Ph 77 - Advanced Physics Laboratory

Department of Physics, California Institute of Technology

- Electronics Track -

Spectral Analysis

FFTs on Oscilloscopes

Most oscilloscopes include an FFT (Fast Fourier Transform) feature for viewing the spectral content of an input signal. To see this with your Keysight 'scope, send in a 10 kHz square wave and press the FFT button. Adjust the various settings to get a result something like that shown in Figure 1, showing the input square wave and the FFT. You can also view the FFT in dBm (logarithmic) units. Have a look at both linear and logarithmic scaling and try changing the horizontal scale to see how this affects the FFT resolution (displayed on the lower right of the screen) and the spectrum.

Exercise 1. Use a lower FFT resolution value to yield spectral peaks that are sharper than in Figure 1. Save a screenshot of the linear-amplitude spectrum in your e-notebook, and save another showing the FFT with the vertical scale in dB. Note how the higher dynamic range of logarithmic units gives a better view of the noise floor and small-amplitude features, while the linear scale gives you a better feel for the relative importance of the different peaks. Use the cursors to measure the frequencies and amplitudes of the first several peaks and compare with expectations from the Fourier decomposition of a square-wave signal. Try a pulse waveform and see how changing the Duty Cycle changes the relative amplitudes of the harmonics. At what duty cycle can you eliminate the 3^{rd},6^{th},9^{th}… harmonics? Include a screenshot with these harmonics missing, using the dB scale.

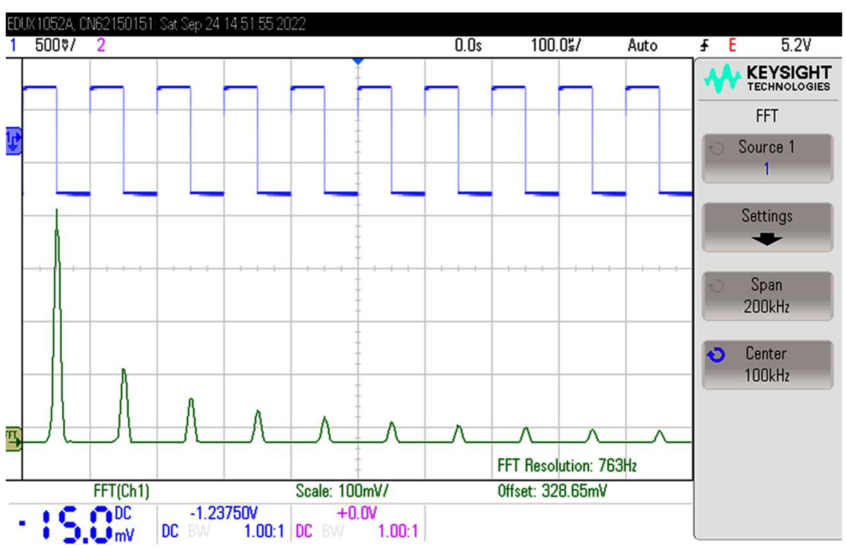


Figure 1. A 10kHz square-wave signal (top trace) and the FFT of the square wave (bottom trace). Here the FFT is shown from 0 to 200 kHz with the amplitude on a linear scale.

