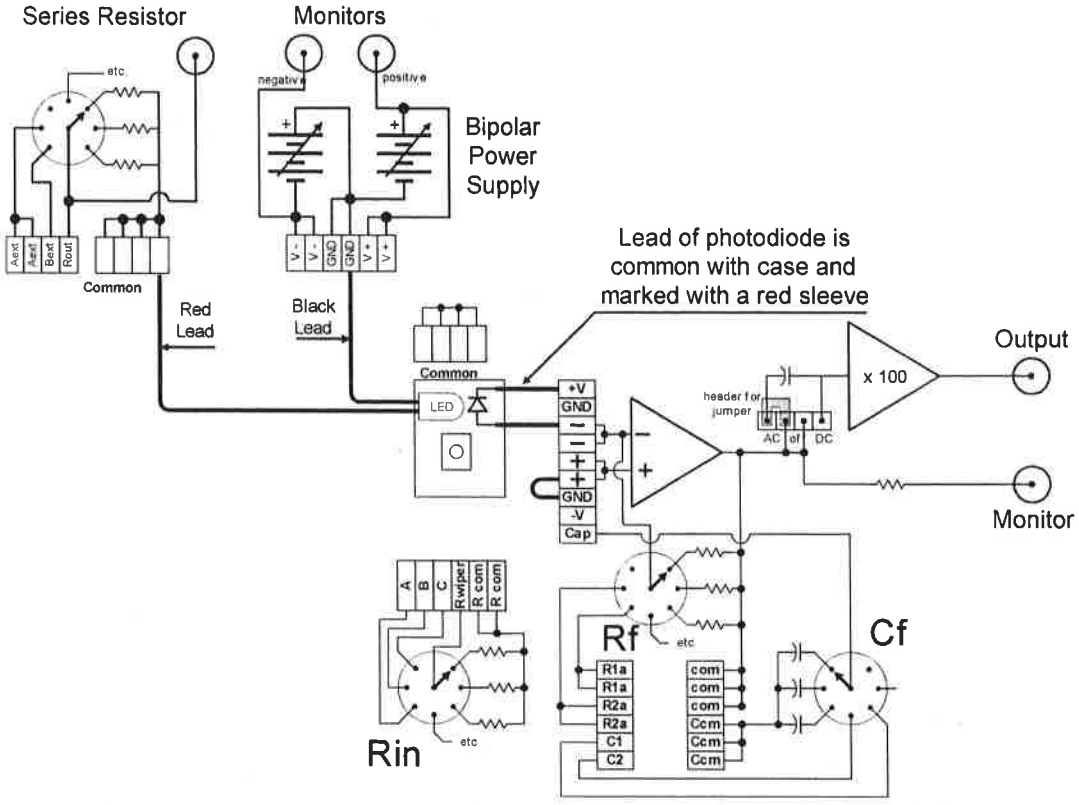


Fig. 3.4b: Wiring diagram for an LED-illuminated photodiode to study step-function response.

What's the point of this exercise? The point is to look at a close-up view of the 'effect' waveform, which shows the result of a sudden rise in the drive to the LED. Zoom in on the time axis until you can see that the 'effect' waveform displays a notable 'ringing', as well as a delay, and a slower risetime, compared to the 'cause' waveform. In fact, you are getting a direct view, in the time domain, of the 'step response' of the input stage. Since you're exciting a linear system, that time-domain view is connected (via a Fourier transform) to the frequency-domain description of the response of the input stage. That frequency-domain description would tell you everything about the bandwidth of the system, the very information desired. In fact, it's not even necessary to perform the actual Fourier transform, since your time-domain view of the step response already contains all the desired information -- it's just encoded in a different way.

Here's a view of the step response of the photodiode/TIA front-end circuit used in Section 3.3 to measure shot noise. Results from your circuit may differ in detail, because this data depends on details of circuit capacitances and op-amp high-frequency response.

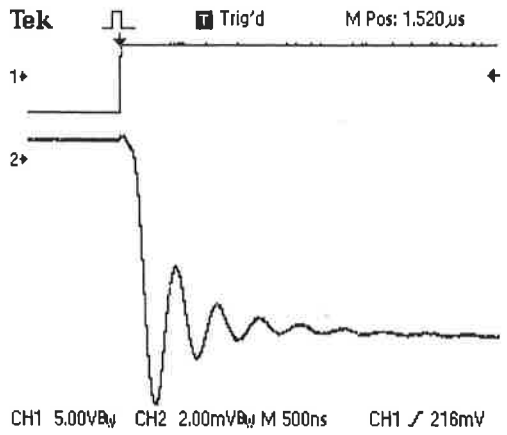


Fig. 3.4c: Step response of the first stage of the pre-amp. Upper trace: the drive waveform for the LED illuminating the photodiode; lower trace: the response measured at the monitor output of the first stage.

But the important features of this step response can be read off at a glance. The first conclusion is that the effect is delayed, relative to the cause, by about 0.2 µs. This is an indirect but valuable measure of the bandwidth of this circuit. The second conclusion is that the circuit, in response to a step function, shows the excitation of damped oscillations. The period of the visible oscillations is about 0.45 µs. This, in turn, is indirect evidence of 'gain peaking', at a frequency ≈2.2 MHz given by the inverse of this period. Your intuition might suggest that if the behavior were less damped, this gain peaking would be even greater.

So the time-domain view of your front-end electronics suggests that this crucial first stage has a current-to-voltage coefficient which is given by R_f at low frequencies, but exhibits a larger response at the peaking frequency near 2 MHz, before dropping toward zero. Fortunately, this localized excess gain presents little trouble for the noise measurements you've made thus far, even with filter bandwidths set to 100 kHz. There are two 'lines of defense' in the circuits you used. First, the low-pass filter in the high-level electronics has a 'corner frequency' of at most 100 kHz = 0.1 MHz. Because the frequency response of such filters is not of a brick-wall character, your noise measurements have some sensitivity to frequencies of 0.2, 0.4, even 0.8 MHz. But the methods of Section 2.1 show that contributions above 800 kHz account for only 0.06% of the total response to white noise. If your shot-noise measurements to date were subject even to a four-fold excess response to noise at the peaking frequency, that would at worst change the this contribution to 0.24%, still small.

In fact, the actual effects of gain peaking are smaller still, because other parts of the noise-processing chain have their gain drop off at various corner frequencies of 1.4 to 1.6 MHz. So the 0.24% effect computed above is an *upper limit* to the effects of gain peaking.

3.4b Eliminating gain peaking

The gain peaking which you have diagnosed by step-response behavior will not always be so harmless. Since the location in frequency, and the degree, of gain peaking will depend on the photodiode capacitance and the feedback resistor R_f, the use of other photodiodes, or other choices of R_f, might put the gain peaking, and the excess noise, into a place in the frequency spectrum where it could compromise the validity of your noise measurements. Because of the general utility of the technique, we introduce here a method for curing the problem.

The only new component needed is a feedback capacitor C_f, connected in parallel with the feedback resistor in the TIA. In your pre-amp, C_f is selected by a front-panel switch. Heretofore this switch, and its bank of capacitors, has deliberately been left disconnected in your first-stage op-amp circuit. But you can now add the necessary jumper wires which will connect a switch-selected C_f into your TIA.

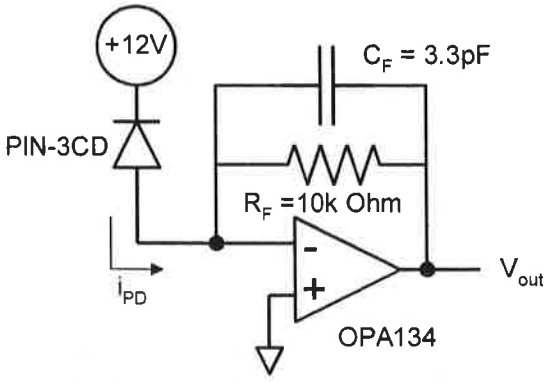


Fig. 3.4d: Schematic diagram for the feedback capacitor added to an i-to-V converter.

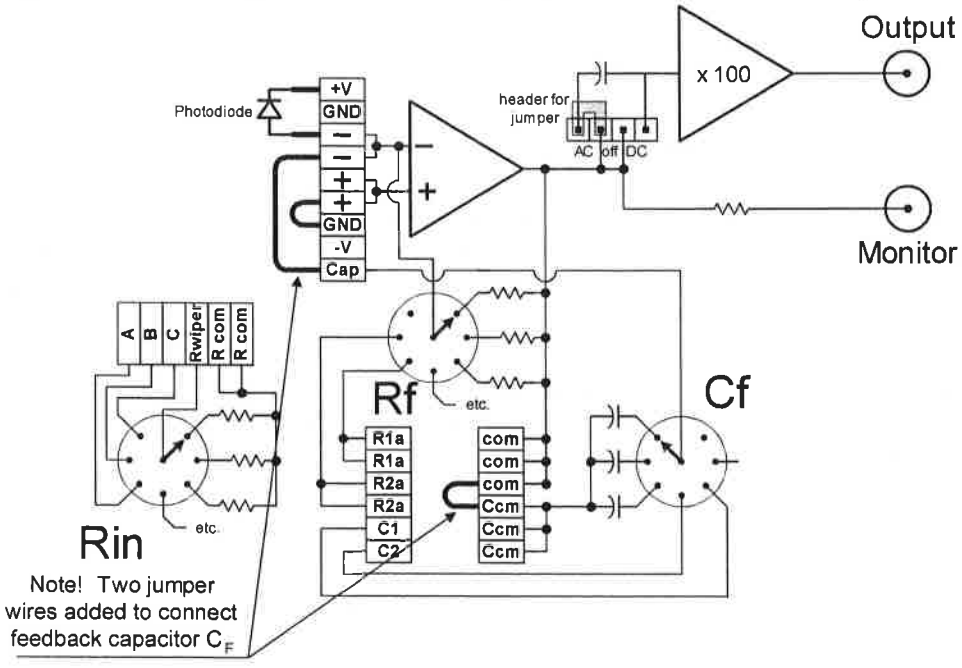


Fig. 3.4e: Wiring diagram for an i-to-V converter including a feedback capacitor C_f in parallel with R_f.

The theory of C_f's effect can be worked out in a model which includes the capacitance that's effectively in parallel with the photodiode, and the op-amp's finite gain and bandwidth. But the theory is messy, and its predictions in any case it depends on parameters that are difficult to measure. So we adopt here a more empirical approach: we'll use a desired shape of step response as a goal, and we'll use a now-adjustable C_f as a means to achieve it.

Our photodiode/TIA circuit now involves two capacitances, and (to a fair approximation) can be modeled by a 'two-pole response'. The same response has been built into the filter sections of the HLE, and the details of the step response of such systems are worked out in Section 6.1. Here too we'd like our finished circuit to have frequency response which is as level as possible, from d.c. to some maximum frequency. The two-pole response of maximal flatness-in-frequency is the Butterworth response. Fortunately, Butterworth response can be easily recognized in the time domain -- it is characterized by less-than-critical damping, and thus shows a bit of overshoot. The ideal Butterworth low-pass response in fact shows a 4-5% of overshoot relative to its eventual asymptotic value.

You can achieve this response, with its slight but detectable overshoot, by adjusting C_f in your TIA.

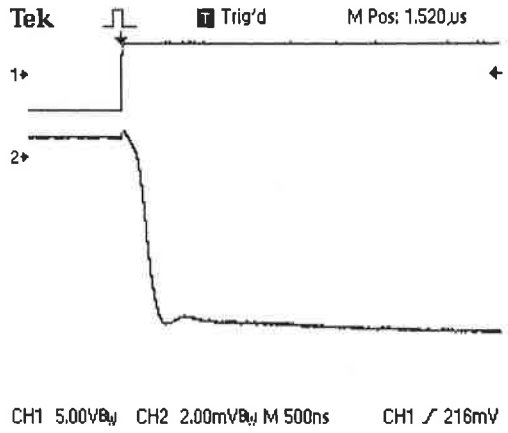


Fig. 3.4f: Step response of the first stage of the pre-amp, after the compensating capacitor has been adjusted to give only a slight overshoot.

[If you'd like a finer resolution of C_f adjustment, you can replace one of the jumper wire you've added by a 3.3-pF or 10-pF capacitor. Now the effective C_f will be the series combination of the added capacitor and the switch-selected value.]

This choice of C_f has given you Butterworth response, or (more intuitively) has damped the oscillations visible in Fig. 3.4c. But in a frequency-domain vocabulary, the result is that you have eliminated gain peaking. What's more, Fig. 3.4f contains additional information on the corner frequency of the now-modified circuit. You know that a two-pole Butterworth system, driven by a step function at $t = 0$, will have some long-term or asymptotic value to which the response will level off. By the methods of Section 6.1, you can show that the time t_{first}, after $t = 0$, at which the response first passes through this level (prior to its modest overshoot) is connected to the corner frequency of the system by the relationship

$$f_{corner} = 0.5303 \,/\, t_{first} \ .$$

So from step-response data for an approximately-Butterworth system as in Fig. 3.4f, you can deduce a corner frequency, and thereby an approximate representation of its frequency response. Instead of being 'flat' for all frequencies, its response drops off according to a gain function

$$G_{\text{pre-amp}}(f) = [\ 1 + (f/f_{\text{corner}})^4\ ]^{-1/2}\ .$$

Now if your indirectly-measured value for f_{corner} is vastly higher than the corner frequency you use in the filtering in the HLE, this 'droop' in high-frequency gain of the pre-amp is not relevant. But if the corner frequency of the low-pass filter you use in the HLE is as much as a quarter of the pre-amp's first-stage corner frequency, then there is a correction of order 1% or more to the equivalent noise bandwidth of your system. Section 2.2 shows you how to compute Δf for the total system of pre-amp plus HLE filtering, and you can use the representation for $G_{\text{pre-amp}}(f)$ above to see what effect this will have.

3.4c Other applications of a compensated TIA

Now that you've expended the effort to compensate your TIA, you've learned a technique of rather general applicability. But particular to your investigation of shot noise, you now have a front end whose performance is more nearly ideal than before. In Section 3.4a we argued that even the previous, sub-optimal, circuit would still give reliable shot-noise measurements out to bandwidths of 100 kHz, but now you can check that claim directly. So you can now repeat your favorite protocol of Section 3.3, and get an even more trustworthy result for the electronic charge e.

You are also now liberated to perform other checks and extensions of shot noise measurements. You can now change R_f, which will certainly change your d.c. sensitivity to photocurrents. (A larger value of R_f will make you more sensitive in the range of very small photocurrents.) For any new value of R_f, you will want to re-check the step response, to get an optimal value of C_f. You will find that for $R_f > 100$ kΩ, you will not need to add any C_f: stray capacitance is enough to compensate your first stage. But you should still use the pulsed-LED method to investigate the first stage's properties. You will probably notice a *single*-pole RC roll-off, and will need to repeat your ENBW calculation accordingly.

Alternatively, you can now try out the use of a photodiode with*out* reverse bias. That will require only that the cathode of your photodiode is moved from the "+V" point, or a +12-Volt potential, to a ground point. The photodiode capacitance will now be larger, and again you'll need to check the step response to pick a best value of C_f. The step response will also give you the corner frequency of the now-compensated circuit, which can use in your noise modeling. If you can take a new set of shot-noise data, and deduce from it a new but consistent value of e, you will have further evidence that your e-value is not an artifact of a particular circuit topology for using the photodiode.

3.5 Sub-shot-noise currents

In Chapter 3 you've seen the statistical argument that shows why (certain) currents display shot noise, and predicts how big that shot noise should be. You've also measured actual shot noise in a photocurrent, to see if the shot noise obeys Schottky's prediction. This section has a special purpose: to prove empirically that it's easy to produce currents whose fluctuations are smaller, **much smaller**, than the standard shot-noise prediction.

It turns out to be very easy to produce (but not so easy to explain) a sub-shot-noise current. Consider, for example, a simple 9-V battery as a voltage source, with a 100-kΩ resistor across its terminals. The current flowing will be i_{dc} = 9 V/(100 kΩ) = 90 μA. You'll be able to prove that the fluctuations in such a current are *not* given by the shot-noise formula, but are much smaller. (This empirical fact does **not** seem to be very generally known!) The immediate implication of such an observation is that electrons in such a circuit are ***not*** moving independently and at random, but instead in some more nearly regular way. The mechanism for this enhanced regularity seems to be the Coulombic interactions of the whole cloud of electrons that is collisionally diffusing through the resistor. If electrons do interact this way, the previous argument of statistically-independent arrivals fails, and the 'shot-noise limit' with it. (See the treatment by Landauer noted in the Bibliography.)

By contrast, the photoelectrons which you've already seen displaying full shot noise *are* produced independently by photon arrivals, and then are removed (by internal electric fields) from the detector's depletion region in so short a time that it becomes fair to think of the electrons as not interacting.

Turning from possibly-unsatisfying theoretical arguments to empirical tests, here's the method for producing and quantifying a sub-shot-noise current. Take *out* the photo-diode from its position in the pre-amp module, but leave intact the current-to-voltage (i-to-V) converter. Recall that you have a feedback resistor, R_f, which sets the i-to-V conversion constant. Now you want to feed into the input point of that circuit a current derived simply from $i = V/R_{in}$. Here V is the stable voltage available from the 0-11 V power supply in your low-level electronics, and R_{in} is an input resistor, selectable via the rotary switch in the pre-amp. (These are the resistors previously serving as sources of Johnson noise.) Figure 3.5 shows the schematic diagram for the input of the circuit.

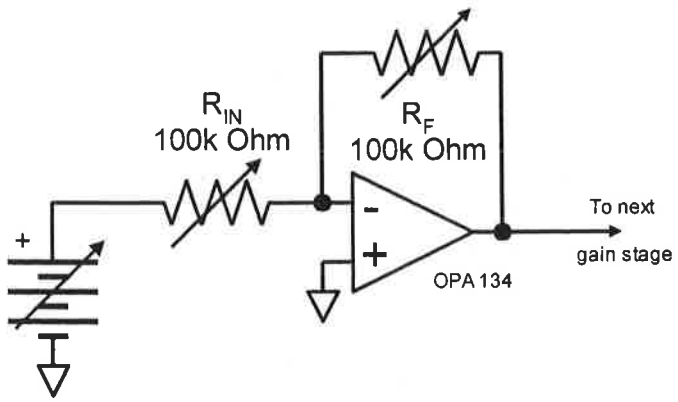


Fig. 3.5: Schematic diagram for testing V/R currents for noise level.

Some good choices are R_f = 100 kΩ, and R_{in} = 100 kΩ also. Now the use of a 6-Volt setting from the adjustable supply delivers to the (virtually-grounded) input of the i-to-V converter a current of $i = V/R_{in}$ = 6.0 V/(100 kΩ) = 60 μA. And the i-to-V converter gives at its output

$$V_{out} = - i_{in} R_f = - (60\ \mu A)(100\ k\Omega) = (-)6.0\ V\ ,$$

and the MONITOR point on the pre-amp's panel will make this voltage available to a multimeter. Exactly as in the photocurrent case, it gives a surrogate for i_{in}'s value.

After this first stage, everything is exactly as it was in the photocurrent case. There's a d.c.-blocking capacitor, there's a further 100-fold gain in the pre-amp, there's a cable to the high-level electronics, and then there's filtering in frequency, and further amplification by gain G_2, and then squaring and time-averaging, finally to deliver a mean-square voltage output. As shown in Section 3.3, the time-average of the squarer's output is given by

$$<V_{sq}(t)> = <\delta i^2(t)> (100\ G_2\ R_f)^2 / (10\ V)\ ,$$

so the noise density in the current is

$$<\delta i^2(t)> /\Delta f = (10\ V) <V_{sq}> / [\ (100\ G_2\ R_f)^2\ \Delta f\ ]\ .$$

Since everything on the right-hand side is measurable, you can quantify the spectral density of current fluctuations in A^2/Hz. Better still, you are doing that for a current whose average value you also know, so you can use the shot-noise formula to compare to the current-noise density which would be predicted for a current displaying full shot noise.

In the previous case of the photo-current's noise, it was easy to separate the effects of actual photocurrent noise from amplifier noise by doing two experiments:

- the 'control group' had the lamp off, giving zero photocurrent, so the $<V_{sq}>$ measured was wholly due to amplifier noise;
- and the 'experimental group' had the lamp on, and gave a larger $<V_{sq}>$ which was made of a sum of mean-square noise in the photocurrent, plus the mean-square noise due to the amplifier.

(This works because uncorrelated noise sources have their mean-square values combine as a simple sum -- see Section 2.3 to see why there's no cross term in the squarer's time-averaged output.)

For the new case of the $i_{in} = V/R_{in}$ current, you want to use the same two-group comparison:

- the 'control group' with V set to near-zero;
- and the 'experimental group' with V set to (say) = 6.0 V.

Subtraction of the two $<V_{sq}>$ results can isolate the $<\delta i^2(t)>$ value for the current of interest (freed of amplifier-noise contributions), and that value can be compared to a similarly measured value of $<\delta i^2(t)>$ for photocurrent of equal size. By this wholly empirical comparison, you can falsify, and *strongly* falsify, the claim that all currents display full shot noise. There is the remarkable implication that 'not all 60-μA currents are created equal', but that they can be distinguished by their noise. There is the further amazing implication that macroscopic currents, of equal average value, in room-temperature electronics, can show distinctive features that are telling you something about the different statistical properties of the arrivals of individual electrons.

Last note: The current $i_{in} = V/R_{in}$ that you've been studying will certainly have another kind of noise on it if the source voltage V is unstable, or has noise on it. In fact, your circuit using R_{in} together with a current-to-voltage converter can alternatively be viewed as an ordinary inverting amplifier, driven by the power supply delivering the input voltage V. (See Appendix A.4 on input-stage op-amp topologies.) So if nothing else adds to the noise, you can re-interpret your results as a measurement of the voltage noise on the d.c. voltage V delivered by your 0-11 V power-supply. That power supply has been built to display very low voltage noise -- under 5 nV/$\sqrt{\text{Hz}}$. (If you get a dramatically larger result, you should look at Appendix A.7, in case there's been damage to the noise-suppressing regulators in your 0-11 V supply.)

3.6 Photodiodes and photocurrent

This section introduces the empirical description of a very valuable semiconductor device, the photodiode. The description applies both to the small photodiodes we use in instrumentation, and the large solar cells used for energy production. Apart from these applications, the photodiode will also be our first choice for generating a macroscopic electric current displaying 'full' shot noise.

A photodiode is first of all a p-n junction diode, so it's a two-terminal device whose d.c. electrical properties can be characterized by a potential difference ΔV applied, and a current i passing through. For a photodiode in the *dark*, the simplest model for a diode response is

$$i = i(\Delta V) = i_0 \, [\exp (\Delta V / V_0) - 1] \ .$$

(Section 3.7 suggests why V_0 should have a magnitude of about 25 mV.) A 'reverse bias' for a diode is the application of a sufficiently negative ΔV, which gives

$$i = i_{\text{reverse}} \approx - \, i_0 \ .$$

Near 'zero bias', ie. for $|\Delta V| << 25$ mV, a series expansion of the exponential model predicts

$$i(\text{small } \Delta V) \approx i_0 \, (\Delta V / V_0) = (i_0 / V_0) \, \Delta V \ ,$$

which has the form $i \propto \Delta V$. This is Ohmic behavior, with an effective resistance of V_0 / i_0, which is typically very large. Finally, for $\Delta V >> 25$ mV, the exponential term dominates the (-1) term in the brackets, and we get

$$i(\text{large positive } \Delta V) \approx i_0 \exp (\Delta V / V_0) \ .$$

We'd call this 'forward conduction', which is the ordinary application of power diodes. But we'll see that photodiodes are typically not operated in this region.

There are relatively direct ways of measuring the $i = i(\Delta V)$ curve (laid out for your optional work in Sec. 3.7), but here we show results for the type of photodiode included in your Noise Fundamentals kit. Fig. 3.6 shows (as its uppermost curve) the behavior described by the diode equation above. There is no way that a plot using a single choice of scale on the vertical axis can encompass the large variation in the currents that could be measured, which range from pA to mA (a variation by factor 10^9).

Thus far, we have a diode model which describes a photodiode *in the dark*. What happens when light falls onto its active surface? The results are quite simple: the former curve described by $i(\Delta V$, dark) is everywhere *shifted vertically downward* in Fig. 3.6, by an amount i_{light} which is directly proportional to the illumination level I. The proportionality constant is called the responsivity r, and it has units of A/W (Amperes of photocurrent per Watt of light incident on the diode). So in this new model,

$$i = i(\Delta V, I) = - \, r \, I + i_0 \, [\exp (\Delta V / V_0) - 1] \ ,$$

which predicts the lower curves shown in Fig. 3.6. This displays the results from illumination levels of 0, I_1, and $2I_1$, where I_1 is some reference level of illumination. Again, Sec. 3.7 describes how you can take such data for your own photodiode, if desired.

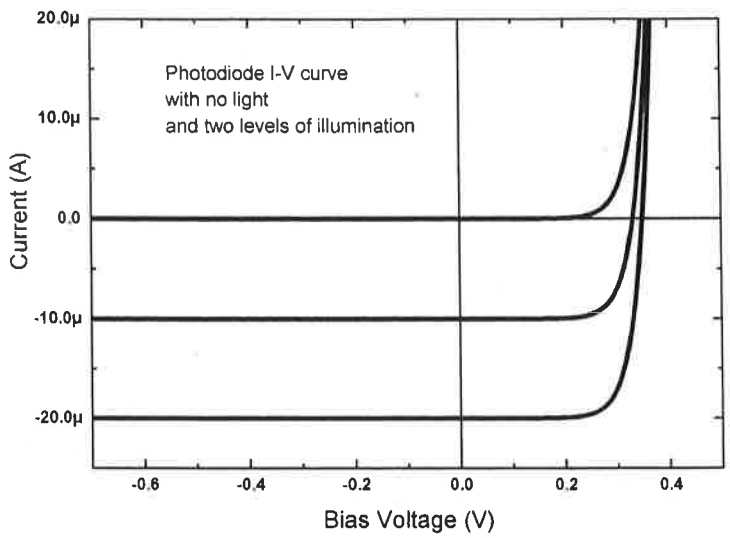


Fig. 3.6: The current passing through a photodiode as a function of potential difference across its terminals. Top curve: photodiode in darkness; middle curve: photodiode in light; bottom curve: photodiode in light twice as bright.

For instrumentation purposes, it is very valuable to measure the 'short-circuit current' of a photodiode. This is the (photo)current which flows when an external circuit enforces $\Delta V = 0$ (as a short circuit would). The model above predicts

$$i_{sc} = i(\Delta V = 0, I) = - r I \ ,$$

which is a current strictly proportional to the level of illumination I. In fact, for light not too intense, this proportionality extends over six or more orders of magnitude, perhaps from 1 nA to 1 mA of photocurrent for your photodiode.

The responsivity r of a photodiode is not just an empirical value provided by the manufacturer, but a constant which can be understood quantum-mechanically. We imagine the illumination I is provided by a 'rainfall' of photons, arriving in quantity N during time τ. If the photons are from a light field of frequency f, then each one delivers energy hf, so the illumination delivered by their average rate of arrival is $I = (hf)(N/\tau)$. Now if hf is large enough (bigger than the semi-conductor's band gap), then the simplest assumption is that each arriving photon is absorbed in the semiconductor, and lifts one electron from the valence band to the conduction band. Once in the conduction band, that electron is free to flow as part of the photocurrent i. So under this assumption of '100% quantum efficiency', the current i will be given by $(-e)(N/\tau)$. Thus we have

$$i = -e\ (N/\tau) = (-e/hf)\ (hf\, N/\tau) = -\ (e/hf)\ I \ ,$$

so the responsivity is $r = e/(hf)$. For illumination by red light of wavelength $\lambda = 650$ nm, we compute frequency $f = c/\lambda = 460$ THz $= 4.6 \times 10^{14}$ Hz, and predict responsivity

$$r = e/(hf) = (1.6 \times 10^{-19}\ \text{C})/(6.6 \times 10^{-34}\ \text{Js} \cdot 4.6 \times 10^{14}\ \text{Hz})$$
$$= 0.53\ \text{C/J} = 0.53\ \text{(C/s)/(J/s)} = 0.53\ \text{A/W} \ .$$

The responsivity of actual silicon photodiodes is embarrassingly close to this computed value. Manufacturer's graphs of the responsivity also displays (approximately) the predicted variation $r \propto 1/f$, or $r \propto \lambda$, with respect to the wavelength of the light used. But for $\lambda \geq 1$ μm, the photon energy is too small for silicon's band gap, and the response drops very rapidly to zero. So the room-temperature thermal photons in which a photodiode is immersed do *not* contribute to the photocurrent.

Now the short-circuit photocurrent i_{sc} shown in Fig. 3.6 is not the only way to use a photodiode. In practice, we often use a 'reverse-biased' photodiode, which has ΔV large and negative -- at the far left in Fig. 3.6. This predicts

$$i(\text{reverse } \Delta V, I) \approx - r I - i_0 \ .$$

The disadvantage is that the (typically very tiny) reverse current i_0 flows in addition to true photocurrent, but the advantage is that the reverse bias applied to the photodiode thickens the depletion layer in the p-n junction. This lowers the device's capacitance, and speeds up its time response. We've seen this effect used in Sec. 3.3 to get, from the improved time of response, the largest possible bandwidth for the photodiode's performance. This enables the circuit to respond uniformly to current fluctuations (ie. to shot noise) out to a maximum possible frequency.

Returning to Fig. 3.6, and temporarily leaving instrumentation in favor of power generation, we note that the maximum-power point of operation lies not at the short-circuit, $\Delta V = 0$, location, but instead at a point that falls in the fourth quadrant of our plot. This desired operating point is located where the magnitude of the product $i(\Delta V) \cdot \Delta V$ is maximized. Thus solar cells in operation need to be connected to an external circuit which imposes the correct potential-difference ΔV, so as to operate at such a point in the $(\Delta V, i)$ plane.

Finally, plots such as Fig. 3.6 also show the 'open-circuit voltage' ΔV_{oc}. This is the ΔV that would be required to enforce $i = 0$, or equivalently, the ΔV that would develop across an illuminated, but dis-connected, photodiode. Some open-circuit voltages can be seen as <u>*x*-axis</u> intercepts in Fig. 3.6. The diode model above predicts that such points are located where

$$0 = i(\Delta V_{oc}, I) = - r I + i_0 [\exp (\Delta V / V_0) - 1] \ ,$$

and if the illumination is not very dim, ie. if $\Delta V >> 25$ mV, this can be solved to give the approximate result

$$\Delta V_{oc} \approx (V_0) \ln [r I / i_0] \ .$$

This approximately-logarithmic response to illumination I is sometime useful, but a photodiode used in this way suffers from slow response in time, and a solar cell used in this way delivers no useful power.

3.7 Photodiodes' current-voltage curves

In previous sections you have used photodiodes as transducers from light to photocurrent, and also as sources of noise. This section drops back from noise measurement, to offer you a way to confirm the current-voltage characteristics, the i-V curve, of such devices. The results will test the models of Section 3.6 for your photodiode, but they apply much more generally to the operation of photovoltaic or solar-cell modules designed for the generation of power.

We seek here to measure the photodiode current, i, as a dependent variable. It depends on two independent variables. The one which you can measure absolutely is the potential difference ΔV maintained across the diode. The other which you can measure relatively is the illumination of the photodiode. This will enable you to measure a series of points along an $i(\Delta V, I)$ curve, with as many points as you like along the ΔV-axis, and repeating for some relatively-known values of illumination, such a $I = 0$, I_1, and $2{\cdot}I_1$.

Our suggestion for how to control the level of light falling on the photodiode is to use the red LED to illuminate it, while both are held in the black plastic structure illustrated in Fig. 3.1d. When that LED is off (and the whole assembly is in the dark, inside the closed-up low-level electronics box), you can assume the illumination level is adequately close to zero. Then you may further assume that for moderate levels of current passing through the LED, its light output is a linear function of the current. So if you can control the LED current to take on the values of 0, 5, and 10 mA, you can assume that three light levels in proportions 0:1:2 fall on the photodiode. (This method will *not* give you the absolute levels of illumination, in W/m^2, however.)

To control the current through the LED requires the use of an ***additional external d.c. power supply***. You can attach that power supply to the LLE's Series-Resistance BNC input, and use the 1-kΩ series resistor to limit the LED current. To have 10 mA flowing through the LED will entail a voltage drop of (10 mA)(1 kΩ) = 10 Volt across the resistor, and there will be another voltage drop of order 2 V across the LED itself. So you'll need a power supply variable in the 0 to +15-V range, and you'll need to have the polarity of the LED arranged so that this positive supply will drive forward current through the LED, and light it up. But the point is that an external current meter can be used to measure the actual current flowing through the LED. By this means you can get LED currents in any proportions you wish, including for example the choices 0, 5, and 10 mA.

Our suggestion for how to measure the current being generated by the photodiode is to use the transimpedance amplifier just as in Section 3.3. If you use a feedback resistor of $R_f = 10$ kΩ, then the output voltage of this TIA first stage will be visible at the MONITOR output of the pre-amp, and its size will be given by $(i)(10\ \text{k}\Omega)$. Thus $V_{mon}/10$ kΩ will give the absolute value of the photodiode current i. But an added feature of the TIA is that its use ensures that one electrode of the photodiode is being held (by feedback) at ground potential. Then the potential measured at the *other* electrode of the photodiode will be sure to give you the potential difference, ΔV, across the photodiode. Finally, we suggest the use of the adjustable bias supplies inside the LLE as a source of this voltage. You will probably want to use, in turn, both positive and negative polarities for the bias you supply to the photodiode.

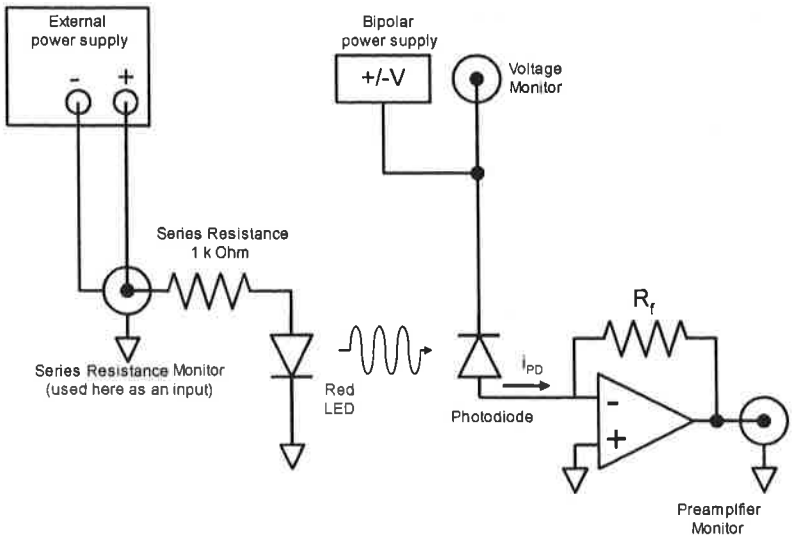


Fig. 3.7a: Schematic diagram for circuits used in measuring the i-V curve of a photodiode.

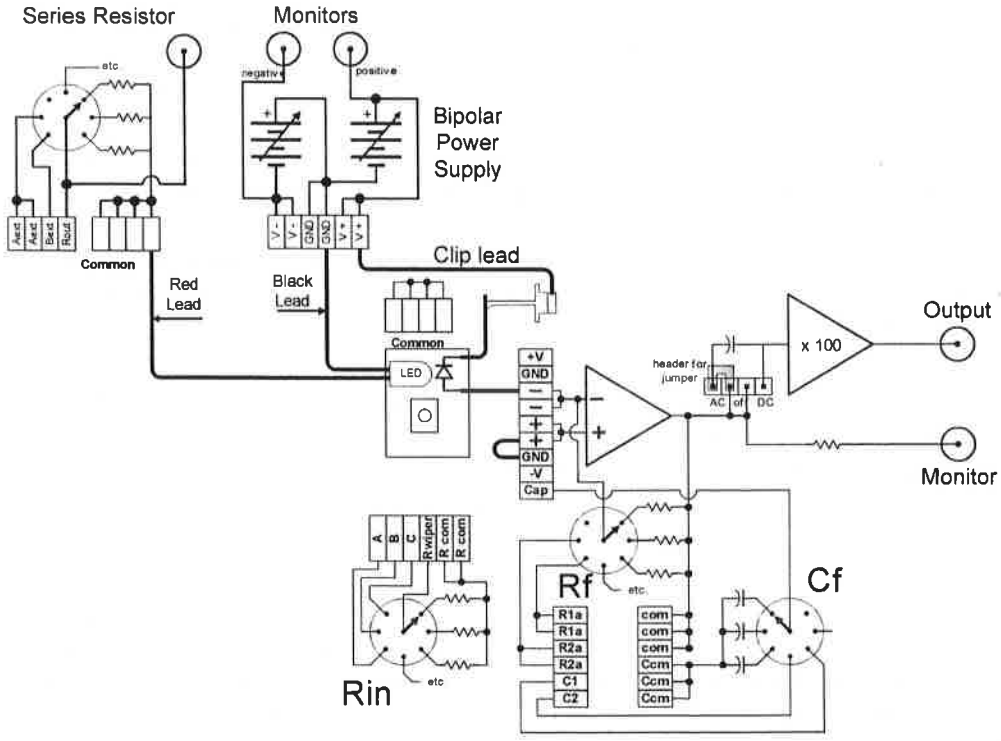


Fig. 3.7b: A wiring diagram for the circuit.

There are conventions on polarity which are used in (some) discussions of photodiodes. The one we adopt here, as in Section 3.1, starts with a photodiode in the *dark*. We call positive the potential difference that can drive a substantial (say, 100 μA) current through the photodiode. (By contrast, we call negative the potential difference that will drive only a tiny reverse current through the photodiode.)

Given this convention on the voltage, or ΔV-axis, we call the substantial current that can be made to flow through a photodiode in the dark a positive forward current. (For the connections you have made between the photodiode and the TIA, this current has produced a negative output at the monitor point.)

Now with these conventions established, you will find that a photodiode, maintained at any constant level of potential difference ΔV, will display a photocurrent which is *negative* by our convention on current. The result is that the point in the i-V plane which represents the point of maximum power generation, ie. the point of maximum conversion of light into electrical energy, lies in the fourth quadrant of the plane.

Now that you know how to control and measure your variables, and what sign convention to use, you can gather data which will fill in points in the i-V plane. You can fit the data for an *un*-illuminated photodiode by the model discussed in Section 3.6,

$$i = i(\Delta V) = i_0 \, [\exp (\Delta V / V_0) - 1] \ .$$

The values of i_0 and V_0 you obtain ought *also* to describe the photodiode when it is illuminated, when your model for the current will need to include a (negative) offset proportional to the illumination level.

If you know enough about p-n junction diodes, you will know that the parameter V_0 is not merely an arbitrary fitting parameter for your diode. It is instead shorthand for a deeper understanding of the junction properties. Insofar as recombination is unlikely in the p-to-n transition regions in the p-i-n photodiode structure, the expected form of the i-V curve is

$$i = i(\Delta V) = i_0 \, [\exp (e \, \Delta V / (k_B \, T)) - 1] \ .$$

For the combination $(e \, \Delta V / (k_B \, T))$ to represent $\Delta V / V_0$, it must be that

$$V_0 = k_B \, T / e \ ,$$

where T is the (absolute) temperature of the photodiode during your measurements. Now you can see that V_0 has been elevated from a mere curve-fitting parameter to an indication of some of the physics of the p-n junction in the photodiode.

The model above will show deficiencies at larger positive currents (above 10 or 100 μA) because it does not yet include the effect of ordinary series resistance of the photodiode structure. You are free to improve the model to include this effect, at the cost of introducing another fitting parameter.

4. Noise as a function of temperature

4.1 Equipment, methods, and issues

This section explains the use of the 'thermal probe' and its associated Dewar vessel. Together they make it possible to measure Johnson noise from a source resistor as a function of its temperature T. The system is designed for use with liquid nitrogen (LN_2) as a coolant, and an electrical heater allows exploration above that base temperature. The probe is suited for use in the 77 K - 400 K range. The lower end is set by the normal boiling point of LN_2; the upper end (which is +127 °C) is set by temperature limits of wires and components in the probe head, and is enforced by the limited power available to the heater.

The motivation for this temperature coverage is of course the theoretically-predicted ($4\ k_B T\ R\ \Delta f$) **temperature** dependence of the mean-square Johnson noise voltage. Using the accessible temperature range, you'll be able to vary this quantity by a factor of 4 or 5.

SAFETY WARNING: The Dewar supplied is made of un-silvered glass to help you see the liquid level inside. Because it's made of glass, it will shatter if you drop it. The disaster will be even more dangerous if the Dewar is full of LN_2 when dropped. So: *do **NOT*** drop the Dewar, and use and store it only in the base built to hold it securely.

SAFETY WARNING: Liquid nitrogen is *very* cold, boiling at about -195 °C. It is dangerous to have it contact your skin, and even more dangerous to undergo skin contact with clothing soaked with LN_2. The hazard is not chemical, but physical. You can suffer frostbite, and permanent nerve and/or tissue damage, from the localized freezing that will occur.

There is also a special electronics issue involved with the use of the probe. The source resistors are now not built into the pre-amp module, but instead a few feet away. They are connected to the first stage of amplification by wires inside the low-level electronics box, and then by a coaxial cable over to, and down into, the probe. We have succeeded in preserving the required electrical grounding and shielding of those remote resistors against external electrical interference; but the inevitable cost is much larger capacitance between the 'live wires' and the shields. This capacitance (about 100 pF) has consequences on the bandwidth of the noise signals. *The Johnson noise is still spectrally 'white' **at its origin**, but its spectrum is already modified by capacitive effects when it reaches the first amplification stage.* Some solutions to the problem will be presented.

4.2 Two-temperature Johnson-noise measurement

This section teaches you how to prepare your system for measuring Johnson noise in 'remote resistors'. It pre-supposes that you've worked through Chapter 7 and Sections 1.1-1.5 on how to measure Johnson noise in 'local resistors'. The goal is to measure Johnson noise at two distinct temperatures: ambient and liquid-nitrogen.

The first thing you'll need to do is to confirm the installation of resistors into the probe. The unit is shipped with resistors R_A = 10 Ω, R_B = 10 kΩ, and R_C = 100 kΩ already installed, as you can see in the Figure 4.2a:

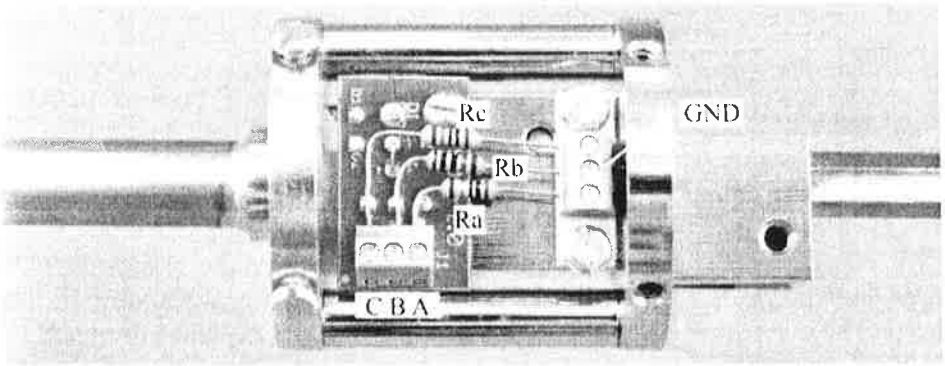


Fig. 4.2a: The interior of the temperature probe, showing the A, B, and C positions of source resistors.

Notice that to get this view, you have to loosen four screws, and remove four more, to slide away the shielding sleeve on the probe. The resistors' 'hot ends' are on the circuit board, and the ends near the copper flange are grounded. Once you've confirmed the resistors are present, you need to close up the shielding sleeve again.

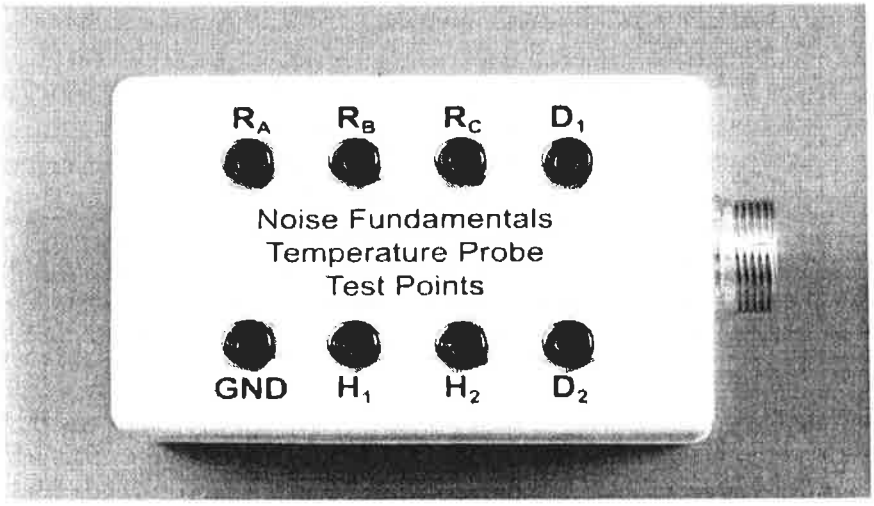


Figure 4.2b The 'breakout box'

To facilitate checking that all the components are properly connected inside the variable temperature probe, we have included with the unit a 'breakout box' shown in Figure 4.2b. This box connects to the variable-temperature probe and has 8 test points, one for each wire lead from the components in the probe to the connector. Points R_A, R_B, R_C and GND can be used to test the resistors with an ohmmeter, since all three resistors have a common ground. (Note that the 10 Ω resistor will likely read 12 Ω because of the resistance of the leads.)

The heater leads are present at H_1 and H_2 which measure about 75 Ω. The diode thermometer should be checked with a multimeter (on the diode-testing scale). It should read about half a volt, if the positive lead is connecter to either D_1 or D_2. Note that there is only one diode thermometer connected at the factory and *both* wires D_1 and D_2 are connected to it.