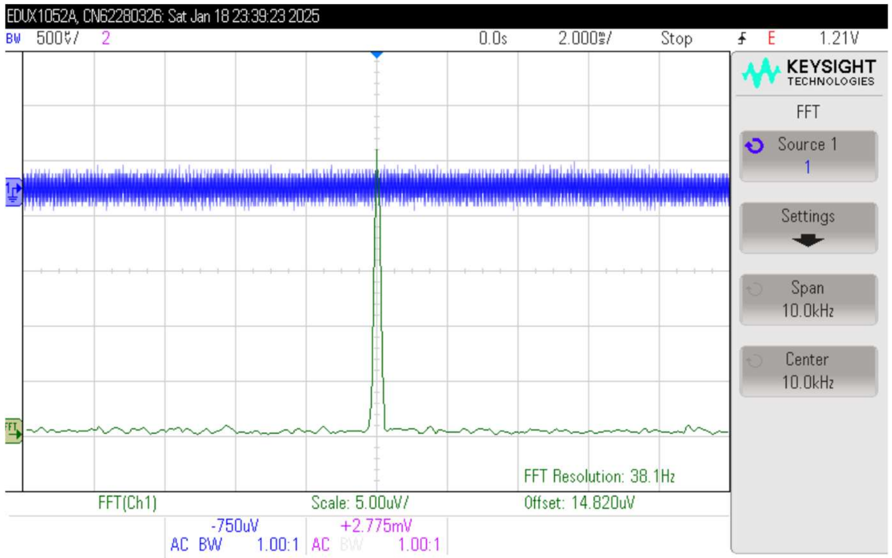


Figure 2. A small-amplitude sine-wave signal looks like noise on the 'scope screen (top trace). But the FFT (bottom trace) shows a clear 10kHz peak that reveals the presence of the buried signal. Here the FFT is set to display a linear amplitude scale.

Exercise 2. Spectral analysis is often useful when looking for small periodic signals buried in noise. To see this in practice, send a 10 kHz sine-wave signal to your oscilloscope with a signal amplitude of 10 μV pk-pk (using a 10,000x attenuator at the 'scope input). You should be able to see this small signal on the 'scope, provided you use external triggering.

Turn on the FFT and zoom in on the peak at 10kHz. (Pro tip: If the FFT walks off the screen when you increase the FFT gain, center the trace vertically first.) Play around with the adjustments until you get a screen that looks something like that in Figure 2. Note the whopping big peak at 10kHz, even though the raw signal was barely visible on ch1. Save a screenshot of your result in your e-notebook.

What this demonstrates is that spectral analysis provides a powerful tool for finding weak periodic signals in noisy data. Even though the signal here is not phase-coherent (because the 'scope is not triggering properly), it still shows up strongly in the FFT trace, allowing you to measure the signal amplitude and frequency. Moreover, if the signal contained many frequency components, you could easily find and measure all the various spectral peaks, provided they are not overlapping in frequency space.

A further extension of this technique is to acquire the FFT signal as you have done and then average many FFT traces together to improve the signal-to-noise ratio. The Keysight EDUX1052A does not include this FFT-averaging feature, unfortunately, but it can be found in many higher-end oscilloscopes.

Exercise 3. Quantify these observations further by comparing the MOOS (Minimum Obviously Observable Signal) for the time-series signal V(t) and the FFT spectrum. In both cases, use a 10kHz sine-wave signal and express the MOOS values as Vpp at the 'scope input. Do not average traces.

For V(t), optimize the settings to view the smallest V(t) you can clearly see on Ch1 without signal averaging, and document your results in your e-notebook, including a screenshot of the MOOS V(t) signal.

For the FFT MOOS, optimize the settings again to view the smallest spectral peak you can see in the FFT. [If you see a peak even with the signal generator signal turned off (try it), then turn off the Sync signal in the signal generator. In some circumstances, the strong Sync signal can affect the 'scope input.] Once again document your MOOS measurement along with a screenshot in your e-notebook.

You should see that the FFT has a much lower MOOS, in part because the signal is confined to a narrow peak. In the time domain, the signal is spread over the entire screen, making it harder to recognize. If your only question is whether a small 10kHz signal is clearly present or not, then the FFT signal is better able to answer that question.

If you buy a sufficiently expensive oscilloscope, then its V(t) and FFT capabilities may be sufficient for all your signal-analysis needs. However, most labs have separate oscilloscopes and spectrum analyzers, as the two instruments require somewhat different optimizations. It is also cheaper to buy separate instruments, and the user interfaces are more intuitive when one instrument does not have to serve many purposes. Our next task, therefore, is to introduce the Spectrum Analyzer (SA).

The Siglent SSA 3021X Plus Spectrum Analyzer

This general-purpose instrument is designed to display electronic spectra from 9kHz to 2.1GHz with a frequency resolution (a.k.a. the resolution bandwidth, RBW) down to 1 Hz. While an oscilloscope is optimized to display electronic signals as a function of time with many measurement capabilities, a spectrum analyzer like the 3021X is designed to display spectra with a host of complementary measurement features. The instrument incorporates fast parallel-processing chips so FFTs can be quickly calculated and averaged to reduce noise. Spectrum analyzers are especially useful in telecommunications, so 2GHz is a low-end model – 10x faster than a low-end oscilloscope. At very high frequencies, digital is not fast enough and spectra are generated using analog circuits. Note also that spectrum analyzers typically have a 50Ω input impedance (necessary when working at GHz frequencies), unlike the 1MΩ input impedance of a typical oscilloscope. To keep things working together smoothly, many oscilloscopes allow you to set the input impedance to either 50Ω or 1MΩ.