Internal wiring brings three 'live wires' from the three source resistors to a junction box atop the probe, and then via a cable to a connector for the Temperature module of your low-level electronics. Before you connect the probe to that module, open up the low-level electronics (by the familiar flip operation) to see what connections you need to make between the Temperature module and the Pre-amp module. The Figure 4.2c shows the necessary connections. In particular, you need three wires connecting the R_A, R_B, and R_C resistors to the A_{ext}, B_{ext}, and C_{ext} positions in the pre-amp. The 'other wire' of each of the three resistors is already grounded to the body of the probe, as the resistors are both thermally and electrically connected to the copper block through the terminal post.

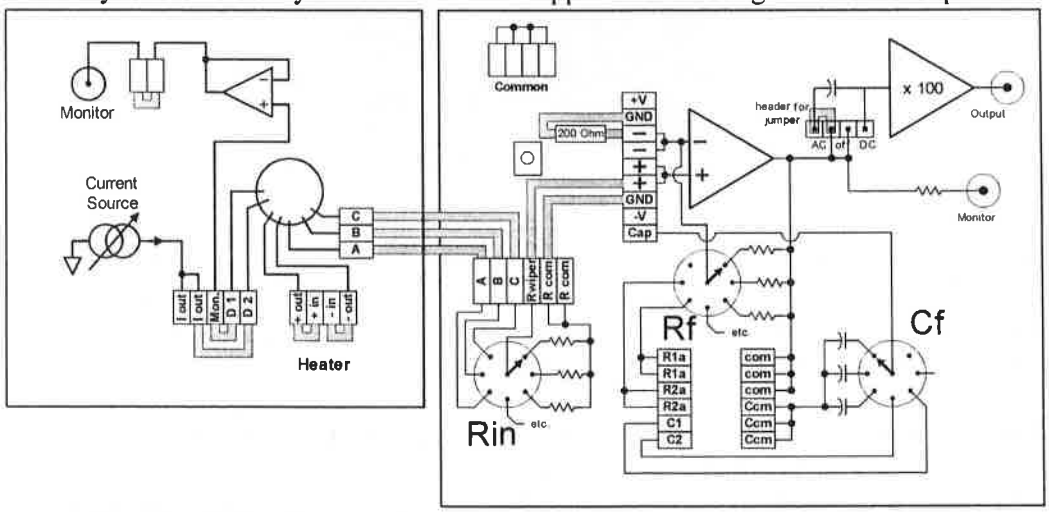


Fig. 4.2c: Wiring diagram of the interior of the low-level electronics, to bring the A, B, and C remote resistors to the A_{ext}, B_{ext}, and C_{ext} positions of the R_{in} selector, and to connect the temperature transducer and heater.

Also shown in the diagram are the connections you'll want to make for the temperature transducer, and the heater, in the probe. You'll need those devices in future sections. Re-flip the low-level electronics into its enclosure, confirm its power is ON, close up the box, connect the probe cable to the thermal module, and install the probe as shown in the photo below. Note that the Dewar is absent, and the whole probe is at ambient temperature.

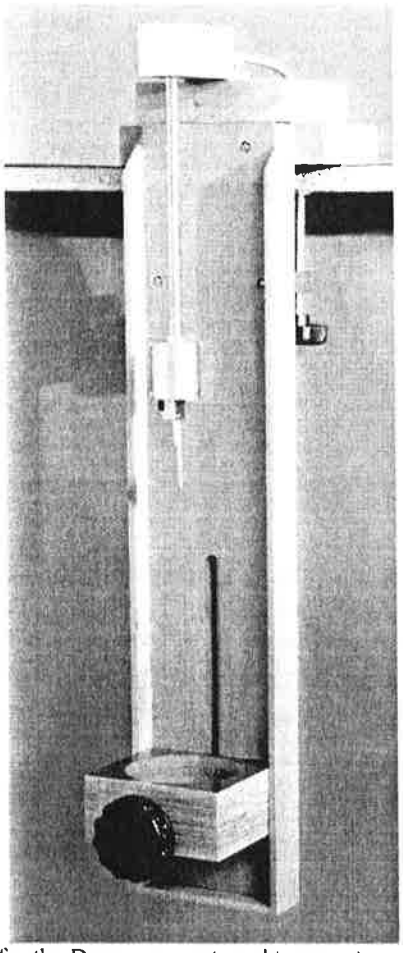


Fig. 4.2d: A mounting for the Dewar support, and temperature probe (with Dewar vessel removed).

Now you should be able to measure (room-temperature) Johnson noise from three remote resistors, just by using the A, B, C, positions of the source-selector (R_{in}) switch on the pre-amp. Other positions of this switch make available the noise from a set of 'local' resistors, including values of 10 Ω, 10 kΩ, and 100 kΩ.

For initial measurements, we suggest a bandwidth of about 10 kHz (set perhaps by using a 1-kHz high-pass, and a 10 kHz low-pass, filter). As usual, you'll need to recall that the standard gain is $G_1 = 600$ in the pre-amp (if that's in its default condition), and you'll need to use a suitable gain G_2 in the main-amp to get the squarer to operate in its optimal regime. As previously, here too you'll need to use the 10-Ω source resistor as a way to get the amplifier-noise contribution, which needs to be subtracted from the mean-square noise measurements.

It is important that you take data from both local and remote 10 kΩ and 100 kΩ source resistors, and also that you try some different bandwidths. Because of the effects of probe capacitance, it is to be expected that the values of $\langle V_J^2(t) \rangle$ you infer will be smaller for the remote, as compared to the local, resistors. The deficiency will be larger for the

larger source resistance, and a broader bandwidth. To see why this is expected, compute an RC time-constant for choices of R of 10 kΩ and 100 kΩ, now assuming $C \approx 100$ pF for connections to the probe resistors. Then compute a corner frequency of the undesired one-pole low-pass filter that results, from $f_c = 1 / (2\pi \tau)$. See Appendix A.8 for how to handle the consequences.

Use these (all room-temperature) results to decide on a measurement strategy that you'll use when the probe is *not* at room temperature. When you've worked that out, it is finally time to cool your probe. The photo above suggests how a (warm, and empty) Dewar can be slid into place, mounted into its movable base, and used to surround the probe. You can lower that base and the Dewar together and pour about 1 liter of LN_2 into the Dewar. [Go back to Section 4.1 and re-read the SAFETY WARNINGS we've posted there -- liquid nitrogen is a tool, or a hazard, **but not a toy**.] Wait for the boiling to subside, slide the black foam insulating cover down onto the Dewar's mouth, and now use the clamp on the Dewar's base to raise the Dewar until the probe makes contact with the LN_2. Here, as in general, the probe's sample chamber should end up at about the mid-height in the Dewar, and (for purposes of *this* experiment, exceptionally) also to end up with its copper bottom plate immersed in the liquid. (In later sections, you'll want only the brass 'cold-finger' on the bottom of the probe to be immersed.) The purpose is to ensure that your resistors really are at the temperature of your boiling LN_2.

When all the extra boiling has settled down, you can repeat your noise measurements, using both the local and the remote resistors, and using the protocol you've established. You may need to change the gain G_2 to keep the squarer in its optimal regime.

There's one more necessary measurement task. Your 'remote resistors' are of 1% tolerance, but that does *not* guarantee that their 77-K resistance matches their nominal value to this accuracy. So you'll want to check their 'cold resistance', ie. their R-values when they're immersed in LN_2. To get access to their electrical properties, we've supplied a 'breakout box', to which you can attach the cable of the probe, to get connections with all the items down inside it (see Appendix A.1). For remote resistors immersed in freely-boiling LN_2, you can not only be pretty sure of their temperature, you can also be quite confident that the diagnostic currents used by an ordinary ohmmeter will not warm the resistors significantly, as you measure their values.

Here's a final note -- you might get to this point, and for the first time have a cold probe immersed in leftover LN_2. Here are some suggestions for what to do at the end of a day's experimentation.

When you're done with your work, it might be a good idea to lower the Dewar's base, remove the Dewar, dispose of the surplus LN_2 by your locally acceptable method, and lay the Dewar down on its side to warm up. (Why is this better than leaving it standing vertically? If you do lay it down, do NOT let it roll away to its doom.) This removal will leave a cold probe in the open air, and you do NOT want to touch it -- contact with cold metal can lead to immediate frostbite, as well as the dreaded 'pump-handle effect'. Instead, leave the probe to hang in ambient air, and warm up spontaneously. If you're in a hurry, or if condensation of water onto the chilly probe is a problem, leave it in open

air, with the heater running, perhaps set to 5 or 6 turns on the dial. Then it will warm and eventually equilibrate to a safe-to-touch temperature, yet far enough above ambient temperature to ensure that it dries out properly.

Historical insight: Once you have data of the sort acquired above, here's one use you can make of your ambient-temperature and LN_2-temperature values for $\langle V_J^2(t) \rangle$. Put yourself back into the era of the Centigrade scale of temperature, on which ice melted at 0 °C and water boiled at 100 °C, by definition. On such a scale, your ambient temperature might be 22 °C, and your LN_2 temperature might be -195 °C (find an old reference which quotes you this value -- how do you suppose that it was established?). Now plot your two $\langle V_J^2(t) \rangle$ points as a function of the Centigrade temperatures at which they were measured. You have only two points, so of course you can fit a line to the two points. The pay-off is to find the $\underline{x}$-axis intercept of this line, as the extrapolated low-temperature point at which Johnson noise vanishes.

What you're doing is 'locating absolute zero' according to a noise-based measurement. It is a non-trivial technical, and intellectual, challenge to test whether Johnson noise extrapolates to zero at the same temperature at which the pressure of an ideal gas extrapolates to zero. Success in such tests suggests that the Kelvin scale is not just absolute, but also physics-wide. From a modern point of view, we depend on such a result to enable us to claim that the T-variable which appears in the Johnson-noise equation really is the absolute temperature, ie. the temperature measured relative to the absolute zero which is established by procedures such as these.

4.3 Temperature measurement and modeling

Once you know how to measure Johnson noise at two temperatures, it's time to think about how to control, and measure, a variety of other temperatures, so as to let T be an independent variable. Section 4.4 deals with temperature control in this apparatus; here we take up a method of temperature measurement.

First recall that the TeachSpin probe is intended for use (only) in the 77 - 400 K range. The lower end of the range is set by the open-air boiling point of LN_2. The upper end of the range is set by temperature tolerances of the devices and wiring in the probe. It is also (about) as high a temperature you can reach with the power-limited heater installed, even if you use the probe within a Dewar filled with nothing but air.

The probe comes with a 'transdiode' or 'diode-connected transistor' as the electrical transducer for temperature. The device is a pnp-type silicon transistor, electrically connected as shown in Figure 4.3a,

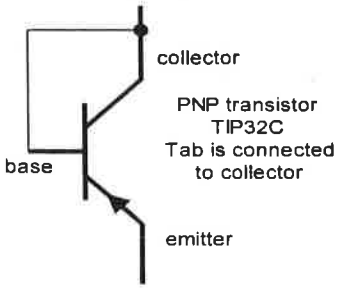


Fig. 4.3a: Schematic diagram for a pnp transistor in its transdiode configuration.

to create a device which very closely approximates the i-V response of an ideal two-terminal (abrupt-junction) p-n diode. The simplest mode of operation of such a device is to run a constant current (perhaps 10 μA) into the transdiode's emitter, to connect its base and collector to ground potential, and to record the potential difference across the diode. This voltage turns out to be a monotonically *decreasing* function of temperature from 77-400 K.

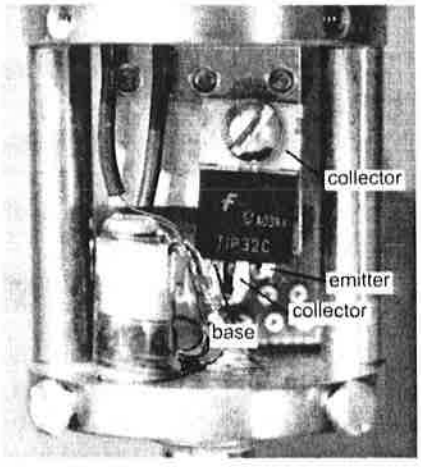


Fig. 4.3b: The transdiode in the temperature probe.

The Temperature Control module of the Low-Level Electronics makes the use of the transdiode as simple as possible. If your wiring matches that of Figure 4.2c, you can choose, by selector switch, the current with which the diode is excited, and read, using a digital voltmeter, an electronically buffered copy of the potential difference that results. (To what extent can you neglect the voltage drop that surely exists in the wires?) For a diode current of 10 μA, that voltage will be about 400 mV at room temperature, changing by about -2 mV/K with increasing temperature (ie. changing by about +2 mV/K with *de*creasing temperature). But the transdiode requires calibration before it can be turned into an actual thermometer.

Here's the simplest model which describes the i-V curve on which you rely for the calibration. The transdiode connection simulates a p-n junction diode with an ideality parameter $\eta = 1$, so that to a fine approximation we can write the diode current i as a function of potential difference ΔV and temperature T as

$$i(\Delta V, T) = i_0(T) \, [\exp (e \, \Delta V/(k_B \, T)) - 1] \, .$$

(See Sconza in the Bibliography.) It turns out that (except at the highest temperatures $T \approx 400$ K, giving the lowest values of $\Delta V \approx 135$ mV), the ratio $(e \, \Delta V/k_B \, T)$ is always large enough such that $e^{e\Delta V/kT} >> 1$, so we can drop the (-1) contribution in the result above. Then we have

$$\ln i(\Delta V, T) = \ln i_0(T) + e \, \Delta V/(k_B \, T) \, ,$$

or

$$\Delta V = (k_B \, T /e) \, [\ln i(\Delta V, T) - \ln i_0(T)] \, .$$

This allows us several ways to use the transducer, varying in their complexity and accuracy:

0) Using a 'look-up' table:
Included with the test documents in your manual is a table of transducer voltages expected at a series of temperatures in the 77 – 400 K range. The table of voltages is particular to the transducer installed in your temperature probe, and it is derived by a series of steps which make it accurate, to perhaps ±1 mV of potential, or about ±½ K of temperature, over the whole range of the table. The numbers in the table depend on several sources of information:

- the temperature calibration, by Lake Shore Cryotronics, of a secondary-standard temperature transducer, their DT-471 sensor;
- the use of that temperature sensor to calibrate, at TeachSpin, a standard version of the TIP32C transdiode;
- the confirmation at TeachSpin that such transistors do differ among themselves, but by a model which is adequately linear in the temperature;
- and finally a room-temperature measurement, ie. a single-point calibration, of the actual transdiode in your temperature probe.

We've used all these steps to generate the table included with your apparatus. Of course you can check this table against measurements on your sensor at two known temperatures: ambient temperature, and liquid-nitrogen temperature. (Learn your local *altitude's* effect on the boiling point of liquid nitrogen before you trust such a check.)

Using this table 'in reverse' to deduce indicated temperature from voltage-as-read will require some sort of interpolation, or fit. You might try making a plot of the tabulated data, and then use a local, two-point interpolation, or alternatively use a global fit motivated by some of the models below.

1) More 'primary' but less precise: According to the model above, we can use the transdiode, with*out* any external calibration, if we compare results at two known currents. So if we're at some fixed but unknown temperature T_x, and we record voltage ΔV_{10} for the use of current $10 \cdot i$, and also record ΔV_1 for the use of current $1 \cdot i$, then we can write

$$\Delta V_{10} = (k_B\, T_x / e)\, [\ln (10\, i) - \ln i_0(T_x)]\, ; \qquad \Delta V_1 = (k_B\, T_x / e)\, [\ln i - \ln i_0(T_x)]\, ,$$

and upon subtracting, we find

$$\delta(\Delta V) \equiv \Delta V_{10} - \Delta V_1 = (k_B\, T_x / e)\, [\ln (10\, i) - \ln i] = (k_B\, T_x / e)\, \ln(10)\, .$$

Hence in this model, we extract the result

$$T_x = [e\, \delta(\Delta V) / k_B]\, (\ln 10)^{-1}\, ,$$

which depends *on two measured voltages and two fundamental constants* (and nothing else, either at TeachSpin or at Lake Shore, except confidence in the diode model above!). Notice that the combination $e\, (k_B \ln 10)^{-1} = 5.04$ K/mV is the constant which maps the experimentally observed difference between two voltages to an inferred absolute temperature. Notice that the Temperature Control module makes it easy to make several i-vs.-$10i$ comparison. Be sure you try all the combinations, at least at room temperature, to see what confidence you can have in your results.

In this technique, you might worry about iR-drops in the wires to the transdiode, and inside the transdiode itself. You can also worry about self-heating in the diode – how much power is it dissipating at your highest currents?

One advantage of the subtraction-of-voltages in this method is that it cancels any d.c. offset voltage of the buffer amplifier which copies your transdiode voltage to your multimeter. Typical offsets are under 0.5 mV, but you might find a ±2-mV offset. In using the non-subtracting methods below, see if you can devise a way to measure the combined offset in your combination of buffer-amplifier plus multimeter.

2) Less 'primary' but more precise: The model we'll now introduce requires a model for the diode parameter $i_0(T)$, which is given (see Sconza in the Bibliography) by

$$\ln i_0 = \ln D - \frac{E_g^{(0)} - \alpha T}{k_B\, T} + (3 + \frac{\gamma}{2}) \ln T.$$

Here $E_g^{(0)}$ is the zero-temperature limiting value of the diode's band gap, and α is the rate of variation of band gap with temperature. The parameters D and γ are presumed to be constants. So in this model, we have

$$\ln i(\Delta V, T) = \frac{e\, \Delta V}{k_B\, T} + \ln D - \frac{E_g^{(0)} - \alpha T}{k_B\, T} + (3 + \frac{\gamma}{2}) \ln T.$$

Now we suppose that we use the transdiode at one fixed setting of current, so that (at any T-value) the left-hand side of this equation is a constant. Gathering all the constants, we are led to

$$\Delta V(i = const) = \frac{k_B T}{e}[const + \frac{E_g^{(0)}}{k_B T} - (3 + \frac{\gamma}{2}) \ln T] = \frac{E_g^{(0)}}{e} + \frac{k_B T}{e}[const - (3 + \frac{\gamma}{2}) \ln T] \quad .$$

This result in turn leads to two sub-models:

2a) Since the function ln T changes much less rapidly than does T itself, the equation above can be approximated as

$$\Delta V(i = 10\ \mu A,\ T) \approx c_1 - c_2\ T\ .$$

It turns out that c_1 and c_2 are positive constants in this way of modeling the transducer. The best feature of such a simple model is that the two unknowns in it can be found empirically by a simple 'two-point calibration'. The use of two adequately-known temperatures, such as ambient and LN_2 temperatures, gives two ΔV readings, and these together fix the values of c_1 and c_2 . Then the model is fully established, and of course can be inverted to give T from a measured value of $\Delta V(i = 10\ \mu A)$.

But this model is imperfect, since the neglect of ln-T terms has been added to any other approximations in the model. So though the model will reproduce (by construction) the right results near 77 and 295 K, it gives errors between the two calibration points, and also in the range above 300 K. The temperature errors can be as large as about 5 K; still they do not undercut the value of ΔV as a temperature *indicator*.

2b) To do a better job requires more information. The theoretical diode model above suggests that an equation

$$\Delta V(i = 10\ \mu A,\ T) \approx d_1 - d_2\ T\ - d_3\ T \ln T\ .$$

ought to fit the data better than the two-parameter model above. Of course it requires *three* items of input data to fix the values of three unknown coefficients. Here are two ways to find such data:

2b.i) If you have access to 'CO_2 snow', easily made just from compressed CO_2 gas, you could try to fill the Dewar with it, and then bury the probe into that 'snow'. The solid/gas phase change of CO_2 at one-atmosphere pressure gives a third temperature fixed point, neatly intermediate between the LN_2 point and ambient temperatures (you'll have to look up its location in temperature). Using it, together with your two previous fixed points, can give you a three-point calibration, ie. it will enable you to solve for all three constants in the model above. Experience suggests that such a temperature scale can yield a ΔV-to-T mapping with temperature errors of order 1 K or less, at least in the 77 - 300 K temperature range.

Some other temperature fixed point could do as well as this dry-ice point. But we do *not* recommend the use of the historic 'steam point' nominally at 373 K. That's because in a bath of steam, water will inevitably condense on and inside the probe, and will be conductive enough to throw into doubt the electrical measurements. (That's in addition to the safety hazards of piping live steam into the Dewar.)

2b,ii) Lacking access to a third fixed point, you can rely on someone *else* supplying you with extra information. If you rely on the behavior of your particular transdiode to follow that of general devices of its type, and you rely on TeachSpin measurements of such devices against supplier-calibrated temperature transducers, then you can try a model

$$\Delta V(i = 10\ \mu\text{A}, T) \approx d_1 - d_2\, T - d_3\, T \ln T\,.$$

Here the value $d_3 = 0.405$ (for T in Kelvin, and ΔV in milliVolts) is the mean result measured for some devices from a batch of nominally-identical transducers (from which batch *your* device was picked), and it is uncertain by about ±0.001.

If you take this coefficient as a given, you're again left with a two-free-parameter model, which you can establish using data from only two temperature fixed points. Notice that in this model, d_1 ought to give $E_g^{(0)}$, a band-gap constant characteristic of the transistor material; similarly, d_3 is expected to be a constant for devices made of a given material, hence for all the devices in a batch. Relative to this sort of three-parameter fit, the residuals are still not zero, and still show systematic variation, but it's variation that's 'worth' only about 1 Kelvin over the 77 - 400 K range.

******

So much on electrical thermometry and calibration -- here's a separate issue. At best, any of these diode models can tell you the temperature (of the p-n junction) of your transducer, but what you really want to know is the temperature of your Johnson-noise source resistors. These devices differ in their degree of thermal anchoring to the copper base of the probe. The first implication is that when the temperature is changing, it might change at different rates for the transducer vs. the resistors. The second implication is that if there's a vertical temperature gradient inside the probe chamber, the transducer and the resistors might be sampling distinct temperatures. A third worry is that the resistors, 'anchored' in temperature by their lower-wire connection to the terminal block, might have an internal temperature raised above that of the block, by virtue of the heat-flow that is reaching them from above, via their electrical 'live wire'.

What are the cures for these worries?

1) You can minimize them by design. In the TeachSpin probe, the heat flow down the 'live wires' is arranged *not* to flow through the source resistors, but to bypass them. The 'heat current' ought mostly to flow into ceramic 'thermal anchors' near the top of the probe body. Have a look at how this is done the next time you open up the probe's shield.

2) You can estimate, or bound, the size of these effects -- by noise thermometry(!). Suppose you've found ways to measure a noise signal, so a quantity proportional to $\langle V_J^2(t)\rangle$ is being displayed on your meter. If you use a large bandwidth and about a 1-s averaging time, you can certainly get better than 1% relative precision (if not accuracy) in this number. Now you can put the source resistor alternately in two conditions:

- entirely 'drowned' in LN_2, and certainly at 77 K, as compared to
- being held near 77 K, perhaps by having the probe's sample volume held above, and its cold finger held in, the volume of the LN_2.

If temperature gradients in the resistors are an issue, this a:b comparison ought to be sensitive enough to reveal, and even to quantify, them. The only drawback of this technique is that boiling LN_2 inside the probe's chamber creates bubbles, and 'microphonics' that might create excess noise.

4.4 Temperature control and management

Section 4.3 describes an electrical way to monitor the temperature of the copper base of the probe, and this section will teach you how that block can be arranged to have a temperature held fixed somewhere other than 77 K or ambient temperature. It's all done with a heater, which you can see is mounted to the copper bar across the bottom of the probe's body.

That heater is a power resistor, of resistance 75 Ω, whose two leads are brought all the way to the Temperature Control module in the Low-Level Electronics. In that module there's a special low-noise power supply capable of delivering a d.c. voltage, variable from 0 to 25 V, controlled by a 10-turn knob. Full power (at 10 turns clockwise) delivers 25 V, giving a heating power $V^2/R = (25\ \text{V})^2/75\ \Omega = 8.3$ W. At 5 (rather than 10) turns on the dial, the voltage is halved, and the power drops to about 2.1 W. Because of the V^2-law, you get rather fine control of the power levels at the low end of the range.

Here's a 'thermal circuit' showing how the heater can control the temperature.

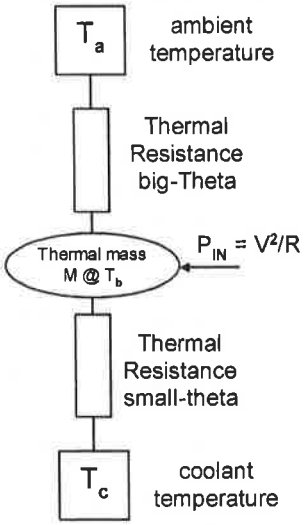


Fig. 4.4a: A 'thermal circuit' showing heat-flow paths from ambient temperature, to the sample block, and to the cold-finger.

There's a heat flow from the heater ($P = V^2/R$), into the copper fin that is under your direct control. There's another heat flow from the fin, down into the LN_2, which is partly under your control -- in particular, you can adjust the height of the Dewar so as to immerse, into the LN_2, either

a) the thin part of the brass 'cold-finger', or
b) the thick part of the brass cold-finger, or
c) the whole copper fin below the probe.

This gives you some control over the 'thermal resistance' θ between the probe and the LN_2. Meanwhile there's a much larger thermal resistance Θ between the probe's body and room temperature -- larger because of the use of a thin-wall stainless-steel tube of poor thermal conductivity to join the probe's body to its warm top.

Since thermal resistance is defined (parallel to $i = V/R$) via

heat flow = (temperature difference)/(thermal resistance),

we can model heat flows as soon as we define three temperatures:

T_a, the ambient temperature at the top of the stainless tube;
T_b, the temperature of the copper block;
T_c, the temperature of the LN_2 coolant, ie. about 77 K.

Steady-state requires that we have the heat flow into the block matching the heat flow out of it, giving us

$$P(heater) + \frac{T_a - T_b}{\Theta} = \frac{T_b - T_c}{\theta} \quad .$$

Solving for the block temperature, we get

$$T_b = (1 + \theta/\Theta)^{-1} \left[T_c + T_a \frac{\theta}{\Theta} + \theta \frac{V^2}{R} \right] \quad .$$

If we temporarily neglect any heat leaks down the stainless tube, we're making the $\Theta \rightarrow \infty$ approximation, and we get

$$T_b \approx T_c + \theta \, (V^2/R) \, .$$

To get temperatures in the 77 - 150 K range, this suggests the use of modest V and small θ (ie., a small thermal resistance $\Rightarrow$ large thermal conductance $\Rightarrow$ immersing the *thick* part of the cold-finger into the LN_2). To get temperatures in the 150 - 300 K range, this suggests the use of larger V and large θ (ie., immersing only the *thin* part of the cold-finger into the LN_2).

To get temperatures *above* ambient, we suggest starting at ambient and then surrounding the probe by a 'dry' Dewar, empty of LN_2. Now the cold-finger is irrelevant, and the 'heat current' from the heater will flow *up* the stainless steel tube, giving a result

$$T_b \approx T_a + \Theta \, (V^2/R) \, .$$

All of these models have entirely neglected convective and radiative heat flow, and none of them can be trusted quantitatively, but still they do offer useful guidance.

If heat flows do *not* match, then we're out of the steady-state case, and the block temperature will change. A typical thermal model for the block, with temperature $T_b(t)$, would have

$$M \frac{dT_b}{dt} = \Sigma P = \frac{V^2}{R} + \frac{T_a - T_b}{\Theta} - \frac{T_b - T_c}{\theta} \quad ,$$

where M is the 'thermal inertia' of the block, in turn the product of its actual mass and its specific heat. In the TeachSpin probe, the total mass of copper (block and fin, cover, top plate, and screws) is about 180 g, and the specific heat of copper (at least near ambient temperature) is 0.34 J/g K, so we get for the thermal inertia

$$M \approx (180 \text{ g}) \, (0.34 \text{ J/g K}) = 60 \text{ J/K}.$$

The differential equation above predicts the steady-state case if net heat flow is zero, and it predicts exponential approach to a new steady state if net heat flow starts out *not* zero. In principle, the time constants of such re-equilibrations can tell you something about θ, given the above estimate for thermal inertia M. In practice, it's easier to use this model in the short-term regime, where it predicts a linear behavior in $T_b(t)$.

Suppose you're done with data-taking at 77 K, after an initial cool-down by immersion, and you want to take data at a next temperature point near 100 K. You might use partial immersion of the thick part of the brass cold-finger into the LN_2. You might monitor the temperature of the copper block, and then apply full power (10 full turns, 25 V, 8.3 W) for a few minutes. Once you see a temperature rise, you can start to estimate a *rate* of temperature rise, and also estimate how long it'll take to reach your target 100 K at this rate. Once you get close to 100 K, you can do two tests, each taking a few minutes:

a) set $V = 0$, to give $P = 0$, and estimate the zero-power cool-down rate dT_b/dt;
b) set $V = 25$ V, to give $P = 8.3$ W, and estimate the full-power warm-up rate dT_b/dt.

Then you can interpolate between these two rates (the first negative, the second positive) to estimate the power level at which you'd get neither warm-up nor cool-down. If your interpolation were to predict that 3 W is the necessary power level, you could say

$$\frac{3\,W}{8.3\,W} = \frac{V^2/R}{(25V)^2/R} = \left(\frac{N}{10}\right)^2 \quad ,$$

and solve to find $N = 6$, so you'd set the dial to 6 turns (out of 10) to get the power level you need.

You'll soon learn to become a 'human servomechanism', adjusting (on about a 1-minute time-scale, making about 1-turn changes in the voltage) in order to stabilize at a target temperature. You'll find that as LN_2 evaporates away, lowering the liquid level on your cold-finger, that you'll either have to raise the Dewar vertically, or reduce the heater setting, to stay at a target temperature.

Remember that with the weak T^{1} dependence of mean-square Johnson noise on absolute temperature, you do *not* need to fixate on temperature fluctuations under a Kelvin in size, and you can tolerate some rate of temperature drift. Remember too that whatever precision you might attain in indicated temperature, this is not the same as temperature accuracy.

5. Calibrations

5.0 Specified accuracy

This section follows a noise signal through the modules of the Noise Fundamentals apparatus, and gives estimates for those uncertainties which affect measurement of the noise. The uncertainty estimates are meant to express **one standard deviation**. Note that sections 5.1 - 5.4 offer ways to perform calibrations that can reduce these uncertainties, if desired.

The first stage in the pre-amp has a gain set by two resistors, $g = 1 + R_f/R_1$. Here R_f is the feedback resistor, chosen by the front-panel selector switch, and R_1 is a resistor wired into a terminal block. As shipped, the apparatus has $R_1 = 200.\ \Omega$, that is intended to be used with $R_f = 1.00$ kΩ. Both are 0.1%-tolerance resistors, so $g = 6.00$ with perhaps 0.2% uncertainty, out to 100 kHz. The gain drops at high frequency, reaching a '-3 dB point' (where the gain has dropped to $6.00/\sqrt{2}$) at about 1.05 MHz.

The next stage of the pre-amp is hard-wired to give gain 100., again set by ratio of resistor values, themselves uncertain by 0.1% at most. So the gain of 100 is uncertain by about 0.4%, at least to 100 kHz. The gain drops at high frequency, reaching it -3 dB point (where the gain has dropped to $100/\sqrt{2}$) at about 1.6 MHz.

Signals are ordinarily sent on to the filter sections. Within the pass-bands of the low- and high-pass filters, their gains are 1.00, to uncertainty 0.3%. Their corner frequencies are uncertain by 1-2%, though the *ratios* of corner frequencies achieved by settings of the selector switches are accurate to 0.3%. (That's because the switches change selections among resistors of 0.1% tolerance, whereas the fixed-value capacitors are of 1% tolerance.) The filters are of Butterworth response, with damping parameter $\gamma \equiv 1/(2Q)$ (see Section 5.2) fixed at the value $1/\sqrt{2}$ to uncertainty 0.2%. The exception is the use of the 33- and 100-kHz corner-frequency settings of the filter. In this case, finite-gain effects in the operational amplifier have the effect of lowering γ, or raising the 'Q', by more than the 0.2% uncertainty quoted above.

Filtered signals encounter more gain in the main amplifier. The gain sections (x1 or x10) and (x10, x15, . . . x100) each give gains correct to 0.2%, set by resistor ratios, at least to 100 kHz. The gains drop at higher frequencies, with a -3 dB point at 1.4 MHz (except that for the use of the highest gain setting (x10^4), there is some 'gain peaking' near 1.5 MHz).

The squarer is subject to zero-offsets (see Section 5.3), but its static squaring accuracy is claimed by the manufacturer to be about 0.2%. Its large-signal bandwidth extends to beyond 3 MHz.

The time-averaging section which drives the output and the panel meter is optimized for correct behavior at d.c., where its gain is 1.00 to 0.1% tolerance. (Separate from this is any d.c. offset.) The filtering behavior is that of two one-pole filters in series (both with buffered outputs). The time constants listed on the panel selector switch are accurate to 5%.

The source resistors (R_{in}) and feedback resistors (R_f) available at the pre-amp's front panel are of 0.1% tolerance (from 10-Ω through 1-MΩ, inclusive); other resistors available at these locations are of 1% tolerance. Resistors in the 'parts box' have the tolerances listed in its table of contents. The 'remote' source resistors intended for use in the Temperature Probe are also of 1% tolerance. The temperature (in)dependence of resistance of these devices is *not* warranted to be within 1%, certainly not down to 77 K. In practice, the source-resistance values have to be measured as functions of temperature, and are typically found to change by the order of a per cent over the whole range 77 - 400 K.

5.1 Calibrating amplifier gains

Noise measurements require knowledge of the pre-amp gain G_1 and the main-amp gain G_2. This section describes how those gains can be measured, ***if calibration is desired.***

Each of the amplifiers has 'sections'. The pre-amp has an input stage that is factory-set to have a gain of 6.00 (though this gain can be changed -- see Appendix A.4). Then follows another gain stage of gain 100. The main amp can be tested modularly as well, since it can be arranged to have G_2 = (1, or 10) x (1, or 10) x (10, or 15, or . . . 100). What follows is a procedure for checking the gain of an amplifier whose nominal gain is (10 or below) or (100 or below). The procedure requires a signal generator, capable of ±10-V sine wave output, and a 2-channel 'scope. The basic idea is to use the special 'Signal Attenuator' in the low-level electronics box to create a precisely-known *attenuation*, by factor either 0.1 or 0.01, and then to amplify that reduced signal back up to near-original strength. The combined schematic and wiring diagram for these experiments is shown in Figure 5.1a.

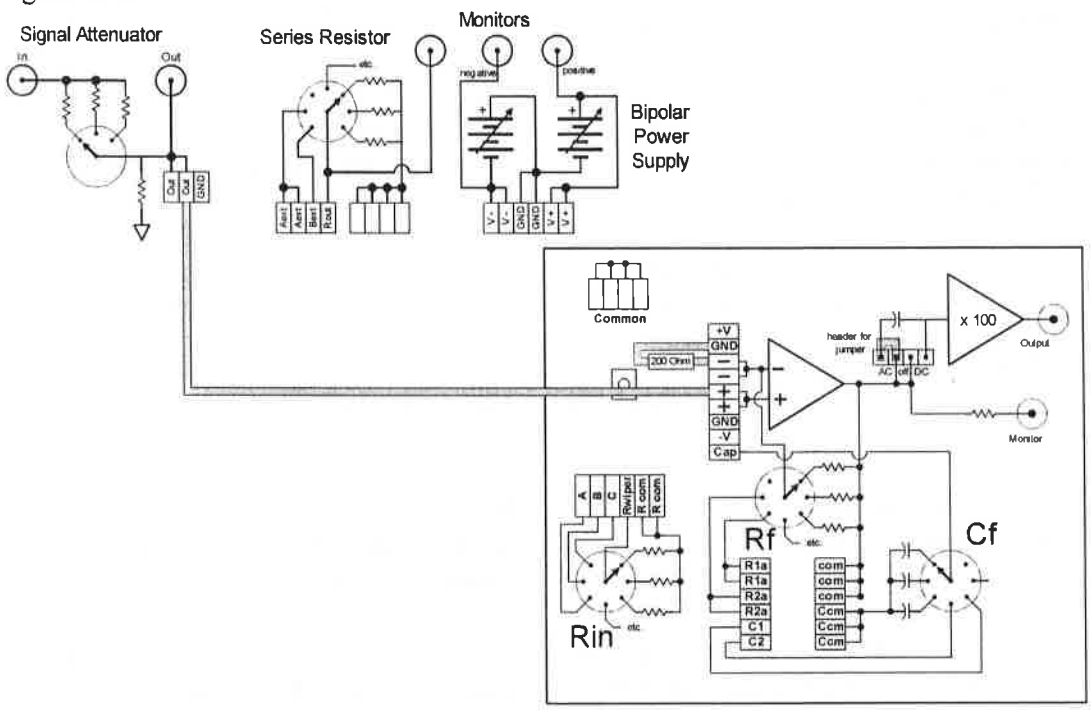


Fig. 5.1a: Schematic and wiring diagram showing the use of the Signal Attenuator for checking amplifier gain. Note that the Attenuator output is carried to the device-under-test by distinct methods: to the pre-amp sections via wires behind the LLE panel, but to HLE modules via external coaxial cable.

If $V_{gen}(t)$ is the generator signal, that's what the 'scope's ch. 1 is arranged to see. The attenuator creates an output of (say) $0.01 \cdot V_{gen}(t)$, where we can trust the coefficient 0.01 in two senses:

- we trust its value to be accurate, because it's created by a purely passive voltage divider, built from resistors of 0.1% tolerance;

- we also trust its value to be frequency-independent to very good precision, because it arises from low-impedance resistive components. Even the typical cable and 'scope capacitances scarcely matter given the 10-Ω output impedance of the divider.

Now if we have an amplifier with nominal gain 100, it's clear we'd choose the 0.01-setting for the attenuator, and then pick a generator drive level such as 1-Volt amplitude. We would measure the input signal at ch.1 of the 'scope, and use the 'scope's ch. 2 to monitor the amplifier output. If the ch.1 and ch.2 (output) traces matched exactly, we'd know $G = 100$ exactly. (This assumes the two channels on the 'scope are matched.)

The value of this procedure is that it's easy to vary the generator frequency from 10 Hz to 100 kHz or beyond, and use the 'scope to execute this procedure at any chosen frequency. In fact, with this method it's easy to find the '-3 dB point' of an amplifier, going up in frequency until the output amplitude drops to $1/\sqrt{2} \approx 0.707$ of its expected value. (At this point, one would expect a sizeable phase shift as well.) And it is important to verify that each stage of amplification is 'flat in frequency' out even as far as 1 MHz.

The only drawback of this procedure is that it depends on the accuracy of the 'scope calibration. In practice, this is rarely guaranteed even at the ±2% level, so to get amplifier gain measurements to better than 1% needs to take this into account. There are at least two ways to deal with this:

First, you could use one single channel of your 'scope, connecting it alternately to the filter input and output, so the 'scope calibration error is a common effect which cancels out in the ratio.

Second, you could use a clever permutation method which can compensate for scale errors in a 'scope:

We suppose that 'scope ch. 1 can be modeled by a scale error, $V^{(1)}_{\text{indicated}} = E_1 \, V^{(1)}_{\text{actual}}$, where ideally $E_1 = 1$ (but actually, E might depart from one). Similarly for ch. 2 there's a scale-error constant E_2. Now in the test configuration shown above,

- 'scope ch. 1 gets signal $V_{\text{gen}}(t)$ and indicates, ie. measures, a result $E_1 \, V_{\text{gen}}(t)$;
- 'scope ch. 2 gets signal $(0.01) \, G \, V_{\text{gen}}(t)$ and measures result $E_2 \, (0.01) \, G \, V_{\text{gen}}(t)$.

The output/input ratio of these two (amplitude) measurements gives a datum,

$$d^{(a)} = [E_2 \, (0.01) \, G \, V_{\text{gen}}] \, / \, [E_1 \, V_{\text{gen}}] = (E_2/E_1) \, (0.01 \, G).$$

Now the gimmick of the permutation method is to do a 'part b' of the experiment, just by swapping the two connections of signals to chs. 1 and 2 of the 'scope. In this new configuration,

- it's 'scope ch. 2 which gets signal $V_{\text{gen}}(t)$ and measures a result $E_2 \, V_{\text{gen}}(t)$;
- and it's 'scope ch. 1 which gets a signal $(0.01) \, G \, V_{\text{gen}}(t)$ and measures a result $E_1 \, (0.01) \, G \, V_{\text{gen}}(t)$.

Now the output/input ratio of the two measurements gives a new datum,

$$d^{(b)} = [E_1\ (0.01)\ G\ V_{gen}] / [E_2\ V_{gen}] = (E_1/E_2)\ (0.01\ G).$$

So the a- and b-methods each produce a datum, and the *product* of these two values is

$$d^{(a)}\ d^{(b)} = (E_2/E_1)\ (0.01\ G) \cdot (E_1/E_2)\ (0.01\ G) = (0.01\ G)^2\ .$$

So *in*dependent of scale errors of the 'scope channels,

$$[d^{(a)}\ d^{(b)}]^{1/2} = 0.01\ G\ ,$$

from which G can be found. (Meanwhile the *ratio* $d^{(b)}/d^{(a)}$ gives $(E_1/E_2)^2$, whose departure from 1 is a measure of 'scope calibration mismatch.)

This method doubles the amount of effort required to measure any G-value, and (since there's no assurance that E_1 and E_2 are constant in frequency), it has to be repeated at each new frequency. It's also important to realize that scale errors like E_1 and E_2 might change with each new range setting of a 'scope's input sensitivity; but the method above is typically measuring signals of similar size at ch. 1 and ch. 2, so there's no need to change range settings when doing the permutation.

One last note: the Signal Attenuator is built to accommodate the calibration of some modules having a 1-kΩ input impedance, and other modules having input impedance of ≥10 kΩ. For its attenuation factor to be reliable to the 0.2% level, it is incumbent on you to know what input impedance the attenuator is 'looking at'. (See Appendix A.1 for the input-impedance values.) If you set the Z_{adjust} switch appropriately for the device you are driving, you will get the attenuation level you expect, to a precision of ≤0.2%. If you forget this issue, an error of a full 1% can result.

5.2 Calibrating filter gains and bandwidths

Section 2.2 describes how a filter's gain profile $G(f)$ sets the equivalent noise bandwidth Δf according to

$$\Delta f \equiv \int_0^\infty G^2(f)\, df \quad .$$

This assumes that the rest of the amplification in the system is frequency-independent. Since the predictions for mean-square noise in the Johnson- and shot-noise cases depend linearly on Δf, it is necessary to be sure of Δf's value to the target precision of the answer. In this section, we discuss how the frequency-dependent gain $G(f)$ can be measured and modeled, to allow this sort of computation of Δf.

We discuss first the construction of the filter sections in the high-level electronics. We've chosen analog operational-amplifier active filters of the 'two-pole state-variable' design. Each filter uses fixed capacitors, and switch-selected resistors, to define the corner frequency. (That's because it's much more feasible to get resistors of precisely-known ratio of values than capacitors.)

A state-variable filter of this type is precisely analogous to a damped simple-harmonic-oscillator system, driven by an input voltage $V_{in}(t)$. The three outputs of the filter (low-pass, band-pass, and high-pass) are analogous to the position, velocity, and acceleration variables of the oscillator. The 'resonant frequency' of the analogous oscillator is determined, in the electronic circuit, to be

$$\omega_0^2 = 1/(R_1\, R_2\, C_1\, C_2)\,,$$

where the R's and C's are resistor and capacitor values in the filter circuits. From this it's plausible to believe that corner frequencies f_c ($= \omega_0/2\pi$) will 'track' to ≈0.2% precision from range to range (since the resistors involved are of 0.1% tolerance). But the actual values of the corner frequencies will depend on the actual capacitor values, and these components have only 1% tolerances. That's why the nominal corner frequencies lead to computed bandwidths with about 2% uncertainty. To get smaller uncertainty is feasible, but it requires more effort. Because of the fixed-capacitor, switched-resistor design, it is probably sufficient to calibrate each of the two filter sections for only one setting of its 6-position corner-frequency selector. (The exception is the choice of 100-kHz corner frequency, where the issue of finite gain in the op-amps complicates the modeling.)

We now introduce simpler, and also more complicated but more precise, ways to measure and model the filter responses. The simpler methods are intended to give 'spot checks' which establish certain filter parameters.

One spot check you can make is to find an empirical value of the 'corner frequency' f_c. Here are two features of f_c which make it detectable; both methods assume sinusoidal excitation of the filter:

- the low-pass response of the filter has gain (-)1 for $f << f_c$, but the gain drops, ideally to value $1/\sqrt{2} \approx 0.7071$, when the frequency is increased to $f = f_c$.
- the band-pass response of the filter has an output vs. input phase shift near 90° for $f << f_c$. But the phase at $f = f_c$ becomes 0°.

Both these effects can be seen by using a sine-wave signal generator to drive the filter input and ch. 1 of a dual-trace oscilloscope, and a selected filter output to drive ch. 2 of the 'scope.

To measure gains independent of phase shift, it's ideal to use the measurement capability of the 'scope, perhaps by using the rms-measure of ch.1 and ch. 2 signals. For any frequency well above 10 Hz, it is also best to use a.c. coupling of the 'scope, to block any d.c. components. The filters are linear systems, so in principle any amplitude of drive could be used to diagnose them; in practice, a drive of $\leq$1 V amplitude is suitable. Of course it's best to pick an amplitude, and a 'scope sensitivity, so that the traces nearly 'fill the display' of the 'scope.

To measure phase shift independent of gain, you can overlay ch.1 and ch. 2 traces. Only when the phase shift is zero will the zero-crossings of the two signals coincide. Or, you can view the same two signals in an XY-display mode. Only when the phase shift is zero will the generic elliptical locus collapse to a line.

For some of these methods, it may pay to choose the d.c.-coupling option at the input of the filter under test. See Appendix A.2 for discussion of this matter -- you might still want to revert to a.c. coupling during actual noise measurements.

Both of the methods above depend on the gains, and the phase shifts, of the two 'scope channels themselves being perfectly matched. In general, this is *not* guaranteed, certainly not to 1% tolerance. So the input/output alternation method, or the ch.1/ch. 2 permutation method, of Section 5.1 can be used to cure the problem. If you have a digital multimeter whose measurement precision for a.c. rms voltages can be trusted to the highest frequency you need, you can use this in place of a 'scope for voltage-magnitude measurements.