To get started with the 3021X, set your signal generator so it has an output impedance of 50Ω (using Utility/Ch1/50Ω) and produce a 100 kHz square wave signal at 10 mVrms, and send this to the RF Input on the 3021X. Now try adjusting these settings to become familiar with what they do:

Auto Tune: Like the name implies, pressing this button should find the dominant signal (in this case, the largest harmonic at 100 kHz) and display it at the center of the screen… nice and simple. If it doesn't work, check your signal generator settings. If you get lost as you start pushing buttons in the next paragraphs, you can always hit the Auto Tune button to start over.

Preset. If you really get lost, this button will send you back to the factory preset settings. The 3021X is also set to wake up with these settings.

Span: The Zoom In and Zoom Out buttons are especially handy here, as these change the span while leaving the Center Frequency fixed. Note that the Start, Center, and Stop frequencies are always displayed at the bottom of the screen. Try zooming out to see more of the square-wave harmonics. As you zoom out, you will see the center frequency change at some point; this is because the start frequency cannot go below zero Hz.

Frequency: Here you can set the Start, Center, and Stop frequencies manually to whatever values you like. Set Start to 0Hz and Stop to 1MHz and you should see the 100 kHz primary frequency along with harmonics at 300 kHz, 500 kHz, etc. Standard spectrum stuff.

BW: Press this button and you can set the Resolution BandWidth (RBW) using the large knob. Try this, and note that you cannot set the RBW to arbitrary values – only values on the 1-3-10 sequence. As you change the RBW, note that several things happen:

1) The widths of the peaks change, reflecting the changing resolution of the spectrum. A smaller RBW means narrower peaks.
2) It takes longer to calculate a spectrum when the RBW is small. This just reflects the fact that there are more points to calculate, which takes time. When the RBW is very small, you can see a small red cross on the screen that shows you how the calculations are progressing. If nothing seems to be happening at all, you might fix this by increasing the RBW to shorten the scan time.
3) The noise floor changes, going down as RBW decreases. This happens because the noise is related to the measurement bandwidth – a lower bandwidth means less noise. (More about that later.)

Amplitude: The most used menu item here is the Reference Level, which helps you center the signal on the screen (try it, using the wheel or entering a number on the keypad). The Scale/Division is another useful display tool. (Best to leave the Attenuator on Auto and leave the Preamp off for now.) We will talk about units below.

Marker: Press this and you see a green dot that follows your spectrum around when you turn the big wheel. And the signal value at the marker position is displayed on the screen. Handy.

Peak: When you press this button, the marker goes to the position of the highest peak (assuming it can find one easily). Try this when observing the spectrum from 0-1 MHz. Next Peak gives you the next highest peak, and the Left and Right buttons do as you might expect. To start over, go to the Peak Peak, which sets the marker on the highest peak. Another nice feature is Peak→CF, which sets the center frequency to whatever peak you are on. This makes it easy to zoom in on a particular peak.

Trace: The names here could be more descriptive, but there seems to be some historical precedent for this terminology. So…

Clear White = display a standard trace that updates as quickly as possible
View = stop updating the trace, so it remains fixed for making measurements
Blank = stop displaying the trace entirely
Average = average N traces

You can also display up to four traces on the screen using the Trace menu. Start by selecting a trace letter, observing a signal, and hitting View to stop updating that trace. Then select another trace letter, observe a different signal, etc. You can blank the traces you don't want to display. This feature is convenient for comparing related signals – for example, a trace with the signal on along with another trace showing the signal-off background.

Be sure to try out the Average feature to see how it cleans up your signal. There is also a counter on the left side of the screen that shows you how many traces have been averaged to make the current display. If the RBW is small, it can take a while to average a lot of traces.

Spectrum analyzers tend to have lots of advanced display and measurement features, but the basics described above will get you far. If you sit down in front of most modern spectrum analyzers, these basic features will probably look much like they do on the 3021X. We will talk about some additional features as they become needed.