Given either of these 'spot check' estimates of f_c, you can make a 1-parameter model of the low-pass filter as a frequency-dependent gain

$$G_{LP}(f) = \frac{1}{\sqrt{1+(f/f_c)^4}} \quad .$$

But there is a more general model of how the filter might behave, given by the general response of a two-pole system,

$$G_{LP}(f) = \frac{g}{\sqrt{(1-(f/f_c)^2)^2 + (2\gamma\, f/f_c)^2}} \quad .$$

This is a 3-parameter model, still with a 'corner frequency' f_c. The new parameters are a 'd.c. gain' or $G_{LP}(f=0)$ value, given by g, and a damping parameter γ. Ideally the filter will have $g = (-)1$ and $\gamma = 1/\sqrt{2}$, in which case the formula above reduces to the previous case. But here are some methods to confirm the values of g and γ.

Since g is the d.c. gain, it can be measured right <u>at</u> d.c. To deal with the issue of d.c. offsets, you can use two d.c. values in succession at the input, and measure a ΔV_{in}; and then look at the two resulting d.c. values in succession at the output, to measure ΔV_{out}. (You'll certainly need to select d.c. coupling at the input of the filter to perform this

measurement.) One advantage of this method is that you can measure all the voltages with d.c. multimeters, whose resolution and accuracy exceed what a 'scope can offer.

If you know the d.c. gain g, and you found the corner frequency f_c via the phase-shift method, there is also an easy 'spot check' of the damping parameter γ. It turns out that the low-pass, band-pass, *and* high-pass outputs all have their gain take on the value $g/(2\gamma)$ at the frequency $f = f_c$. So a 'scope-based measurement of the gain at this one frequency will provide a value for γ.

If you use these spot-check methods to measure f_c (and g and γ), you can of course repeat them for each setting of the filter's corner frequency. If you find a deviation of order 1~2% between a nominal corner frequency on the selector switch, and the measured value of f_c, this is not unexpected. But the *ratios* of nominal corner frequencies (as in 33 kHz/10 kHz) and of measured corner frequencies (as in 33.23 kHz/10.07 kHz) ought to be equal to about 0.2%, since these ratios are controlled by resistance values of 0.1% tolerance.

Once you've taken the trouble to set up a generator-and-meter or generator-and-'scope technology to measure spot values of the gain $G(f)$, you're also in position to do a survey. It turns out that you can sample the $G(f)$ function adequately if you take values of $G(f)$ for several frequencies f well under f_c, several more for f in the vicinity of f_c, and several more for f well above f_c. (It is the choice of a filter design that follows a simple and computable model which permits you this luxury.) Try plotting the theoretical function $G_{LP}(f)$, for several choices of g, γ, and f_c values, preferably on a log-log scale, to see where these parameter choices make what difference. Such a plot will also show you why covering the range $0.1 \cdot f_c$ to $10 \cdot f_c$ is probably sufficient.

Given a collection of $G(f)$ data points (taken for any of the three outputs of the filter), you can perform a least-squares fit of the data to the theoretically-expected forms, which are

$$G_{LP}(f) = \frac{g}{\sqrt{(1-(f/f_c)^2)^2 + (2\gamma\, f/f_c)^2}} \quad ;$$

$$G_{BP}(f) = \frac{g\,(f/f_c)}{\sqrt{(1-(f/f_c)^2)^2 + (2\gamma\, f/f_c)^2}} \quad ;$$

$$G_{HP}(f) = \frac{g\,(f/f_c)^2}{\sqrt{(1-(f/f_c)^2)^2 + (2\gamma\, f/f_c)^2}} \quad .$$

Ideally, the values of g, γ, and f_c in all three functions is the same -- but this would bear checking. You might, by this fitting method, be able to attain 0.1% precision in each of the values g, γ, and f_c. (Assuring 0.1% *accuracy* is harder!). If you measure the 33-kHz and 100-kHz filters carefully, you might see γ drop below its intended value of $1/\sqrt{2}$; this is an effect attributable to the finiteness of the open-loop gain of the op-amps used in the filters.

There is a single-point calibration of the filters included on your data sheet that comes with your instrument. For an ideal two-pole filter of Butterworth design, the intersection of all three response curves occurs at a single frequency, f_c. The gain at each of the outputs at this frequency is given by $1/(2\gamma)$, which is the Q-value of the filter. In practice, the three gain curves do *not* intersect at one common frequency. We have listed, as the nominal corner frequency, the point of intersection of the (fast-varying) low-pass and high-pass gain functions. And we have listed, as the nominal Q, the (locally flat) gain of the band-pass filter at this particular frequency. Numbers thus defined seem to agree, to better than 0.5%, with the corner frequencies and Q's obtained from full best-fit models of each of the filter response functions (where separate fitting parameters for f_c and γ are used in separate models for the three filter outputs).

Finally, it's worth remembering why you <u>care</u> about the filters' $G(f)$ functions. The goal in these calibrations is to make possible the computation of the integral of the (net) $G^2(f)$ function to get the equivalent noise bandwidth. In general, you might be using two filters, perhaps a high-pass filter with corner down at f_1, and another low-pass filter with a corner up at f_2, and it's the product of the filter functions, $G_{HP}(f; f_1) \cdot G_{LP}(f; f_2)$, that you need to square and integrate to get the equivalent bandwidth. That kind of integration is best done numerically. But here's an analytic result that has the value of providing a test case for checking such integration routines, and also of showing you how the results will scale with the parameters in the model. In terms of the three-parameter models given above, the integrals over all frequencies of both $G_{LP}^2(f)$ and of $G_{BP}^2(f)$ can be done exactly, and they turn out to be equal:

$$\int_0^\infty G_{LP}^{\,2}(f)\, df = \int_0^\infty G_{BP}^{\,2}(f)\, df = g^2 f_c \frac{\pi}{4\gamma} \quad .$$

For the intended values $g = 1$ and $\gamma = 1/\sqrt{2}$, these reduce to the result

$$f_c \cdot \pi/(2\sqrt{2}) \approx 1.1108 \cdot f_c$$

previously quoted. But for the general case, this result will teach you that a target precision of 1% for the bandwidth will require achieving about ¼% precision in d.c. gain g, and about ½% precision in measuring the corner frequency f_c and the damping parameter γ.

5.3 Calibrating the squarer

In almost all the noise measurements described in this manual, the Multiplier module in the High-Level Electronics is used in its AxA mode as a 'squarer'. The squarer is essential in forming the mean-square value of a signal, and you've seen it modeled by the expression

$$V_{out} = (V_{in})^2/(10\ V)\ .$$

In practice, you might worry that this model leaves out several effects, such as d.c. offsets at the input, d.c. offsets at the output, scale-factor errors, and square-law fidelity. Here's how to check on such matters.

The simplest model for an input-side d.c. offset would be

$$V_{out} = (V_{in} - a)^2/(10\ V)\ .$$

In practice, the use of a.c. coupling in earlier stages of the signal chain makes this issue nearly irrelevant, but it'll appear soon in a more general model that you can test.

The simplest model for an output-side d.c. offset would be

$$V_{out} = b + (V_{in})^2/(10\ V)\ ,$$

and this effect can*not* be ignored. That's because the typical use of the squarer is to deliver a signal whose d.c. average value is what you're recording to form a measure of the mean-square noise. Fortunately, in this model, it's easy to find the value of b -- you just use the AC/GND/DC switch at the A-input, and set it temporarily to the ground position. This enforces $V_{in} = 0$, and the V_{out} which you read at the output now tells you the value of b. This offset ought thereafter to be subtracted from all the squarer's output values.

The scale factor is the coefficient we've written via a denominator of 10-Volt value in the model, and errors in its value can be represented by

$$V_{out} = k\,(V_{in})^2/(10\ V)\ ,$$

where we expect $k = 1$ (but we need to correct every measured value of the mean-square voltage, if $k \neq 1$). Perhaps the best way to check this is to use d.c. coupling at the A-input, and then send a variable d.c. voltage into the A-input, recording what d.c. value you get at the MONITOR or the final output. The model we'd use to fit the data is

$$V_{out} = b + k\,(V_{in} - a)^2/(10\ V)\ ,$$

which is just a quadratic function, easily accommodated by least-squares fitting. It is important to take data with V_{in} varying on both sides of zero, and for it to cover the full ±10-Volt range. The coefficient of the quadratic term in the fit gives $k/(10\ V)$, according to the model above. If the fit tells you that $k = 1.01$, then all your mean-square values are 1% too large, due just to this scale-factor error.

The same data-set allows a check of the square-law variation. Perhaps the best check is to plot the { (V_{in}, V_{out}) } data set, to plot atop that the best-fit parabola, and then to plot the 'residuals', ie. the differences between the data and the model. If there are systematic departures from a power-2 or square-law dependence, this 'residuals' plot would be the place to see them.

Finally, there are issues of speed of response, or bandwidth, of the squarer. Thus far, all the calibrations have been conducted at d.c., ie. at negligible frequency. But the squarer is in practice used at frequencies up to 100 kHz and beyond. The manufacturer claims a bandwidth of 10 MHz for the device itself, though the drive circuitry in the NF-1 limits this. Here's a way to test that high-frequency response. You can drive the A-input with a sinusoid of 10-V amplitude, and variable frequency. You can also send that drive signal to a 'scope's ch. 1 input, meanwhile sending the squarer's output to the ch. 2 input. Now using an XY-display on the 'scope will show a parabola as the locus of V_{out} vs. V_{in} values.

If you set the generator to frequency 1 Hz, you'll see a slowly-moving spot on the 'scope's display. Going to 10 and then 100 Hz will seem to give a continuous parabola on the display. The qualitative test of bandwidth is to change the drive frequency from (say) 100 Hz to 100 kHz, to see if any part of the parabola changes. If you see a forward trace and a return trace of the parabola which fail to overlap, you'll see a 'double line' replacing the previous single curve. Even this fault would really tell you only about phase shifts. To persuade yourself that this test can reveal the squarer's limitations, increase the frequency to 1 or even 10 MHz, and see if you can detect what happens.

5.4. The 'noise calibrator'

This section describes the 'noise calibrator', a built-in part of your high-level electronics box, whose output emerges at a ***BNC jack on the rear panel***. This is a source of (pseudo-) noise whose frequency spectrum, and whose spectral density, can be very well known. The purpose of this source is to make possible various kinds of 'reality checks' on your noise measurements.

Ordinarily it is *inactive* (lest it generate a stray signal that might interfere with some noise experiment). The noise calibrator can be activated by the toggle switch on the rear panel. A front-panel red LED indicator will remind you that the noise calibrator is now running.

Convey the noise-calibrator's rear-panel output, via a BNC cable, to an oscilloscope. You should see a signal lying in the ±800 mV range, and you might first view it using 50 μs/div on your time axis. A series of sweeps of your 'scope will show a waveform something like this figure:

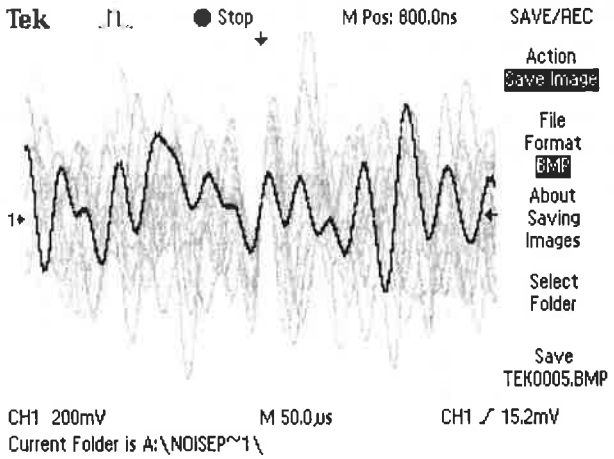


Fig. 5.4a: Sample waveforms from the Noise Calibrator. Vertical scale 200 mV/div, horizontal scale 50 μs/div, triggering on positive-going zero-crossings.

To get a better look at this waveform, you might want to set a trigger level of +500 mV, and slope positive. Then you can look for what happens in the 250 μs (5 divisions) before, and after, the occurrence of relatively rare positive-going large excursions of the signal.

This waveform has ***some*** of the characteristics of noise -- it looks random and non-periodic. But it is ***not broadband*** noise; instead, it has been constructed to give a (simulacrum of a) white-noise spectrum in the 0 - 32 kHz range, but to have almost no content above 32 kHz. This makes the waveform look quite different (less 'fuzzy') compared to the (unfiltered) noise waveforms you've been seeing, which have frequency content out to >1 MHz, and consequently have time variation on a <1 μs time scale.

The waveform you're seeing might have a measurable d.c. average value (near zero), and it also has a root-mean-square deviation from zero (or from its d.c. average). There are three possible ways to measure this rms value:

1) You can use the main-amp and squarer built into the NF-1 instrument, and quantify the noise-calibrator output just as you have previously quantified noise waveforms.
2) You can use a digital multimeter, if it is rated to compute 'true rms' values and for frequencies as high as 32 kHz.
3) Some digital oscilloscopes will measure this 'rms value'. To get a good measurement, you'll need to pick a vertical-axis setting which ensures that only the rarest of large excursions, if any, go beyond the measurement range of your 'scope's display. Also, since there are noise components out to 32 kHz (but no higher), you'll succeed, provided you acquire samples at a rate of 10^5 samples per second (or higher). If your 'scope gives you a choice, ask for 'full screen rms' and not 'cycle rms'. The values you get will display fluctuations, which you (or the 'scope) can average to give a converged value. The reading R that you get can be written as

$$< [V_{\text{calib}}(t)]^2 >^{1/2} = R \, ,$$

whose units are Volts. (It might be *called* 'rms Volts', but there is only one kind of Volt, equal to one Joule per Coulomb; what the notation means is that the rms measure of the waveform is being found, in units of Volts.)

By any of these methods, you should find a value near $R = 213$ mV. Clearly R^2 is a number, in Volts-squared, which gives $< V_{\text{calib}}^2(t) >$, the mean-square value of the output.

Now since the spectrum of the noise is designed to be uniform in frequency (in the range $0 < f < 32$ kHz), but limited in coverage. (If you have a 'scope with FFT capability, and know how to use it, you can view the frequency spectrum of the noise-calibrator output, to see how it is restricted to ≤32 kHz.) Given knowledge of the whiteness of the noise spectrum, and its rms value, and its frequency range, you may conclude that the 'noise power spectral density' (within that frequency range) obeys

$$S = < V_{\text{calib}}^2(t) > / \Delta f = R^2 / (32 \text{ x } 10^3 \text{ Hz}) \, .$$

If you measure a value $R = 0.213$ V, then $S = 1.42 \text{ x } 10^{-6}$ V^2/Hz. Most noise practitioners want to quote or remember not S itself but rather its square root, often called the 'voltage noise density', $\sqrt{S} = D = 1.19 \text{ x } 10^{-3}$ V/$\sqrt{\text{Hz}}$ = 1.19 mV/$\sqrt{\text{Hz}}$. (Now you know where those exotic units, Volts-per-root-Hertz, come from.) Of course, both these densities apply in the $0 < f < 32$ kHz band, and drop to much lower values beyond this frequency range.

What's the noise calibrator *for*? It is to give you a chance to check your understanding of analog filtering and squaring, or of digital recording and mathematical transformation, of 'real noise' signals. In the same spirit, if you had devised some complicated circuit alleged to measure voltage, you might like to attach to it a plain old dry cell to see if a result such as 1.51 Volts would emerge; the dry cell would be serving as your 'voltage calibrator'.

******

Here's a first example you can try. Take the noise calibrator output, and send it into a single filter section (*no* pre-amp required!); set it to be a low-pass filter of corner frequency 3.3 kHz. Then send that output into the main gain stage, and the amplified signal $V_A(t)$ into the squarer. Here's how an experienced practitioner would reason to predict the result: Start with a voltage noise density $D(f)$ equal to 1.19 mV/√Hz, say, out to 32 kHz, and equal to 0 beyond there. Then send it through a filter, whose gain function $G(f)$ varies from 1 (below the corner) toward 0 (above the corner). After further amplification by gain G_2, you expect an amplifier-output signal $V_A(t)$ whose noise density is now given by $D(f)\ G(f)\ G_2$. The mean-square measure of this signal is

$$\left\langle V_A^2(t) \right\rangle = \int_0^\infty S(f)df = \int_0^\infty [D(f)\ G(f)\ G_2]^2\ df = \int_0^{32\,kHz} D^2\ G^2(f)\ G_2^2\ df \quad .$$

Since the filter gain function $G^2(f)$ drops rapidly beyond 3.3 kHz, the upper limit of the integral can be extended to infinity with little error, leaving the prediction

$$\langle V_A^2(t) \rangle = D^2\ G_2^2 \int G^2(f)\ df \, .$$

The integral that remains is given by the filter's equivalent noise bandwidth, given by (1.1108) (3.3 kHz) = 3665 Hz, by the methods of Section 2.2. Notice too that what's emerged is D^2, which is just S, the noise power density -- the exotic V/√Hz unit has turned into the mundane V^2/Hz unit. Meanwhile, the squarer's time-averaged output is given by

$$\langle V_{sq} \rangle = \langle V_A^2(t) \rangle \ / \ (10\ V) \, ,$$

so it should be given by

$$\begin{aligned} \langle V_{sq} \rangle &= (10\ V)^{-1}\ (1.19\ mV/\sqrt{Hz})^2\ G_2^2\ (3665\ Hz) \\ &= (10\ V)^{-1}\ (1.42 \times 10^{-6}\ V^2/Hz)\ G_2^2\ (3665\ Hz) \\ &= (519 \times 10^{-6}\ V)\ G_2^2 \, . \end{aligned}$$

If you pick $G_2 = 40$ (by using the x1, x1, and x40 settings), you should get $\langle V_{sq} \rangle = 0.830$ V. Try this out, and test the results for some other settings of G_2 as well.

******

For a second example, suppose you keep using this 3.3 kHz low-pass filter, and pick $G_2 = 40$ for the gain stage. Now consider the filtered and amplified, but *not* squared, signal emerging from the main amplifier, $V_A(t)$. It should have a spectral distribution given by a voltage noise density

$$D(f)\ G(f)\ G_2 = (1.19\ mV/\sqrt{Hz})\ G(f)\ 40 = 48\ mV/\sqrt{Hz}\ [1 + (f/3.3\ kHz)^4]^{-1/2} \, .$$

So if you send this voltage $V_A(t)$ into a signal-acquisition system and use software to compute the spectral density, your answer had better come out to match this: 48 mV/√Hz for $f << 3.3$ kHz, and dropping beyond the 3.3 kHz corner, before plummeting to a near-zero value beyond 32 kHz. When you see (in Appendix A.10) how intricate such a software calculation can be, you'll appreciate the value of this sort of cross-check in revealing errors by factors of 2, √2, π, 2π, 1024, and the like.

******

Finally, a word about the noise-calibrator waveform. It's been optimized for stability of signal strength, uniformity of spectral density (in the $0 < f < 32$ kHz band), and absence of spectral density (for $f > 32$ kHz). To achieve these goals, it's actually been made to be a crypto-periodic waveform. Hence if you acquire data over a long enough time, you'll see it repeat. It follows that if you measure its frequency spectrum at high enough resolution, you'll see the apparently uniform spectral density break up into a resolved set of delta-functions. These occur at integer multiples of the fundamental frequency f_1, which is in turn the inverse of the 'repeat time' or period T of the waveform. So this source is really a 'multi-sine' source, the superposition of lots of sinusoids, all equally spaced in frequency and (ideally) equal in amplitude. For experiments *not* capable of the spectral resolution that's required to resolve the 'teeth in the comb' in frequency space, this multi-sine source acts just like 'actual', though band-limited, noise.

For a parallel example, consider a crystal of atoms. You think it has a mass density -- in one dimension, a number given in kilograms per meter. But in fact, if you were to measure with high enough *spatial* resolution, you'd see the 1-d mass-density function $\rho(x)$ change from a constant value to a series of delta-functions, one at the location of each atom. Yet you could still be justified in treating the mass density as a constant function, for purposes of solving any problem *not* involving spatial resolution at the nanometer scale.

5.5. What are the 'right' units for measuring noise?

You've seen 'noise power density' usually denoted by S, with units of V^2/Hz, and elsewhere you've seen the 'voltage noise density' given by D or $\sqrt{S}$, with units of $V/\sqrt{Hz}$. Which of these is the proper measure for noise? The answer: it depends with what you're *doing* with the noise measure.

If you have a signal, such as that from our Noise Calibrator (which gives $\sqrt{S} \approx 1.2$ $mV/\sqrt{Hz}$ from d.c. to 32 kHz), and you send it through an ordinary gain-of-10 amplifier, then $\sqrt{S}$ is the sensible measure for the noise. An amplifier which has gain of 10 for any sinusoid will also multiply voltage noise density by 10-fold, here from the voltage noise density of 1.2 $mV/\sqrt{Hz}$ to 12 $mV/\sqrt{Hz}$. So for 'multiplication by a constant', $\sqrt{S}$ in $V/\sqrt{Hz}$ is the right measure to use for noise.

But if you have two independent noise signals, and you add those signals, different considerations apply. The absence of correlations between the two signals $V_1(t)$ and $V_2(t)$ is defined by $< V_1(t)V_2(t) > = 0$, and under those conditions we get

$$< [V_1(t) + V_2(t)\ ]^2 > = < [V_1(t)]^2 > + 0 + < [V_2(t)]^2 >$$

where the cross term vanishes in the time average. This tells us that voltages-squared are the noise measures which are additive, and it follows that 'noise power densities' of the form $< V^2(t) > /\Delta f$ are additive too. So if you have Johnson noise of size $D_J = \sqrt{S_J} = 13$ $nV/\sqrt{Hz}$ from a resistor, and also have amplifier input noise of $D_A = \sqrt{S_A} = 8$ $nV/\sqrt{Hz}$, what you get is *not* a density of (8+13) = 21 $nV/\sqrt{Hz}$. Instead, you form the noise *power* densities

$$S_J = 169 \times 10^{-18}\ V^2/Hz,\ S_A = 64 \times 10^{-18}\ V^2/Hz, \text{ and then}$$

$$S_{net} = (169+64) \times 10^{-18}\ V^2/Hz = 233 \times 10^{-18}\ V^2/Hz, \text{ or } \sqrt{S_{net}} \approx 15\ nV/\sqrt{Hz}.$$

So for 'addition of two noise sources' of uncorrelated noise, S in V^2/Hz is the right measure to use for noise.

6. Further projects

The projects that follow are not in any particular order. Each of them represents an extension of some technique covered in Chapters 1 - 5, and each refers back to the natural preparation that should be performed first.

6.1. Time-domain characterization of the filter sections

You recall from Section 2.2 that it is necessary to know a filter's gain function $G(f)$ in order to compute its equivalent noise bandwidth Δf. In that section, you used 'paper models' for the $G(f)$ functions, and saw how integration of the $G^2(f)$ function over frequency gave the bandwidth Δf. In section 5.2, you saw one way to measure $G(f)$ values on a frequency-by-frequency basis, and how to fit $G(f)$ values to a 1- or 3-parameter model, thereby to calibrate the bandwidth values for your actual filters. This section shows you an alternative to this sort of 'frequency-domain' measurement.

The new method lies in the 'time domain', and makes use of the wonderful connections that exist between time- and frequency-response of *any* linear system. We'll describe the method here only for low-pass filters, and we'll assume a 2-pole state-variable response, but the method is actually of much greater generality. Briefly put, we pick a low-pass filter, choose its corner frequency, and diagnose its properties by seeing its response to a drive by *square*-wave excitation. We pick that square-wave frequency to lie well below the corner frequency, so the response we see will really display the 'step response' of the filter. Then we show that modeling the step response can give all the parameters needed to model the filter, and thereby *also* predict its frequency response. Notice that in this method we get *all* the information on the filter using only ***one square-wave*** frequency.

Consider a pair of 'scope traces of a square-wave input, and a filtered-output, from a low-pass filter set for corner frequency 10 kHz, and driven with a 1-kHz square wave, Figure 6.1a. Notice that the filter output is a version of the *inverted* input (ie. the filter has a gain near -1 for low frequencies). But notice that the output lacks what the input has, namely sharp edges. That's because the filter is expressly set *not* to pass those high-frequency components which the square wave must contain if it is to possess sharp edges.

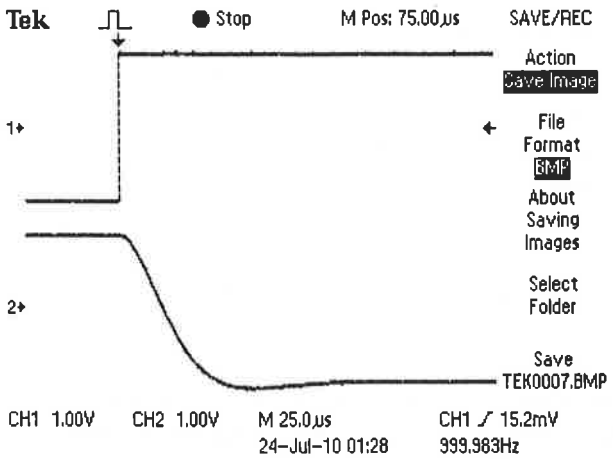


Fig. 6.1a: Upper trace: part of a 1-kHz square wave input to a filter; Lower trace: the output of a low-pass filter, set for corner frequency 10 kHz. Vertical scales 1 V/div, horizontal scale 25 µs/div, triggering on rising edge of upper trace.

This sort of display gives the 'step response' of the filter in the time domain. Acquire the step and response waveforms on two channels of your 'scope, triggering on the drive, and getting a high density of sampling points on the relevant part of the response function. Finally, you'll need a way to get the digital values of the samples into a computer environment where you can model them by fitting.

Once you get a step response, what can you do with it? There are very general Fourier-transform methods that could be applied, but here we'll use a differential-equations method, which is appropriate for the two-pole design of the electronics in the filter. We suppose that $V_{in}(t)$ -- above, taken to be a square wave -- is the 'drive term' in a second-order differential equation, of which $V_{out}(t)$ is the solution. If we're driving the experimental system in its linear regime (say by keeping V_{in} and V_{out} in the ±5-V range), we expect that differential equation describing it to be linear and to have constant coefficients, of the form

$$A\frac{d^2V_{out}}{dt^2} + B\frac{dV_{out}}{dt} + C\,V_{out}(t) = V_{in}(t) \quad .$$

Now defining $B/A = 2\,\gamma\,\omega_0$ and $C/A = \omega_0^2$, we get

$$\frac{d^2V_{out}}{dt^2} + 2\gamma\,\omega_0\frac{dV_{out}}{dt} + \omega_0^2\,V_{out}(t) = \frac{1}{A}V_{in}(t) \quad ,$$

which has the form of a damped and driven simple harmonic oscillator. Such an oscillator has an undamped frequency ω_0, a damped frequency $\omega_0\sqrt{(1-\gamma^2)}$, and a dimensionless damping constant γ (where $\gamma\rightarrow 0$ for undamped, and $\gamma\rightarrow 1$ for critical damping). In this standard form, it's easy to show that a sinusoidal drive of the form $a\cdot\cos(\omega t)$ gives a sinusoidal response of the form $G\cdot a\cdot\cos(\omega t - \phi)$, where the 'gain' obeys

$$G = G(\omega) = \frac{1/A}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2\gamma\,\omega\,\omega_0)^2}} \quad .$$

That provides a 3-parameter model (using A, ω_0, and γ) for the frequency-domain sort of data you took in Section 5.2.

But the differential equation above can instead be solved for a drive by a unit step, say with $V_{in}(t)$ changing, at time t=0, from value 0 to 1. The computed response is

$$V_{out}(t) = 1 - e^{-\gamma\omega_0 t}\,[\cos(\omega_0 t\sqrt{1-\gamma^2}) + \frac{\gamma}{\sqrt{1-\gamma^2}}\sin(\omega_0 t\sqrt{1-\gamma^2})] \quad .$$

To apply this to the low-pass filters at hand, imagine a square wave input which 'steps down', at t=0, from a level (+a) to (-a). The output will 'step up', from a level of (-g·a) toward (+g·a), with a transient behaving for t>0 as

$$V_{out}(t) = -g\,a + 2\,g\,a\,\{1 - e^{-\gamma\omega_0 t}\,[\cos(\omega_0 t\sqrt{1-\gamma^2}) + \frac{\gamma}{\sqrt{1-\gamma^2}}\sin(\omega_0 t\sqrt{1-\gamma^2})]\} \quad .$$

This is a complicated expression , but it really depends on only three parameters: a 'gain' g (which is intended to be 1), a 'damping' γ (which is intended to be $1/\sqrt{2}$), and a 'corner frequency' ω_0 (which is intended to be $2\pi \cdot f_c$, where f_c is the corner frequency you select). So if you acquire the 'step-response' data and fit it to an expression of this form, you will get out, as best-fit parameter values, good estimates for g, γ, and ω_0.

And with them, you can immediately be quite confident that the frequency-domain gain function (of that same filter, for the same setting) will be given by

$$G(f) = \frac{g\,f_c^{\,2}}{\sqrt{(f_c^{\,2} - f^2)^2 + (2\gamma\,f\,f_c)^2}} = \frac{g}{\sqrt{(1-(f/f_c)^2)^2 + (2\gamma\,f/f_c)^2}} \quad ,$$

where f_c is given by the ω_0-value from your fit, divided by 2π. For the intended case of $\gamma = 1/\sqrt{2}$ and $g = 1$, it's easy to show that this reduces to the ideal Butterworth response

$$G(f) = \frac{1}{\sqrt{1+(f/f_c)^4}} \quad ,$$

that we've used in previous paper modeling. But whether the response matches the Butterworth function or not, the step-response method will have given you the parameters for a three-parameter model which fully characterizes the frequency response curve. And from these three parameters, you can compute the filter's equivalent noise bandwidth by methods of Sections 2.2 and 5.2.

6.2 Narrow-band measurement of noise density - the 'lock-in' method

You've seen in Section 2.1 how to measure a 'noise density', and in Section 2.3 how to measure more nearly a 'local density' by using a filter with a well-defined center frequency. But in this section, you'll see a little-known method for making a very local-in-frequency measurement of noise density, with an equivalent noise bandwidth which can easily be as small as a Hertz, and which is *in*dependent of the frequency at which you locate it.

The method is closely related to the way a lock-in amplifier works, but it does not require a separate lock-in. We'll assume that you have some noise measurement in progress, and that you've used high-pass and low-pass filtering to isolate a broad swathe in frequency space, perhaps 0.1 - 100 kHz. You've usually brought these signals to the main-amplifier section of the high-level electronics, and used a main-amp gain which brings the noise signal to a level of about 3 V (rms measure). Now the biggest difference of this new method is to use the 'squarer' in a new way -- as a multiplier instead. So you bring the amplified noise signal $V_A(t)$ to the A-input of the squarer as usual, but you select the AxB, or multiplier, function of this module.

Now if you want to measure the truly local value of the spectral density of $V_A(t)$ at a target frequency f_t, what you need is to bring a sinusoid, of frequency f_t and of amplitude 10 V, to the B-input of the multiplier. In the language of lock-in detection, you'd call this the 'reference input'. If there were actually present a signal at frequency f_t buried under the noise of $V_A(t)$, the multiplier would now reveal it by producing an output whose time average included a non-zero d.c. value. But if $V_A(t)$ is 'pure noise', the multiplier output will *lack* any such d.c. average value.

In the absence of a d.c. signal, what then can you measure? You want to quantify the *fluctuations around zero* in the multiplier's output, by taking the time-average of it with time-constant τ (your choice, 0.01 - 3 seconds), capturing that on a 'scope, and finding the rms value of that 'scope signal. Here's the analysis which shows that this rms value will tell you, quantitatively, the noise density in the neighborhood of f_t in the A-channel signal.

Start with a raw noise signal $V_n(t)$, and a gain G_1 from the pre-amp, and G_2 from the main amp. Set the gain $G(f)$ from the filter section(s) so as to pass (with $G(f) \approx 1$) the noise in the vicinity of the target frequency. So the main-amp output is given by

$$V_A(t) = G_2 \cdot 1 \cdot G_1 \cdot V_n(t) \, .$$

Now we temporarily represent the noise signal $V_n(t)$ by a Fourier series, valid over the (presumed long) time interval $0 < t < T$:

$$V_n(t) = \sum_{i=1}^{N} A_i \cos(2\pi \cdot if_1 \cdot t - \phi_i) \quad .$$

Here T might describe the full duration of your experiment, a minute or even an hour. In this series, the fundamental frequency is $f_1 \equiv 1/T$, and we note the absence of a d.c. term,

and the presence of all the harmonics of the fundamental. The terms in the Fourier series are equally spaced in frequency by interval

$$\delta f \equiv f_{i+1} - f_i = (i+1)f_1 - (i)f_1 = 1\, f_1 = 1/T\,.$$

If the noise $V_n(t)$ is spectrally white, we can take all the amplitudes A_i to be equal, and use a single A-value. The phases ϕ_i are presumably random, and their values will turn out not to matter.

Now $V_n(t)$ has a spectral density, and we compute it <u>on paper</u> by the usual means: we filter $V_n(t)$ to some bandwidth Δf, we square that filtered signal, and then time-average to find the mean-square value. In filtering our model noise signal to a bandwidth Δf (assuming for simplicity a sharp-edged filter bandwidth), we transmit only a fraction of all N terms that appear in the Fourier sum; the number of terms which pass through the filter is given by

$$\text{(filter's bandwidth) / (terms' spacing)} = \Delta f / \delta f\,.$$

In computing the square of the filtered signal, the cross terms vanish upon time averaging, so only the squares of individual terms survive. Using the number of surviving terms, we get for the sum's value

$$< [V_n^{(F)}(t)]^2 > = A^2\, (\Delta f/\delta f) \,<\cos^2(2\pi \cdot i f_1 . t - \phi_i)>\,,$$

and as the cosine-squared terms all average to 1/2, we find spectral density of noise

$$S_n = < [V_N^{(F)}(t)]^2 > / \Delta f = A^2\, (1/\delta f)\, (1/2) = A^2/(2\, \delta f)\,.$$

The result relates the observable spectral density of the noise to the parameters A and δf of our paper noise model.

That model can now be followed to, and through, the multiplier in a paper description of our actual measurement. We have at the A-input of the multiplier a model for the (amplified) noise,

$$V_A(t) = G_2 \cdot 1 \cdot G_1\, V_N(t) = (G_2\, G_1\, A) \sum_{i=1}^{N} \cos(2\pi \cdot i f_1 \cdot t - \varphi_i) \quad ,$$

and at the B-input we have our 'reference signal' at target frequency f_t,

$$V_B(t) = (10V) \cdot \cos(2\pi\, f_t\, t - 0) \quad .$$

The multiplier produces the scaled product,

$$V_{multt}(t) = V_A(t)\, V_B(t)/(10V) = G_2\, G_1\, A \sum_i \cos(2\pi \cdot i f_1 \cdot t - \phi_i) \cos(2\pi\, f_t\, t) \quad .$$

Here again we have a product of cosines, which can be written as a sum of two new cosines:

$$V_{multt}(t) = G_2\, G_1\, A\, \frac{1}{2} \sum_i \left[\cos(2\pi \cdot (i f_1 - f_t) \cdot t - \phi_i) + \cos(2\pi \cdot (i f_1 + f_t) \cdot t) \right] \quad .$$

Of these terms, those at the 'sum frequencies' $(i f_1 + f_t)$ will certainly **not** pass the final time-averaging stage, so we'll drop them. Those at the 'difference frequencies' include terms close to zero frequency, and these are the ones that *will* contribute to the

measurable fluctuations of the output. (We do assume that there's no term at *zero* frequency -- that would constitute lock-in detection of a signal, at target frequency f_t, in the noise.)

Now the terms of difference frequencies $\partial f_i = (i f_1 - f_t)$ pass to the final time-averaging filter, which is a low-pass filter of gain $g(f)$, so the filtered multiplier output can be written as

$$V_{mult}^{(F)}(t) = G_2\, G_1\, A\, \frac{1}{2} \sum_{i \in F}\, g(\partial f_i) \cos(2\pi \cdot \partial f_i \cdot t - \phi_i) \quad .$$

(The filtering also creates changes in the phase constants ϕ_i, which we ignore.) On a 'scope, this signal seems to wander about; but what we'll measure is its rms value. To compute that on paper, we need to square it, giving

$$[V_{mult}^{(F)}(t)]^2 = (G_2\, G_1\, A/2)^2 \sum_{i,j \in F}\, g(\partial f_i)\, g(\partial f_j) \cos(2\pi \cdot \partial f_i \cdot t - \phi_i) \cos(2\pi \cdot \partial f_j \cdot t - \phi_j) \quad .$$

In this result, as usual, the cross terms do not survive time averaging, so what does survive can be written as

$$\left\langle [V_{mult}^{(F)}(t)]^2 \right\rangle = (G_2\, G_1\, A/2)^2 \sum_{i \in F}\, g^2(\partial f_i) \left\langle \cos^2(2\pi \cdot \partial f_i \cdot t - \phi_i) \right\rangle = (G_2\, G_1\, A/2)^2 \sum_{i \in F} g^2(\partial f_i) (\frac{1}{2}) \quad .$$

The sum over all difference-frequencies that pass the filter include terms with both positive and negative values of $\partial f_i = (i f_1 - f_t)$. But in the limit of a long-T experiment, both form sets of closely-spaced 'combs' which equably sample the $g^2(f)$-function's values in the sum. So the result is the same if we double it, but take the sum only over positive frequencies ∂f_i :

$$\left\langle [V_{mult}^{(F)}(t)]^2 \right\rangle = (G_2\, G_1\, A/2)^2\, 2 \sum_{\partial fi > 0} g^2(\partial f_i) (\frac{1}{2}) \quad .$$

To evaluate that sum, we notice that a related sum is the very Riemann sum which would go, in the $\partial f_i \to 0$ limit, to an integral:

$$\sum_{\partial fi > 0} g^2(\partial f_i)\, \Delta(\partial f_i) \to \int_0^\infty g^2(f)\, df \quad .$$

The spacing $\Delta(\partial f_i)$ in the comb of frequencies is just the original δf, so we have

$$\sum_{\partial fi > 0} g^2(\partial f_i) \to \frac{1}{\delta f} \int_0^\infty g^2(f)\, df \quad ,$$

and finally can write as the mean-square fluctuation

$$(V_{rms})^2 = (G_1\, G_2\, A/2)^2\, \frac{1}{2}\, 2 \frac{1}{\delta f} \int_0^\infty g^2(f)\, df \quad .$$

In this result, we recognize the combination $A^2/(2\,\delta f)$, which in our noise model gives the spectral density of the input noise S_N, so we have

$$(V_{rms})^2 = (G_1\, G_2\, /2)^2\, 2S \int_0^\infty g^2(f)\, df \quad .$$

The final-filter function for a double one-pole filter is $g(f) = \{1/\sqrt{[1+(f/f_c)^2]}\}^2$, with corner frequency given by $f_c = 1/(2\pi\,\tau)$, so the integral gives

$$\int_0^\infty g^2(f) df = \int_0^\infty \frac{df}{[1+(f/f_c)^2]^2} = \frac{\pi}{4} f_c = \frac{1}{8\tau} \quad .$$

Hence the observable output is predicted to have mean value zero, but fluctuations characterized by rms measure V_{rms}, where

$$(V_{rms})^2 = \frac{G_1^{\,2}\, G_2^{\,2}}{4} 2\, S \frac{1}{8\tau} = G_1^{\,2}\, G_2^{\,2}\, S \frac{1}{16\tau} \quad .$$

Since the 'observable' which you might read off a 'scope is the rms measure of the fluctuations at the output, V_{rms} itself, we write this as

$$V_{rms} = \frac{1}{4} G_1\, G_2\, \sqrt{S}\, \tau^{-1/2} \quad .$$

Sure enough, that relates the spectral density of noise at the input, S, to an observable quantity at the output. In fact, the rms output value is related to the *voltage* noise density $D = \sqrt{S}$. This argument also predicts a $\tau^{-1/2}$ averaging-time dependence of the fluctuations in a general lock-in amplifier's output, but here we care chiefly about the dependence on S.

To be concrete, suppose we start with a 1-MΩ resistor as a source of Johnson noise, which gives a noise power density of $\langle V_J^2 \rangle / \Delta f = 4\, k_B\, T\, R = 1.63 \times 10^{-14}$ V^2/Hz at room temperature. This is the source's S-value, and its square root, the voltage noise density D, is 128 nV/√Hz. After pre-amplification by gain $G_1 = 600$, that voltage noise density is up to 76.6 μV/√Hz. This noise can now be low-pass filtered, to restrict its bandwidth to (say) about 100 kHz. Then a main-amp gain of $G_2 = 100$ will produce a noise density of 7.66 mV/√Hz, flat out to about 100 kHz, which gives a mean-square measure of the output of the main amplifier of

$$\langle V_A^2(t) \rangle \approx (7.66 \text{ mV}/\sqrt{\text{Hz}})^2 \times 100 \text{ kHz} \approx 6 \text{ V}^2 ,$$

which is of a size so as *not* to saturate the A-input of the multiplier. (That is to say, 100 is about the right level of gain G_2 to use in this example.)

Now we've assumed the B-input of the multiplier gets a sine wave of 10-V amplitude (or 20 V pk-to-pk), at any target frequency f_t we choose (in the < 100 kHz range); so we are enabled to predict that the output of the squarer-and-averager will exhibit

- mean value *zero*, but
- fluctuations about the mean of zero, with rms measure of fluctuations given by

$$V_{rms} = (1/4)(600)(100)(128 \text{ nV}/\sqrt{\text{Hz}})\, \tau^{-1/2} .$$

If we pick a minimal averaging time $\tau = 0.01$ s, the prediction is

$$V_{rms} = 19.2 \text{ mV}.$$

This number is directly connected to the voltage noise density of the source, at the target frequency selected.

The calculation above leaves out two important effects: amplifier noise, and front-end bandwidth. For a source resistor as large as 1 MΩ, the Johnson-noise density of the resistor dominates considerably over the effective input noise of the amplifier, but you would still want to find a way to measure the amplifier noise, and subtract it. Also, for a source resistor as large as 1 MΩ, the effect of about 10 pF of input capacitance gives an RC time-constant at the input of about 10 μs, or a corner frequency of about 16 kHz. So

for so large a source resistance, you should expect less than the 'full Johnson noise' if you use this method to measure S at frequencies above a few kHz.

The advantage of this method is that you can measure $\sqrt{S}$ at any target frequency f_t you choose. Furthermore, the method is almost perfectly 'local', in the sense that the noise is being sampled in a neighborhood of f_t, with an equivalent bandwidth $(8\ \tau)^{-1}$, which is 12.5 Hz in our example. (Note also that this bandwidth is *in*dependent of the choice of f_t.) The bandwidth could be reduced even farther: for the maximal choice of $\tau = 3$ s, the bandwidth is down all the way to 0.04 Hz.

There are real *disadvantages* to the use of so narrow a bandwidth. Even if we go back to our $\tau = 0.01$ s choice, it's clear that an oscilloscope view of the averaged output will show a value that's zero on average, which wanders about zero with typical ±20-mV excursions, and which gives an effectively 'new value' every 0.01 seconds or so. Hence in 10 seconds of measurement time, you can get on the order of 10 s/0.01 s = 10^3 statistically independent values. It also follows that $(10^3)^{-1/2} \sim 0.03$ or 3% gives the level of fluctuations you expect in successive computations of V_{rms}, each of them based on such a 10-s block of data. The time required to get a stable value of V_{rms} will rise, dramatically, if you choose a longer τ-value. So there's a trade-off between the degree of 'locality' in frequency space at which you measure the noise density, and the time required to get a value with good statistical reliability.

In addition to the considerable time required to establish a noise-density value, the result you get applies at only *one* target frequency. If you want a reading of $S = S(f)$ at another target frequency, you need a fresh investment of as long an averaging time. One of the appealing features of Fourier methods (see Appendix A.10) is that they can give spot values of noise density for a whole collection of frequencies *all at once*. Fourier methods are *not* free from the need to perform sufficient averaging to give statistically stable noise-density values, but they offer the advantage of simultaneous 'parallel processing' of a whole range of frequencies.

Finally, this method is in principle capable of making accurate values of input spectral density. The issues of statistical precision are dealt with above; for issues of accuracy, there's the usual need to know the gain factors G_1 and G_2 of the amplifiers. In place of needing to know a filter bandwidth Δf, in this lock-in method there's the need to know the final-filter time constant τ. If the nominal (±5%) value given on the panel is not good enough for you, perhaps this τ-value is best measured in the time domain. If you drive the multiplier as a squarer, by a square-wave voltage alternating (say) between 0 and +10 V, then its output will 'step' between levels of 0 and 10 Volts. Then if you capture that waveform, appearing at the OUTPUT jack in the Output module of the high-level electronics, you should see a low-pass version of that square wave. For this sort of filter, a unit step (from 0 to 1, at time t = 0) at the input should produce, at the output, a waveform described by

$$V_{out}(t) = 1 - (1 + t/\tau)\exp(-t/\tau)\ .$$

So capturing the output waveform and fitting it to a function of this form can give you the precise results for the τ-values that you need in this approach.

6.3. Noise from *pn*-junction devices

You've now seen currents in Section 3.3 (from an illuminated photo-diode) displaying full shot noise, and other currents in Section 3.5 (from a humble $i = V / R$) displaying much less noise than Schott's formula predicts. This section is intended to show that the use of photons is *not* required to generate a current displaying shot noise. Instead, here we'll get currents very similar to $i = V / R$, except that in three cases they'll involve a p-n junction, a *forward*-biased diode, as an element in the current-generating circuit.

Here's a series of schematic diagrams of ways to produce currents; in each case the current is delivered to the input point of an i-to-V converter.

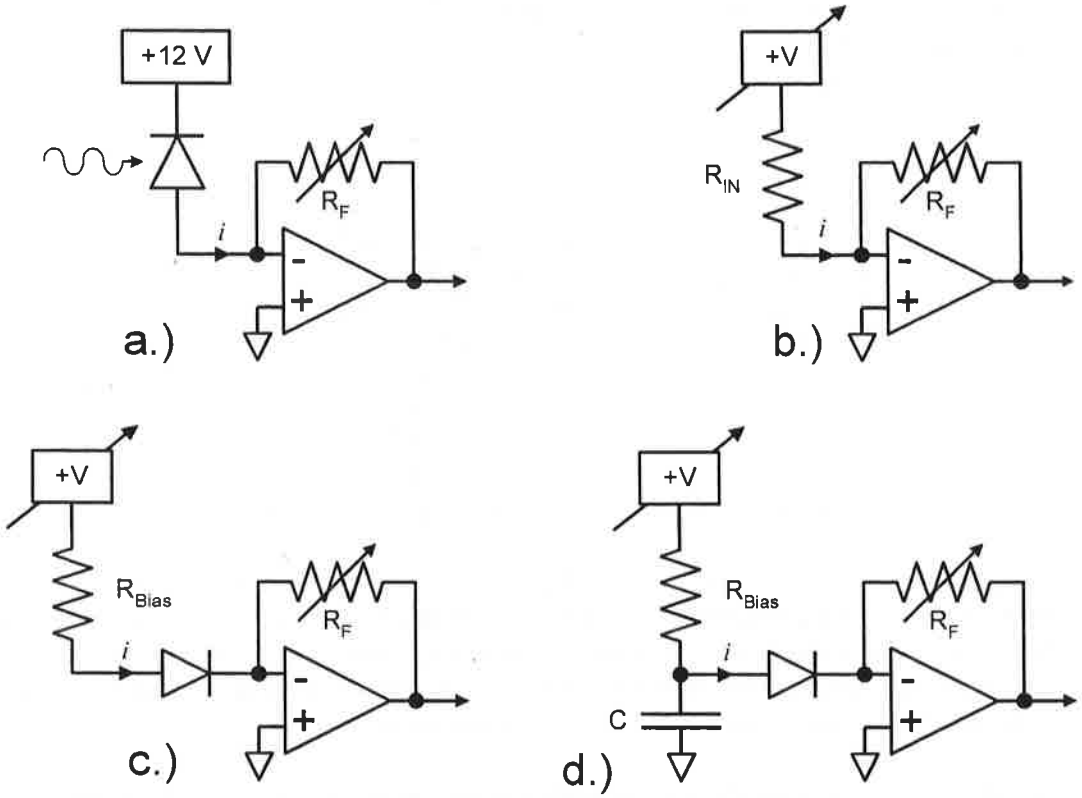


Fig. 6.3a: Four ways to get a current suitable for noise testing. a) A photocurrent. b) A V/R current. c) A forward current from a p-n junction diode. d) The same as c), except for an added capacitor.

Of these circuits, a) is the one with an illuminated photodiode which delivers 'full shot noise', and it's being driven by 'thermal photons'. By contrast, circuits b) - d) involve no photons at all. Of these, circuit b) is the one proven in Section 3.5 to be capable of delivering a current with much *less* than full shot noise. Circuit c) can easily be arranged to deliver the same current as b), although to do so, the 'bias supply' V_b will need to set about 0.6-0.7 V more positive to make up for the 'diode drop' in potential. Clearly, circuit c) does put a pn-junction in series with the current, but you can show that it does *not* restore full shot noise to the current.

To get back the full shot noise, you'll need circuit d), which has just one new component. If you think that adding a capacitor to c) would give an RC-filter that would lower the noise, you'd be wrong -- adding this capacitor **raises**(!) the noise, back up to the shot-noise level.

Let's confirm these facts empirically. We suggest the use of an ordinary silicon diode (such as a 1N4148 from the parts bin), and the choice of R_{bias} = 100 kΩ, so supply voltage $V \approx 6.6$ V would give about 0.6-V drop across the diode, and 6.0-V across the resistor, again ensuring that $i \approx 60$ μA. Under these circumstances, the diode has $\Delta V =$ 0.6 V, and $i = 60$ μA, so it displays an 'effective resistance' of $\Delta V/i = 10$ kΩ. In fact, for purposes of noise generation, what we need is the 'dynamic resistance' given by dV/di, whose value turns out (from the diode equation) to be about 50 mV/i ≈ 800 Ω. So from an a.c.-circuit point of view, the diode is acting like a 60-μA current source with an 800 Ω resistor in parallel with it.

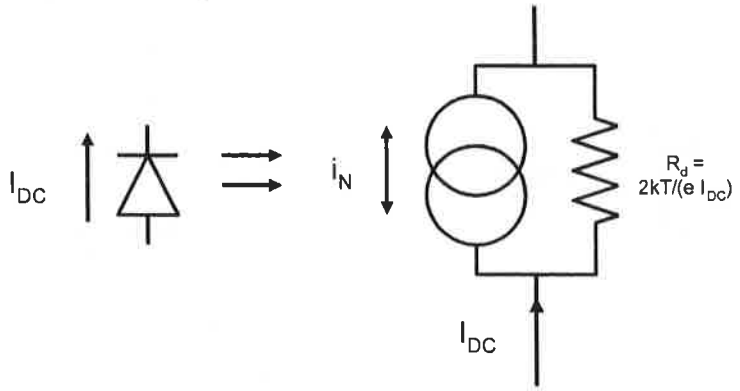


Fig. 6.3b: An equivalent circuit for a forward-biased diode as a current source.

But in circuit c), any a.c.-component of current generated by the junction is heavily diminished, because it flows preferentially through 800 Ω of shunt resistance, rather than the 100-kΩ path to the i-to-V converter. By contrast, circuit d) *does* have a low-impedance path for a.c. noise current to flow -- from ground, through the capacitor and diode, into the virtual-ground at the i-to-V converter's input.

This also tells us we want the a.c. impedance of the capacitor, $(2\pi f C)^{-1}$, to be smaller than 800 Ω, at even the lowest frequencies we care about. Suppose we've set our bandwidth (by downstream filtering) to the 1 kHz-to-100 kHz band, so that 1 kHz is that lowest frequency of interest. If we set $(2\pi f C)^{-1} = 100$ Ω at $f = 1$ kHz, we get $C \approx 10$ μF.

So finally you have a circuit you can build, namely d) with choices such as $R = 100$ kΩ, $C = 10$ μF, $V_b = 6.6$ V, and single silicon diode (which has to be inserted with the right polarity!), and using in the i-to-V converter the choice $R_f = 100$ kΩ. With this circuit, you should be able to monitor the current you achieve as before, by measuring a d.c. voltage at the pre-amp's MONITOR point. And you can also test, by familiar means, to see if it has the 'full shot noise' predicted for a current this large. The remarkable thing is that just removing the capacitor C will take you back *down* to a sub-shot-noise current, and adding it back in will *increase* the noise again.

To do this quantitatively over a wide range of diode currents reveals some new complications. As the diode current rises, its dynamic impedance falls. Eventually, the diode's dynamic impedance gets as low as the feedback resistor's value. Now, for purposes of op-amp input noise voltage, the input stage acts like a non-inverting amplifier, with gain 2. This 'noise gain' will continue to rise as the diode current rises, and at some point, the amplified op-amp noise voltage will become a significant fraction of the total noise. The onset of this effect becomes noticeable when the MONITOR output of the pre-amp's first stage reaches 100 mV in magnitude. So to confirm the existence of shot noise in the p-n junction diode, you'll need to work with fairly small currents, or you'll have to lower the feedback resistor (as you raise the diode current) to keep the potential difference across the feedback resistor under 100 mV.

There is also a finite-bandwidth consequence of this changing noise gain. Suppose you were to use a feedback resistor of 100 kΩ and a diode forward current of 50 μA; that would entail a diode dynamic resistance of order 50 mV/50 μA or about 1 kΩ. The 'noise gain' of the input stage is then about 100, but that noise gain cannot extend to very high bandwidth. The bandwidth limit will be limited by the gain-bandwidth product of the op-amp used (about 8 MHz for the OPA134 used) to about GBP/gain, or 8 MHz/100, or about 80 kHz. So if you want quantitatively correct noise measurements, you should use filter sections to restrict your noise measurements to bandwidth 10 kHz or less.

Given the need to restrict yourself to rather small diode currents, you will find it important to correct your noise readings for 'amplifier noise', by subtracting out the noise that occurs for zero diode current. Similarly at zero diode current, note the d.c. offset of the MONITOR output, and later subtract that from you other monitor readings.

But if you can achieve all of this experimentally, it'll be clear that one can get full shot noise from a circuit making no appeal to photons, or to boson statistics. It's also clear that while a p-n junction is necessary to get full shot noise, it's not sufficient: one also need the low-impedance path provided in circuit d).

When you think you've understood this, go back to circuit a): it has the necessary junction, but not the capacitor! Yet it gave full shot noise -- what's going on here? [Hint: circuits c) and d) have a forward-biased diode, of low dynamic impedance, while circuit a) has a reverse-biased diode -- what dynamic impedance does that give it? and why does that matter?]

6.4. Johnson noise vs. temperature: noise thermometry

Here's an emerging technology which applies the measurement of Johnson noise to the establishment of an absolute temperature scale. This technique is being actively pursued at national standards laboratories, and you're now it a position to understand it, and try out a simplified version of the technique.

Suppose we really trust the Nyquist prediction for Johnson noise,

$$<V_J^2(t)> = 4\, k_B\, T\, R\, \Delta f\,.$$

It's clear that if you can measure source resistance R, bandwidth Δf, and the mean-square Johnson noise voltage $<V_J^2(t)>$, this equation can establish the absolute temperature converted to energy units, ie. the product $k_B\, T$.

Now the SI system of units defines the Kelvin scale of temperature by means of ***only one*** 'fixed point', assigning $T \equiv 273.16$ K to the triple point of water. If you can subject your source resistor to this known temperature (by putting it into a 'triple point cell'), and if you measure the other quantities, then you can establish the value of k_B in SI units.[1]

Once that's known, the source resistor can be put at any other (unknown) temperature T_x, and measure T_x in absolute SI units, using

$$T_x = <V_J^2(t)> \;/\; (4\, k_B\, R\, \Delta f)\,.$$

This technique has actually been used to establish absolute temperatures all the way down to about 20 mK (see the Spietz reference in the Bibliography), where other techniques of thermometry are very hard to calibrate. Another application is to establish, with high accuracy, some 'ordinary temperatures' such as the triple point of nitrogen (near 63 K) or the freezing point of zinc (near 693 K).

The hard part is to make this technique work to the state-of-the-art precision in 'primary thermometry'. The goal might be part-per-million precision and accuracy, and this is very hard to attain for several reasons:

1) Measuring $<V_J^2(t)>$ is hard for *statistical* reasons (apart from any others). To get precision 10^{-6} in measuring a fluctuating quantity takes something like 10^{12} samples, where one 'sample' is a fresh reading of $<V_J^2(t)>$. The inverse of the electronic bandwidth sets the timescale for producing a fresh reading, so a choice of $\Delta f = 10$ kHz ensures you have to wait about 10^{-4} seconds to get a fresh reading, so that you can only get 10^4 independent readings per second. For an averaging time of 1 second, you're getting 10^4 samples, and thus you ought to expect part-in-10^2, ie. 1%, statistical fluctuations, in your result. To improve this to precision 10^{-6} takes truly heroic measures: amplifier chains with bandwidth of a full 1 GHz, and averaging times of 10^{+3} s (about 20 minutes), would just suffice to give the 10^{12} independent readings necessary.

---

[1] In current practice, the best values for k_B come from R/N_A, where the gas constant R and Avogadro's number N_A present their own measurement challenges.

2) Even supposing that $<V_J^2(t)>$ can be measured to the target precision, *and* that it can be corrected for amplifier noise to the same accuracy, there's still the need to know the 'other factors' k_B, R, and Δf to similar precision. Of these tasks, we suppose that some international consortium of standards labs has established k_B (find the accepted result at http://physics.nist.gov/cuu/constants/index.html), and we suppose that the same sort of labs can measure resistance to part-per-million accuracy too. The really hard part is in establishing the bandwidth Δf to such precision -- Section 5.2 reveals the complications which arise even at the 1% level.

So the true genius of noise thermometry is to use the same electronics to measure, alternatively and using the identical bandwidth, first Johnson noise and then shot noise. Let's see how that trick could work, temporarily ignoring the (highly non-trivial) issue of amplifier noise.

We have a sense resistor R_s at unknown temperature T_x, producing Johnson noise of mean-square value

$$<V_J^2(t)> = 4\, k_B\, T_x\, R_s\, \Delta f\,.$$

Next we create a current, displaying full shot noise, and having d.c. average value i_{dc}, and let it enter an i-to-V converter with conversion constant R_f, giving us a shot-noise voltage obeying

$$V_{sn}(t) = R_f\, i_{sn}(t)\,, \qquad \text{so} \qquad <V_{sn}^2(t)> = R_f^2 \cdot 2\, e\, i_{dc}\, \Delta f.$$

We use identical downstream filtering in the two measurements, to ensure that the same Δf value applies to the two results.

Now imagine a balancing operation: we alternate between the Johnson-noise and shot-noise sources, varying the current i_{dc} in the latter, until mean-square shot noise *balances* mean-square Johnson noise. Then i_{dc} has been arranged to obey

$$<V_{sn}^2(t)> = <V_J^2(t)>\,, \qquad \text{so} \qquad R_f^2\, 2\, e\, i_{dc}\, \Delta f = 4\, k_B\, T_x\, R_s\, \Delta f\ .$$

The bandwidth Δf cancels out, and so do all the gain factors -- so we do *not* need to measure them; we need only be certain they're the *same* for both measurements. What's left can be written as

$$k_B\, T_x / e = 2\, R_f^2\, i_{dc}\, /(4\, R_s) = (1/2)\, (R_f / R_s)\, (i_{dc}\, R_f)\ .$$

This makes it clear that what we're really measuring is the combination $(k_B\, T/e)$, which is the absolute temperature converted to voltage units. (On this scale, you can check that room temperature maps to about 25 mV.) On the right-hand side, you have the factors (1/2), the resistance ratio (R_f / R_s), and the iR-product $i_{dc}\, R_f$, which is a voltage. That is to say, the measurement of an unknown temperature T has changed to the measurement of $(k_B\, T/e)$, but has turned into a problem of electrical measurements only.

Now amplifier noise is *not* negligible; nor should you trust the necessary equality-of-bandwidth if one side of the balance involves an i-to-V converter and the other doesn't. So here's a circuit in which you *can* conduct a noise-temperature measurement (though not of part-per-million precision!) because it should maintain equal amplifier noise, and equal bandwidth, under all conditions.

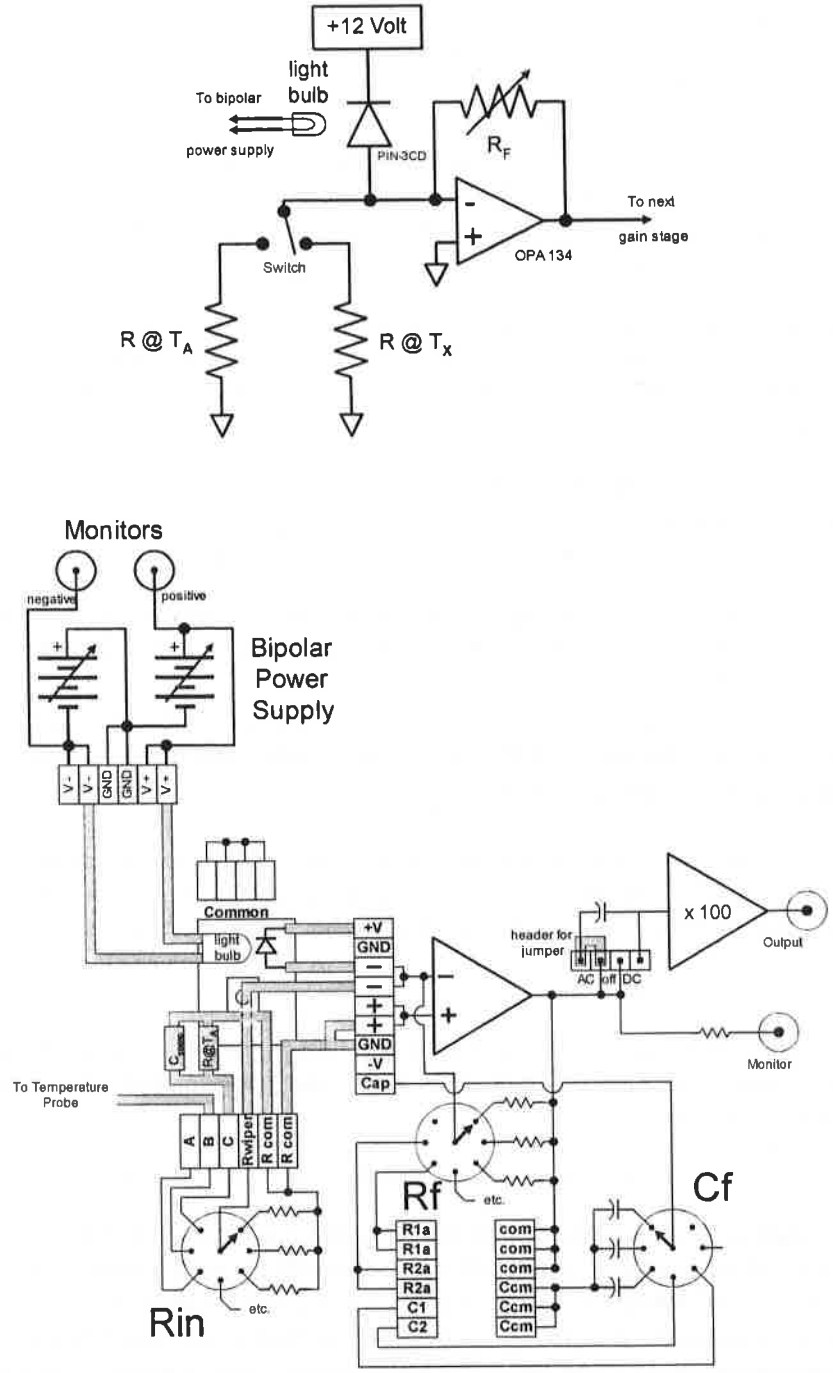


Fig. 6.4: Schematic diagram, and wiring diagram, of input circuitry for noise thermometry.

This circuit has the reverse-biased photodiode, and the i-to-V converter with feedback resistor R_f, just as in Section 3.3. What's added is the (two) input resistors, assumed both to be of resistance R, but at two possibly distinct temperatures. We'll use T_a for the ambient temperature, and T_x for the (assumed lower) unknown temperature of the remote one of the two input resistors.

First we analyze this circuit for d.c. current. The reverse-biased photodiode, when illuminated, generates a current i_{dc} as previously. What's new is the presence of a (switch-selected) input resistor R_{in} -- but here's why no d.c. current will flow to ground through it. The op-amp's feedback sees to it that the potentials at the two op-amp input terminals stay the same, ie. both stay fixed at zero potential. But that also ensures that the potential difference across R_{in} is zero; hence it conducts no current. As a result, the whole of the photodiode's current i_{dc} flows through R_f as before, ensuring that a monitor of the op-amp's output voltage will still give

$$V_{mon} = (-)\, i_{dc}\, R_f\,.$$

Next we analyze this circuit for noise currents. The photocurrent, i_{dc}, is presumed to have the usual shot-noise density, giving $<\delta i^2> = 2\, e\, i_{dc}\, \Delta f$ as before. Both resistors also generate Johnson *current* noise of size V_J/R, or of $<\delta i^2> = <V_J^2>/R^2$, giving mean-square fluctuations

$$<\delta i^2> = 4\, k_B\, T\, R\, \Delta f / R^2 = 4\, k_B\, T\, R^{-1}\, \Delta f\,.$$

So the usual addition of uncorrelated mean-squares gives a net mean-square noise current

$$<\delta i^2> = 2\, e\, i_{dc}\, \Delta f + (4\, k_B\, T_a / R_f)\, \Delta f + (4\, k_B\, T_x / R_{in})\, \Delta f\,.$$

This noise current appears at the output with mean-square fluctuations in V_{mon} given by

$$<\delta V_{mon}^2> = R_f^2\, <\delta i^2>\,,$$

But there is also the noise voltage of the amplifier itself to be considered. We take some stated amplifier-noise density V_A, but in this circuit, that voltage noise is subject to a gain G which happens to be $G = 1 + R_f/R_{in}$. So the total noise at the monitor point has mean-square value

$$<\delta V_{mon}^2> = (G\, V_A)^2\, \Delta f + R_f^2\, [2\, e\, i_{dc}\, \Delta f + (4\, k_B\, T_a / R_f)\, \Delta f + (4\, k_B\, T_x / R_{in})\, \Delta f]\,.$$

Now we balance between an 'a' and a 'b' situation:

a) We use the ambient-temperature resistor R_{in}, so we can write $T_x = T_a$, but we turn the lamp off, so $i_{dc} = 0$. This gives

$$<\delta V_{mon}^2>|_a = (G\, V_A)^2\, \Delta f + R_f^2\, [2\, e\, 0\, \Delta f + (4\, k_B\, T_a / R_f)\, \Delta f + (4\, k_B\, T_a / R_{in})\, \Delta f]\,.$$

b) Now we switch to the colder input resistor, at some T_x below ambient, and the noise level drops. But we make up for the smaller Johnson noise by turning the lamp back on to create some shot noise, giving

$$<\delta V_{mon}^2>|_b = (G\, V_A)^2\, \Delta f + R_f^2\, [2\, e\, i_{dc}\, \Delta f + (4\, k_B\, T_a / R_f)\, \Delta f + (4\, k_B\, T_x / R_{in})\, \Delta f]\,.$$

Note we assume that the feedback resistor stays at the ambient temperature T_a, and that the amplifier noise doesn't change either.

If we dial up the lamp so as to get 'noise balance', then $<\delta V_{mon}^2>$ has the *same value* in the 'a' and 'b' situations. So subtracting the two equations above, we get lots of cancellations: first the amplifier noise cancels, and then the factor R_f^2 cancels, and next the Johnson noise due to R_f cancels, and finally the bandwidth Δf cancels too.

All that's left is

$$2\,e\,0 + 4\,k_B\,T_a / R_{in} = 2\,e\,i_{dc} + (4\,k_B\,T_x / R_{in})$$

which can be re-arranged to give

$$(k_B/e)(T_a - T_x) = (1/2)\,R_{in}\,i_{dc} = (1/2)\,R_{in}\,(-V_{mon}/R_f)\,.$$

Here we've used the d.c. value of the monitor voltage as a surrogate for i_{dc}. So the result is a direct measurement of a *difference* in absolute temperatures (converted into voltage units by the k_B/e factor), given in terms of (1/2), the resistance ratio (R_{in}/R_f), and a single measured d.c. voltage V_{mon}. This is a form of noise thermometry you can actually try out!

There are details, the most important of which, is to ensure that the two R_{in}'s really do act alike. The chief difficulty is that one of them will likely be a 'remote resistor' in the probe, while the other is a 'local resistor' in the pre-amp. In practice, the remote resistor has in effect about 90 pF of capacitance in parallel with it, while the local resistor has less than 10 pF. It's feasible to restore equality of bandwidth by adding, in parallel with the local R_{in}, an actual capacitor to make up the difference. (Such a compensating capacitor is shown, as C_{comp} and in parallel with the local resistor, in the wiring diagram in Fig. 6.4.) The harder task is to know when the right amount of capacitance has been added. There's a first test you can do when both the local and the remote resistor are at ambient temperature. Another test is to achieve noise balance, with the use of (say) 10-kHz bandwidth in downstream filtering, and then to change to 100-kHz bandwidth. Only if the capacitive matching is correct will the noise still be in balance.

7. Practical guide to Johnson-noise measurements

7.1 Introduction

As you are now keenly aware, the subject of noise and noise measurements with NF1-A can be complicated. This section was written with the express intent of helping you get started making real noise measurements. In this section, we are not going to explain the physics of your measurements. That is done in the preceding sections. Rather, here we give detailed instructions, in a step-by-step manner, on how to configure the apparatus for several explicit measurements. We hope that if you follow these directions, you will be able to make these measurements and compare your results with the same measurements made with your unit at the factory. In this section, we will put in 'genuine' values for the measurements you will make. These measurements were made on a unit we keep at the factory. Since the electronics cannot be identical, they will not be exactly the same as your results. The measurements we report should, however, be close (within a few percent) to what you obtain.

All of the measurements you will be making will be on Johnson noise (as in Chapter 1), using the pre-amplifier configuration installed at the factory before the unit was shipped. Shot noise (as in Chapter 3) requires rewiring of the low-level electronics, with which you will be more comfortable after you have had experience measuring Johnson noise.

We will also do some sample calculations using the generic raw data. These calculations are essential for extracting both Johnson and shot noise from the raw signals. Again, the *physics* of these calculations is carefully explained in the other sections of the manual.

7.2 Test measurement of Johnson noise

Let's start by examining the 'default' configuration of the low level electronics (LLE). Figure 7.2a shows photographs of the unit. The left is an outside view. The right photo shows it detached and 'flipped' so that the pre-amplifier section is on the lower *right* side.

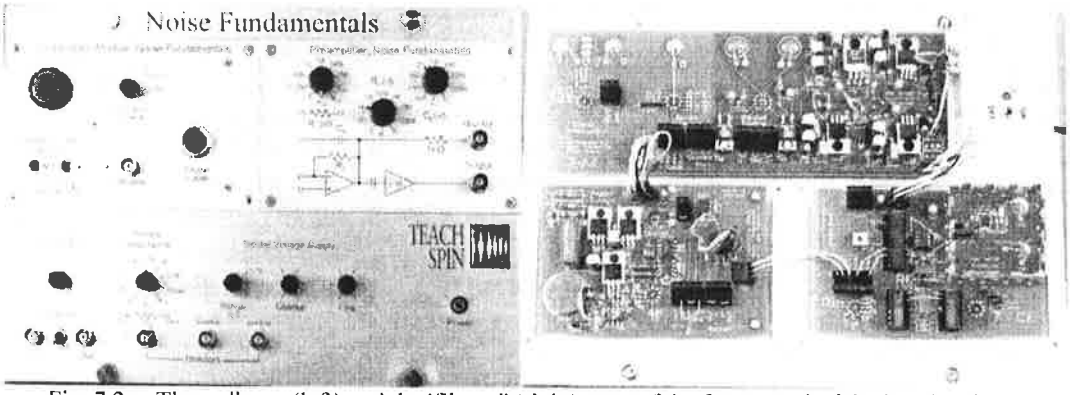


Fig. 7.2a: The ordinary (left) and the 'flipped' (right) state of the front panel of the low-level electronics. Note that the LLE's power cord will emerge from the right side of the panel in both orientations, provided you execute the flip in a top-over-bottom manner.

To 'flip' the front panel, loosen the four stainless steel thumb screws (two at the top and two on the bottom) that hold the thick aluminum front panel to the black-finished steel enclosure. With the broad side of the black enclosure resting on a table, flip the front panel top-over-bottom (*NOT* left over right) and then set it back into the steel enclosure. The enclosure was designed to support the aluminum frame in a way that provides convenient access to all of the electronics inside the LLE. (The power cable for the LLE leaves the panel from its *right*-hand side in both frames of Figure 7.2a.)

Now, you should confirm that the LLE is indeed configured in the so-called 'default' configuration. This is the configuration optimized for Johnson noise experiments. The schematic diagram of the Johnson noise configuration is shown in Figure 7.2b.

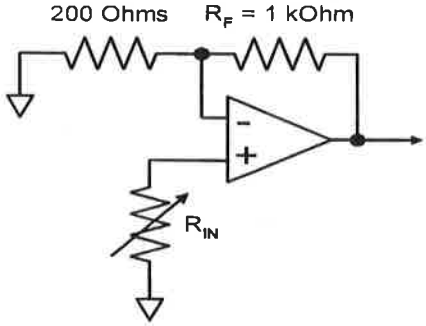


Fig 7.2b: Johnson noise preamplifier schematic

The wiring diagram for this configuration is shown in Figure 7.2c.

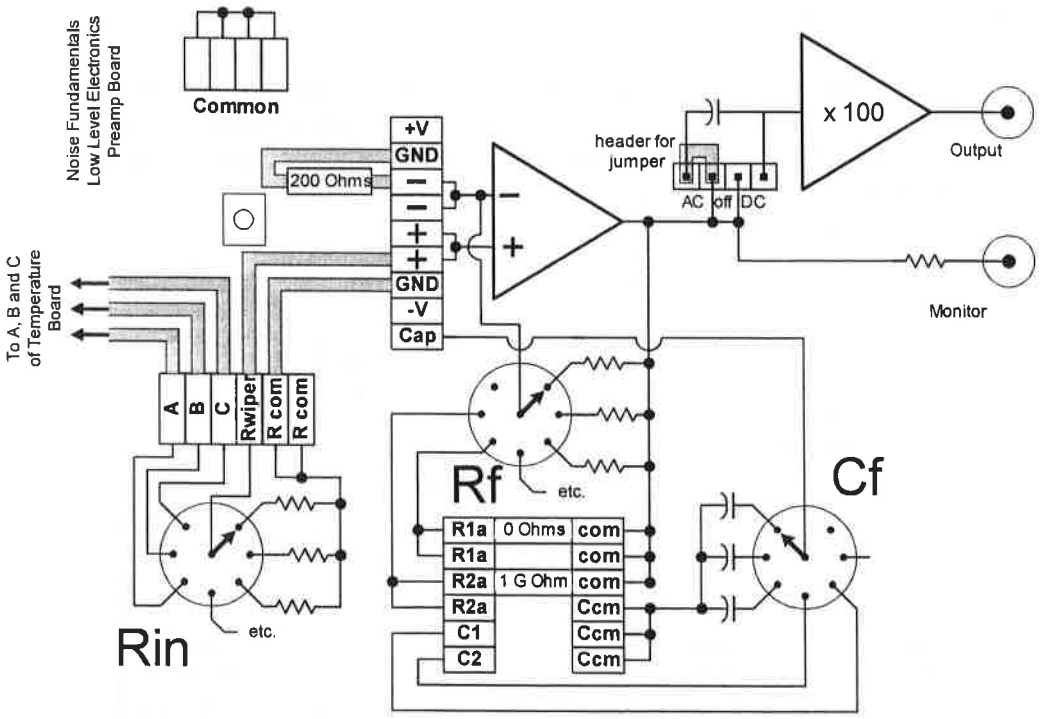


Fig. 7.2c: Wiring diagram of the default condition of the interior of the low-level electronics.

Note that in addition to the signal wiring shown here, there is also power wiring. Connectors convey regulated power from a fixed power-conditioning board to the printed-circuit boards of the temperature-control and pre-amplifier modules. Be sure that the switch on the aluminum block located in the upper right corner is in the **ON** position. This switch controls the d.c. power to the entire LLE unit.

After you have checked the configuration of the LLE, flip the aluminum panel back into its original orientation, set it in place, and tighten the four thumb screws that attach it to the steel enclosure.

Now, check all of the settings on the LLE.

Pre-amplifier:	R_{in} = 10 kΩ, R_f = 1 kΩ, C_f = min (not used); 200-Ω resistor: hard-wired into place
Temperature Module:	Heater Voltage = 0.0, Current Source = any (not used), Probe Cable: input covered with shielding cap
Lower Half of Panel:	None of these settings matter for these measurements.

Connect the low-level electronics to the high-level electronics. This is done via two connections. One connection is made with the power cable (permanently connected to the LLE) to the front connector on the lower left side of the high-level electronics (HLE) box. The other connection is the signal line, which is connected with a blue BNC cable from the x100 Gain Output of the pre-amplifier to an appropriate input on the HLE. The cable diagram is shown in Figure 7.2d

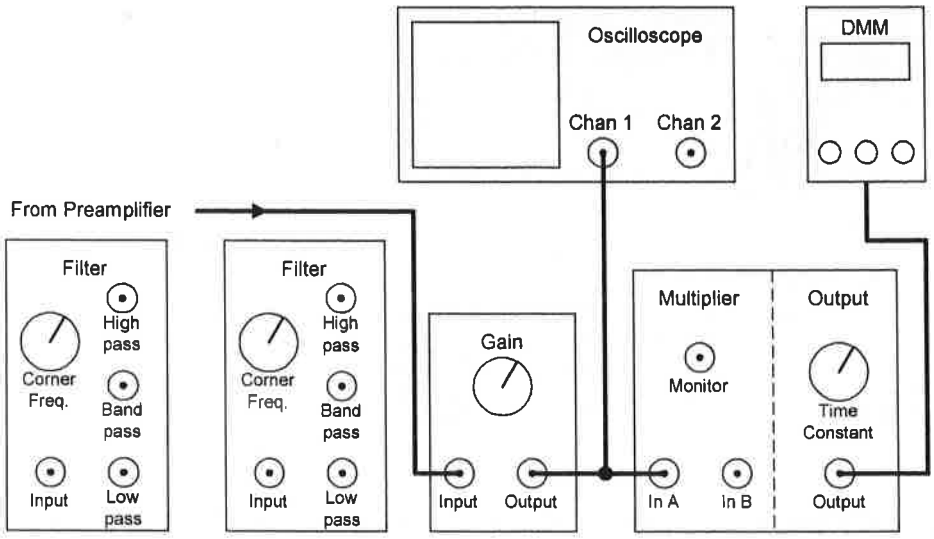


Fig 7.2d: Cable Diagram for Johnson noise measurement at full bandwidth

To get the whole unit powered up, find the ±15-Volt d.c. switching supply and connect its output line directly to the rear panel of the HLE. When you use the a.c. power cord to connect this power supply to the a.c. line, the HLE will be powered up – there is no separate power switch. The connection you have previously made between the units will also power up the LLE (with carefully conditioned d.c. power) from the HLE. When the power is on for the entire unit, you should see the green LEDs on both the LLE and the HLE light up. If both of the LEDs do not light up, check your connections, and also make sure the switch on the *inside* of the LLE is in the ON position.

Before you make your first measurement, it would be advisable to make sure your equipment is in an environment which reduces the possibility of outside electromagnetic signals interfering with your measurements. Keep the LLE unit away from your oscilloscope, from any power supply, power transformers, or any operating electronic instruments. If you can, turn off any fluorescent lights. You may later want to identify the various sources of interference that can affect your measurements, but for now you want to eliminate any real or potential source of such interference. Similarly, when it comes time to use the source resistors in your temperature probe, you need to be sure the probe has its cylindrical shield in place (covering the enclosed resistors), and you will need to screw its cable snugly to the previously capped-off input connector on the Temperature Module of the low-level electronics.

You are now ready to make your first measurement. The noise signal you will see is from an internal 10-kΩ resistor. As shown, the signal on the BNC cable from the LLE is connected directly to the GAIN module.

Gain: Fine Adjust: 30; Switch (top): x10; Switch (bottom): x1; (for Gain = 300)
Input Switch: AC

Multiplier: Input Switch: AC
Multiplier Switch: AxA

Output: Time Constant: 1 second
Meter Scale: 0 – 2 V

Oscilloscope: Sweep: 2 μs/div
Vertical: 2 V/div
Trigger: Channel 1, Normal, Rising, 11 volts
Acquire: Average 128
(used to get qACF, quasi-auto-correlation function – see 'scope traces below, or consult Appendix A.11)

Generic Data

Result (internal 10 kΩ source, full bandwidth, HLE gain 300): output 0.7421 Volts

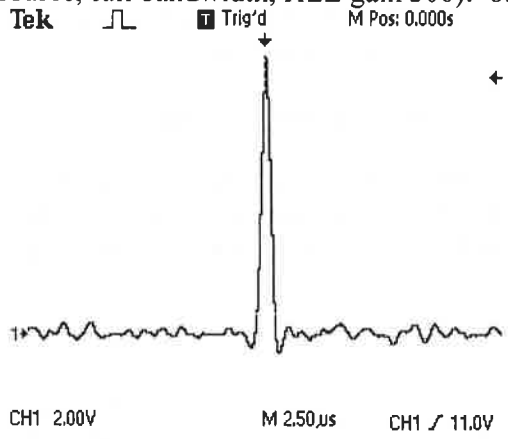


Fig. 7.2e: qACF (see Appendix A.11) -- internal 10 kΩ resistor, full bandwidth

Now make ONLY ONE change. On the preamplifier board, rotate the switch marked R_{in} to 10 Ω. Measure the output voltage.

Result (internal 10 Ω source, full bandwidth, HLE gain 300): output 0.2123 Volts

Check that the resistors mounted in the variable temperature probe are the same as those installed at the factory. This is done by plugging the probe cable into the Breakout Box, and testing the connections with an ohmmeter:

$$R_A = 10\ \Omega;\quad R_B = 10\ k\Omega;\quad R_C = 100\ k\Omega\ .$$

Now connect the cable from the variable-temperature probe to the connector on the LLE. To connect to the 10-kΩ resistor inside this probe, rotate the switch marked $R_{in}(\Omega)$ to B_{ext}. Repeat the measurements made above.

Result (probe 10 kΩ source, full bandwidth, HLE gain 300): output 0.3465 Volts

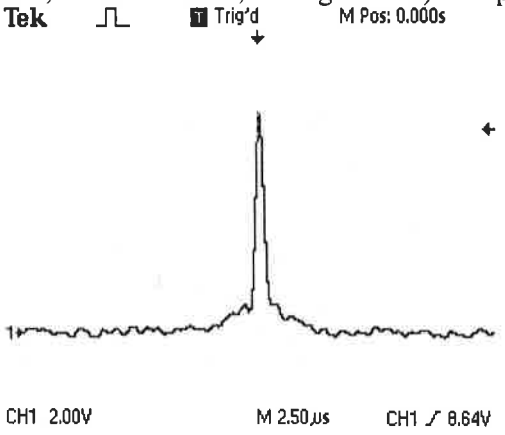


Figure 7.2f: qACF -- probe 10 kΩ resistor, full bandwidth

Now connect a 10-Ω resistor inside the variable temperature probe by rotating $R_{in}(\Omega)$ to A_{ext}. Repeat the measurement of the output voltage.

Result (probe 10 Ω source, full bandwidth, HLE gain 300): output 0.2168 Volts

Now you will once again observe the noise signal from a 10 kΩ resistor, but this time you will reduce the bandwidth of the amplification by the controlled use of filtering. First of all, change the cabling of the HLE to that shown in Figure 7.2g.

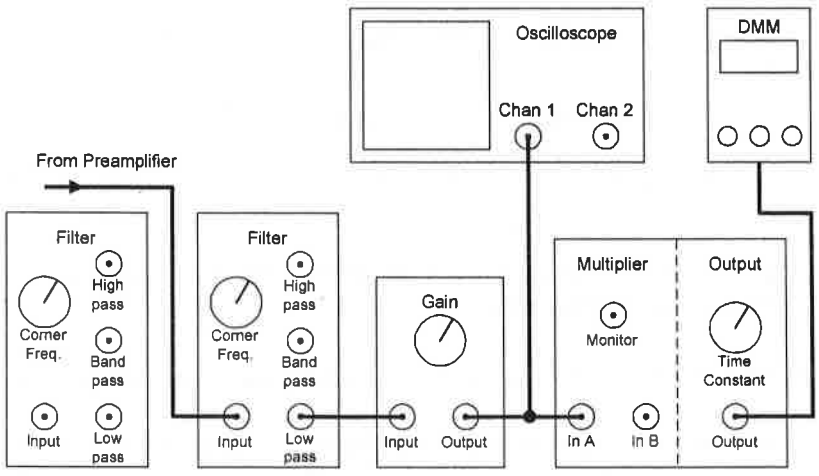


Figure 7.2g: Cabling diagram for the Johnson noise of 10-kΩ resistor at reduced bandwidth.

HLE
Gain: Fine adjust: 100, Switch (top): x10, Switch (bottom): x1, (for Gain 1,000)
Input switch: AC
Filter: Input: AC
Output: Low pass
Corner frequency: 100 kHz

LLE
Rotate $R_{in}(\Omega)$ to B_{ext}
All other settings on LLE the same

Result (probe 10 kΩ source, 100 kHz bandwidth, HLE gain 1000): output 0.7950 Volts

Oscilloscope: Sweep – 10 μs/div
Vertical – 2 V/div

Triggering – same

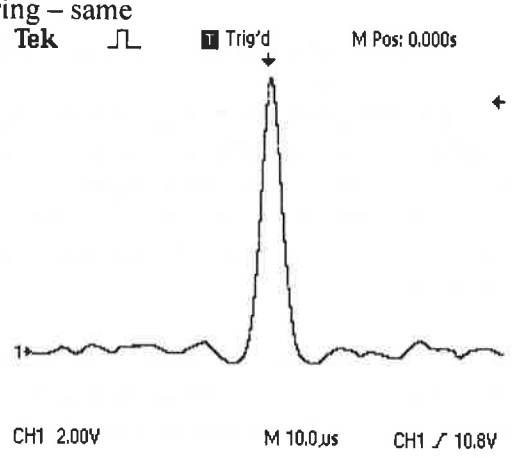


Fig 7.2h: qACF -- probe 10-kΩ resistor, 100 kHz bandwidth

Make only one change in the set-up. Rotate the $R_{in}(\Omega)$ switch from B_{ext} to A_{ext}. This connects a 10-Ω resistor to the preamplifier. Measure the voltage out

Result (probe 10 Ω source, 100 kHz bandwidth, HLE gain 1000): output 0.2465 Volts

Finally, repeat the above measurements for 10-kΩ resistor that is mounted inside the pre-amplifier. Rotate $R_{in}(\Omega)$ to 10 k. All other settings on LLE, HLE, and oscilloscope remain the same.

Result (internal 10 kΩ source, 100 kHz bandwidth, HLE gain 1000):output 0.9260 Volts

The qACF for this configuration is shown in Figure 7.2i.

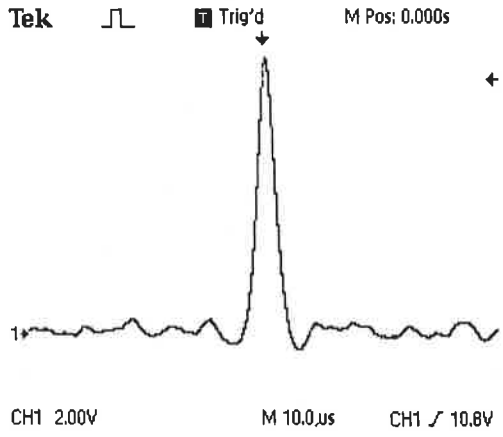


Fig 7.2i: qACF -- internal 10-kΩ resistor, 100 kHz bandwidth

Make ONE change: Rotate $R_{in}(\Omega)$ from 10 k, to 10. That makes a 10-Ω resistor next to the pre-amplifier the 'signal' source.

Result (internal 10 Ω source, 100 kHz bandwidth, HLE gain 1000): output 0.2461 Volts

You have made four measurements, all of which were directed toward the Johnson noise of a 10-kΩ resistor, but under different conditions. We will not try to explain why these measurements do not produce the same output noise signal, since they are all from the same resistance, and use the same pre-amplifier. But the differences are important, and you may already understand them. If not, you will understand them after reading the other sections on Johnson noise.

Now that you have made eight measurements of an average output voltage, under different conditions, it will help to put them in some organized format. Creating your own organization for the data you will be collecting is extremely helpful. It is easy to forget the parameters used in your measurements, and thus easy to make mistakes. In Table 7.2, you can see how we chose to organize the raw data.

Voltage (meter out)	Source resistor	Location	LLE gain	HLE Gain	Bandwidth
0.7421 V	10 kΩ	Internal	600	300	Full
0.2123 V	10 Ω	Internal	600	300	Full
0.3465 V	10 kΩ	Probe	600	300	Full
0.2168 V	10 Ω	Probe	600	300	Full
0.9260 V	10 kΩ	Internal	600	1,000	100 kHz LP
0.2461 V	10 Ω	Internal	600	1,000	100 kHz LP
0.7950 V	10 kΩ	Probe	600	1,000	100 kHz LP
0.2465 V	10 Ω	Probe	600	1,000	100 kHz LP

Table 7.2: Tabulated measurements of raw data

7.3 Sample Calculation of noise density

Table 7.2 has all the raw data for the measurements we made at TeachSpin. By now you should have your own version of Table 7.2 of your measurements. But the motivation for these measurements is to determine the Johnson noise voltage *from the resistor.* That noise signal itself originates far back at the beginning of an amplifier – filter -- multiplier chain. In order to get an accurate measurement of this noise signal, we will need to take into account the noise generated in this amplifier chain.

Let's start with the multiplier. The master equation for Johnson noise (Section 1.0) involves the square of the signal voltage. The multiplier provides the electronic means of squaring the input signal voltage, but be careful! The output of the multiplier is itself a voltage, and that voltage is *proportional* to the square of the input voltage. To obtain the value of the square of the input signal (with the correct proportionality and units) one must multiply the multiplier output voltage by the factor [10 Volts]. Mathematically we can write this relationship as

$$V_{out}(t) = [V_{in}(t)]^2 / (10\ \text{V})\ \text{, so } < [V_{in}(t)]^2 > = (\text{measured average } V_{out}) \text{ x } [10\ \text{Volts}]\ .$$

Read the previous paragraph again to make sure you understand it. You may want to put d.c. test voltages, or a sine-wave test signal, into the just the multiplier to check your comprehension.

Since noise density is defined as $< V_n^2 > / \Delta f$, we need now to calculate an accurate bandwidth Δf . We call this 'the equivalent noise bandwidth' or ENBW. Unfortunately, this is not the same as the -3-dB points of the filter discussed in Section 2 and in great detail in Section 5.2. In Section 5.2, we derived the expression for the ENBW to be

$$ENBW = f_c \frac{\pi}{4\gamma} = f_c \frac{\pi Q}{2} \quad ,$$

where f_c is the corner frequency and Q is the quality factor of the filter. For the 100-kHz bandwidth low-pass filter, the values we will use in this sample calculation are

$$f_c = 98.82\ kHz \quad and \quad Q = 0.738$$

Thus ENBW = 114.6 kHz (for the 100kHz filter)

The values you should use in *your calculations* are given in the *custom data sheet at the front of this manual.* These values were determined using your apparatus, and apply specifically to your instrument.

The noise signal from the resistor is amplified in both the high level electronics (HLE) and the low level electronics (LLE) and those are the numbers you recorded in your version of Table 7.2.