Learning how to use a spectrum analyzer is like learning any instrument – it takes time to become proficient. As you use it more, you start to remember where the features are and what they do, and soon it becomes as easy as driving a car. Our goal in this class is not to make you an expert in spectral analysis, but just to give you some familiarity with the tools of the trade. That way, when you first walk into a physics research lab, you will at least recognize some of the hardware and be comfortable working with it. (If you are a theory person, then you can at least use this knowledge to impress your experimentalist friends.)

Spectral Analysis Units

If you are going to measure anything other than frequencies using a spectrum analyzer (SA), then you are going to need some amplitude units. While the vertical scale is mainly volts on oscilloscopes, the coin of the realm on spectrum analyzers is usually dBm. The definition is

$$x(\text{in dBm}) = 10\log_{10}\left(\frac{P}{1\text{ mW}}\right) \tag{1}$$

where P is the power in the signal. More accurately, P is the power in one RBW. If you are looking at a strong sine-wave signal, then the peak signal will just be the power in the sine wave. If you are looking at noise, then bandwidth matters. Typically the power depends on the load being driven, and this is the 50Ω input impedance of the 3021X.

To see this in action, set up your signal generator to produce a 100kHz sine wave with an amplitude of 10 mVrms, and make sure the output impedance is set to 50Ω (Utility/Output/Load/50Ω). Double-check the amplitude before you proceed, as changing the output impedance of the signal generator will change the output amplitude setting. (Truly, there are far too many pesky little details to contend with when doing science … or anything else.) Send this signal to the SA, which has a matching 50Ω input impedance.

With this input signal, the power entering the spectrum analyzer is $V^2/R = 2\ \mu\text{W}$, and plugging this into the above gives $x = -27\text{ dBm}$. View this signal on the 3021X and use the Peak function to measure the peak height. You should get the same number. Note also that the peak height does not change if you change the RBW or other settings.

If you go to the Amplitude menu, you can view some other units you have available. Most of these are not much used, but check out the Volts and Watts units. Select these and the peak height should agree with the numbers above.

If you increase your sine wave amplitude to 100 mVrms, you will see the peak height increase by 20dB. This is because dBm is a power measurement. Increasing the voltage by 10x means the power goes up by 100x, and 100x is 20dB. If you want to work in terms of input voltage, then the formula above becomes

$$x(\text{in dBm}) = 20\log_{10}\left(\frac{V}{224\text{ mV}}\right) \tag{2}$$

where V is the input voltage (assuming a 50Ω load). This is something you just have to remember – dBm and other "dB" units usually refer to power, not voltage. (There are plenty of online calculators to do the conversions for you, if you wish.)

Exercise 4. Set up to view a 100kHz sine wave with 100mVrms on the oscilloscope, and press Mod (on the signal generator) with the parameters: Type = AM, Source = Internal, Depth = 100%, AM Freq = 5 kHz, and Shape = Triangle. Use external triggering on the 'scope, and set the signal generator to produce a sync signal appropriate for AM modulation (Utility/Sync/Type=Mod-Ch1). You should see a stably triggered signal on the 'scope, showing a sinusoidal waveform with a triangular amplitude modulation. (If you are not sure, ask someone.) Take a screenshot of the signal for your e-notebook.

Then send it to the 3021X and reproduce the result in Figure 3, again for your e-notebook. (Note that most of the important SA settings are visible in Figure 3. If you want these settings to be visible in your e-notebook also, make sure the screenshots are large enough to make out these details. In Word, for example, using default settings, the image resolution will be reduced if the image is not made large on the page.)

This exercise just gives you some practice using the SA, while also demonstrating how powerful this instrument is for analyzing signals. As you can see, the oscilloscope and spectrum analyzer provide somewhat complementary information when observing electronic signals, so both are useful instruments to have in the lab. We will look more at signal modulation below.