Now we are ready to calculate the equivalent signal at the start of the chain. We will call this equivalent signal 'the input'. The noise density is given by

$$S = \frac{V_{meter} \cdot 10\,Volts}{[(HLE\ gain)^2 (LLE\ gain)^2]} \cdot \frac{1}{ENBW} \quad .$$

For the 10-kΩ internal resistor at 100-kHz bandwidth, this gives

$$S = \frac{(0.9260\,V)(10.0\,V)}{(1000)^2 (600)^2} \cdot \frac{1}{114.6 x 10^3\ Hz} = 2.244 x 10^{-16}\ V^2 / Hz$$

This does not yet give the Johnson noise density of the 10-kΩ resistor, because we have yet to deal with the fact that the electronics also contributed to this noise signal. But since we are dealing with noise power density, we can subtract off this amplifier's contribution from our signal. The noise of a 10-Ω resistor adds very little to the noise of the amplifier in the chain, so we can (for the sake of these calculations) consider the noise signal from the 10-Ω resistor experiments to come *entirely* from the amplifiers. Thus we use these 10-Ω noise density measurements to subtract the amplifier noise density from our 10-kΩ measurements to obtain a determination of the Johnson noise density for the 10-kΩ resistor itself. These data and the results of these calculations are given in Tables 7.3a and 7.3b.

Resistance (source)	ENBW	Noise density ($Volt^2/Hz$)	Noise density (amplifier noise subtracted)	Voltage noise density $(Volt^2/Hz)^{1/2}$
10 kΩ (pre-amp)	114.6 kHz	2.244×10^{-16}	1.648×10^{-16}	12.8 $nV/(Hz)^{1/2}$
10 Ω (pre-amp)	114.6 kHz	0.5965×10^{-16}		7.72 $nV/(Hz)^{1/2}$
10 kΩ (probe)	114.6 kHz	1.927×10^{-16}	1.330×10^{-16}	11.5 $nV/(Hz)^{1/2}$
10 Ω (probe)	114.6 kHz	0.5975×10^{-16}		7.73 $nV/(Hz)^{1/2}$

Table 7.3a: Data taken at 100 kHz bandwidth

In the last column we have listed the square root of the S-values, sometimes called the 'voltage noise density' in its own units of 'Volts per square root of a Hertz', or V/√Hz. This is sometimes shortened to 'Volts per root Hertz'. This can be useful if you want to check the amplifier noise against the specifications of the OPA-134 op-amp used in the input stage of the pre-amp, which lists the noise as 8 nV/√Hz (in good agreement with our measurements).

Resistance (source)	ENBW	Noise density ($Volt^2/Hz$)	Noise density (amplifier noise subtracted)	Voltage noise density $(Volt^2/Hz)^{1/2}$
10 kΩ (preamp)	1.1 MHz	2.082×10^{-16}	1.487×10^{-16}	12.2 $nV/(Hz)^{1/2}$
10 Ω (preamp)	1.1 MHz	0.5956×10^{-16}		7.72 $nV/(Hz)^{1/2}$
10 kΩ (probe)	1.1 MHz	0.9782×10^{-16}	0.3699×10^{-16}	6.08 $nV/(Hz)^{1/2}$
10 Ω (probe)	1.1 MHz	0.6083×10^{-16}		7.80 $nV/(Hz)^{1/2}$

Table 7.3b: Data taken at full bandwidth. The ENBW for this data is only an estimate. The full bandwidth depends on several factors, including the pre-amplifier configuration

8. Practical diagnostics

8.0. Introduction

We all owe a debt of gratitude to the engineers and scientists who developed the modern operational amplifier and other linear analog integrated circuits. These devices make it possible for students to execute reliable and affordable low-noise measurements without heroic efforts. But even with these remarkable low-noise components, the experimenter must be aware of several important and sometimes subtle effects that have important consequences for the accuracy of noise measurements. This section attempts to arm you with experimental techniques that will aid you in diagnosing the characteristics of your experimental set-up.

8.1. Three experimental parameters

The measurement of noise consists of determining three parameters: 1) output voltage, 2) amplifier gain and, 3) amplifier bandwidth. Consider the measurement of each one.

1) Output Voltage: The NF1-A unit provides two places where one can measure the amplified noise voltage; before the multiplier (squarer) where it is an a.c. measurement, and after the multiplier at the low-pass filter output, where it is a d.c. signal. The instrument was designed for the students to measure the d.c. voltage after the multiplier, which is why an expensive and accurate analog multiplier has been included in the HLE. It is assumed that most advanced labs will already have an accurate modern digital multimeter to make these d.c. measurements.

But you can intercept these amplified noise signals *before* the multiplier and measure the rms a.c. voltage with a digital voltmeter, a digital oscilloscope, or some computer-based data acquisition system. But be aware of the limitation of accuracy of the a.c. voltage measurements on these instruments. Consult their specifications. We certainly recommend that the students compare rms voltage measurements made with an a.c. true-rms voltmeter, and with the multiplier serving as squarer. They should be in reasonable agreement.

2) Amplifier Gain: The voltage gain of an operational-amplifier circuit is determined by the feedback resistors. The NF1-A unit has 0.1% precision resistors in the various amplifier stages, so one might assume that the gain is accurate to 0.1%. But this accuracy assumes that the op-amp itself has an open-loop gain that is very large at all of the frequencies that it is amplifying. This assumption sometimes breaks down at high frequencies (> 100 kHz), and thus the gain at high frequencies may not be the same as it is at intermediate frequencies (1 kHz).

At low frequencies (< 25 Hz), there is another concern. Usually the amplifier chain is configured with a.c. coupling, to avoid the problem of amplifying the ever-present d.c. offset of the op-amps. The a.c. coupling is, in effect, a high-pass filter, and thus the gain at very low frequencies (~ 1 Hz) drops off precipitously. This ordinarily (but certainly not always) does not cause a problem with a noise measurement, since there is ordinarily only a tiny portion of the bandwidth at these extremely low frequencies.

For most measurements you will make on a noise signal, the gain registered on the front panel of the HLE will be accurate to 0.2%. Students can always use a signal generator, a good calibrated attenuator, and an a.c. voltmeter or oscilloscope to check these specifications.

3) Amplifier Bandwidth: This critical importance of bandwidth in the filter sections is discussed in great detail in Section 2. But bandwidth limits of other amplifier sections, previously modeled with 'flat' gain constants G_1 or G_2, is also important. Capacitance in the first stages of amplification is the major source of the problem, because it can cause a decrease in the high-frequency gain for voltage-noise experiments, or a 'gain peaking' (increase in gain) in current-noise experiments. We have adopted the rather draconian technique of providing amplifier gain to about 1 MHz, and then limiting the effective bandwidth to 100 kHz or less using the filter sections. Under these circumstances, errors due to the amplifier sections' bandwidths make corrections under 1% to the equivalent noise bandwidth of the system as a whole. But this also reduces the noise signal, since we have the filters 'throw away' about 90% of the available signal. This is an example of excluding (high frequency noise) data because it is subject to systematic errors, even at the price of raising the statistical errors in the data that remain. Statistical error, after all, can always be reduced merely by taking more time to form stable averages.

But the moral of this story is that your measurements of noise are no more accurate than the poorest accuracy in any one of these three parameters: voltage, gain, and bandwidth. And of the three, bandwidth is most difficult to characterize and control.

8.2. The qACF and the qFFT

This section describes an important technique for diagnosing the amplifier chain's bandwidth characteristics. It requires a digital oscilloscope with a mathematics menu that includes a fast Fourier transform (FFT) capability. This is now a common and affordable instrument that should be a part of any advanced laboratory test equipment. We will refer to this instrument as DSO, a digital storage oscilloscope.

We will be showing you various DSO images. Figure 8.2 shows some typical images with boxed text identifying the important parameters. Many times the images may appear similar, so you will need to pay attention to the differences in scales that are displayed at the top and bottom of the images.

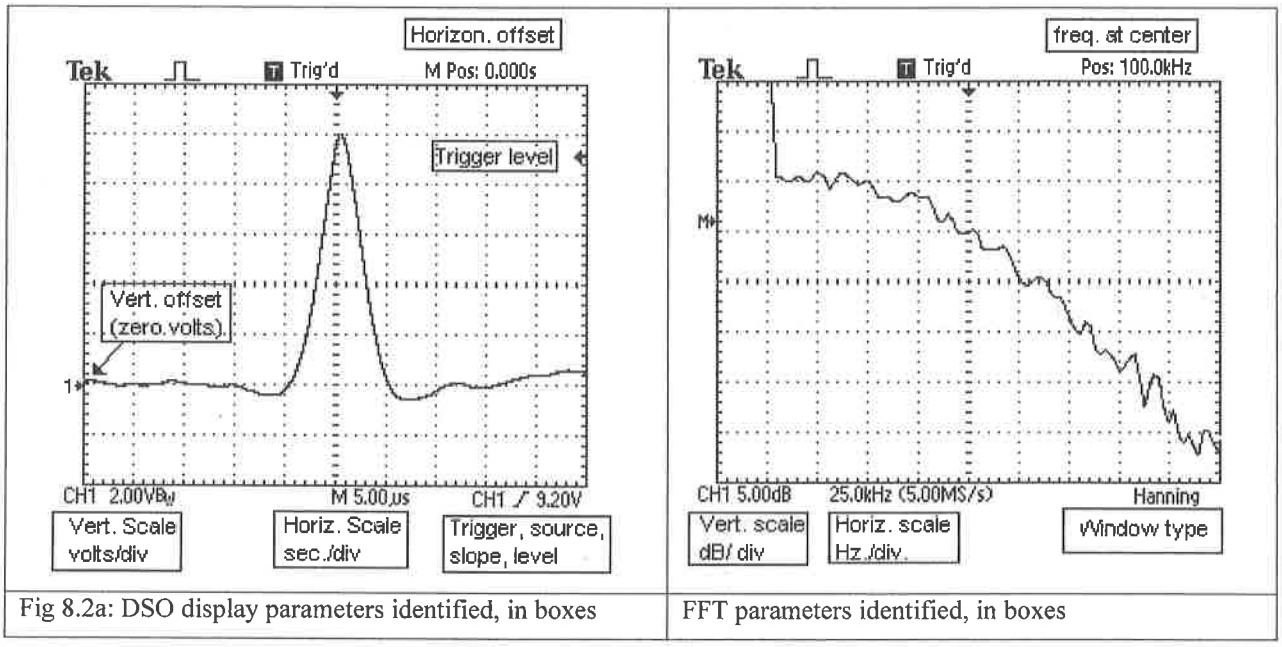


Fig 8.2a: DSO display parameters identified, in boxes

FFT parameters identified, in boxes

First let's examine voltage, or Johnson, noise signals with the unit configured in the so-called 'default' setting, as it came from the factory. We will be examining noise from both the internal and variable temperature probe resistors. Figure 8.2b shows the cabling diagram of the set up.

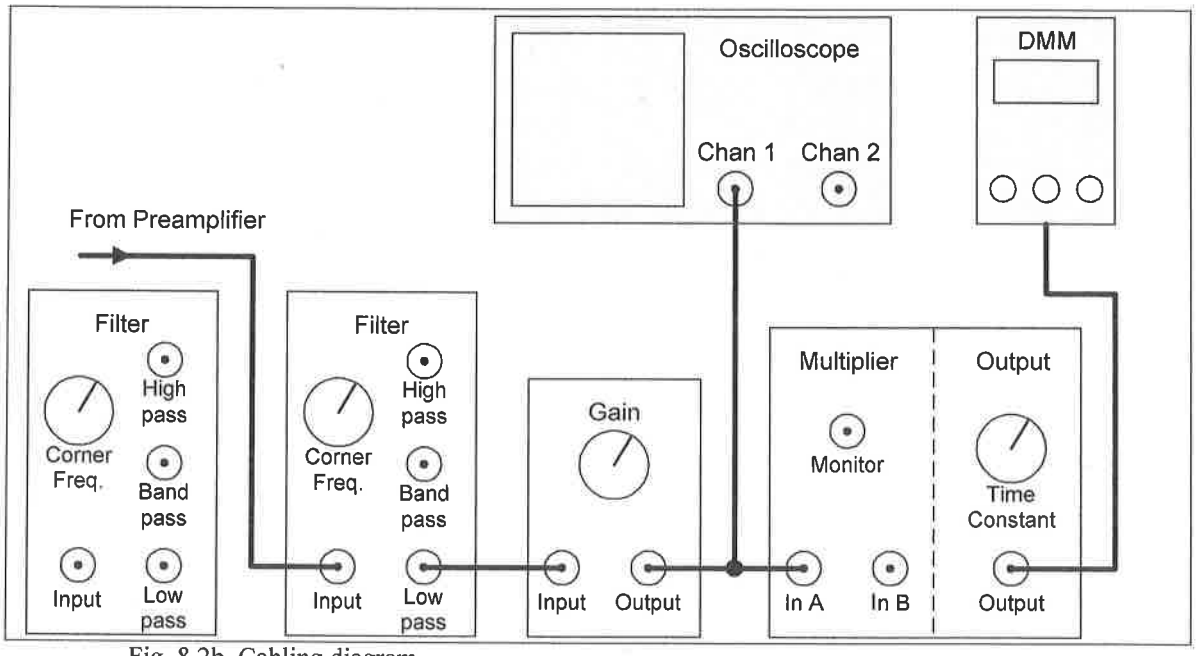


Fig. 8.2b Cabling diagram

Configure your apparatus as follows:

High Level Gain 800 — Low-pass filter, 100 kHz

R_{in} is 10 kΩ internal — AC coupled on all inputs

DSO Settings

Trigger: Channel 1, Normal, edge, positive slope

Level ~ 9 Volts

Time Base 5 μs/div — Persistence: 1 Second

Vertical Gain 2 V/div — Acquire: Sample

You should observe a signal something very similar to Figure 8.2c, seen some pages ahead. (This may require some adjustment of the triggering level).

Now change your scope so that it averages the signals using the maximum number of averages (in our case, 128) and you should observe a signal like Figure 8.2d. Change the time base by a factor of 10, to 50 μs/div (slower sweep), and you observe Figure 8.2e and 8.2f, again one with persistence and the other with signal averaging.

The averaged signals in d) and f) represent what we are calling the quasi Auto-Correlation Function (qACF), which is described in more detail in Appendix A.11. There is important information in these scope traces which we will soon discuss, particularly in the signals that have been averaged.

Now let's examine the FFT of these same signals, that is the signals that the DSO accepts for the trigger level we have established. (Note that trigger level is high compared to some average voltage level of the noise signal, so we are only sampling waveforms triggered by a larger-than-average voltage excursions). We call these qFFT. Figure 8.2g is the qFFT of the

signals observed in Figure 8.2c, that is at 5 μs/div, persistence 1 second, but no signal averaging. Of course, now you are looking at an FFT trace where the vertical axis is logarithmic, in 10-dB units, and the horizontal axis is in frequency units (250 kHz/div, and where the arrow at the top indicates the location of the 1-MHz point).

The mathematical details of the qACF, and qFFT are discussed in Appendix A.11. We have added the quasi- prefix to these names because the averaging process of the DSO makes these signals a bit different from the actual auto-correlation function of the time-varying voltage signal. For reasons discussed in the Appendix, the half-power (or -3-dB) point of filter or amplifier response shows up as the -6-dB point in 'scope displays of the quasi-ACF.

The math options on different DSOs have different characteristics that you will need to master. We will describe the ones on our Tektronix DSO. There are various ways to compute the FFT of a signal, and we have chosen the 'Hanning Window'. One can also expand the frequency scale by applying the 'zoom' function, but the display will record the new frequency scale at the bottom so you don't compromise the calibration.

All the data shown in Figures 8.2c - 8.2j comes from the same noise signal. Nothing in the signal has changed. It is still all from an internal 10-kΩ resistor, 100-kHz low-pass filter, 800 high-level gain, and all inputs are a.c. coupled, but we have adjusted the scope to analyze this data in different ways. Compare c) with d), both in the time-domain, and there is no surprise, except that one may just be able to detect in d) a small 'undershoot' after the main decay signal that is certainly not obvious in c) (which is not averaged).

The real surprise comes when in comparing g) with h). Clearly Figure 8.2h shows a significantly faster roll off in frequency response than Figure 8.2g. Although the noise signal source and amplification is the same for both figures, Figure 8.2h is the FFT of a time-averaged signal, which clearly is not the same as the averaged-by-eye-and-persistence FFT of a signal.

But these two images do not give us a clear picture of what is going on at lower frequencies. Since we set the 2-pole filters for 100-kHz low-pass, we should expect a flat response at say 1 kHz, and the response to decrease to the -6-dB point at 100 kHz. To observe this low-frequency response, we need to slow down the sweep, or time base, to 50 μs/div, which is what we did for the four figures in the lower half: e), f), i), and j).

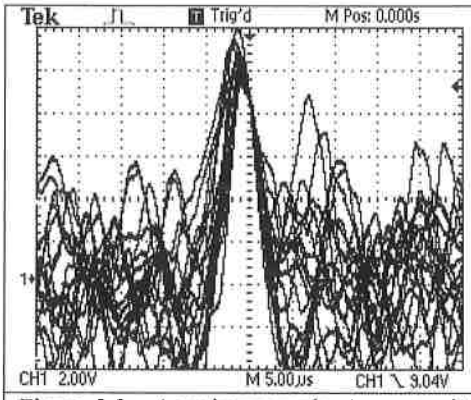


Figure 8.2c: Acquire=sample, 1 sec persistence

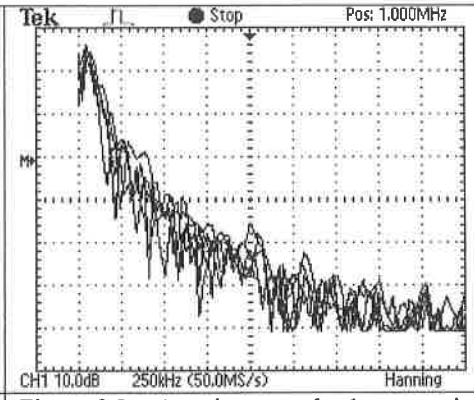


Figure 8.2g: Acquire=sample, 1 sec persistence

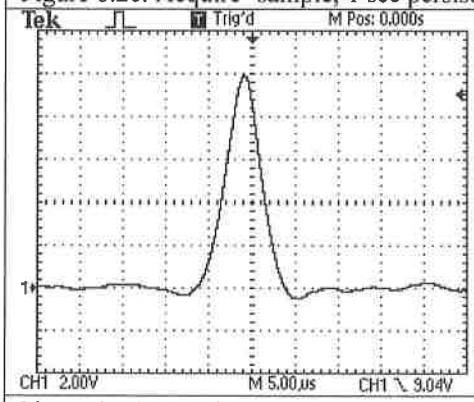


Figure 8.2d: Acquire=average, num = 128

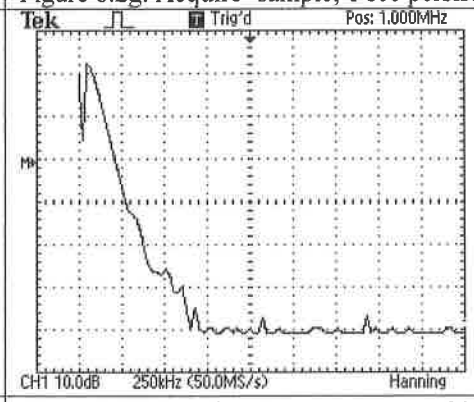


Figure 8.2h: Acquire=average, num = 128

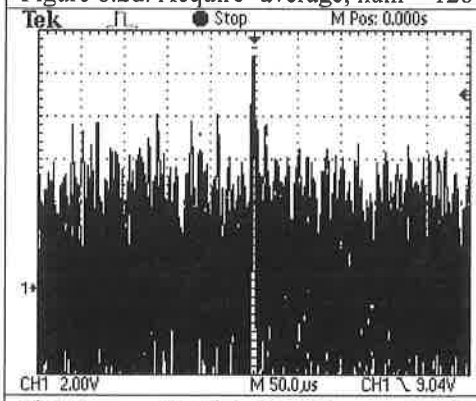


Figure 8.2e: Acquire=sample, 1 sec persistence

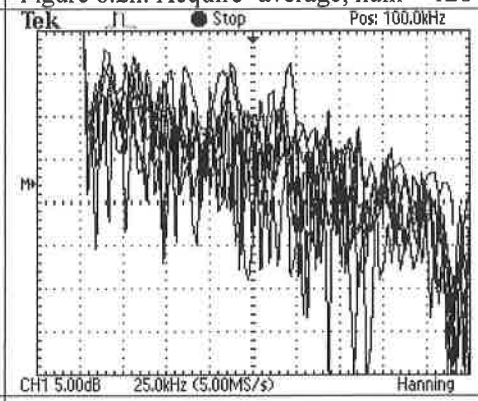


Figure 8.2i: Acquire=sample, 1 sec persistence

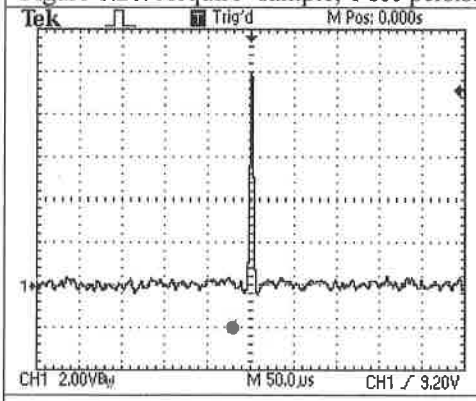


Figure 8.2f: Acquire=average, num = 128

Figure 8.2j: Acquire=average, num = 128

Looking at Figure 8.2e is truly unimpressive, but its averaged signal in Figure 8.2f is more enlightening. We observe only one narrow sharp spike, and nothing obvious in the baseline. If there were low-frequency interference signals in this experiment, they would show up on this baseline. We will see some of those signals later. If we look now at the FFT at this resolution, where the frequency scale is now 25 kHz/div, we can clearly characterize the gain profile on the averaged signal. At low frequency (<50 KHz) the gain is constant, and it drops off about -6 dB at the 100 kHz (center) mark. It also displays a smooth, regular, decrease in gain as the frequency increases, showing no anomalous bumps or dips.

Now let's examine the 'main peak' of the qACF in more detail. In Figure 8.2k, l we have expanded the time scale to 2.5 μs/div, but changed the trigger slope, so that k) triggers on a rising slope, and l) a falling slope. Note the horizontal shift in the peaks.

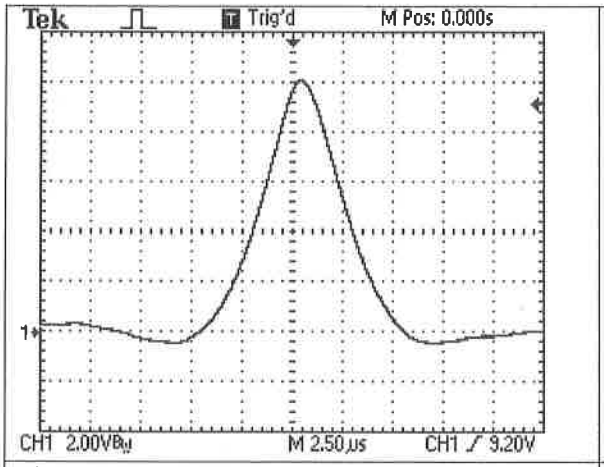


Figure 8.2k: qACF triggered on the rising slope

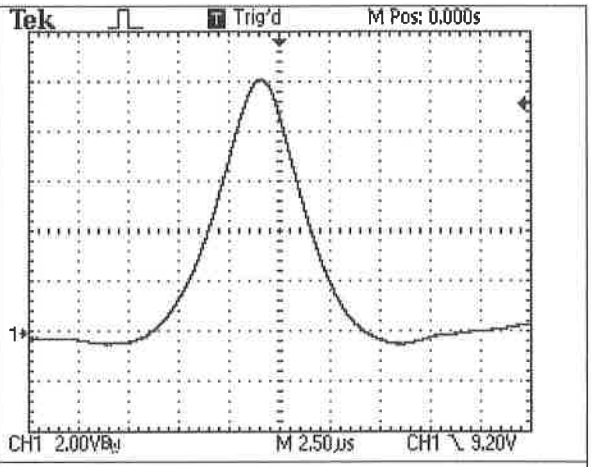


Figure 8.2l: QACF triggered on the falling slope

Try changing the trigger level. If you set the level very high the scope will not trigger at all, since the amplifier rails are at ±12 Volts. But if you trigger at, or very near, the zero crossing and signal-average, you get a display like 8.2m. Notice the main peak is bipolar, which is what you might expect. When you examine the FFT of these averaged signals you get a display shown in Figure 8.2n. This shows a peculiar gain vs. frequency spectrum, not at all like the one shown in Figure 8.2j. The big difference is the low frequency gain, which drops to near zero in 8.2n. This is not a faithful representation of the amplifier's configuration, and that is all due to the fact that we are triggering near zero level and averaging the signal before the scope calculates the FFT. You might also observe that the signals display much greater fluctuation.

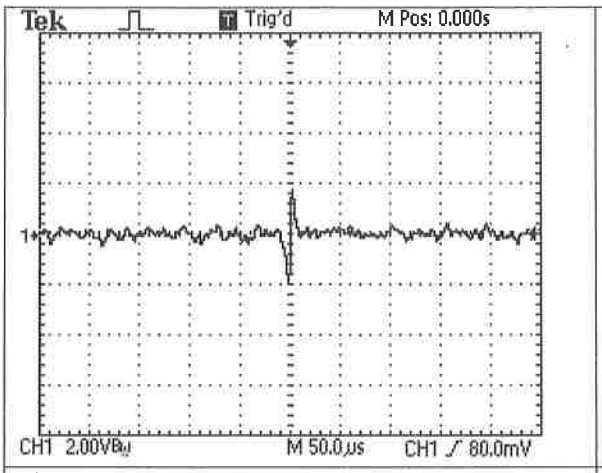


Figure 8.2m: qACF triggered at zero volts

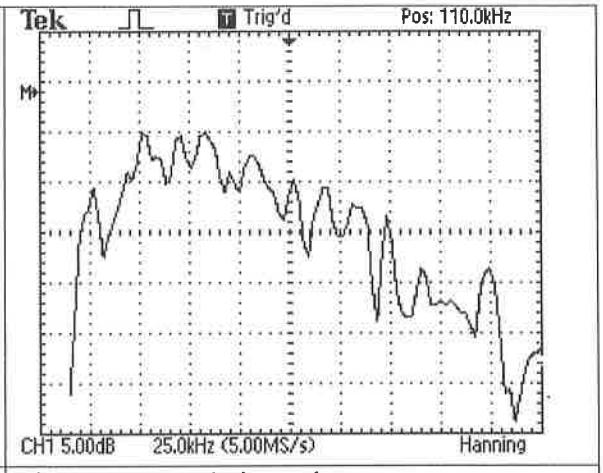


Figure 8.2n: qFFT triggered at zero.

We will now look at some other qACFs. This is only meant to give you a taste of what other time- and frequency-domain spectra can look like. Suppose we now take the filtered signal from the 100-kHz *band*-pass output filter circuit, rather than the 100-kHz low-pass. We will still look at the Johnson noise of the 10-kΩ resistor internal to the LLE. Now we will return to the 'correct' triggering level, ie. one that is considerably above the rms value of the noise signal (in this case ~ 9 Volts). Both the qACF and its qFFT are shown in Figures 8.2o,p. You should note that we are now showing a qFFT that is taken at slower horizontal scale (sec/div) than corresponding qACF.

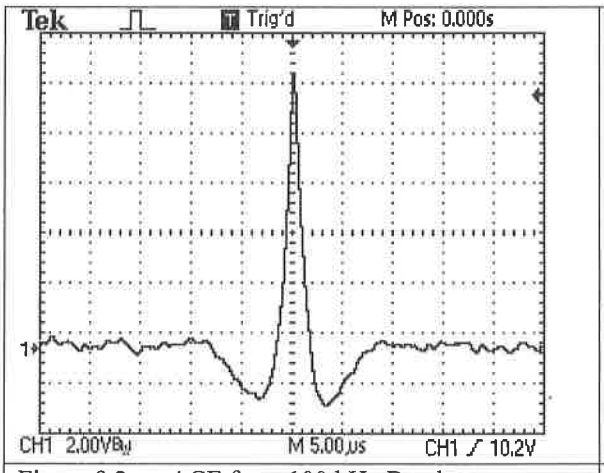


Figure 8.2o: qACF from 100 kHz Band pass

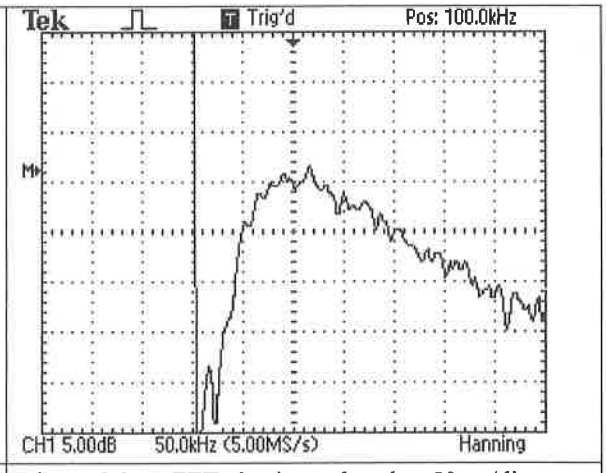


Figure 8.2p: qFFT, horizontal scale =50 µs/div

Figures 8.2q, r show the signals taken from the 100-kHz high-pass filter. (Note changes to a HLE gain (G_2) of 300, and to a faster sweep rate / frequency range to best show off the characteristics of the amplifier chain.)

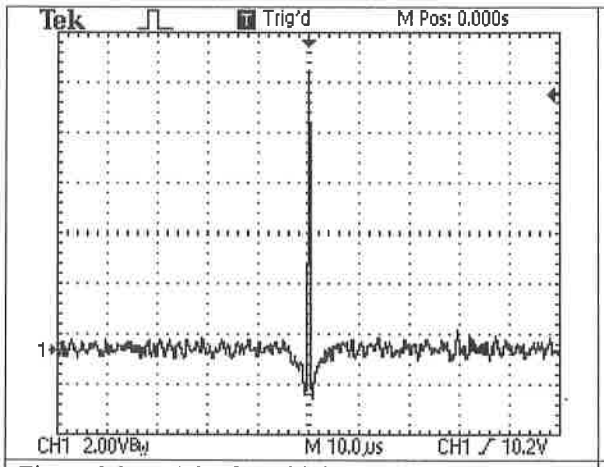


Figure 8.2q: qACF from high pass, G_2 =300

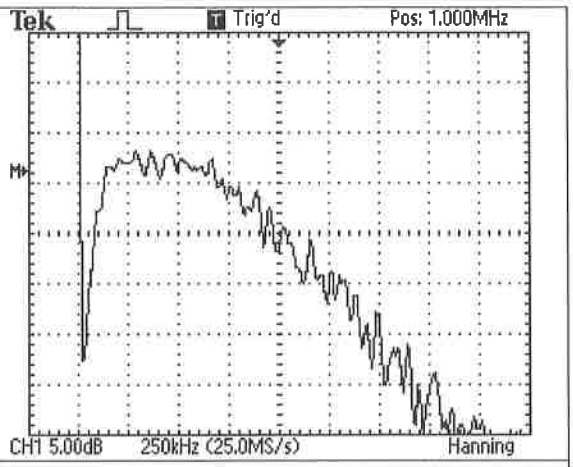


Figure 8.2r: qFFT from high pass, 10 µs/div

One might expect an 'ideal' high-pass filter to pass frequencies out to infinity. This is of course not possible. At some high frequency, the gain decreases. The high-pass output thus looks similar in character to the band-pass signal, though the details are different.

8.3. Full-bandwidth signals

Now that you have some experience with known-bandwidth noise signals, let's consider studying the qACFs and qFFTs of so-called 'full bandwidth' signals, ie. signals which are not passed through any filter. First, let's use the same Johnson-noise source, the 10-kΩ internal resistor R_{in}. Figure 8.3a shows the qACF with HLE gain of 300 and a trigger level of +10V, and Fig 8.3b show the qFFT of that same averaged signal with a 250 kHz/div horizontal axis. Recall that the arrow at the top indicates the position on the horizontal axis of the frequency recorded at the top right, marked Pos: 1.000 MHz, that is, 1 MHz. Notice that the gain (or the noise power) is reasonably constant to about 500 kHz, but then begins to decrease, and it decreases by 6 dB at about 1 MHz. This -6-dB point indicates that the full bandwidth is 1 MHz, as specified for the apparatus. The width of the qACF will also tell us the approximate bandwidth, but it is hard to estimate that from Fig. 8.3a. But it certainly is reasonable to estimate that width as 1 μs, or the reciprocal of the bandwidth of 1 MHz.

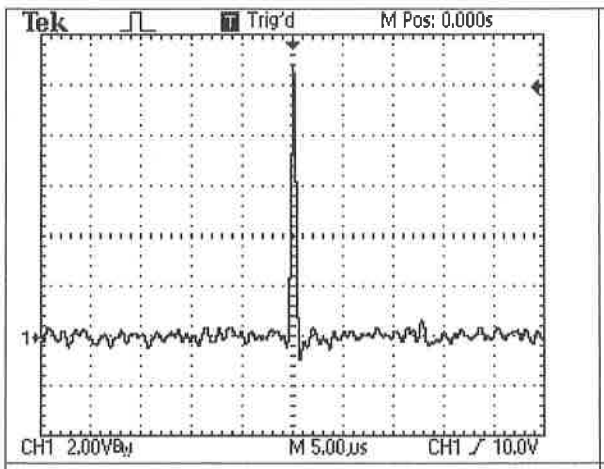


Figure 8.3a

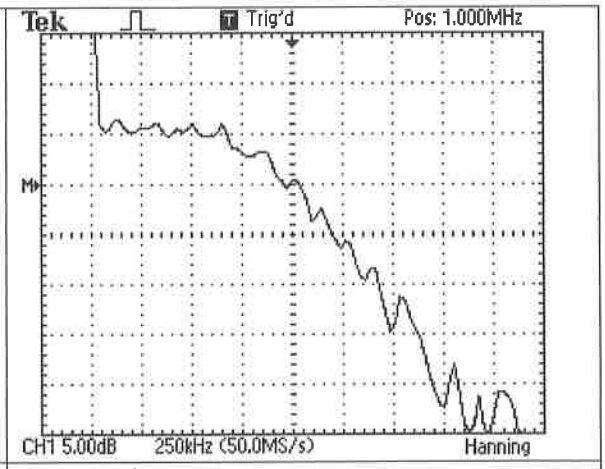


Figure 8.3b

Let's examine the same Johnson noise source, a 10-kΩ resistor, but this time the resistor is the one mounted in the variable temperature probe. Thus, you must connect the probe cable to the preamplifier box, but leave all the other cable connections the same. Figure 8.3c is the qACF and Figure 8.3d is the qFFT of this signal. Both these screens look different from the ones taken for the internal resistor of the same value. What has changed?

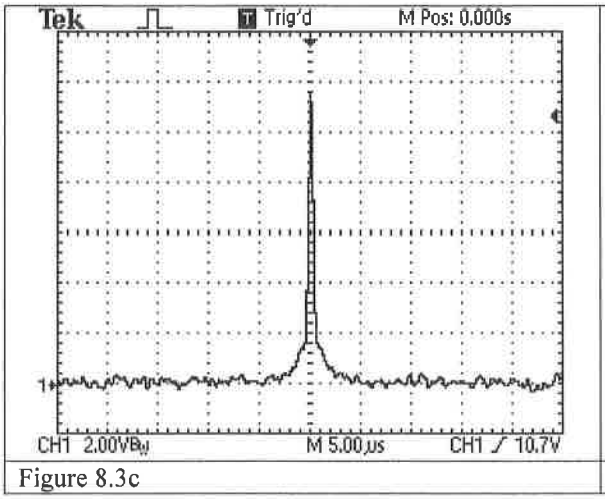


Figure 8.3c

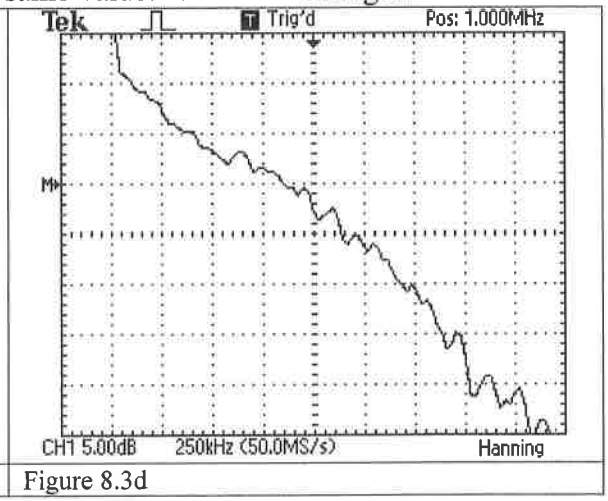


Figure 8.3d

There now exists an extra cable from the 10-kΩ source resistor to the input of the pre-amplifier. That cable adds a small resistance, a small inductance, and a small capacitance to our circuit. It turns out that the so-called 'small' capacitance, about 100 pF, has a significant impact on the noise signal. Consider the qACF in Figure 8.3c. The broad feature at the bottom of the central peak is due to the Johnson noise signal from the 10-kΩ resistor, but with the bandwidth reduced by the capacitance of the probe. The central sharp spike is the wide-band noise from voltage noise of the first-stage op-amp in the pre-amplifier, which is not reduced by the cable capacitance.

To better understand the qFFT under these conditions, switch to the 100-kΩ probe resistor as the Johnson noise source. (R_{in} switch position set at C_{ext}). Again we do not use any filter; these are full-bandwidth experiments. Figure 8.3e and f shows the data.

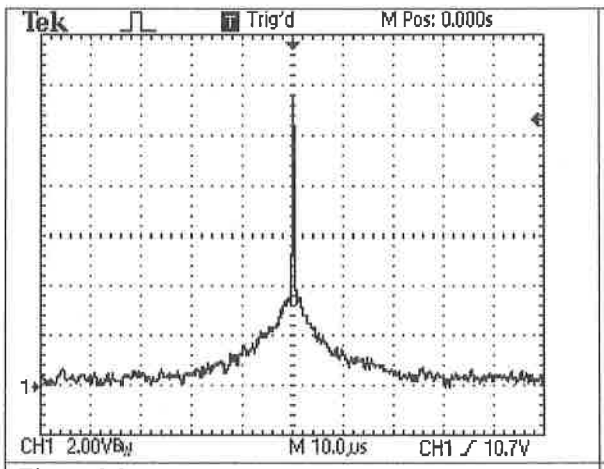


Figure 8.3e

Figure 8.3f

Note the change in the time base of the qACF from the previous 5 to 10 µs/div. There is now a much broader feature under the spike, again due to the 100-kΩ Johnson noise at bandwidth reduced by capacitive effects, and there is the spike due to the op-amp voltage noise. The qFFT now clearly shows a flat region of about 500 kHz, then a rising noise signal at lower frequencies. Be careful in interpreting this qFFT spectrum. The final noise signal, as monitored at the output of the low-pass filter applied to the multiplier output, is less than the total noise from the same-value resistor connection internally within the preamp. That is, the final noise signal is less from the probe resistor then from the internal resistor. *Simply put, the probe cable capacitance has reduced the bandwidth of the noise signal.* But this capacitance has not only reduced the bandwidth, it has changed the spectrum of frequencies in the noise signal.

To observe the lower frequency part of this spectrum, we again change the time base for both the qACF and the qFFT. Figures 8.3g and h show this data.

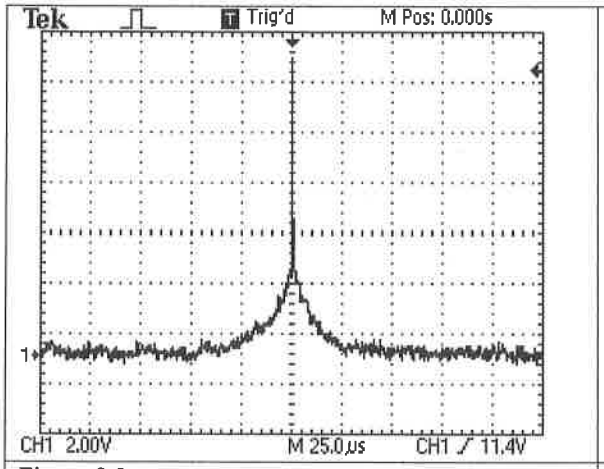


Figure 8.3g

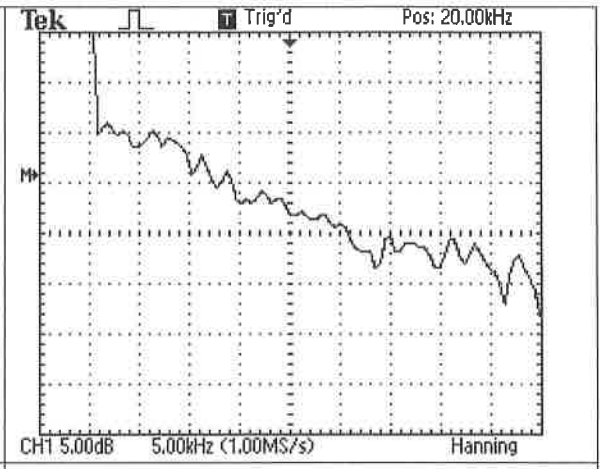


Figure 8.3h: qFFT, 250 µs/div

It is hard to see any 'flat spot' or plateau in the qFFT noise spectrum, but the qFFT does show a decrease by 6 dB at a frequency near 16 kHz. The noise is now being 'rolled off' by a single-pole RC filter. (R = 100 kΩ and C = 100 pF does indeed entail a 16-kHz corner.)

You might want to create an intentional single-pole RC filter and send your noise signal through it. Let's choose a 10-kΩ resistor and a 1-nF capacitor, which will give a time constant of about 10 µs and a -3-dB frequency corner at 16 kHz -- about the same as the RC time-constant for the 100-kΩ probe resistor plus the probe-cable capacitance. The schematic and wiring diagram for this filter is shown in Figures 8.3i and 8.3j. A note of caution: if you use a very big capacitors and small resistors to get the same time constant, you may run into oscillations. Op-amps are (typically) not designed to drive large capacitive loads.

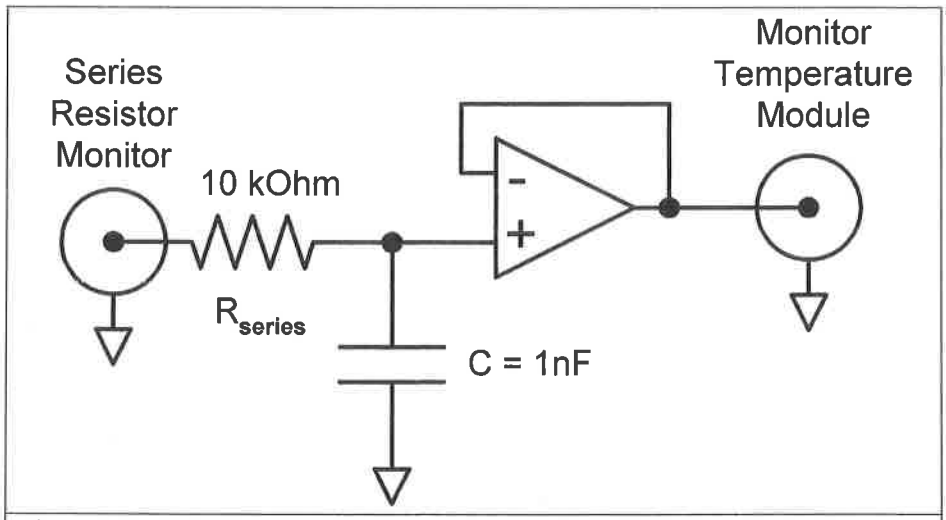


Figure 8.3i: An RC low pass filter using the op-amp buffer built in the LLE box.

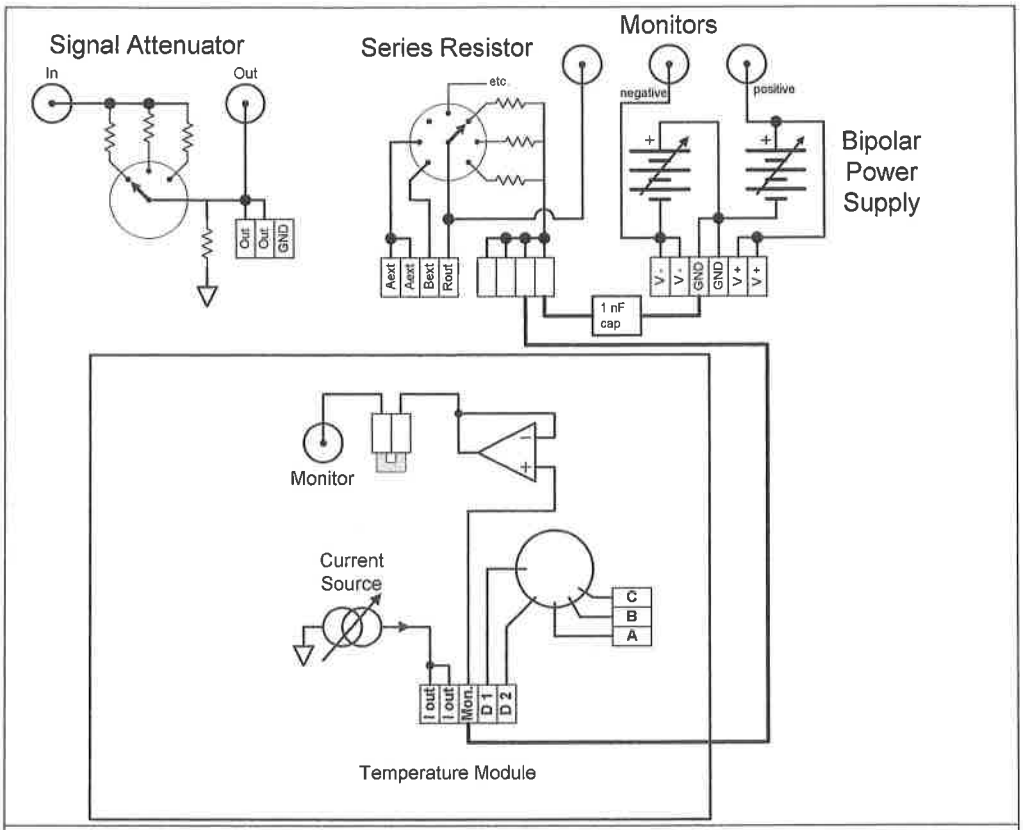


Figure 8.3j: Wiring diagram RC low-pass filter using the op-amp buffer built in the LLE box.

Now examine carefully the qACF and the qFFT of the noise signal with this single-pole filter (Figure 8.3k and l).

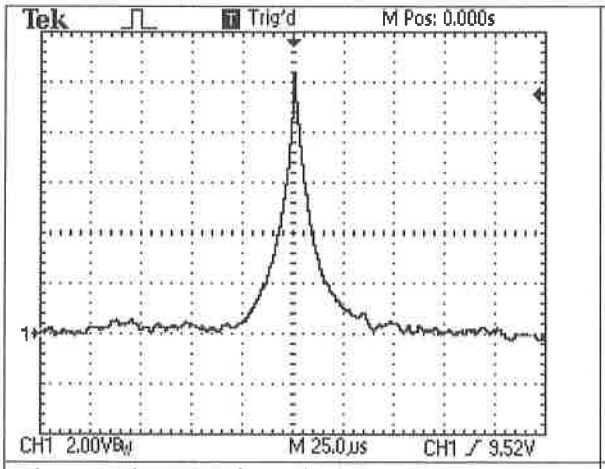


Figure 8-3k: qACF from single RC low-pass

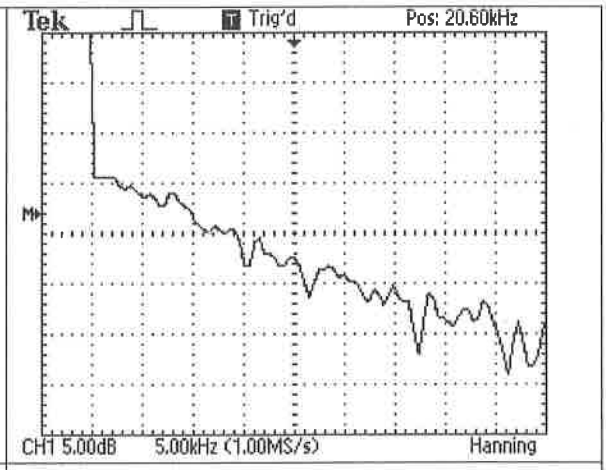


Figure 8.3 l: qFFT, 250 µs/div

You should compare these two signals with those in Figures 8.3g and 8.3h.

8.4. Observing Interference

We live in a world of electromagnetic radiation. At TeachSpin we have done our best (within reasonable limits on cost) to shield this equipment from these fields, but you want to be confident they have not somehow invaded the apparatus.

To get some idea of the electric fields in your laboratory you might try the following simple experiment. Cut a 15 cm (6 inch) piece of bare wire (found in your parts kit, called 'buss wire'), and stick it in the center conductor of the input BNC connector on your oscilloscope. If your scope has both 50-Ω and 1-MΩ input impedances, make sure you use the 1-MΩ choice. Now adjust the 'scope gain and triggering until you see a reasonably stable signal. Figure 8.4a shows what we see with the fluorescent lights turned on in our lab. Figure 8.4b shows the FFT of this signal.

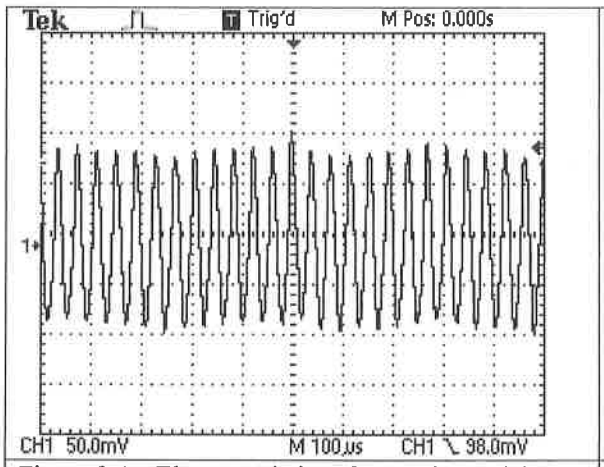


Figure 8.4a: Electrostatic interference in our lab

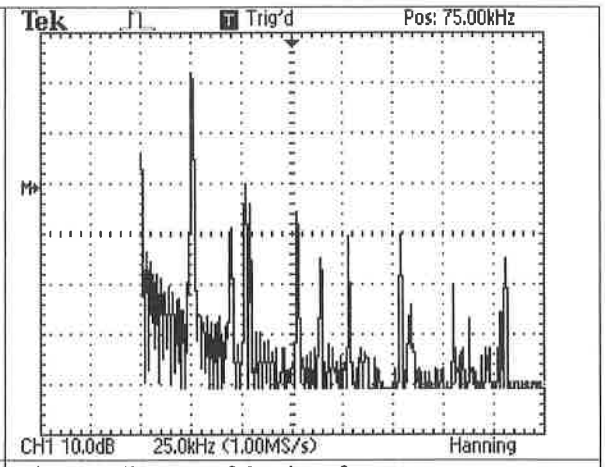


Figure 8.4b: FFT of that interference

You might notice that this signal increases as you move your hand towards the wire. Does it get much larger? Why? What happens if you put one hand on the wire and the other hand on the outside (grounded) cylinder of the input BNC connector? Keeping one hand on the wire, raise your other hand over your head and point in various directions, including toward fluorescent lights. In the lab at TeachSpin, these fluorescent lights are the major source of 25-kHz electric fields. Your lights might produce a different frequency of electric fields.

With the lights *off*, we change to a five-fold higher sensitivity on the 'scope, and observe a different spectrum of electric fields shown in Figures 8.4c and d.

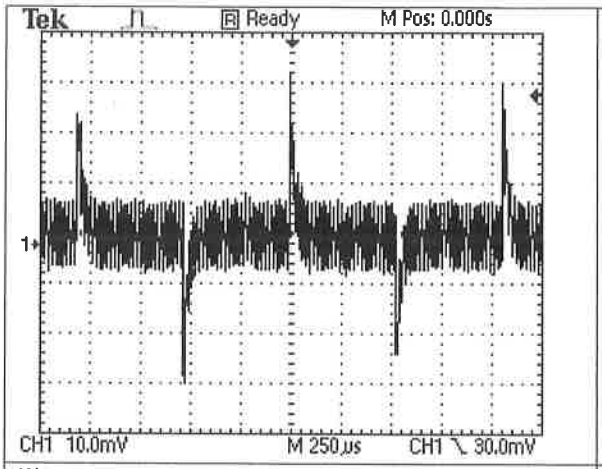


Figure 8.4c: qACF of electro-static interference with fluorescent lights turned off

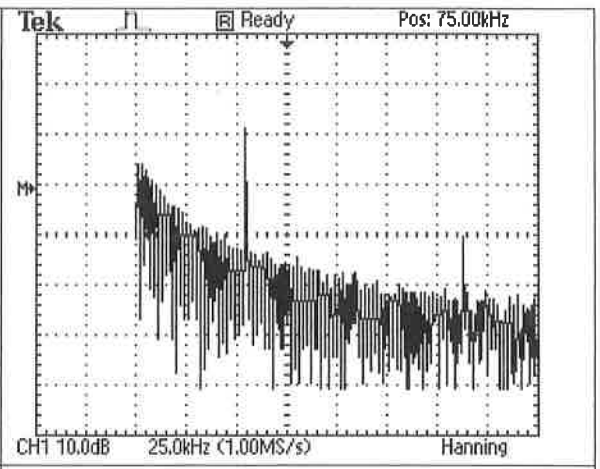


Figure 8.4d: qFFT of same.

What effect does an interference signal have on our noise measurements? To study this, we will again measure the Johnson noise from a 10-kΩ internal resistor and pass the signal through a 100-kHz low-pass filter. To allow some of this external electric field into the unit, we remove the variable-temperature probe cable from the LLE and **leave the shielding cap off the probe connector.** Now set up your oscilloscope to examine both the qACF and the qFFT. In Figures 8.4e and f we show these measurements made in TeachSpin's lab with the fluorescent lights on.

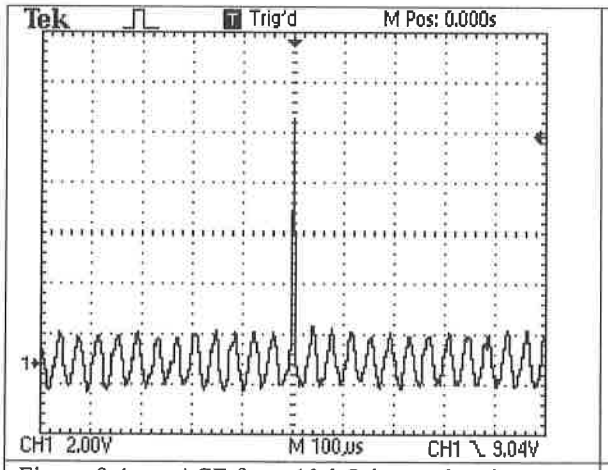


Figure 8.4e: qACF from 10-kΩ internal resistor. Interference effects with probe cap removed

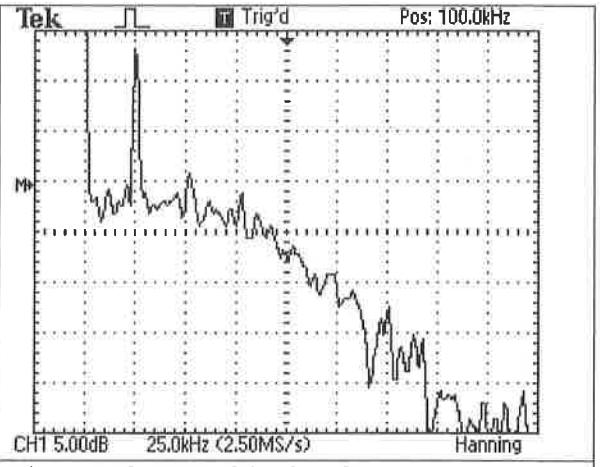


Figure 8.4f qFFT of that interference

It is interesting to note that although this interference signal is so obvious on the qACF, and shows up as a distinct and narrow peak in the qFFT, it only contributed about a 15% change to the final noise signal as monitored at the output with a voltmeter. Why is this? Try changing the noise source resistor, R_{in}, to 100 kΩ or to 1 MΩ. What do you observe? Can you explain this?

Leaving R_{in} = 1 MΩ, you might explore other possible leakage paths for electric fields. Replace the cap on the variable temperature probe connector, but now take a short buss-wire antenna, and place it in turn into the center connector of each of the seven open BNC connectors on the LLE panel. See if you can identify which ones, if any, are more likely to

form leakage paths into the unit. Those connectors are good candidates for the use of a cover cap, of the sort provided in the spare parts boxes.

You can inject a controlled amount of 'interference' into the unit by connecting a signal generator into the monitor *out*put in the preamplifier. The cabling diagram is shown in figure 8.4g.

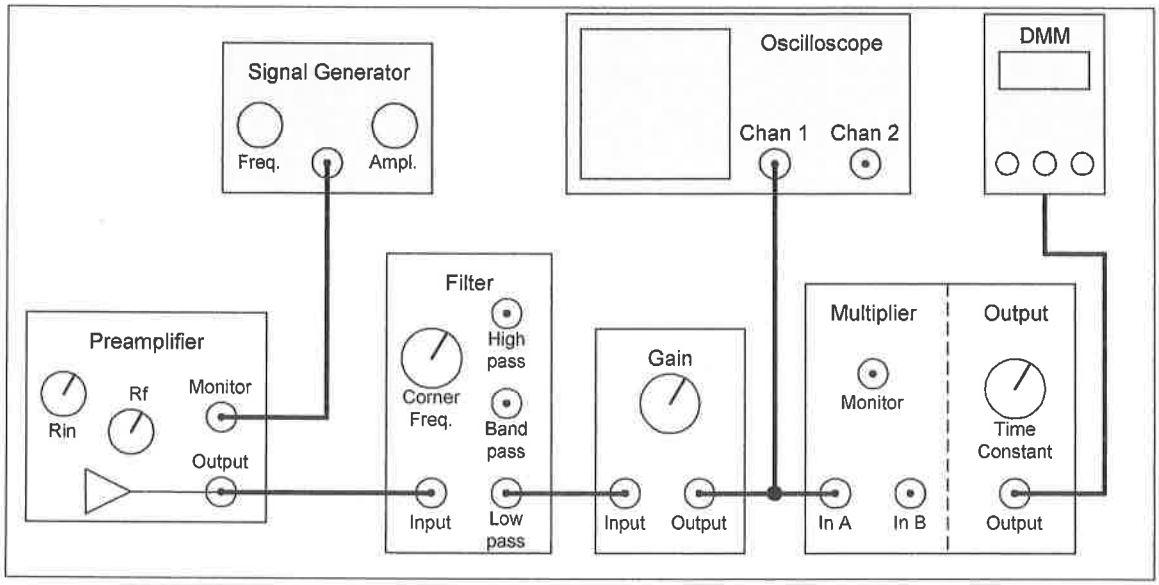


Fig. 8.4g Cabling diagram for injecting a 'dummy' interference signal into the LLE.

For these measurements we have returned to R_{in} = 10 kΩ, and no filtering ie. full bandwidth. The signal generator is set to deliver a 10-kHz, 200-mV peak-to-peak sine wave. The two qACFs are only different in the trigger levels set on the scope: Fig. 8.4h is at 11 Volts and Fig.8.4i is at 5.5 Volts.

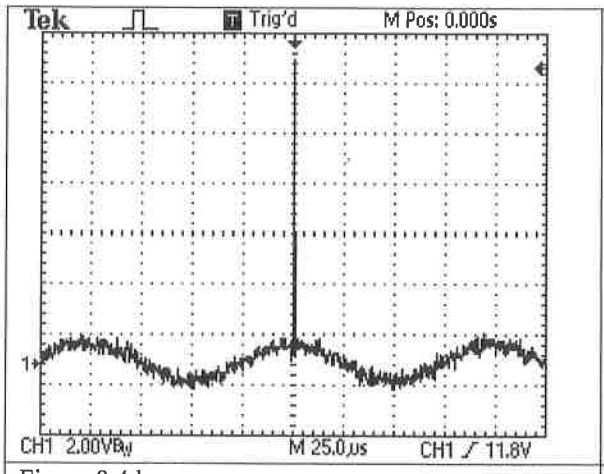


Figure 8.4 h

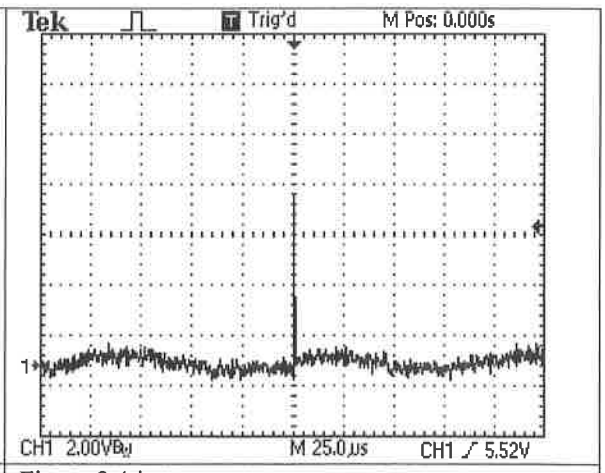


Figure 8.4 i

You can easily see that the interference signal is more easily observed in the qACF with the higher level trigger. You'll find that getting a 'good' qACF is a bit of a compromise, To be able to see underlying interference effects you'd like to trigger at as large a voltage as

possible, but the noise signal does not always reach to the top and so you have to wait a longer time for each new trace.

8.5. Magnetic-field interference

It is more difficult to couple a.c. magnetic fields into the unit in any controlled way. If you bring a soldering gun very near the pre-amplifier, you will likely see large 60-Hz (or 50-Hz) interference. Switching power supplies can also generate large time-varying magnetic field, though typically at much larger frequencies.

To observe time-varying magnetic-field effects, we created a large loop of wire (and exploited Faraday's Law of induction). The short piece of wire that runs between the R_{in} switch wiper and the non-inverting input of the op-amp has been replaced with a 45-cm (18-inch) wire, formed into a large loop. Then the switching supply for Noise Fundamentals has been placed face-down on top of the LLE box. What we observe is shown in Figure 8.5a.

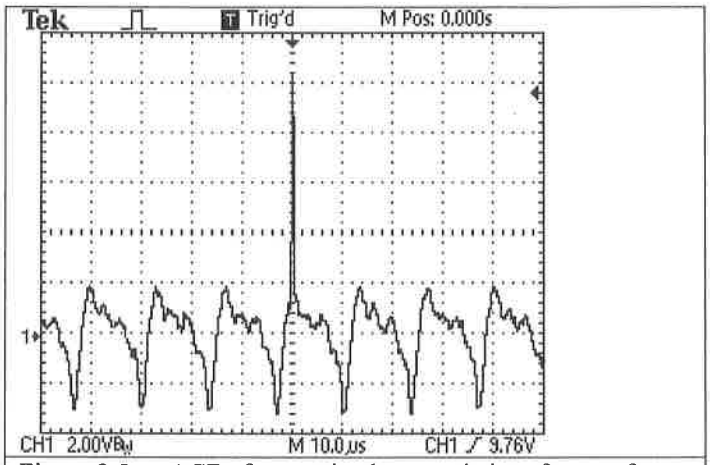


Figure 8.5a: qACF of a contrived magnetic interference from NF power supply.

Note that the periodicity of the qACF is about 14 µs, revealing the use of a switching frequency of order 70 kHz inside that power-supply module.

We did not observe changes in the signal as we varied R_{in}, but this may not be the case for pre-amp topologies used in measuring currents, as in the case of shot noise.

Bibliography

Historical works:

'Thermal Agitation of Electric Charge in Conductors', H. Nyquist, Phys. Rev. **32**, 110-113 (1928),

- the first theory of the noise that Johnson had measured, and reported on in the previous paper in the journal

'On the Theory of Electrical Fluctuations', R. Furth, Proc. Roy. Soc. A**192**, 593-615 (1948)

- early connections between Johnson and shot noise

Tutorial works:

'Undergraduate Experiment on Thermal and Shot Noise', James A. Earl, Am. J. Phys. **34**, 575-579 (1966)

'Random-Walk Model of Thermal Noise for Students in Elementary Physics', Richard W. Henry, Am. J. Phys. **41**, 1361-1363 (1973)

'An Experiment on Electronic Noise in the Freshman Laboratory', D. L. Livesey and D. L. McLoed, Am. J. Phys. **41**, 1364-1367 (1973)

'Undergraduate experiment on noise thermometry', P. Kittel et al., Am. J. Phys. **46**, 94-100 (1978)

'Comments on: Undergraduate experiment on noise thermometry', W. T. Vetterling and M. Andelman, Am. J. Phys. **47**, 382-384 (1979)

'A tutorial approach to the thermal noise in metals', A. J. Dekker et al., Am. J. Phys. **59**, 609-614 (1991)

'Shot-noise measurements of the electron charge: An undergraduate experiment', D. R. Spiegel and R. J. Helmer, Am. J. Phys. **63**, 554-560 (1995)

- the oldest of these paper using recognizably 'modern' electronics

'Two student experiments on electrical fluctuations', Yaakov Kraftmakher, Am. J. Phys. **63**, 932-935 (1995)

'Experiment on the physics of the PN junction', A. Sconza et al., Am. J. Phys. **62**, 66-70 (1994)

- good coverage of the 'transdiode' model for a p-n junction

Advanced works:

'Solid-state shot noise', Rolf Landauer, Phys. Rev. B**47**, 16427-16432 (1993)

- an early treatment of quantum effects is mesoscopic systems, and another recognition of the connections between Johnson and shot noise

'Quantum shot noise', Carlo Beenakker and Christian Schönenberger, Physics Today, pp. 37 ff., May 2003

- whose last line is "the field as a whole has progressed far enough to prove Landauer right: 'There *is* a signal in the noise.' "

'Linearity of 1/*f* Noise Mechanisms', Richard F. Voss, Phys. Rev. Letters **40**, 913-916 (1978)

- a look at 'conditional voltage distributions', our inspiration for the quasi-ACF

'Simple schemes for measuring autocorrelation functions', Z. Kam et al., Rev. Sci. Instr. **46**, 269-277 (1975)

- defines the conditions under which our quasi-ACF (= their 'comparator-trigger autocorrelator') agrees with the actual autocorrelation function

'Observation of Hot-Electron Shot Noise in a Metallic Resistor', Andrew H. Steinbach et al., Phys. Rev. Letters **76**, 3806-3809 (1996)

- shot noise in an all-metal mesoscopic device, unified with Johnson noise

'Primary Electronic Thermometry Using the Shot Noise of a Tunnel Junction', Lafe Spietz et al., Science **300**, 1929-1932 (2003)

- how primary thermometry using noise is approached in modern times

Appendices

Appendix A.1. Technical specifications

Low-level Electronics:

Pre-amplifier module:

First stage is user-configurable (see Appendix A.4.)

As shipped, first stage is

FET-input operational amplifier, non-inverting mode

gain (using R_f = 1. kΩ) is (1 + R_f/200. Ω) = 6.00

-3 dB bandwidth > 1.0 MHz

input impedance > 100 MΩ

Next stage are fixed-configuration

gain 100.

-3 dB bandwidth > 1.6 MHz

Temperature module

Current source

accuracy <1%, 10 nA to 1 mA settings

Transducer voltage buffer

gain 1.00 to <0.1% error, ±2 mV d.c. offset

Heater power supply

0 - 25 V (for floating loads), 330 mA current capability

Signal Attenuator

input impedance: variable, 100 Ω to 10 kΩ

output impedance: 10 Ω (for use driving Z_{in} = 1 kΩ stages)

or 10 Ω less 1% (for driving Z_{in} ≥ 10 kΩ stages)

-3 dB bandwidth > 10 MHz

Bipolar Power Supplies

output: (±) 10 mV to 11 Volts

noise: < 5 nV/√Hz, typically < 2 nV/√Hz

current capability: 250 mA maximum

High-level Electronics:

Filter sections
- state-variable 2-pole Butterworth design
- input impedance 10 kΩ

Main amplifier
- two stages, of gain x1 or x10 selectable
- one stage, gain variable from x10 to x100
- -3 dB bandwidth > 1.4 MHz
- slew rate ≈ 20 V/μs
- input impedance 1 kΩ

Multiplier
- scaling factor for output: $V_{out} = V_A V_B /(10.0\ V)$
- input impedances of A and B channels: 50 kΩ
- d.c. offset: under ±10 mV

Output stage
- hard-wired d.c. coupling to output of multiplier
- two successive (buffered) stages of 1-pole, low-pass filters

(back panel) Noise Calibrator
- output level ≈ 212 ± 2 mV, rms measure
- noise power is located >99% in $0 < f < 32$ kHz range
- spectral density uniform to ± 2 % in $0 < f < 32$ kHz range

The 'Break-out Box' for the Thermal Probe:

In normal operation, the cable from the Thermal Probe into the Temperature-Control module connects all the devices in the probe to circuits in the module. It does so in a shielded, all-grounded, low-noise environment. But there are times when you want ordinary access to connections in the probe -- for example, if you want to use an ohmmeter to diagnose the resistance (at ambient, or at LN_2, temperatures) of the resistors mounted in the probe. To do that, you can disconnect the Probe's cable from the Temperature-Control module, and connect it instead to the plastic breakout box. Now you won't have full shielding, but you *will* have test-probe access to wires:

GND	labels ground, ie. the shell and body of the probe, including the copper fin at its bottom
R_A, R_B, R_C	label the three source resistors' live ends (each has its other end grounded)
D_1, D_2	label the two wires from the temperature-monitoring transdiode (see Section 4.3); this transistor has its collector and base leads grounded, so the 'live wires' D_1 and D_2 connect to the emitter
H_1, H_2	label the two ends of the 75-Ω heater on the lower fin of the probe; this resistor has *neither* end grounded

Appendix A.2. The matter of a.c. or d.c. coupling

Real amplifiers are subject to 'd.c. offsets', such that a potential difference of zero at the input can still lead to a non-zero steady d.c. value at the output. Because the overall gain of the Noise Fundamentals system can be as high as (600) x (10^4) = 6 x 10^6, even an effective 1 μV offset at the pre-amplifier's input stage would lead to a full 6-Volt offset at the main amplifier's output. The amplified noise voltages would be lying atop that d.c. offset, and this would create unacceptable errors. So at many stages of the electronic signal chain, there is the option to use a.c. coupling between the stages.

Every a.c.-coupled connection (including that selection at the input of test instruments) is actually a high-pass filter, with a corner frequency typically located at 10 Hz or so. Thus the d.c. component of any signal is entirely blocked, and high-frequency a.c. signals are entirely passed, by the filter. But a.c. signals of frequencies below 10 Hz can be considerably attenuated, as well as phase-shifted, by the filter in question. This attenuation *matters* if the study of low-frequency noise is of interest to you.

What follows is a description of the a.c. vs. d.c. coupling options, stage by stage, in the Noise Fundamentals signal path.

The pre-amplifier's first stage is always d.c. coupled, as that's a necessity in shot-noise measurements. A MONITOR output allows a view of the d.c. output level of the first stage; any a.c. signal or noise is lying super-imposed on that d.c. level.

The gain-100 stage in the pre-amp can be a.c.- or d.c.-coupled to the first stage's output. As shipped, the coupling is a.c., with a high-pass corner at frequency 16 Hz. The change to d.c. coupling can be made via a moveable jumper on the printed-circuit board inside the pre-amp.

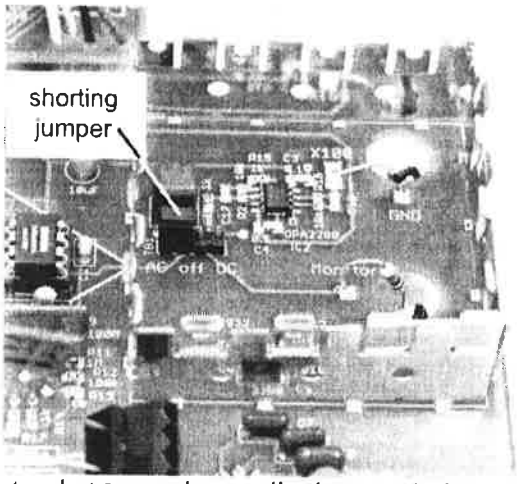


Fig. A.2a: How to select a.c. vs. d.c. coupling between the input stage, and the gain-of-100 stage, of the pre-amp.

In the high-level electronics, the two filter sections can be a.c.- or d.c.-coupled by front-panel switches. In the a.c.-coupled mode, there's a high-pass corner at frequency 1.6 Hz. In the main-amplifier section, the input can be a.c.- or d.c.-coupled by front-panel switch. In the a.c.-coupled mode, there's a high-pass corner at frequency 16 Hz.

The multiplier's inputs can both be grounded, or configured with a.c. or d.c. coupling. In the a.c.-coupled mode, there's a high-pass corner at frequency 1.0 Hz.

The output stage is internally connected, by d.c. coupling, to the multiplier's output, and its averaging action is optimized for accuracy all the way down to d.c.

The Noise Calibrator output is d.c. coupled, to preserve the flatness of its noise spectrum down near zero frequency. As a result, there may be a milliVolt-level d.c. offset in its average value.

Finally, a word about the consequences of a d.c. offset on an a.c. noise voltage going into the squarer. Recall that in typical operation, gains are chosen so that the signal reaching the squarer has an rms measure of about 3 Volts. Suppose that the actual signal entering the squarer is

$$V_A(t) = D + \Sigma_i A_i \cos(2\pi f_i t + \phi_i) ,$$

where here D represents the d.c. offset, and the sum is a Fourier representation of all the component frequencies in the signal (or noise). Since the squarer gives output

$$V_{sq}(t) = [V_A(t)]^2/(10\ \mathrm{V}),$$

the instantaneous output of the squarer contains lots of terms:

$$\begin{aligned} V_{sq}(t) = (10\ \mathrm{V})^{-1} \{\ & D^2 + \Sigma_i A_i^2 \cos^2(2\pi f_i t + \phi_i) \\ & + \Sigma_i D A_i \cos(2\pi f_i t + \phi_i) \\ & + \Sigma_{i,j} A_i A_j \cos(2\pi f_i t + \phi_i) \cos(2\pi f_j t + \phi_j)\ \}. \end{aligned}$$

Upon taking the time average, the terms in the last two lines average to zero, while the cosine-squared terms average to 1/2. So what you observe as the time-averaged output is

$$\langle V_{sq}(t)\rangle = (10\ \mathrm{V})^{-1} \{\ D^2 + \Sigma_i A_i^2\ (1/2) + 0 + 0\ \}.$$

The result is that the expected and intended output,

$$\langle V_{sq}(t)\rangle = \langle\ [\text{a.c. part of } V_A(t)]^2\ \rangle / (10\ \mathrm{V}) ,$$

is polluted by an error of

$$\delta\langle V_{sq}(t)\rangle = D^2 / (10\ \mathrm{V}) .$$

So if the output of the main amplifier has an offset of even 100 mV, lying underneath the typically 3-Volt (rms) noise signal, and if the squarer is used in its d.c.-coupled mode at the A-input, this error will be $(0.1\ \mathrm{V})^2/(10\ \mathrm{V}) = 0.001$ V, relative to an output due to the intended noise signal of $(3\ \mathrm{V})^2/(10\ \mathrm{V}) = 0.9$ V.

This offset of 1 part in 900, or 0.1%, will not be caught or corrected by switching the input of the squarer to the ground (GND) position, since in this position the squarer does not get to see the d.c. component that might be present in the main-amp's output.

The moral of this story: unless you have reason or need to study noise below about 20 Hz, use a.c. coupling throughout. If you do use d.c. coupling at various places in the system, monitor the signal (being sure to use an oscilloscope set for d.c. coupling at *its* input!) at every point in the signal chain, to ensure

- that nowhere is the d.c. level sufficient to saturate the next stage, and
- that the d.c. average level underlying the input to the squarer is under 100 mV or thereabouts.

Appendix A.3. Operational-amplifier circuits and noise

This Section describes how the amplifier noise in operational amplifiers can be modeled, and goes on to discuss the implications for experimentation.

We start with the open-loop model for a (bare) op-amp, with two inputs (called inverting and non-inverting, but labeled by - and + respectively):

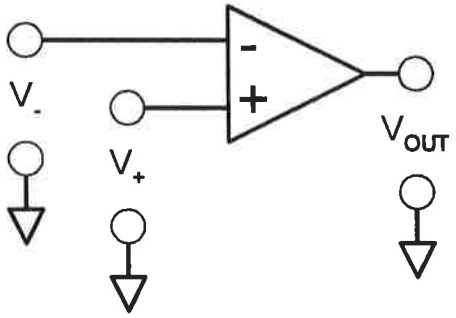


Fig. A.3a: An operational amplifier without feedback, showing labeling of inputs.

The noise-free model behavior is $V_{out} = A\ (V_+ - V_-)$, where A is the (typically huge, but frequency-dependent) open-loop gain. In this model, we neglect issues such as input offset and linearity limits.

Suppose that this device is used in the voltage-follower mode, which would ordinarily give $V_{out} = V_{in}$. Now we model the noise behavior of this amplifier. We imagine a referred-to-input voltage noise density V_n as an actual broadband white-noise emf, functionally in series with one of the amplifier inputs; and we also imagine that the very real d.c. bias current emerging from both inputs has atop it a white noise current source, with current noise density i_n. If the signal source is a resistor R, we have a circuit

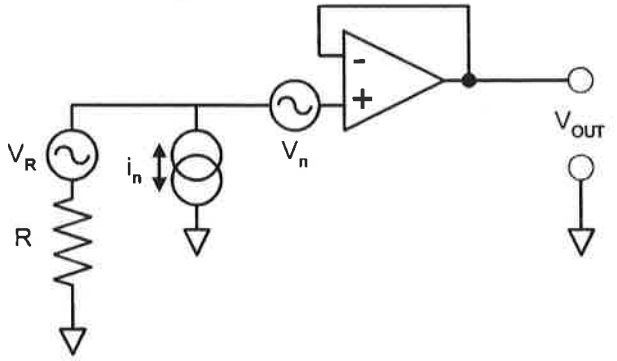


Fig. A.3b: An operational amplifier as voltage follower, showing model noise sources.

The noise behavior of this circuit includes three terms:

- amplifier voltage noise V_n is effectively applied to the non-inverting input, and (in this circuit) appears with gain (+)1 at the output.
- the Johnson noise of the resistor V_R , of noise power density $4\ k_B\ T\ R$, is also applied to the non-inverting input, and also appears with gain (+)1 at the output.

- current noise i_n , which has nowhere to go but through resistor R, where it causes a voltage drop across R which acts just like a voltage noise signal.

So the output has fluctuating voltages from three sources, presumed to be uncorrelated. As usual, the mean-square fluctuations of V_{out} simply add, to give

$$<V_{out}^2> = V_n^2\,\Delta f + 4\,k_B\,T\,R\,\Delta f + (i_n\,R)^2\,\Delta f.$$

Thus the noise density at the output can be written as

$$<V_{out}^2> / \Delta f = V_n^2 + 4\,k_B\,T\cdot R + i_n^2\cdot R^2.$$

This is a quadratic function of R, and it is well imagined in a log-log plot vs. R. For small source resistance R, the V_n^2 term dominates; for large R, the $i_n^2\,R^2$ term dominates. These two terms make equal contributions when $V_n^2 = i_n^2\,R^2$ or at $R = V_n/i_n$. But in addition to these R^0 and R^2 terms, there is an R^1 contribution from Johnson noise, which can *exceed* the other two (amplifier-noise) terms in an intermediate-R region. If you're trying to study Johnson noise, you'd like the $4\,k_B\,T\,R$ term to dominate both the V_n^2 and $i_n^2\,R^2$ terms, at least in the neighborhood of this R-value.

To be concrete, suppose that a generic (FET-input) op-amp is characterized by input voltage-noise density $V_n = 10$ nV/√Hz and input current noise density $i_n = 10$ fA/√Hz. The quotient $V_n/i_n = 10$ nV/10 fA = 10^{-8} V/10^{-14} A = 10^6 Ω defines the 'sweet spot' at the crossing of the R^0 and R^2 lines in the plot. So at source resistance $R = 1$ MΩ, the terms V_n^2 and $i_n^2\,R^2$ both contribute 10^{-16} V²/Hz to the noise power density. That defines the amplifier-noise baseline, against which the Johnson noise has to compete. For a 1 MΩ source resistor, that gives a density

$$4\,k_B\,T\,R \sim (1.6\times10^{-20}\text{ J})(10^6\ \Omega) \sim 1.6\times10^{-14}\text{ V}^2/\text{Hz} = 160\times10^{-16}\text{ V}^2/\text{Hz}.$$

Sure enough, at (and around) this source impedance, Johnson noise dominates, 160-fold in power, over both voltage noise and current noise in the amplifier.

The numbers for V_n and i_n picked above are typical for rather generic FET-input op-amps. But there are also op-amps whose front-end components are BJT-based, bipolar junction transistors. Such devices can offer smaller voltage noise (eg. 3 nV/√Hz), but they display much larger current noise (eg. 1 pA/√Hz). So a BJT-input op-amp would have its 'sweet spot' in the vicinity of a source resistance $R = V_n/i_n = 3$ nV/1 pA = 3×10^3 Ω = 3 kΩ, where both terms contribute $V_n^2 = (i_n\,R)^2 = 10^{-17}$ V²/Hz. Relative to that amplifier-noise baseline, a 3-kΩ source resistor generates Johnson noise of a spectral power density $(1.6\times10^{-20}\text{ J})(3\times10^3\ \Omega) = 4.8\times10^{-17}$ V²/Hz. Again, there's a zone in which Johnson noise dominates over both forms of amplifier noise, though not by so large a factor as in the FET-based example. Then again, the absolute amplifier noise density is lower.

These examples teach us some lessons. If we have a voltage source, it'll have some characteristic source impedance. If that impedance is low ($< 10^4$ Ω), then a BJT-input op-amp is better suited; if that impedance is high ($>10^5$ Ω), then an FET-input op-amp is better suited. If (as in Johnson-noise experimental investigations) the source impedance has to vary over a wide range, then we have to understand that amplifier voltage noise, or current noise, might dominate over the source's Johnson noise in some regions of R-space.

There's another lesson to be learned. It might be that source resistance near 1 MΩ is best suited to the noise characteristics of a FET-input op-amp, but that does *not* tell us what bandwidth we can achieve. Given even 10 pF of input capacitance, a 1-MΩ source impedance gives $2\pi\, R\, C \sim 10^{-4}$ s, and a 'corner frequency' of about 10^4 Hz, 10 kHz, beyond which point the noise will roll off badly. So the Johnson noise of a 1-MΩ resistor is indeed detectable, but an experimenter might be well advised to use only the 0-1 kHz bandwidth in which to detect it.

There are finer points, too. The values V_n and i_n provided by the manufacturer are typically quoted as densities near 1 kHz. In practice, V_n tends to rise at lower frequencies (excess or $1/f$ voltage noise near d.c.). In practice, i_n tends to rise, badly for FET-input op-amps, at higher frequencies. So in addition to the 'sweet spot' of source resistance, an amplifier can have a range or region in frequency space for which it offers its lowest-noise performance. The clever experimenter (using, for example, lock-in detection) will want to ensure that the signal being investigated has been arranged to lie near the optimal location on both the source-impedance and the signal-frequency axes.

Defining 'noise temperature' and 'noise figure' of an amplifier

The noise model above also allows us to define a figure-of-merit for an amplifier called the 'noise temperature' T_n. We imagine that we have a sensing resistor R, at a temperature T_R, and we seek to detect temperature changes in T_R via Johnson-noise measurements. For our model amplifier, the noise at the amplifier's output will be the same as if we'd used an ideal (noiseless) amplifier, whose input was driven by a noise density

$$S = \langle V^2 \rangle / \Delta f = V_n^2 + 4\, k_B\, T_R \cdot R + i_n^2 \cdot R^2 .$$

We've seen that the Johnson-noise term dominates the amplifier-noise terms most dramatically if we pick R's value at the 'sweet spot', choosing $R = V_n/i_n$. In this case we get

$$S = V_n^2 + 4\, k_B\, T_R \cdot R + i_n^2 \cdot (V_n/i_n)^2 = 2\, V_n^2 + 4\, k_B\, T_R\, R .$$

If we had the sense resistor at absolute zero ($T_R = 0$), we'd get the first term only; it's the net amplifier-noise contribution. Now we <u>define</u> the noise temperature of the amplifier, T_n, to be that resistor temperature at which the second (Johnson-noise) term would rise to be equal in value to the first term. So by this definition, raising the resistor temperature from 0 to T_n will raise S from $2V_n^2$ to double this value. This definition gives us

$$2\, V_n^2 \equiv 4\, k_B\, T_n\, R \text{, or } T_n = (2\, V_n^2) / [4\, k_B\, R] = (V_n^2) / [2\, k_B\, (V_n/i_n)] ,$$

so finally the amplifier noise temperature is given by

$$T_n = (V_n\, i_n) / (2\, k_B) .$$

To be concrete, we suppose that an FET-input op-amp will give us noise performance (at least in the vicinity of 1 kHz) characterized by $V_n \approx 8$ nV/√Hz and $i_n \approx 6$ fA/√Hz. Then we get

$$V_n\, i_n = (8 \times 10^{-9}\ \text{V}/\sqrt{\text{Hz}})\ (6 \times 10^{-15}\ \text{A}/\sqrt{\text{Hz}}) = 48 \times 10^{-24}\ \text{W/Hz} ,$$

and we get an amplifier noise temperature of

$$T_n = (V_n\, i_n) / (2\, k_B) = (48 \times 10^{-24}\ \text{W/Hz}) / (2 \cdot 1.38 \times 10^{-23}\ \text{J/K}) = 1.7\ \text{K}\ .$$

This is remarkable performance for an amplifier whose *physical* temperature is 300 K.

It does *not* follow that a ΔT of 1.7 K is the smallest change in temperature that this resistor/amplifier combination can detect. We've defined T_n such that (compared to a resistor at $T_R = 0$), a resistor at $T = T_n$ will *double* the value of measurable noise density S. Rather than such a 100% rise in noise density $\langle V^2 \rangle / \Delta f$, it is certainly possible to detect a 10% or even a 1% increase in S. The smallest temperature change you could detect by this system would ultimately depend on

a) how stable your system would be against systematic variations, and
b) how long you were willing to wait, in averaging down the statistical fluctuations in the noise you observe.

One example of the state-of-the-art in such ΔT measurements comes from the microwave radiometry of the cosmic (blackbody) background radiation by various satellite missions. Those measurements of the 2.7-K blackbody radiation are conducted with microwave amplifiers whose noise temperatures are of order 60 K, yet they have by now resolved *micro*Kelvin variations in the blackbody temperature (variations with respect to angle, not as a function of time). But they required about a *year* of averaging time to achieve this.

Noise temperature is the preferred measure of amplifier noise performance in radio and microwave regions of the spectrum, because such amplifiers are optimized for source impedance of a fixed value (typically 50 Ω). With such an R-value matching the quotient V_n/i_n, and a noise temperature given via the product of V_n and i_n, it's clear that an assumed R-value, and a quoted noise temperature T_n, fully characterize the noise performance of the amplifier. In the world of operational amplifiers, there's no need to stick to a single source resistance, so rather than specify an amplifier by optimum source resistance and a noise temperature, the two parameters V_n and i_n are quoted instead.

In regimes where impedances *are* assumed, and noise temperatures alone therefore suffice to characterize amplifiers' noise, another figure of merit often quoted is the 'noise figure', defied by a temperature ratio, and transformed to decibel (dB) units via

$$NF = 10\ \log_{10}\ (1 + T_n/290\ \text{K})\ .$$

Our op-amp example above, with $T_n = 1.7$ K, gives $NF = 0.025$ dB, which is (very roughly speaking) a measure of how much worse is the signal-to-noise ratio at the output of such an amplifier, compared to the signal-to-noise ratio at the input.

Appendix A.4. Front-end amplifier choices and consequences

The pre-amp module in the low-level electronics part of TeachSpin's Noise Fundamentals has a 'front end', or first stage, which is user-configurable. In particular, the operational-amplifier chip for the first stage can be changed, and so can the 'topology' or choice of circuit. Here's a summary of what can be changed, and why you'd want to change it.

The choice of chip is basically between an FET- or BJT-input op-amp chip. The unit is shipped with an FET-input op-amp in place, with performance of the sort described in Appendix A.3. The voltage-noise level of the input stage is not as low as it could be made, but the range of source impedances for which this choice is adapted have led us to choose it as the default condition of the pre-amp.

If you want the lowest in amplifier voltage-noise levels, and are willing to work in the range of source impedances under about 10 or 100 kΩ, then it can help to use a BJT-input op-amp chip. The substitution is easy to make, as we've provided a pin-compatible integrated circuit in the spare-parts bin. You'll need to know how to use a 'chip puller' to removed the as-shipped input-stage chip from its socket, and you'll need to be able to recognize the pin-1 end of the 8-pin dual-inline package of the new chip to orient it properly in the socket. (Any op-amp with a '741 pinout' and tolerating ±12-V supplies may be used.)

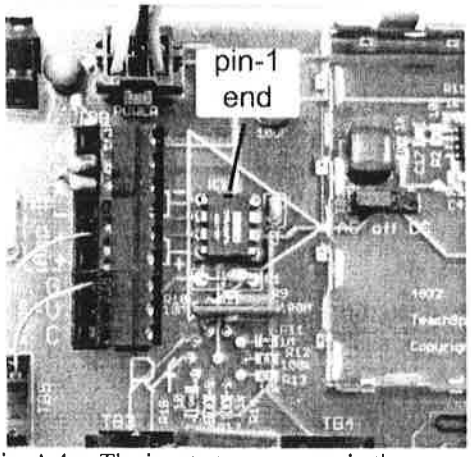


Fig. A.4a: The input-stage op-amp in the pre-amp, with the pin-1 end of chip (and socket) indicated by semi-circular 'dimples'.

You can store the op-amp chip that's <u>not</u> in use in the conductive foam in the spare-parts box.

Whether you use one chip or the other, there remains the choice of circuit topology for the first stage of the pre-amp. The unit is shipped with the configuration of a non-inverting amplifier, whose chief benefit is its very high input impedance:

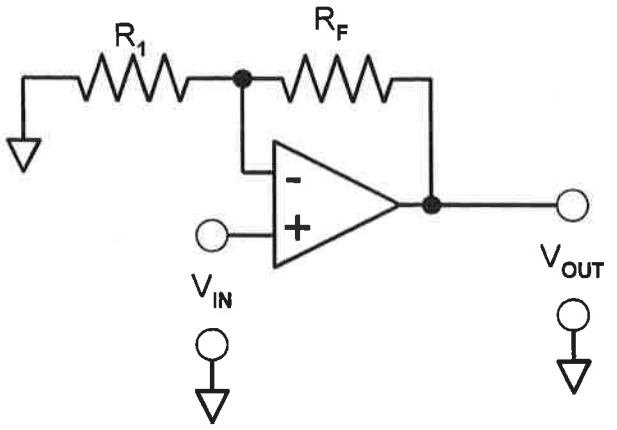


Fig. A.4b: A non-inverting amplifier topology.

This amplifier has d.c. gain $g = 1 + R_f/R_1$, which takes on the value 1 + (1.00 kΩ/200. Ω) = 6.00 in the as-shipped condition. Note that R_f is selected via the front-panel selector switch, while R_1 is a resistor attached to the terminal blocks.

A second topology retains R_f, omits R_1, and acts as a current-to-voltage converter:

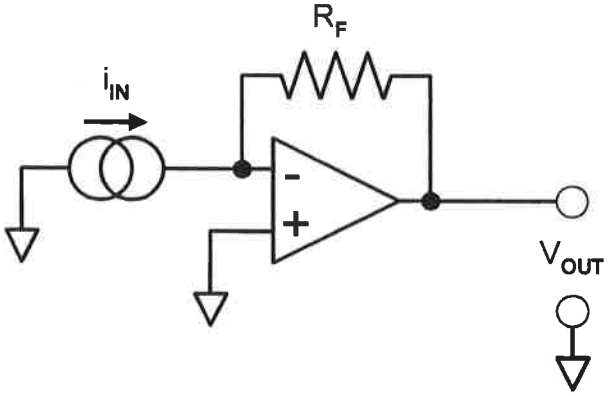


Fig. A.4c: A current-to-voltage converter topology.

A third topology is another voltage amplifier, this one inverting in character:

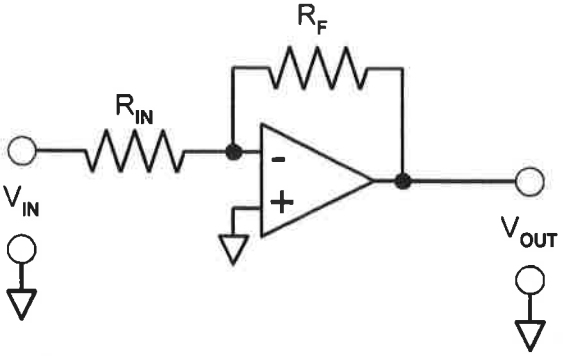


Fig. A.4d: An inverting-amplifier topology.

Here the gain is $g = -R_f/R_{in}$, and the main disadvantage is the relatively low input impedance of the circuit.

Beyond these textbook results, it is now necessary to consider the noise performance of these circuits. We take up this topic at two levels of treatment: first, the low-frequency behavior, and second, the behavior at higher frequencies (where capacitances, and op-amp bandwidth limits, start to matter).

The simpler treatment of the non-inverting voltage amplifier of Fig. A.4b is to model the op-amp voltage noise V_n as a series emf in (one of) the amplifier inputs, and to add Johnson noise as a model emf in series with each resistor. (This model omits the op-amp current noise.)

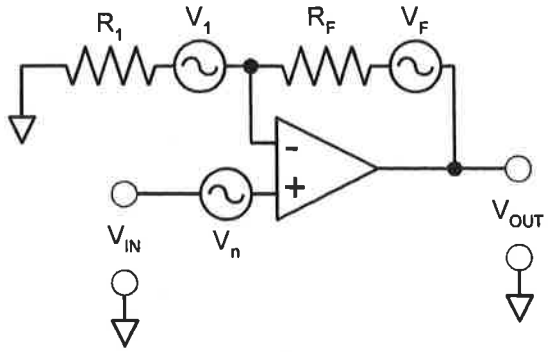


Fig. A.4e: A noise model for the non-inverting amplifier.

The result is to give

$$V_{out} = g\,(V_{in} + V_n) - V_f + V_1\,(R_f/R_1)\,,$$

where g is still the d.c. gain given by $1 + R_f/R_1$. Relative to the expected output $g\,V_{in}$, the actual output displays noise of mean-square size

$$\langle \delta V_{out}^2 \rangle = g^2 \langle V_n^2 \rangle + 4\,k_B\,T\,R_f\,\Delta f + (R_f/R_1)^2\,4\,k_B\,T\,R_1\,\Delta f\,.$$

If the amplifier noise is modeled by a voltage noise density D, or noise power density $S = D^2$, this gives output noise density

$$\langle \delta V_{out}^2 \rangle / \Delta f = g^2\,S + g \cdot 4\,k_B\,T\,R_f\,.$$

Typical values applicable to experiments in Section 1 are $g = 6$ and $R_f = 1$ kΩ; the choice of an FET-input op-amp might give $D = 8$ nV/√Hz or $S = 64 \times 10^{-18}$ V²/Hz. Then we get

$$\langle \delta V_{out}^2 \rangle / \Delta f = 6^2 \cdot (64 \times 10^{-18}\ V^2/Hz) + 6 \cdot (1.63 \times 10^{-20}\ J)(10^3\ \Omega)$$

$$= (2304 + 98) \times 10^{-18}\ V^2/Hz\,,$$

which shows that the circuit's output noise density is dominated by the op-amp's own voltage noise. The Johnson noise of the two resistors adds a small, and constant, correction -- that's why the resistors were chosen to have small values. The whole noise budget can be treated as 'amplifier noise', and subtracted by the methods of Section 1.3.

A similarly simple treatment of the i-to-V converter of Fig. A.4c is to consider the circuit with amplifier voltage noise, and resistor Johnson noise, added.

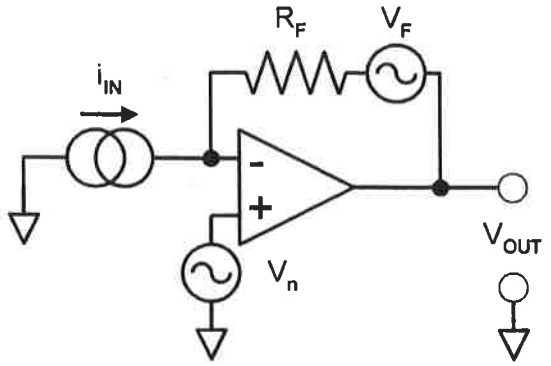


Fig. A.4f: A noise model for the current-to-voltage converter.

This model gives

$$V_{out} = - i_{in} R_f + V_n + V_f ,$$

and it gives output noise, relative to the expected d.c. value, of

$$<\delta V_{out}^2>/\Delta f = <V_n^2>/\Delta f + <\varepsilon_f^2>/\Delta f = S + 4 k_B T R_f ,$$

still ignoring the op-amp current noise. For shot-noise measurements typical of Section 3, the feedback resistor is neither fixed nor small, so we'll consider the exemplary case of $R_f = 10^7\ \Omega$. Then at room temperature we find $4 k_B T R_f = (1.63 \times 10^{-20}\ J)(10^7\ \Omega) = 1.63 \times 10^{-13}\ V^2/Hz = 0.163 \times 10^{-12}\ V^2/Hz$, which dominates, by far, the op-amp voltage noise contribution of $S = (8\ nV/\sqrt{Hz})^2 = 64 \times 10^{-18}\ V^2/Hz = 0.000\ 064 \times 10^{-12}\ V^2/Hz$.

Given so large a Johnson-noise contribution from the feedback resistor, it's worth comparing its effect with the expected shot noise of the input current. A feedback resistor of $10^7\ \Omega$ is an appropriate choice for an input current in the vicinity of $i_{dc} \approx 0.5\ \mu A$, and it will give a d.c. output of $(-)i_{dc} R_f = (-)5$ V. Such a d.c. current allows a computation of expected shot-noise current noise $(2 e i_{dc} \Delta f)^{1/2}$, or a current noise density $\sqrt{(2 e i_{dc})} = 4 \times 10^{-13}\ A/\sqrt{Hz}$. The i-to-V converter maps this to an output voltage noise density larger by the factor R_f, giving $4 \times 10^{-6}\ V/\sqrt{Hz}$, or a noise power density of $16 \times 10^{-12}\ V^2/Hz$. *This exceeds, by 100-fold, the Johnson-noise contribution of the resistor, which in turn exceeds, by far, the voltage-noise contribution of the op-amp.*

That completes a 'first level' treatment of expected noise levels; at this level, resistors' Johnson noise and amplifier voltage noise have been included, but capacitance of devices, and bandwidth limits of op-amps, have not been included. We now take up some examples where these effects are considered.

We return first to the non-inverting topology of Fig. A.4b, but now include the effect of source capacitance C_{in}, in parallel with a source of impedance R_{in}.

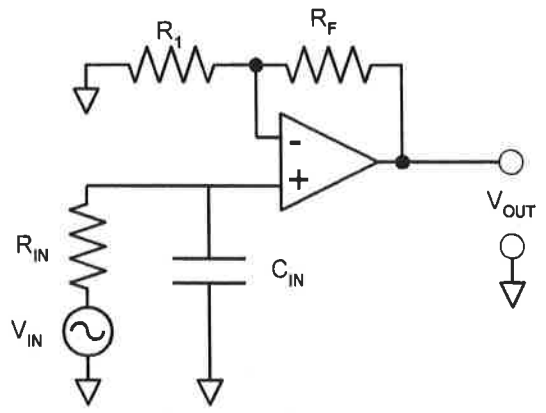


Fig. A.4g: The effects of input capacitance in the non-inverting amplifier.

We're still ignoring any capacitance that might be in parallel with R_f or R_1, because the input capacitance C_{in} is typically the effect that becomes important first. In this circuit, any Johnson (or other) emf in series with R_{in} is RC-filtered by the R_{in} C_{in} combination, which puts a bandwidth 'corner' at $f_c = (2\pi\ R_{in}\ C_{in})^{-1}$. The result is that the output noise spectrum drops *below* the white-noise limit at and above f_c, with consequences that are explored quantitatively in Appendix A.8. The input capacitance does *not* reduce the equivalent bandwidth of the op-amp voltage noise or the Johnson noise of the resistors R_1 and R_f.

A second example of the effects of capacitance is in the i-to-V converter topology of Fig. A.4c, now shown with an actual current source, having parallel capacitance C_{in}. Also shown is a user-selectable capacitance C_f in parallel with the feedback resistor R_f.

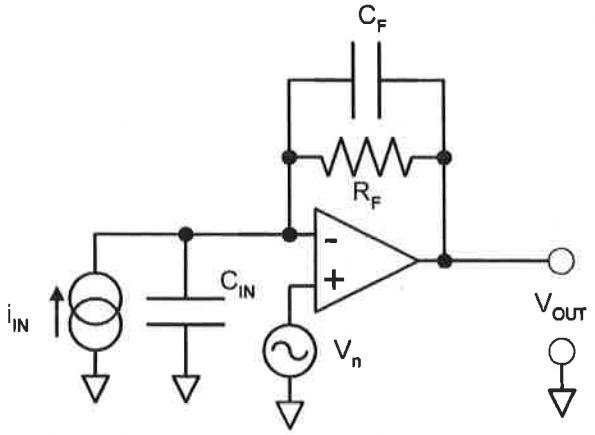


Fig. A.4h: The effects of capacitance in the current-to-voltage converter.

In the applications of Section 3, the current source may be a photodiode, and R_f is chosen to lie in the range $(10^3 - 10^7)\ \Omega$, depending on the light level. Temporarily ignoring the presence of C_f, the novelty in this circuit is the frequency-dependent gain applicable to amplifier voltage noise. At low frequencies, the op-amp acts like a voltage follower for noise signal V_n, and thus gives a 'noise gain' of 1. But starting at corner frequency $f_c \approx (2\pi\ R_f\ C_{in})^{-1}$, this noise gain starts to *rise* with frequency. This gives an excess gain for

amplifier noise which peaks near $f_p = (f_c f_m)^{1/2}$, where f_m is the open-loop unity-gain frequency of the amplifier.

To be concrete about this, we estimate $C_{in} \approx 20$ pF as the combined capacitance C_{in} of the input circuitry and the reverse-biased photodiode. If we're using an intermediate resistance value $R_f \approx 10^5\ \Omega$, then $2\pi\ R_f\ C_{in} \approx 10^{-5}$ s, and so $f_c \approx 0.1$ MHz. Given an op-amp with 'gain-bandwidth product' of $f_m \approx 10$ MHz, this tells us that there is excess voltage noise in the 0.1 ~ 10 MHz range, with about a 10-fold excess in the vicinity of $f_p \approx 1$ MHz.

This sort of 'noise peaking' is certainly visible, by using a 'scope-based FFT to look at an amplified version of V_{out}, obtained even with a photodiode in the dark.

Once such noise peaking is detected, the noise peak near f_p is readily reduced, by user selection of the feedback capacitor C_f. An approximate treatment suggests a value of C_f obeying

$$\frac{f_m}{f_p} \approx \frac{C_f + C_{in}}{C_f} \quad ,$$

which in this example gives the numbers

$$10\ \text{MHz} / 1\ \text{MHz} \approx 1 + C_{in}/C_f, \qquad \text{or} \qquad C_{in}/C_f \approx 9, \text{ or } C_f \approx C_{in}/9 \approx 2\ \text{pF}.$$

In practice, a slightly larger value of C_f might be used; the use of a generic C_f will give an i-to-V converter whose response drops below its low-frequency limiting value of $-\Delta V_{out}/\Delta i_{in}$ of R_f, starting at a corner frequency near $(2\pi\ R_f\ C_f)^{-1}$.

Appendix A.5. Grounding, shielding and screening, and interference

There are many sources of noise, some of them fundamental and some of them just a nuisance. Happily, electronic noise arises in obedience to Maxwell's Equations, and this provides some guidance on ways to diagnose and suppress undesired forms of noise.

1) Grounding

There is the complicated matter of grounding, ie. in establishing the point, assigned to have $V \equiv 0$, relative to which all potentials are measured.

In the Noise Fundamentals apparatus, local ground is exhibited by the front and back panels of the HLE, by the aluminum front panel of the LLE and the module panels installed onto it, and by the metal of the thermal probe (if that is being used). The HLE and LLE grounds are connected by both the power supply cable, and the shield of the coaxial cable, that is connecting them. Maintaining good electrical contact among all these objects is important for good electrostatic screening.

All of these grounds are connected to the third-wire ground of the a.c. power line through a 10-Ω resistor, which is in place to limit ground-loop currents that might otherwise be induced in low-resistance closed paths through which there exists a time-varying magnetic flux. The separate ground lines in the power, and signal, cables interconnecting the HLE and LLE potentially form such a ground loop. To keep such effects minimal, it is useful to keep these two cables close together, perhaps by loosely twisting them around each other.

The common $V = 0$ level is thus present at the shells of all the front-panel BNC jacks on both low- and high-level electronics. That $V = 0$ level is not changed by attaching any isolated device, such as a battery-powered multimeter, to any such BNC jack. But issues *can* arise with the use of any line-powered instrument such as an oscilloscope, whose input connectors typically have their *own* idea about what ground is, established by their own connection to the line supply.

Hence this advice: make the most sensitive noise measurements with multimeters connected to, but 'scopes disconnected from, the apparatus. Try a measurement sometime with, vs. without, a 'scope connection, while monitoring the mean-square output with a DMM, to see if ground issues are affecting your noise measurements. If they are, and if you need the 'scope connection, you can minimize this effect by plugging the power cords from Noise Fundamentals, and from your 'scope, into the same outlet fixture (and **NOT** using two outlets whose 'common ground' is established somewhere unknown or far away).

2a) Shielding and screening

Though these terms are not always distinguished, for this discussion we'll use shielding and screening as the names for methods of blocking the effects of external magnetic and electric fields, respectively.

Here's a way to see the effects of imperfect screening. Set up a Johnson-noise or equivalent exercise in which low-level signals are produced in the pre-amp module and sent to the high-level electronics. To favor the detection of rather low-frequency noise, use both filter sections as 1-kHz low-pass filters, and pass along the signal to the main amp. Use a gain of about 400 there, and look at the main-amp output with a 'scope. Set the 'scope to 10 ms/div, and arrange for it to trigger synchronously with your local a.c. power line. You should see 'desired noise' on the 'scope; pick a vertical-axis sensitivity which keeps the noise within range.

Now change the 'scope to the averaging mode. The noise level should drop, by about $\sqrt{N}$, where N is the number of averages you're taking. What you're looking for is residual structure, signals of fixed phase with respect to a period of 16.7 or 20.0 ms (depending on whether your power is supplied at 60 or at 50 Hz). If you don't see such interference, that's good news. But to generate some (so that you can learn to recognize it), power up a soldering gun or other transformer-containing appliance, and now hold that appliance somewhere near the pre-amp. You should now be able to view some 60- or 50-Hz interference on the 'scope. Try re-orienting and re-positioning that transformer, and testing its effects near the high-level electronics too.

Such signals as you're now seeing are due to Faraday's-Law emfs, due to rates of change of magnetic flux. The magnetic fields leaking out of the transformer core are the source, and their fields are coupling to circuit loops inside the pre-amp. Once you've seen that this can happen, you'll learn

- to imagine all such sources, and move them away if possible, especially from your pre-amp. (Remember that any line-powered instrument can be a source, too.) Take advantage of the r^{-3} drop-off of magnetic fields.
- why the pre-amp box is made of thick aluminum, and of steel. Good conductors can shield against a.c. magnetic fields by virtue of the a.c. currents induced in the shield material; good ferromagnetic materials can shield well against d.c. magnetic fields (and less well against a.c. fields.) Shielding against low-frequency a.c. magnetic fields is the hardest.
- that the LLE has the greatest sensitivity to a.c. magnetic fields, and that the size of this sensitivity will be greater if your pre-amp circuits present a larger-area loop to the magnetic field. This applies particularly to the use of the temperature probe, with its wire connection between the Temperature to the Pre-amp modules.

2b) **Screening proper**

Now that you've tested for B-field effects, almost always related to line frequencies, you're ready to think about E-field effects. If you've never observed these, it's time for you to do so. You need only a 'scope and a paper clip. Unbend the paper clip to form a single stiff wire, and poke one end of the wire into the center conductor of the BNC input of your 'scope. You've built a sort of antenna, which is resistively coupled into the 'scope, but capacitively coupled to the outside world.

Set your 'scope for 1-MΩ input impedance and for automatic triggering, and look for a signal, without averaging. The signal will grow, perhaps to large (> 20-mV) size, if you

touch the paper clip. (While you're touching that lead, your body is one capacitor electrode -- where's the other?) You will likely see, amid all the other interference, some sinusoidal signals somewhere in the 25-75 kHz range. Choose the appropriate sweep speed on the 'scope, and try to trigger on these signals. The typical source of such signals are fluorescent light fixtures, computer monitors, liquid-crystal displays (including your 'scope's own display!), and lots of other devices using internal sweeps, scans, and oscillators. If your 'scope has enough bandwidth, you might also see some signals near 100 MHz, due to local broadcasts of FM signals. This might also teach you to use the reduced-bandwidth option your 'scope may offer, so long as you're doing <1-MHz noise investigations.

All of these signals are capacitively coupled, so they depend on *E*-field lines terminating on your antenna and inducing charges there. Such effects are easily 'screened' by interposing a grounded conductor to serve as an alternative, and harmless, place for those *E*-field lines to end. That's why coaxial cables have a grounded outer conductor, and why the Noise Fundamentals pre-amp and probe have grounded metal exteriors. That's why the incomplete coverage of the braided outer conductor of most coaxial cables makes them somewhat 'leaky' – signals can leak out, and interference can leak in.

To see that such things matter, here's a way to see what can 'leak through a screen'. Remove your paper clip from your 'scope, and devote the 'scope again to looking at the noise signal emerging from the main amp. (Use a configuration like that of Section 1.1.) Now here's a way partially to *defeat* the screening of your pre-amp. Remove one (of the four) screws which hold the pre-amp module into the low-level electronics. Now build an 'antenna' from a few inches of plastic-insulated wire. Strip away a cm of insulation from one end of that wire, and hold that bare-metal end. Try lowering the insulated end of the wire, down through the now-open screw hole, so its bottom end is protruding into the pre-amp's internal spaces. You should see the effect of the failure of screening, as fluctuating potentials on your fingers are now capacitively coupled into the pre-amp's circuits. Once you've seen this effect, you'll understand that screening needs to be complete to be effective. You'll understand the construction of the probe better, too.

3) **Interference**

Grounding, shielding, and screening are all defenses against interference, ie. the injection, into your desired noise signal path, of other kinds of signals generated elsewhere. Once you've seen ways to detect and defeat such undesired noise, you should get suspicious of what sources can generate it.

If your 'scope is at full bandwidth, you might see effects due to local FM stations, and then worry about nearer sources of radio-frequency noise. If these are weak enough, their high frequencies puts them out-of-band for your 0-1 MHz noise investigations. But if they're strong enough, then non-linearities can make their effects show up even in the <100kHz band. So if you can identify and turn off such sources, do so.

We've mentioned the interference in the 10 - 100 kHz band that's generated by sources like the solid-state ballasts in modern fluorescent lights. Many other devices containing switching power supplies can also generate interference in this vicinity. Typically this

interference lies at, or near, one single frequency. The 'dimmers' sometimes used between the line supply and incandescent lights are another source of interference, typically at harmonics of the line frequency, but extending to very high frequencies. Your high-gain noise electronics and your 'scope are the tools for detecting such effects, but your environment is unique, and it'll take some creativity and imagination to identify all the forms of interference which might be troubling you.

There's one more source of interference that you can identify, test, and avoid: it's called 'microphonics', and it shows up as signals generated by mechanical motions of conductors due to vibrations. These motions can either cause a rate-of-change of magnetic flux, or a variation in position, changing a capacitance which maps a charge into a changing potential. Either way, you can detect microphonics by watching a 'scope view of noise while tapping suspected parts of an apparatus. When performing shot-noise experiments with the light bulb, you'll see this effect if you tap the pre-amp near that black plastic block containing the bulb. When using the temperature probe, you may also see microphonics during episodes of boiling of liquid nitrogen, especially when it fills the probe. Clearly this effect puts a premium on building rigid circuits, and then not bumping them.

Appendix A.6. Trouble-shooting

You might be familiar with trouble-shooting, which ought to be a semi-systematic method of following a signal through stages of electronics, trying to identify the place where something goes wrong. Trouble-shooting a *noise* apparatus is harder -- not only because there's no 'signal' to track, but also because it can be hard to distinguish the noise you care about from extraneous noise. So here are some suggestions to follow in case things seem not to be working as they should.

a) Connections

The first step in trouble-shooting is to review your interconnections. Have you plugged all your tools into the a.c. line? Do you have the right cable going to a 'scope input? Are you using the correct output from the Low-level Electronics (LLE)? Have you included all the required cables interconnecting sections of the High-Level Electronics (HLE)?

Next, check the connections you have made inside the LLE, particularly in configuring the pre-amp's first stage to your measurement needs. Pull gently on interconnecting wires and component leads to ensure they are held firmly in their terminal blocks. There are also some ground connections that you must make in configuring the pre-amp. Finally, have someone else look over your connections, to see if they match the wiring diagram, and the schematic diagram, you are trying to emulate.

b) Power

Start with your a.c. line cord, and look for a green LED on the cord-transformer itself. Then look for the green LED on the front left of the HLE, and another green LED on the front panel of the LLE. All should be lit when your cord is connected. If the LLE is showing an un-lit LED, suspect that you might have forgotten to turn back the internal toggle switch that's accessible when you open up the LLE and 'flip' the front panel. Check the power supplies for the operational-amplifiers inside the pre-amp, by measuring (relative to ground) the potentials at the two far ends of the terminal block at the input of the first-stage op-amp. Those points should show potentials of (±) 13-14 Volts. As a further check, use a voltmeter to check that the auxiliary ±11-V power supplies in the LLE are working -- there are monitor points on the front panel for this purpose.

c) Signal integrity

The modules in the HLE can be tested independently, by injecting signals from a waveform generator at an input, and looking for outputs with a 'scope. If you can spot a signal at the input of a module, and nothing emerges from that module, you have identified a problem with that module.

Recall that the filter sections give gain near 1 when you're 'in band', but can give gains $<<1$ when you're far outside their pass-bands. Recall that the main amplifier can have its gain set in the range 10 - 10,000. At the lower end of this range, it's easy to use a 0.5-V

amplitude sine wave in, to get a 5-V amplitude sine wave out. At the high end, the gain is so large that any input amplitude > 1 mV will saturate the output.

d) **Saturation**

In normal operation, you have access to the noise signal at many points in the amplification chain: early in the pre-amp, and again at its output; and then again at every interstage connection in the HLE. You can use a 'scope, with its input set to d.c. coupling, to look for three kinds of pathologies:

First-stage effects: you always have d.c.-coupled access to the output of the first-stage op-amp in the pre-amp. This output is connected (through a 1-kΩ resistor) to the MONITOR BNC jack on the Pre-amp panel. If this output shows a potential near (±)12 Volts, that suggests that the first stage has 'railed out', most likely because of an incorrect wiring of this first stage. Under these conditions, the first stage can*not* be faithfully transmitting noise to subsequent stages.

d.c. offsets: If you use d.c. coupling between stages, then a gain stage can turn an input of (1 V of offset, + noise) into an output of (say) 10 x (1 V of offset, + noise) = 10 V of offset, + 10 x noise. Much more of this, and the d.c. offset will drive the noise into the 'rails', the upper and lower voltage limits, of about ±12 V. Once a d.c. offset has caused a signal to 'rail out', any noise atop the d.c. offset is wiped out.

The third thing to look for is any evidence of noise that's gotten too large. Supposing that the use of a.c. coupling between stages (see Appendix A.2) has dealt with d.c. offsets, yet still there remains the problem of saturation. If a noise input falls in the ±2-V range, a gain-of-10 amplifier ought to produce output noise in the ±20-V range. But in this apparatus it *won't*: the largest positive and negative excursions will be 'clipped' at the levels near ±12 Volts imposed by the range of linearity of the amplifiers. That clipping not only removes part of the energy which should be in the noise, it also generates distortion, which puts energy at unpredictable locations in frequency space.

e) **Excess noise**

There will be times when you suspect that you're getting more noise than you should be. Here are some possible causes:

First, you should monitor the squarer's time-averaged output with a digital multimeter (DMM) as a measure of the noise. Then you should disconnect any and all ground-reference test instruments (like oscilloscopes) from interacting with the apparatus. If the DMM reading changes, then you can suspect some interference (typically, a ground loop) is contributing to what you're seeing. See Appendix A.5 for details.

If you have a 'scope attached, and have shown by this test that it's *not* contributing to the measured noise, you can now use that very 'scope to look for interference in what should

be a pristine noise signal. Appendix A.5 teaches you the use of a.c. line triggering, plus averaging, to see the effects, if any, of all 60- (or 50-)Hz periodic effects.

You may have seen, in Appendix A.5, that a dominant form of electrostatically-coupled interference in your lab is at 25 or 48 kHz, or some other medium-high frequency, generated by fluorescent light fixtures, etc. Here's a 'scope-based way to search for contamination of your noise signal by interference from such a source:

- use the 'paper clip method' of Appendix A.5 to get a view of that interference on ch. 1 of your 'scope, and *trigger* on that interference; also pick a time base giving several cycles' view of the interference. Now put your noise signal into ch. 2 of the 'scope. Use signal averaging. If the noise signal is contaminated by this sort of interference, upon this kind of triggered signal averaging, ch. 2's signal will average *not* to zero, but to a non-zero trace revealing the interference which is contaminating the noise.
- or, use the 'scope-based FFT on the paper-clip pick-up signal to establish where (in frequency space) the interference is located; now switch to an FFT of the noise signal, and look for a peak, a location of excess noise, atop the expected white-noise background.

Turning off and on all the room lights, while monitoring with a multimeter the squarer's averaged output $<V^2>$, can sometimes show that electrostatic interference is contributing to the total noise detected. If you do see such an effect, suspect that you have a problem in some part of the LLE with imperfect screening against electrostatic effects.

f) Suppressing interference

If you find interference by one of these tests, you might wonder how it's getting into your system. Start by suspecting an entry point in the LLE. Check that all four thumbscrews (at top and bottom of the main LLE panel) are snugged down (finger-tight). Then check that all eight flat-head screws holding the Modules' panels to the frame are in place, and tightened. If you're not using the Thermal Probe, be sure that you 'cap off' the connector where its cable would enter. Use BNC 'shielding caps' as well at the two most crucial locations: the pre-amp's Monitor output, and the Series Resistor's Monitor position. This should deal with the potential paths for capacitively-coupled interference to get into the pre-amplifier.

Appendix A.7. Test and repair of the d.c. power supplies

The low-level electronics box, in addition to 'hosting' the pre-amp and temperature-control modules, provides you with some utilities. Among these are the two 11-Volt power supplies. (Note that one range switch, and one variable knob, control one supply with 0 to +11-Volt output, and also control another, with 0 to -11-Volt output.) These supplies have a current capability of 250 mA, yet they have been crafted to have noise levels under 5 nV/√Hz, all the way from d.c. to >1 MHz. Separate from these two utility supplies, but of very similar design, are the power supplies which run all the operational amplifiers in the pre-amp and temperature-control modules in the low-level electronics.

Building voltage supplies as 'quiet' as this takes careful regulation, which rejects voltage variation at all sorts of frequencies, and therefore has to react in a time << 1 μs. The regulators, in turn, are protected against damage in case the power supply outputs are short-circuited. But the protection cannot be instantaneous, so there are implications:

1) Please wire items to these ±11-V supplies only with the power switched OFF inside the Low-level Electronics. When you've 'flipped' the low-level panel to work on its inside, there's a toggle switch visible on the power-entry box, with a *red* LED to remind you when the power is on.

2) **Try to try not to short-circuit these power supplies !**

3) When you turn back on the power inside the LLE box, you can check that the red power-on LED comes back on, and you can check that two fault-mode red LEDs on the power-regulating printed-circuit board *don't* light up.

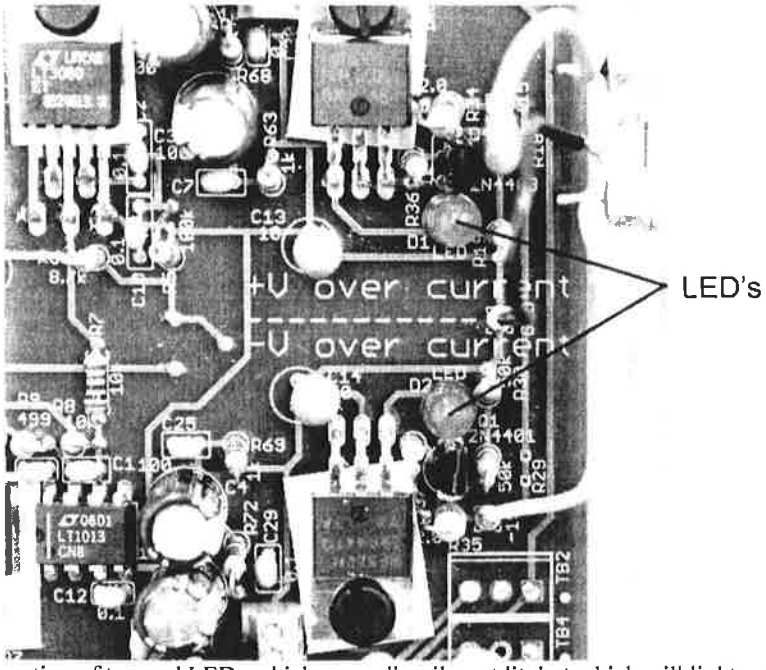


Fig. A.7a: Location of two red LEDs which are ordinarily not lit, but which will light up (though not very brightly) in case of a fault in the power-regulating circuits.

When you've flipped the panel back to its right-side-out configuration, you can re-check the proper operation of its ±11-V power supplies.

4) Some "insults" to the power supplies could result in the last-stage pass transistors failing, and failing to an open-circuit condition. Under this mode, you'll see no ±11-V capability, and you'll have to replace the pass transistors. The procedure for **Replacing** is described at the end of this s ection.

5) Other insults to the power supplies can cause the pass transistors to fail to a *short*-circuit condition. Under this mode, they will pass d.c. current, but will *fail* to remove the high-frequency fluctuations in the voltage they supply. So you'll still get an apparently useful 0 to +11-V, and/or a 0 to -11-V, output, perhaps with even a bit more voltage range than you got before. But the output you get will now be much noisier, giving voltage noise density perhaps >200 nV/√Hz instead of typically <2 nV/√Hz at the outputs. You'll need to measure this noise level to diagnose this problem, but you can't see this excess noise on a 'scope. (That's because 200 nV/√Hz of noise density, extending all the way from d.c. to 1 MHz, still gives a net voltage fluctuation of only 200 μV, rms measure.) Here's a circuit that will do the noise measurement you need -- it'll pass all noise components above about 1 Hz to the pre-amp, which is configured just as in Section 1.1:

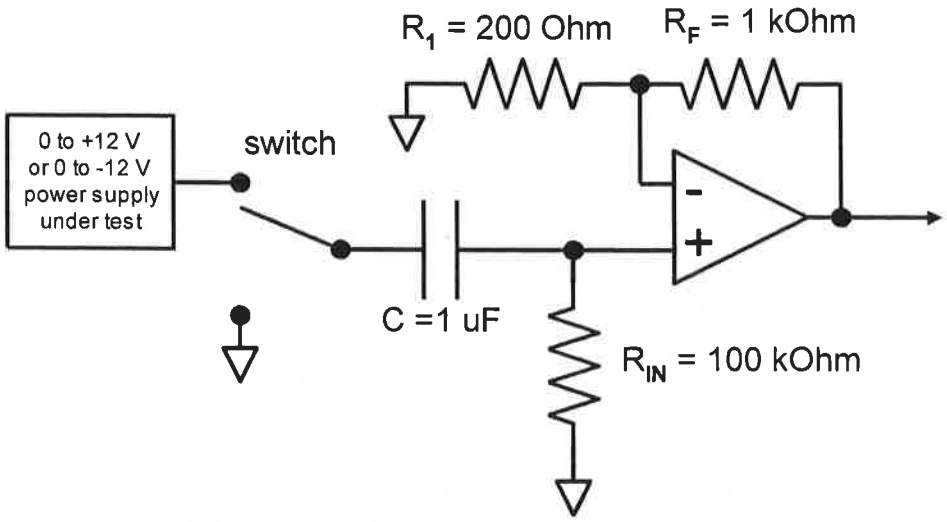


Fig. A.7b: A circuit for a.c.-coupling one of the 0-11 V power supplies to the input-stage op-amp of the pre-amp, with input stage configured for gain 6.00. (This gives overall pre-amp gain G_1 = 600.) Note that a switch setting can give you input connected to ground, rather than the positive or negative supply, as a 'control group' or check.

You'll need to understand the gain of the pre-amp, and the main-amp, and the bandwidth of your filtering, to get this method to work quantitatively. If you confirm the high-noise state of output, you'll need to replace the pass transistors.

Replacing the pass transistors

Conduct this operation with the low-level electronics' power turned **OFF**.

The transistors you might need to replace are labeled on the silk-screened printed-circuit board of the power-regulating part of the low-level electronics:

(for the 0 to +11-V supply)	Q2	type 2N4401
(for the 0 to -11-V supply)	Q6	type 2N4403
(for the op-amp + supply)	Q3	type 2N4401
(for the op-amp - supply)	Q5	type 2N4403

Before you take out a suspect transistor, make a sketch of its orientation of its black plastic package, so that you can orient the three leads of the replacement part correctly. Note that the orientations of Q5 and Q6 differ from those of Q2 and Q3. When you're ready, unscrew the three terminals to remove a transistor's leads; get a replacement device from your spare-parts bin; bend its leads to match those of the suspect device; and insert and screw it into the terminal block.

After you've replaced one transistor, here's a two-part test to see if it is working correctly: restore power to the low-level electronics, and

- use a DMM to test the d.c. level, variable or fixed, of the power supply you've repaired, for proper operation;
- then re-wire the noise-level test above to see if the power supply output has now gotten 'quiet' -- for this test, you'll have to re-flip the low-level electronics' front panel and close up the box.

Appendix A.8. Limits to the Johnson noise spectrum

We claim that Johnson noise is white, ie. that it delivers equal amounts of energy into frequency bins of equal width. But the frequency axis extends from zero to infinity, so the total energy summed over all frequencies would seem be infinite. Clearly the Johnson noise spectrum must drop to zero at some high frequency; else we'd have an 'ultraviolet catastrophe', just as in blackbody radiation.

Nyquist's derivation of Johnson noise shows that not only the disease, but also the cure, has the same form in both problems. In blackbody radiation, the electromagnetic spectral energy density $\rho(f)$ (with units of Joules of energy, per cubic meter of volume, per Hertz of bandwidth) has a frequency dependence of the form

$$\rho(f) \propto f^3 \, [\exp(hf / k_B T) - 1]^{-1}$$

which can be written as

$$\rho(f) \propto f^2 \frac{f}{e^{hf/k_B T} - 1} \quad .$$

The factor f^2 is appropriate to a 3-d calculation, and it turns into a factor f^0 in Nyquist's 1-dimensional calculation of electromagnetic energy in a transmission line joining two resistors. The second factor has the same origin, and the same consequences, in blackbody radiation and in Johnson noise. It's a factor which goes to a constant at low frequencies, but drops exponentially like $\exp(-hf / k_B T)$ once we have $hf >> k_B T$. That result puts quantum mechanics into our electronics problem, and it cures our ultraviolet catastrophe. It also tells us that (if nothing else were to limit the spectrum) Johnson noise can only extend out to about $f_{max} \approx k_B T / h$. (What upper frequency limit does that set, for room-temperature experiments? How about at $T = 20$ mK?) Short of this quantum cut-off, and certainly in the range relevant to tabletop electronics, we have $hf << k_B T$, and using $hf/(k_B T) << 1$ allows us to write the spectral energy density appropriate to one dimension as

$$\rho(f) \propto f^0 \frac{f}{e^{hf/k_B T} - 1} \approx f^0 \frac{f}{(1 + \frac{hf}{k_B T}) - 1} \approx f^0 \frac{k_B T}{h} \quad .$$

Notice this result is linear in temperature T, and that it is also independent of frequency. So this is the origin of the overall $f^0 T^1$ or 'white', but temperature-dependent, Johnson-noise spectrum.

In practice, Johnson noise nearly always drops *below* the book-value density at *much* smaller frequencies than the quantum limit mentioned above. We model a real resistor, which displays Johnson noise, as the series combination of a Johnson-noise emf and an ideal (noiseless) resistor. What we'd like to measure, with an ideal voltmeter, is the Johnson noise voltage $V_{IN}(t)$, using the circuit in Figure A.8a.

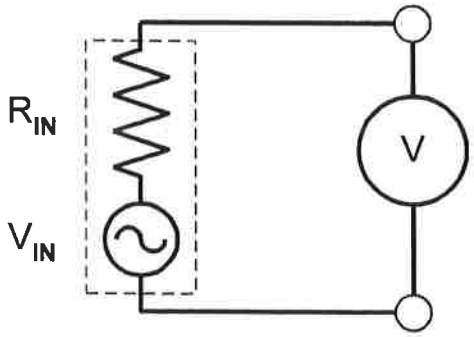


Figure A.8a

But in practice, there is capacitance, say between the two wires, or at the input of the voltmeter, so in reality the circuit we have (as shown in Fig. A.8b, in equivalent circuits) is formed by the source resistor's own resistance, and the stray (or voltmeter) capacitance. A filter like this has a 'corner frequency' given by $f_c = (2\pi R_{in} C)^{-1}$, and this corner is of real concern. If you use a source resistor of R_{in} = 100 kΩ and have even 10 pF of stray capacitance, you have $R_{in} C = (10^5\ \Omega)(10^{-11}\ \mathrm{F}) = 10^{-6}$ s, so $2\pi R_{in} C \approx 10^{-5}$ s, and $f_c \approx 10^{+5}$ Hz. That is to say, the Johnson noise spectrum can easily be rolling off at 100 kHz and above.

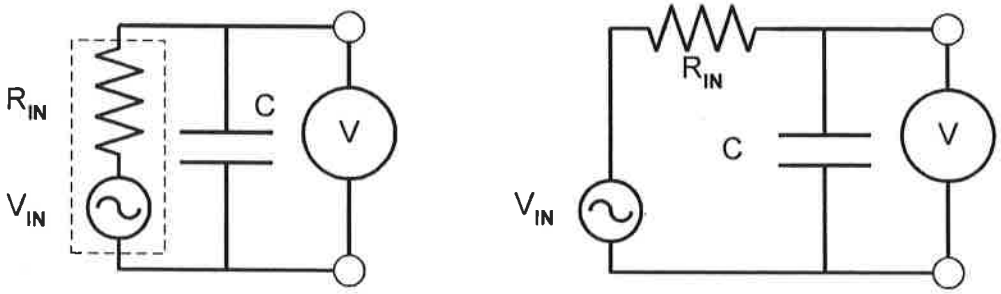


Figure A.8b

The problem is worse for larger source resistance, and much worse when the temperature probe of Section 4 is used -- there, the need for a coaxial cable raises C to about 100 pF. That's why the use of bandwidths Δf *not* extending to high frequencies is important for getting accurate values of mean-square Johnson noise.

A model for equivalent noise bandwidth, under these circumstances, is the usual integral of the square of the gain function, but now with three factors in it:

- a possible high-pass response at (low) frequency f_1 , and
- a low-pass response at (higher) frequency f_2 ; both of these taken to be ideal Butterworth functions, but now these are supplemented by
- a one-pole low-pass roll-off response at the corner frequency f_c determined by capacitive effects.

So the complete gain function becomes

$$G(f)=[\frac{(f/f_1)^2}{\sqrt{1+(f/f_1)^4}}][\frac{1}{\sqrt{1+(f/f_2)^4}}][\frac{1}{\sqrt{1+(f/f_c)^2}}] \quad .$$

You'll find (by numerical integration) that if the capacitively-caused corner f_c lies at frequency $10 f_2$ or higher, the equivalent noise bandwidth is decreased by less than 1% due to this effect.

But there is another interesting limiting case. Suppose that we use no high-pass filter at f_1, and no low-pass filter at f_2, but that the bandwidth is limited *only* by the capacitive roll-off at the source. The noise is then born with a density uniform in frequency,

$$S = <V^2(t)> / \Delta f = 4\, k_B\, T\, R\,,$$

so the mean-square value of the emerging signal would be

$$\left\langle V^2(t)\right\rangle = \sum \frac{\left\langle V^2(t)\right\rangle}{\Delta f}\,\Delta f \rightarrow \int_0^\infty 4\, k_B\, T\, R\, df \quad .$$

That would be infinite! but only because we left out the factor $G^2(f)$ in the integrand. In the RC-filter case at hand, $G(f)$ is the magnitude of the RC-filter's transfer function, which is given by

$$G(f) = [1 + (f/f_c)^2]^{-1/2}\,.$$

So instead of an ultraviolet catastrophe, we get

$$\left\langle V^2(t)\right\rangle = \int_0^\infty 4\, k_B\, T\, R\, \frac{1}{1+(f/f_c)^2}\, df = 4\, k_B\, T\, R \int_0^\infty \frac{df}{1+(f/f_c)^2} = 4\, k_B\, T\, R\, \frac{\pi}{2}\, f_c \quad .$$

Now the use of $f_c = (2\pi\, R\, C)^{-1}$ gives a neatly finite result,

$$<V^2(t)> = 4\, k_B\, T\, R\, (\pi/2)\, (2\pi\, R\, C\,)^{-1} = k_B\, T\, /C\,.$$

If you look at the circuit we're effectively using, you'll see that $V(t)$ is not only the voltage across the meter, it's also the potential difference across the capacitor C. That capacitor stores energy in amount $U_{cap} = C\, V^2/2$, so, though the instantaneous value is fluctuating, the time-averaged value of stored energy is non-zero, and given by

$$<U_{cap}> = (1/2)\, C\, <V^2(t)> = (C/2) \cdot k_B\, T\, /C = (1/2)\, k_B\, T\,,$$

which is a lovely illustration of the equipartition theorem. In fact, it shows that dissipation in a resistor (attached in parallel to a capacitor) comes accompanied by thermal fluctuations which prevent the resistor from discharging the capacitor all the way to zero. Instead, those fluctuations are the very mechanism responsible for the energy that, on average, is present in the capacitor.

The most remarkable feature of this result is that the measurable answer for $< V^2(t)>$ *depends not at all* upon the value of the resistance R, yet the resistor is nevertheless the source of the mean-square voltage being measured. In fact, you can measure a result

which depends on the resistor's temperature T, but not on its resistance R! The reason is that R's value turns up in two places, and cancels in this result: doubling R would double the Johnson noise power density, but it would also halve the equivalent bandwidth of the circuit, leading to the disappearance of R's value from the result. Perhaps you can think of a project, using a thermistor or a photoresistor to give a resistor of externally-controllable R-value, which tests this remarkable prediction.

Appendix A.9. Gaussian noise vs. white noise

You've repeatedly seen the words 'white noise', and you have perhaps also heard of Gaussian noise. Both the Johnson noise, and the shot noise, that you've been studying are ***both*** white and Gaussian in character. *But these are two separate attributes of noise*, and this section discusses the distinction.

We'll start by assuming you have a Johnson- or shot-noise source, amplified by the pre-amp, unfiltered but further amplified by the main amp to give a broadband noise signal of about 3 Volts (in rms measure).

That noise is white if it delivers equal amounts of power in any two frequency bands of equal width. That is to say: if after the main amp, an ideal sharp-edged filter were to pass all (but only) frequencies in the band $(f_0 - \Delta f/2, f_0 + \Delta f/2)$, then the mean-square value for the resulting filtered signal would be linear in the choice of bandwidth Δf, but independent of the choice of band-center f_0. Whiteness of noise in a frequency-domain stipulation, summarized by saying *that spectral density S(f) is in fact frequency-independent*. The broadband noise you'd get in the set-up mentioned above is very near to white in the 1-100 kHz range. In practice, there might be excess low-frequency noise visible below 1 kHz, originating in the amplifiers; and there would also be some roll-off of noise at high frequencies, perhaps below 100 kHz or beyond 1 MHz (see Appendix A.8 for details).

By contrast to this frequency-domain view, the Gaussian nature of noise is specified wholly in the time domain. Think back to that broadband noise signal emerging from the main amplifier, which (in the absence of high- or low-pass filtering) has frequency content out to about 1 MHz. Suppose you sample and digitize that voltage, at a collection of random times (or equivalently, at a collection of times separated by more than the autocorrelation time of the source, which is here about 1 μs), and produce a long list of instantaneous voltage values $\{V_i\}$. Now you can make a histogram of that list, and the noise is Gaussian only if that histogram matches a Gaussian distribution. (Have you lost the rare outliers that the Gaussian distribution predicts? If so, is that because your analog voltage signal has 'run into the rails' at some point? If so, reduce the rms measure of the source noise, or the gain of the main amplifier, until even rare events will fit into your range.)

Noise can be Gaussian as a consequence of the independent operation of many independent sources; in that case, Gaussian behavior is to be expected because of the central limit theorem. Noise can be white as a consequence of processes of short or zero autocorrelation time (see Appendix A.11.) So it's no accident that some fundamental kinds of noise are both white and Gaussian.

But noise can be white and not Gaussian at all. For example, if you deliver a single pulse of fixed amplitude and brief duration, its Fourier spectrum is white (out to a frequency about equal to the reciprocal of that brief duration). Now if you have a succession of such pulses, all of identical polarity, amplitude, and still-brief duration, but occurring at random (Poisson-distributed) times, the noise this represents is still spectrally white. But its voltage histogram is *nothing* like Gaussian -- instead, it would consist of only two values, corresponding to the pulse-absent and pulse-present conditions.

Similarly, noise can be Gaussian but not white. The Noise Calibrator built into the high-level electronics has a voltage histogram which is very close to Gaussian; that's due to the central limit theorem and the use of lots of sinusoids in its construction. As is happens, the noise is also white, by design, in the 0 - 32 kHz band. But the way such pseudo-noise sources are built would allow for any desired shape of $S(f)$-function, including 'pink noise' with extra energy at low frequencies.

Finally, there's a connection between the voltage histogram of a time-domain signal and its rms measure. If $p(V)$ gives the probability of getting a particular value V for the voltage, then

$$\int p(V)\, dV = 1$$

expresses the normalization condition for probability. Similarly

$$\int V\, p(V)\, dV$$

would be the way to compute the d.c. average value of the signal (if any), and

$$\int V^2\, p(V)\, dV$$

would give the mean value of the square of the voltage. The rms measure of the signal is the square root of this,

$$V_{\rm rms} \equiv [\int V^2\, p(V)\, dV]^{1/2} .$$

Of course the rms measure is alternatively given by a calculation in the frequency domain. By definition of (single-sided) spectral density, we have

$$\langle V^2(t) \rangle = \int_0^\infty S(f)\, df \quad ,$$

so we can also write

$$V_{rms} = [\int_0^\infty S(f)\, df\ ]^{1/2} \quad .$$

The particular form of $p(V)$ for a Gaussian noise signal of rms measure A is given by

$$p(V) = \frac{1}{A\sqrt{2\pi}} \exp(-\frac{V^2}{2A^2}) \quad .$$

Thus a noise signal of 3-Volt rms measure has parameter $A = 3$ Volts, and a Gaussian distribution of voltage values which has relative size 1 (at $V = 0$), $e^{-1/2} \approx 0.607$ (at $V = 3$ Volts), $e^{-2} \approx 0.135$ (at $V = 6$ Volts), and $e^{-4.5} \approx 0.011$ (at $V = 9$ Volts). You may use an integration on the formula above to find, for example, the proportion of all voltage samples which are expected to have $|V| > 10$ Volts.

Appendix A.10. Fourier methods for quantifying noise

This section takes up the possibility of getting the frequency spectrum of noise by computational processing of an amplified noise signal, captured in the time domain. We assume an ordinary noise experiment, complete with pre-amp, filter sections(s), and main amplifier, except that in this method, we give up the use of the squarer, and instead acquire the main-amp output as a voltage-vs.-time waveform.

a) Via oscilloscopes

You've monitored the Noise Fundamentals main-amp output on an oscilloscope on many occasions, but have always viewed the waveform itself -- that's the 'time-domain' $V_A(t)$ signal. But many 'scopes offer an 'FFT' or fast-Fourier-transform utility, intended to show you a 'frequency-domain' view instead, of the spectral content of the $V_A(t)$ signal. We'll see below some details on how such things are computed from a sampled and digitized version of $V_A(t)$. Here, let's mention typical limitations of 'scope-based FFT presentations:

1) There's nothing to enforce on the user the choice of an adequate sampling rate, and the wrong choice will lead to a grossly deceptive view of the frequency content of the signal. (This effect involves the 'aliasing' of spectral content to wholly other locations in frequency space.) The requirements for sampling rates which *will* give a display faithful to the waveform's actual spectral content are given in section b) below.

There's also nothing to prevent the user from choosing too sensitive a vertical scale on the 'scope, in which case an input signal which saturates the digitization range of the 'scope can have its spectral content splattered about unpredictably in frequency.

2) The horizontal scale of spectral displays is given correctly by oscilloscopes' FFT routines, but the vertical axis is typically left in arbitrary units. It's also traditionally plotted on a logarithmic or decibel (dB) scale, with 10 dB/div meaning that every vertical division signifying a ten-fold increase in spectral power. But reading the actual spectral power, in absolute V^2/Hz units, is a capability reserved for special 'spectrum analyzers'. One of the main goals of the sections below is to lead readers through a treatment of actual computation, by Fourier means, of results for noise power-density spectra, whose units and normalization can be understood and trusted quantitatively and in detail.

b) Sampling

This is possible given a digital sampling instrument, such as an oscilloscope, which can acquire a long series of (perhaps 10^3 to 10^5) voltage 'samples', all acquired at some uniform spacing in time. The reciprocal of this inter-sampling spacing is called the 'sampling rate', and it is critical that this rate be high enough.

How high is high enough? This comes from Shannon's 'sampling theorem', which says that *if* a waveform contains only frequency content below a maximum frequency f_{max}, *then* a sampling rate $\geq 2\, f_{max}$ is adequate. In such a case, in fact, the samples alone permit a complete reconstruction of the signal (including its unseen portions between the sampling points!). So if TeachSpin's Noise Calibrator output has frequency content (only) in the 0 - 32 kHz range, a sampling rate of ≥64 kSa/s (kilo Samples per second) would suffice. In practice, we might have a 'scope arranged to acquire one sample every 10 μs, giving an adequate sampling rate of 10^5 Sa/s = 100 kSa/s.

Now generic noise signals *lack* such an obvious upper-frequency limit, so for faithful sampling, it's important to limit their spectral coverage, by using a low-pass filter before the sampling. (You may have heard this called an 'anti-aliasing' filter.) But typical filters do not impose a sharp upper edge to a spectrum. You've seen in Section 2.2 that the TeachSpin low-pass filters pass some spectral energy out to $\approx 10 \cdot f_c$, where f_c is their nominal corner frequency. So if you use a 100-kHz low-pass filter, there's enough energy out to ≈1 MHz (and a bit more beyond) that you'd want to sample at 2 MSa/s. Note that at this sampling rate, an array of 10^5 samples will fill up in just 50 ms of time. Note also that if you use a lower corner frequency in your filter, you can afford a lower sampling rate.

c) **Scaling**

Suppose from a waveform $V(t)$ you have a collection of samples, $\{V(t_k)\}$, where the t_k are the sampling instants, separated by fixed sampling interval Δt. If there are N such samples, we could lay them out in the $-T/2 < t < T/2$ interval according to

$$t_k = -\,T/2 + (k) \cdot \Delta t\,, \quad \text{for } k = 0 \text{ to } N - 1.$$

Here T is the total duration of your sampling, and $N\,\Delta t = T$ relates N, T, and Δt.

Now if you had captured the actual continuous waveform $V(t)$, you'd reach for Fourier transforms, which we'll quote here in their complex-exponential form, and in ordinary (not angular) frequencies. In that notation, the Fourier-transform pair is

$$\tilde{V}(f) = \int_{-\infty}^{\infty} V(t)\, e^{2\pi i f t}\, dt \quad ; \quad V(t) = \int_{-\infty}^{\infty} \tilde{V}(f)\, e^{-2\pi i f t}\, df \quad ,$$

which together form a theorem, under certain conditions. But noise signals which go on indefinitely do *not* meet those conditions, since they're of constant power, rather than of finite energy. Yet we can define a scaled version of the voltage signal which preserves the frequency content of $V(t)$, via

$$W_T(t) \equiv (1/\sqrt{T})\, V(t)\,, \text{ for } -T/2 < t < T/2\,; \qquad \text{but } \equiv 0 \text{ elsewhere.}$$

This claims that $W_T(t) = 0$ outside your sampling duration (which might be true, for all that you've recorded). With this definition, we have

$$\int_{-\infty}^{\infty} |W_T(t)|^2\, dt = \int_{-T/2}^{T/2} \left| \frac{1}{\sqrt{T}} V(t) \right|^2 dt = \frac{1}{T} \int_{-T/2}^{T/2} |V(t)|^2\, dt \quad .$$

Experimental voltage signals are real-valued, so this right-hand side clearly defines $<V^2(t)>$, the mean-square value of the noise voltage, which we presume is finite. Then the left-hand side shows that $W_T(t)$ is a square-integrable function to which Fourier's Integral Theorem *does* apply, allowing us to define its transform as

$$\tilde{W}_T(f) = \int_{-\infty}^{\infty} W_T(t)\, e^{2\pi i f t}\, dt = \int_{-T/2}^{T/2} \frac{1}{\sqrt{T}} V(t)\, e^{2\pi i f t}\, dt \quad .$$

The inverse transformation is given by

$$W_T(t) = \int_{-\infty}^{\infty} \tilde{W}_T(f)\, e^{-2\pi i f t}\, df \quad .$$

Because these W-functions are a Fourier-Transform pair, they satisfy Parseval's Theorem,

$$\int_{-\infty}^{\infty} |W_T(t)|^2\, dt = \int_{-\infty}^{\infty} \left|\tilde{W}_T(f)\right|^2\, df \quad ,$$

and now we can see that *both* sides of this equation have value $<V^2(t)>$, the mean-square noise voltage. So a physicist's viewpoint on this equality is to think of a noise source of some mean-square strength, and then to see that this given quantity (proportional to noise power) can be dis-aggregated either according to its time of occurrence (on the left), or according to its spectral distribution (on the right).

d) **Frequency content**

The Fourier transform $\tilde{W}_T(f)$ is defined on the whole line, $-\infty < f < \infty$, so it seems to contain both positive and negative frequencies. In practice, since the original signal $V(t)$ is real-valued, $\tilde{W}_T(f)$ can be shown to obey

$$\tilde{W}_T(-f) = \tilde{W}_T(+f)^* \quad ,$$

where the * stands for complex conjugation. So the information in $\tilde{W}_T$ for positive frequencies alone is sufficient to describe the whole function. It's easy to show that

$$\int_{-\infty}^{\infty} \left|\tilde{W}_T(f)\right|^2\, df = \left\langle V^2(t) \right\rangle = 2\int_0^{\infty} \left|\tilde{W}_T(f)\right|^2\, df \quad ,$$

so integrals over positive frequencies alone can tell you the full mean-square measure of the noise.

In practice, the discrete Fourier-transform methods described below are best conducted by keeping $\tilde{W}_T(f)$ as a complex function, and extracting its spectral content at the end of the computation by adding together the 'positive and negative' frequency contributions.

e) **Spectral density function**

So given a noise waveform $V(t)$, observed for a duration T, it's feasible to define a scaled function $W_T(t)$, and to compute its Fourier transform $\tilde{W}_T(f)$. Then we might define a noise power spectral density

$$S(f) = 2\, |\, \tilde{W}_T(f)\, |^2 \, ,$$

which is a computable function obeying the desired normalization

$$\int_0^{\infty} S(f)\, df = \left\langle V^2(t) \right\rangle \quad .$$

Thus the mean-square value of a voltage-noise function has been dis-aggregated into its frequency content. We'd call this $S(f)$-function the 'single-sided spectral density of noise power'. It turns out to have units of V^2/Hz, so integrating it over frequency gives Volts-squared, the correct units for $\langle V^2(t) \rangle$. So this is the computational route from $V(t)$ to a spectral density $S(f)$.

The only deficiency in this procedure is that it lacks any proper limit as $T \to \infty$. If you have an actual recording of the waveform of a noise source, and process it for ever-wider $-T/2 < t < T/2$ windows of observation, you'll find that the $S(f)$ functions computed by the above procedure gives you ever-higher spectral resolution, and shows you ever-finer details of apparent frequency variation of $S(f)$. All of this highly resolved structure is *irreproducible*, and would show up differently on a second try, for the same noise source. In practice, spot values of $S(f)$ produced by this procedure aren't convergent or useful, but wide-band or even narrow-band integrals like

$$\int_{f_1}^{f_2} S(f)\, df$$

are useful, and they *do* converge to well-behaved limits as $T \to \infty$. We'll use this fact below to motivate spectral-averaging of computed $S(f)$ values.

f) Discrete Fourier transforms

It should be clear that actual Fourier integrals cannot in fact be computed unless you were to have access to continuously-varying functions like $V(t)$. In practice, we have to be content with a finite collection of samples, such as the set $\{V(t_k)\}$ measured at N sampling points t_k separated by intervals Δt. But this very finiteness allows us to change from the integral transforms to 'discrete Fourier transform' sums instead, as follows.

We give ourselves a time window of duration T, which might be the full duration of the experiment, so that (for all we know to the contrary), a signal $V(t)$ might actually repeat, with period T, outside our window of observation. That's a convenient assumption, since any complex-valued function with period T can be written as a sum of complex exponentials of particular frequencies. We'll choose an indexing in which $f_0 = 0$ is the 'd.c.' term, $f_1 = 1/T$ is the 'fundamental' frequency, and write

$$f_n = (n)\,(1/T) = (n)\,(N\,\Delta t)^{-1}\,, \quad \text{for } n = 0,\, 1,\, 2,\, \ldots$$

Then there exists a Fourier series for the assumed-periodic $V(t)$-function,

$$V(t) = \Sigma_n \,(\text{coefficient } \#n)\, \exp(-2\pi\, i\, f_n\, t)\,.$$

Under our convenient fiction of the periodicity of $V(t)$, we can just as well take a time window $0 \le t < T$, defining the N sampling points spaced by interval Δt as

$$t_k = (k)\,(\Delta t)\,, \quad k = 0 \text{ to } N - 1, \text{ still with } \Delta t = T/N\,.$$

To achieve a perfect fit to the N sampled data-points $\{V(t_k)\}$, it turns out that we require exactly N (complex-valued) coefficients, which we choose to write as

$$V(t_k) = \sum_{n=0}^{N-1} [\tilde{V}(f_n)]\, e^{-2\pi i f_n t_k} = \sum_{n=0}^{N-1} \tilde{V}(f_n) \exp[-2\pi i (nk)/N] \quad .$$

This sum is called the 'forward DFT', and it maps the N frequency-domain entries $\{\tilde{V}(f_n)\}$ to the N time-domain entries $\{V(t_k)\}$. This mapping is also exactly invertible -- the transformation going the 'other way' is called the 'inverse DFT', and is given by

$$\tilde{V}(f_n) = \frac{1}{N} \sum_{j=1}^{N} V(t_k) \exp[+2\pi i (kn)/N] \quad .$$

These two equations form a discrete-Fourier-transform (DFT) pair, and they are of extreme computational interest because of the amazingly efficient Cooley-Tukey 'fast Fourier transform' or FFT algorithms which have been devised to evaluate them. We've written the transforms with indices k, $n = 0$ to $N - 1$, and matched the notation and the normalization used by the open-source program Sage in its `fft()` and `inv_fft()` functions.

So here's what we actually do to get power spectral density. We want values of

$$\tilde{W}_T(f) = \int_0^T \frac{1}{\sqrt{T}} V(t)\, e^{2\pi i f t}\, dt \quad ,$$

which we compute by changing the integral to the Riemann sum we'd use to approximate it,

$$\tilde{W}_T(f) = \sum_{k=1}^{N-1} \frac{1}{\sqrt{T}} V(t_k)\, e^{2\pi i f t_k} \cdot \Delta t \quad .$$

For fictionally-periodic $V(t)$, and for a finite number of samples of it, we're content to know $\tilde{W}_T(f)$ at the frequency values f_n given above, yielding

$$\tilde{W}_T(f_n) = \sum_{k=0}^{N-1} \frac{1}{\sqrt{T}} V(t_k)\, e^{2\pi i f_n t_k}\, \Delta t = \frac{\Delta t}{\sqrt{T}} N \bullet \frac{1}{N} \sum_{k=0}^{N-1} V(t_k) \exp[2\pi i (kn)/N] \quad .$$

The factor preceding the bold dot is just $\sqrt{T}$, while the whole quantity appearing after the dot is precisely a value from an output array of N complex numbers, the result of the inverse DFT on the input array $\{V(t_j)\}$. The normalization needed after doing the DFT is just multiplication by that $\sqrt{T}$ factor, which gives $\tilde{W}_T(f_n)$-values their proper units of $V \cdot \sqrt{s} = V/\sqrt{Hz}$, so we have

$$\tilde{W}_T(f_n) = \sqrt{T} \bullet \textit{entry n of inv\_fft of the } \{V(t_k)\} \textit{ array} \quad .$$

Next, the absolute squares of these values give $|\tilde{W}_T(f_n)|^2$ values, with units of V^2/Hz, which are very closely related to the desired spectral density $S(f)$. We need only to remember three things:

1) Given an list of sampled voltage values, indexed by $k = 0$ to $N - 1$, namely $\{V(t_k)\}$, we need only an inverse-DFT algorithm which produces the output list, $\{\tilde{V}(f_n)\}$, a list indexed by n running from 0 to $N - 1$. The DFT algorithm needs to know the value of N (the length of the input and output lists), but it does not 'need to know' anything about the value of Δt or T. And by this stage of the computation, all reference to the W and $\tilde{W}$ functions can be dropped – the inverse DFT algorithm can be applied directly to the list of $V(t_j)$ values.

Given our choice of indexing, the frequency values associated with the index n are $f_n = (n)\cdot 1/T$, so f_0 is the d.c. term, $f_1 = 1/T$ is the 'fundamental frequency', $f_2 = 2/T$, and so on. To get single-sided spectral densities, we need to account for the 'negative frequencies' too, and (since $\tilde{W}_T(f)$ turns out to be periodic in f) these can be found in the upper half of the list of N values of f_n. In fact, to get the spectral density at a p-for-particular frequency, where integer p maps to frequency $f = p\cdot 1/T$, we take

$$S(f = p\cdot 1/T) = \left|\tilde{W}_T(f_{n=p})\right|^2 + \left|\tilde{W}_T(f_{n=N-p})\right|^2 = T\cdot[\left|\tilde{V}(f_{n=p})\right|^2 + \left|\tilde{V}(f_{n=N-p})\right|^2] \quad .$$

For a real-valued function $V(t)$, the DFT will produce results giving equal contributions from the two absolute-squares shown. (This is the discrete version of our previous result $S(f) = 2\,|\,\tilde{W}_T(f)\,|^2$). The frequencies which collectively account for all of the noise power include $p = 0$ (the d.c. term), and then from $p = 1$ to $(N/2)$-1. So the maximum frequency at which we get back spectral-density data is

$$f_{max} = (\frac{N}{2} - 1)\cdot\frac{1}{T} = (\frac{N}{2} - 1)\cdot\frac{1}{N\,\Delta t} = \frac{N-2}{N}\cdot\frac{1/2}{\Delta t} \quad .$$

Since $1/\Delta t$ is the sampling frequency, we see that our spectral coverage is from d.c. to (just below) half the sampling frequency. (That's why digital audio uses a sampling frequency of 44.1 kHz, so as to cover completely the audible frequency range from d.c. to about 20 kHz.)

2) The $S(f_n)$-values thus computed obey a sum rule, which results from the discrete-Fourier-transform version of Parseval's Theorem:

$$\sum_{n=0}^{N-1}\left|\tilde{V}(f_n)\right|^2 = \frac{1}{N}\sum_{k=0}^{N-1}\left|V(t_k)\right|^2 \quad .$$

This can be manipulated to give

$$\sum_{n=0}^{N/2-1}\frac{1}{T}S(f_n) = \left\langle V^2(t)\right\rangle$$

whose units match: $(1/s)\cdot V^2/Hz$ on the left, and V^2 on the right. This equality can be used as a valuable check on normalizations and DFT algorithms. It is also the finite-sum version of

$$\int_0^\infty S(f)\, df = \langle V^2(t) \rangle \quad .$$

which (under the assumption of adequately dense sampling) has the Riemann-sum approximation

$$\sum_{n=0}^{N/2-1} S(f_n)\, \Delta(f_n) = \langle V^2(t) \rangle \quad .$$

Since $f_n = (n) \cdot 1/T$, we see $\Delta f_n = 1/T$, so this result agrees with that above.

3) The $S(f)$-values thus computed will suffer from the excess spectral resolution previously mentioned, and will display a 100% scatter, with rms deviation equal to their mean. The only cure for this scatter is averaging. One way is to take M multiple successive samplings of the noise stream, each of duration T, to process each of them separately, and then to average together M multiple versions of $S(f)$. Another way is to lengthen the (single) observation window by some integer factor F, to get one long list of $S(f)$ values with high spectral resolution, and then to give up this resolution by averaging together F adjacent frequency-content readings of $S(f)$ to revert to the original spectral resolution. The extra observations, by factor M or F, will give $\sqrt{M}$ or $\sqrt{F}$ less scatter of the $S(f)$ values that result.

As an example of the latter, suppose you take $N = 2^{16} = 65{,}536$ samples of $V(t)$. (Powers of 2 are convenient, since DFT algorithms reach their highest efficiency for such array lengths.) From the methods above, you'd get back $S(f)$ values at 2^{15} distinct frequencies. If you want to end up with local $S(f)$ estimates with scatter of order 10% or less, you'll need to average together >100 $S(f)$ values. So you might average together groups of 128 high-resolution $S(f)$ values to get S-values of lower frequency resolution. But you'd still have 256 distinct $S(f)$ averages, so your frequency span would be covered with better than 1% spectral resolution.

When you finally have a table of $S(f)$ values, it is conventional to take the square root of each, converting 'power spectral density' $S(f)$ in units of V^2/Hz into 'voltage spectral density' $D(f)$ in units of $V/\sqrt{Hz}$. If your electronic signal chain has included pre-amp gain G_1 and main-amp gain G_2, then the voltage noise density at the input of the pre-amp is smaller, by factor $G_1 \cdot G_2$, than the result that you have computed via Fourier methods.

g) Confirmation

You can test the success of your computational route to spectral density by working with the Noise Calibrator signal (see Section 5.4.) You can send it, unfiltered, right into the main amplifier, set to its minimum gain, $G_2 = 10$. If you sample the amplifier's output at 100 kSa/s (so that $\Delta t = 10$ μs), you will end up with $S(f)$-values in the 0 - 50 kHz range. How many samples you can acquire depends on the storage capabilities of your 'scope, but even if you take only 10^3 samples at a time, you can make multiple 'runs' of your

experiment to make the *M*-fold averaging method above work for you. Your sampling and computational route should reproduce a result close to

$$S(f) = (1.19\ \text{mV}/\sqrt{\text{Hz}})^2 = 1.42 \times 10^{-6}\ \text{V}^2/\text{Hz}$$

in the $0 < f < 32$ kHz range, and much smaller values beyond the 32-kHz limit of the noise source.

Appendix A.11. The autocorrelation function of noise

This introduces you to an alternative, and very revealing, method for viewing and thinking about noise signals. It provides a real-time method for understanding the spectrum of noise signals, or the bandwidth of circuits. We present it here first with an oscilloscope exercise you can do, and then describe its connection to the 'autocorrelation function' which is mathematically connected to the spectral distribution of noise power.

a) Observing a 'qausi-autocorrelation function'

This exercise requires only a source of noise in your Noise Fundamentals experiment, and a digital sampling oscilloscope. To do this experiment does not require the squarer, but you should set up your noise source, the pre-amp, a low-pass filter, and the main amplifier. Use a 33-kHz corner frequency for the low-pass filter, and use enough gain in the main amp to get its output up to 2~3 Volts (rms measure). Bring that signal to a 'scope, and use a vertical sensitivity that covers a ±8-V range (to accommodate the range of the noise), and use a horizontal scale of 25 μs/division.

Now if you've been seeing the noise in an automatic triggering mode, it's time to switch to a 'normal' mode, where you choose a trigger level (try a level around +6 Volts) and a slope (try positive slope, ie. trigger on signals rising through the +6-V level). You should see plenty of trigger events, since you're triggering on not-too-infrequent positive excursions of the noise. For a first look at these events, try the 'persistence' mode on your 'scope, and look for a picture like this:

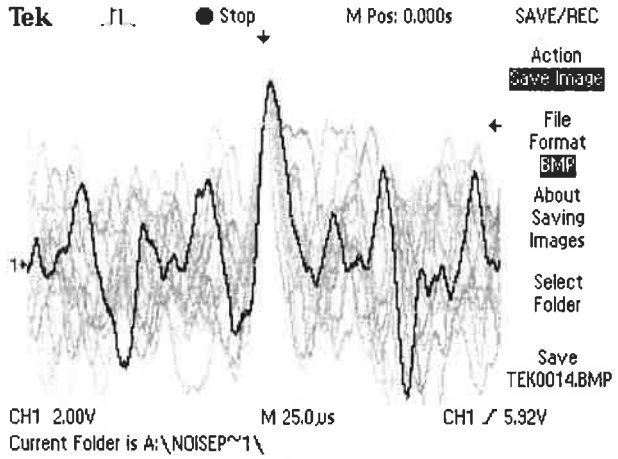


Fig. A.11a: An example of noise waveforms. Vertical scale 2 V/div, horizontal scale 25 μs/div, triggering on positive-going crossings of the +6-V level.

Notice that the trigger point has been centered on the horizontal axis. Note that every trace has the property of passing through the trigger point, both in time and in voltage. But also note that after the +6-Volt excursion, the generic trace shows signs of 'reversion toward the mean' of zero, within some finite time.

To see this in detail, change from the persistence to the 'averaging' mode of your 'scope, asking for the average of (say) 128 or 256 traces. You'll see the averaged view of reversion to the mean, with a result resembling this trace:

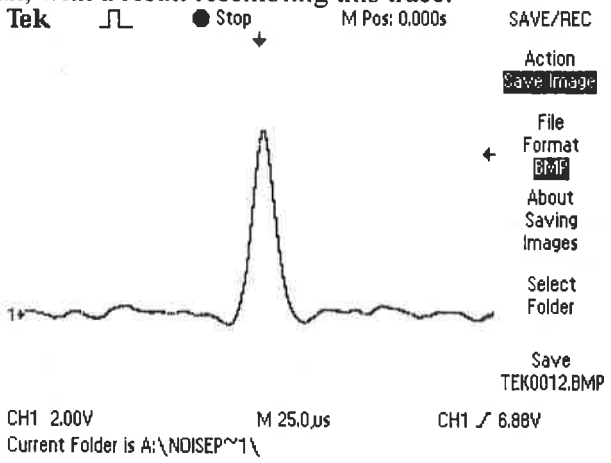


Fig. A.11b: Signal-averaged waveform in the same arrangement as above . Vertical scale 2 V/div, horizontal scale 25 µs/div, triggering on positive-going crossings of the +6-V level.

Because every individual trace passes through the trigger point, so does the average. But far enough downstream (or upstream) in time, the average value of the noise signal becomes zero again. What you have is a visual depiction of the 'autocorrelation time', which is an answer to the question 'How long does a typical positive excursion of noise last?'

To see that this pattern has something to do with your noise signal's spectral distribution, here are two comparison tests you can try:

i) change between 33-kHz and 10-kHz settings for the corner frequency of your low-pass filter. (The smaller bandwidth will give less noise power, so you may want to lower the trigger-level setting.) What you will see is a *longer* duration of the average positive excursion.

ii) change between a 33-kHz low-pass filter and a 33-kHz *band*-pass filter. (These have the same equivalent noise bandwidth, so there'll be no need to change the trigger point.) What you will see in a change in the *shape* of the result, which is due the different spectral composition of the signal you're seeing.

What are you seeing? Officially, it's called the 'conditional probability distribution', which answers the question: 'What is the ensemble-average value of $V(\tau)$, given that $V(\tau=0)$ = +6 Volts?' [Fine point: what you're seeing is conditional on a trigger level of +6 Volts and a positive slope. Change to triggering on a *negative* slope to see how little difference this makes.] Happily, for Gaussian noise, this easy-to-see result on your oscilloscope is directly proportional to the autocorrelation function of the noise waveform.

b) **Introducing the autocorrelation function**

What lies behind the value of this oscilloscope display is its connection to the autocorrelation function or ACF called $C(\tau)$, which is defined for any signal $V(t)$. For a real-valued function, we define

$$C(\tau) = < V(t)\ V(t\text{-}\tau) >$$

where the <brackets> stand for time averaging, and where we're assuming that $V(t)$ has statistical properties which are independent of time. In this expression, τ is called the 'lag time', and this expression measures something about how different $V(t)$ and $V(t\text{-}\tau)$ can be.

It's easy to see that $C(0) = < V(t)\ V(t\text{-}0) > = < [V(t)]^2 >$ gives the mean-square measure of the noise; this shows that $C(0)$ is always positive, and also shows that $C(\tau)$'s units are Volts-squared. There are also some properties of the official autocorrelation function $C(\tau)$ which are similar to those of the 'quasi-ACF' which you're viewing on the 'scope. The first of them (established via the Cauchy-Schwarz inequality) is that $|C(\tau)| \leq C(0)$ for any choice of τ. It's also feasible to show that $C(\tau) = C(-\tau)$, which shows that $C(\tau)$ is a function symmetrical about τ=0 (where it thus has an absolute maximum).

The definition of $C(\tau)$ also makes it clear why the function drops off with time, and why it distinguishes the regimes of short, vs. long, compared to some autocorrelation timescale. If τ is short enough, $V(t)$ and $V(t\text{-}\tau)$ will be similar, hence much more likely to be of the same (as opposed to opposite) signs. So the product $V(t)\ V(t\text{-}\tau)$ will be more probably positive than negative, so its time-average will be positive. By contrast, is τ is long enough, the present value $V(t)$ will be *uncorrelated* with its value 'τ ago' in the past, $V(t\text{-}\tau)$. So at those times when $V(t)$ is positive, $V(t\text{-}\tau)$ will be as likely to be negative as positive. Hence the product $V(t)\ V(t\text{-}\tau)$ will also be as likely to be negative as positive, and so its time average will be near zero.

So the shape of the function $C(\tau)$, like the quasi-ACF you saw on your 'scope, tells you about the degree to which the signal has some 'staying power' or even 'memory'. That is in general not the memory (if any) in the original source of the noise, but rather due jointly to the noise source and the bandwidth that might have been imposed upon its signal by subsequent filtering. In fact, there's an inverse relation between bandwidth and autocorrelation time: a truly white-noise signal with bandwidth to $f = \infty$ would have *zero* autocorrelation time, and act like a system with no memory at all. But the smaller the bandwidth, the longer the autocorrelation time; by the time you get to a pure sinusoid or any other periodic signal, the ACF shows non-zero correlations at arbitrarily long lag times.

And there's more than this informal connection between spectral distribution and autocorrelation function. It turns out that $C(\tau)$ on the one hand, and the power spectral density $S(f)$ on the other hand, are closely related as a Fourier transform pair. So knowing either function of this pair fully determines the other. For example, if we had a sharp-edged spectral distribution of noise, with $S(f)$ a constant from d.c. up to some

maximum frequency f_m (but zero beyond that point), then taking the Fourier transform of $S(f)$ allows us to predict

$$C(\tau) \propto \sin(2\pi f_m \tau) / (2\pi f_m \tau) ,$$

which has the sinc-function 'wiggles', and displays its first zero-crossings at lags $\tau = \pm 1/(2 f_m)$. The 'wiggles' disappear for spectral distributions lacking as sharp a cutoff as in this example, but the width (which was $1/f_m$, between innermost zeroes, for this sharp-edged distribution) will continue to be inversely proportional to the bandwidth of the spectral distribution.

It also follows that two sources with distinct spectral distributions (such as the low-pass, vs. the band-pass, filtered versions of white noise) must have distinguishable $C(\tau)$ functions. It also follows that careful measurement of $C(\tau)$'s values can provide a quantitatively reliable way to compute $S(f)$. The place to pursue this connection is a presentation of the Wiener-Khinchin theorem in signal processing.

c) Best use of a 'scope's FFT-capability

If you have used the FFT utility of your 'scope to view the frequency spectrum of a noise signal, you may have been horrified at the fluctuations of the spectrum. A white-noise spectrum should give an $S(f)$-function which is a flat line, but in practice you've seen a host of jagged spikes and dips downward from the level you've expected. Appendix A.10 deals with some of the cures to this problem that you can impose if you compute your own FFTs off-line, but here's a capability which you can execute on your 'scope directly.

Ideally, you could ask for the FFT, and then the 'averaging' mode. What you'd *want* is an average of many successive spectral distributions. But what you'll *get* is the Fourier transform of (a bunch of noise waveforms all averaged together). That doesn't work right -- averaging the noise together (first) tends to wash away its strength, so your signal and the consequent FFT disappears.

So here's what to do instead. You set a trigger level as above, and average the time waveforms that appear at this trigger level, to get what we've called the quasi-ACF. Now you are using the averaging mode of your 'scope in a way which does not average the time-domain signal away toward zero; instead, you're getting a version of the auto-correlation function $C(\tau)$. *Then* you ask for the FFT, and you'll get the FFT of $C(\tau)$, and what you'll get from the computation is closely related to $S(f)$ -- by the Wiener-Khinchin theorem. What will be *displayed* may be $| S(f) |^2$, so it's hard to use this display with quantitative certainty about the scale of its vertical axis. (The half-power points might be depicted at -6-dB down, for example.) But you *will* get a display of spectral content with markedly smaller scatter, vertically, than you'd get in the direct FFT of the input waveform.

d) Using the quasi-ACF for analysis

For present purposes, here's another result which is easy to prove about $C(\tau)$, and also easy to observe using the quasi-ACF method on your 'scope. Suppose that a signal $V(t)$ is really the superposition of signals from two distinct sources,

$$V(t) = V_a(t) + V_b(t) .$$

You could call one of these signal, and the other noise; or it might be that one is the desired noise, while the other is *un*desired noise. Here's the $C(\tau)$ you get in this case:

$$\begin{aligned} C(\tau) &= \langle V(t)\, V(t\text{-}\tau)\rangle \\ &= \langle [V_a(t) + V_b(t)]\, [V_a(t\text{-}\tau) + V_b(t\text{-}\tau)] \rangle \\ &= \langle V_a(t)\, V_a(t\text{-}\tau) \rangle + \text{two cross terms} + \langle V_b(t)\, V_b(t\text{-}\tau) \rangle . \end{aligned}$$

The cross terms include $\langle V_a(t)\, V_b(t\text{-}\tau) \rangle$, which is *zero* for any and all τ-values, provided only that 'a' and 'b' stand for physically separate, ie. uncorrelated, sources of noise. So in this case,

$$C(\tau) = C_a(\tau) + C_b(\tau) ,$$

which shows that the ACF you'd observe is simply the sum of the ACFs you'd observe from the two sources separately.

You can get a great view of this process with your 'scope-based quasi-ACF. Try grounding the input of your low-pass filter section, set it to a 33-kHz corner frequency, and send its output to the main amp, set for maximum gain of 10^4. Observe the raw noise signal at the output of the main amp, and you'll see on a 'scope an entirely uninformative noise waveform, with no hint that it's composed of two kinds of noise. There's noise generated in the filter (with spectrum rolling off at around 33 kHz), and there's noise generated in the main amp itself (whose bandwidth extends to ≈1.5 MHz). Use your view of the net noise to choose the right vertical sensitivity, and to choose a good trigger level, and averaging, to produce a quasi-ACF. Now use sweep speed about 5 μs/div and look at the result, which should resemble this:

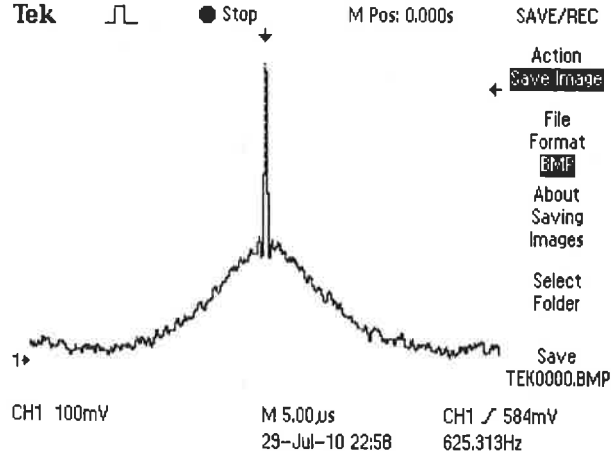


Fig. A.11c: The 'quasi-autocorrelation function' revealing the presence of two kinds of noise in the filter-plus-main-amp combination discussed above.

You can view this with a variety of timescale settings on your scope, to get a good view of both the narrow peak, and the broad hill atop which it's standing. You can even see some structure in that narrow peak -- look for some little valleys on either side of the narrow tall peak. The value of this exercise is the mental separation it permits. The narrow peak is the part of $C(\tau)$ due to a signal of short autocorrelation time, which must be of large bandwidth -- we identify that as main-amplifier noise. The full width at the base of the sharp peak is about 0.6 µs, which corresponds to a noise spectrum extending up to $f_m \approx 1.7$ MHz. The broad hill underlying the peak is the part of $C(\tau)$ due to a signal of *long* autocorrelation time; that must be of small bandwidth, and we claim it's due to noise born in the filter section. In fact the width of the broad hill is of order 20 µs, which is consistent with a frequency spectrum extending to about 50 kHz. And in agreement with the bit of theory above, we can now understand why these two pieces of the autocorrelation function simply add up to give the shape observed.

To test these claims further, you can change the bandwidth chosen on the filter -- what effect should that have? Or, you can send some white noise into the filter's input, which does not change the amount of noise that's actually generated within the main amp -- how will this show up? Or, you could imagine some interference (see Appendix A.5) from fluorescent-light ballasts, of frequency perhaps 25 or 48 kHz and approximately sinusoidal in character, is underlying your noise -- can you compute what *third* contribution to $C(\tau)$ that would create? As a practitioner, you can gain some instinctive knowledge from the easily-acquired, real-time quasi-ACF display on your 'scope, and use it as a key to diagnosing many kinds of experimental pathologies.

Appendix A.12. Fluctuations in measured noise: The Dicke limit

Noise signals are random, and as a result, measurements of 'noise power' display statistical fluctuations. This Appendix explains some nomenclature for these fluctuations, and describes and justifies the expected size of the fluctuations.

Let's imagine Johnson noise, or shot noise, measured by a now-familiar arrangements. We have an original noise voltage $V_n(t)$, characterized by zero mean but a non-zero mean-square. We pre-amplify it (by gain G_1), we filter it to bandwidth Δf (using filter gain function $G(f)$), we further amplify it (with gain G_2), and thus form a filtered and amplified noise voltage $V_{in}(t)$ as input to a squaring circuit. The mean of $V_{in}(t)$ is still zero.

When we square $V_{in}(t)$ to get $V_{out}(t) = [\ V_{in}(t)\ ]^2 / (10\ \mathrm{V})$, we finally get a signal whose mean is *not* zero. So when we average it over averaging time τ, we get a non-zero average

$$V_{meter} = < V_{out}(t) > \propto < [V_n(t)\ ]^2 > \ ,$$

and we can call that concrete meter reading a 'measure of the noise power'. But we can also easily see the visible fluctuations in the meter position, which we can call 'fluctuations in the measured noise power'. (They are sometimes called the 'noise in the noise', or 'second noise'.)

How big do we expect those fluctuations to be? This question was first addressed by Dicke, in an appendix to the paper [Rev. Sci. Instrum. **17**, 268 (1946)] which introduced lock-in detection, and which used it to measure room-temperature blackbody radiation in the microwave region of the spectrum. The result is therefore called the 'Dicke radiometer limit', usually expressed as a characteristic fluctuation δT observed in an instrument whose output gives T, a radiometrically-measured source temperature. Dicke's result can be written in terms of the bandwidth Δf and the averaging time τ as

$$\delta T / T \approx (\Delta f\ \tau)^{-1/2} \ .$$

This result applies more generally than just to temperature measurement by radiometry, and it also applies to noise-power measurements as conducted in Noise Fundamentals. If $V_{meter}(t)$ is the instantaneous voltage applied to the meter, which is traceably connected to the mean-square noise signal $< V_n^2 >$ at the source, then fluctuations in the meter output are also given by

$$\delta V_{meter} / < V_{meter} > = const \cdot (\Delta f\ \tau)^{-1/2} \ .$$

Here δV_{meter} can be taken to be the standard deviation of a sample of (independent) readings of the meter. The equivalent noise bandwidth used in the filtering chain provides the factor Δf, and the averaging time used between the squarer and the meter provides the factor τ. Finally, the constant is of order 1; its numerical value depends on

just what kind of time-averaging is used. (The TeachSpin equipment uses two successive one-pole filters, each of time constant τ, and the predicted value of the constant is about one-half.)

So if we measure noise using coverage limited by a low-pass filter of corner frequency 100 kHz, we have $\Delta f \approx 114$ kHz. If we choose a $\tau = 0.1$-s averaging time at the meter, we get

$$\delta V_{\text{meter}} / < V_{\text{meter}} > = const \cdot (114 \times 10^3 \text{ /s} \cdot 0.1 \text{ s})^{-1/2} = const \cdot 0.009 ,$$

so we expect fluctuations of order 0.9% in the meter reading. Of course we'd have to make $V_{\text{meter}}(t)$ readings at a time spacing of $\geq \tau$, or at least 0.1 s apart in time, for them to represent statistically-independent readings, in computing the fluctuation level. δV_{meter} as a standard deviation.

The dependence of this Dicke limit on Δf and τ is easily visible. Keeping τ fixed at 0.1 s, to give a set of readings with which the analog meter can 'keep up', you can try out he effect of changing from 100 kHz, to 10 kHz, to 1 kHz for the corner frequency of the low-pass filter in the high-level electronics. (Of course, when you reduce the bandwidth, you'll want to raise the gain G_2 to keep the average meter reading near 1 Volt.) What you should see is visibly *larger* fluctuations in the meter's position, since the Dicke equation predicts fluctuations, about the 1-Volt average, of order 0.9%, growing to 3% and then 9% as you reduce the bandwidth. So for the smallest *statistical* fluctuations in any noise measurement, it's always best to use the largest possible bandwidth Δf. (Of course, there may be growing *systematic* errors, such as the effects of capacitive roll-off, which accompany such a choice of larger Δf.)

A simple explanation of the reason for a Dicke limit also explains the $\tau^{-1/2}$ dependence. We know that the output of the amplifier/filter chain is limited to bandwidth Δf. It follows that the autocorrelation time of this signal is of order $1/\Delta f$. Thus the use of 100-kHz bandwidth gives a filtered noise signal with an autocorrelation time of about 10 μs. Hence there's a 'fresh value', or a statistically-independent measure of $<V_{\text{in}}(t)^2>$, available every 10 μs. If we use a $\tau = 0.1$ s averaging time, the number of statistically-independent measures of noise power we can make during that time is

$$N \approx (0.1 \text{ s of time}) / (10 \ \mu\text{s per fresh measurement}) = 10^4 .$$

The mean of all these 10^4 measurements is what the meter reveals via its average reading. But since those 10^4 individual readings are each of them random and independent, we expect fractional fluctuations of the meter reading to be of order $N^{-1/2}$. This hand-waving argument in fact gives

$$\delta V_{\text{meter}} / < V_{\text{meter}} > = 1 \, N^{-1/2} = 1 \, \{ \tau \, / (\Delta f)^{-1} \}^{-1/2} = 1 \, (\Delta f \tau)^{-1/2} ,$$

just as discussed above.

In a computer data-logging environment, it's easy to get a large sample of meter-reading voltages. For any choice of τ, it's easy to test if successive readings, taken at time spacing τ (or better, 2τ), display statistical independence (ie. absence of correlation). It's also easy to compute $<V_{meter}>$, and to form the histogram of V_{meter} readings (expected to be distributed about their mean in Gaussian fashion). The standard deviation of all the readings displayed in the histogram defines the characteristic scale of fluctuations, δV_{meter}. Then the Dicke limit can be tested empirically for its Δf and τ dependence.

The Dicke limit also imposes stiff requirements on any noise-based experiment that seeks to attain really high precision. If shot noise were to be used in search of a part-per-million measurement of e, and if all systematic effects were fully under control, this limit would ultimately require some meter reading to display

$$\delta V_{meter} / < V_{meter} > = 10^{-6} \ ,$$

and that, in turn, would require $(\Delta f \tau)^{-1/2} = 10^{-6}$, or $(\Delta f \tau) = 10^{+12}$. For a bandwidth of $\Delta f \approx 100$ kHz $= 10^5$ /s, that would require a total averaging time of $\tau = 10^7$ s, or about four months! (One method for doing this would be to set the meter-averaging switch to a 1-second time constant, and take one reading every second until 10^7 readings had been collected and averaged.) This provides another example of the desirability, at least on statistical grounds, of using the largest possible bandwidth Δf.

TEACH SPIN

Instruments Designed for Teaching

THE "CARE AND FEEDING" OF TEACHSPIN'S WOOD COMPONENTS

TeachSpin made a conscious choice to use finished hardwood in most of its instruments. The reason for this is both aesthetic and practical. Wood is not only pleasing to the eye, it is also non-magnetic, durable, dependable, and tough. It requires only a minimum amount of care. Do not use harsh chemicals on it, just clean it with a damp cloth. A light coating of paste wax on the wooden pieces every few years will also be helpful.

Wood, however, is considered a "living" material. It shrinks and expands with local environmental conditions. In a heated dry climate (your lab in winter) it may shrink, only to expand in the more humid conditions of summer. This may cause some metal screws to loosen over time. We tighten everything at the factory before shipping, but you may find it necessary to tighten the screws once or twice a year. This is not a defect in workmanship, but rather a fact of life with wood.

If you need further help, please do not hesitate to contact us. We want you to be completely satisfied with these beautiful and classic instruments.

Tri-Main Building · 2495 Main Street · Suite 409 · Buffalo, NY, 14214-2153

Phone/Fax 716-885-4701 · www.teachspin.com

02/2004-SN

TEACH SPIN

Instruments Designed for Teaching

Warranty
for
TeachSpin Instruments

**** Do not attempt to repair this instrument while it is under warranty. ****

This instrument is warranted for a period of **two (2) years** from the date it is delivered. For breakdowns due to defects in components, workmanship, or ordinary use, TeachSpin will pay for all labor and parts to repair the instrument to original working specifications.

For the first year of the warranty, TeachSpin will also pay all shipping costs. For the second year of the warranty, the shipping costs will be the responsibility of the owner.

This warranty is void under the following circumstances:

1) The instrument has been dropped, mutilated, or damaged by impact or extreme heat.
2) Repairs not authorized by TeachSpin, Inc. have been attempted.
3) The instrument has been subjected to high voltages, plugged into excess AC voltages, or otherwise electrically abused.

TeachSpin Inc. makes no expressed warranty other than the warranty set forth herein, and all implied warranties are excluded. TeachSpin, Inc.'s liability for any defective product is limited to the repair or replacement of the product at our discretion.

TeachSpin, Inc. shall not be liable for:

1) Damage to other properties caused by any defects, damages caused by inconvenience, loss of use of the product, commercial loss, or loss of teaching time.
2) Damage caused by operating the unit without regard to explicit instructions and warnings in the TeachSpin manual.
3) Malfunction of accessory instruments such as commercial stand-alone power supplies, signal generators, electronic counters, etc., which TeachSpin has supplied as a convenience for its customers. Those instruments are subject to the individual manufacturer warranties and any problems should be referred directly to those manufacturers.
4) Any other damages, whether incidental, consequential, or otherwise.

Tri-Main Center-Suite 409 2495 Main Street Buffalo, NY 14214-2153
Phone 716-885-4701 www.teachspin.com Fax: 716-836-1077