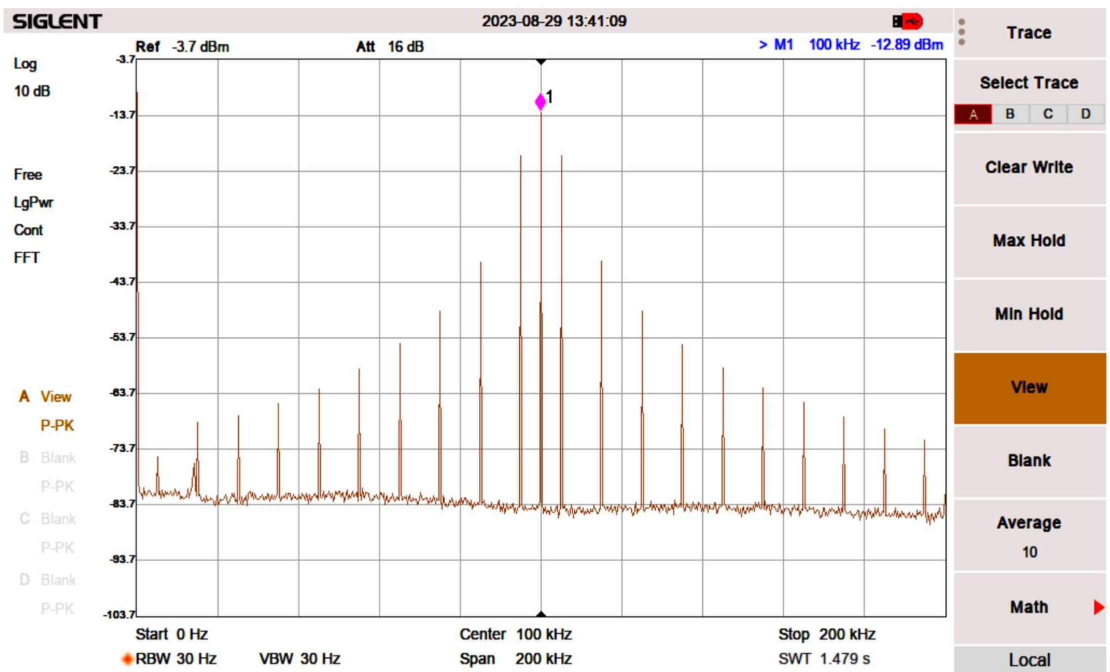


Figure 3. Spectrum of a 100kHz sine-wave signal including amplitude modulation, showing a complex set of sidebands flanking the 100kHz carrier.

Transfer Functions

Another highly useful feature of the 3021X and other spectrum analyzers is to measure the "transfer function" of an electronic device. This is defined as the ratio Vout/Vin (or Pout/Pin for power) as a function of frequency. A basic spectrum analyzer like the 3021X will display amplitude ratios only, while high-end models will provide phase information as well.

For example, try setting up the simple set of cables shown in Figure 5, using one of the red coils of BNC cable in the lab. Press the Preset button on the 3021X to get back to factory settings (and

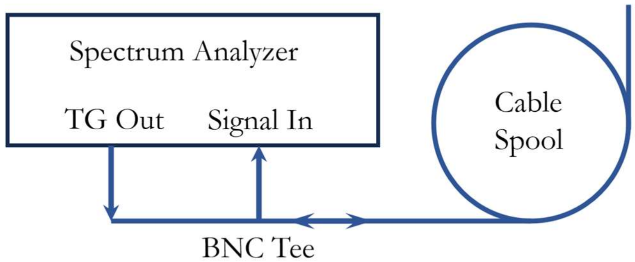


Figure 5. A connection diagram for observing the spectral signature of signal reflections from a long length of unterminated BNC cable. There are two red coils of cable in the lab, each coil about 30cm in diameter, containing 100ft of cable.

then do not press Auto Tune, as the instrument can become confused if there is not strong input signal to lock onto). Configure the SA to 0-50MHz, RBW=100kHz, and you should just see noise.

Now press TG (Tracking Generator), set TG to on, and you should see a spectrum something like that shown in Figure 4. If you remove the long BNC cable from the circuit, the periodic signal goes away. What's happening here is that the TG sends out a sinusoidal signal, and part of it goes into the long BNC cable. That signal reflects from the end of the cable and then comes back to the BNC T, where it interferes with the direct TG signal at that point. At some frequencies, you get constructive interference, at other frequencies you get destructive interference. The spectrum thus tells you something about the travel-time of the signal in the long cable. If you add a 50Ω "terminator" to the far end of the cable, this resistor absorbs the signal and prevents reflections. If you add a 75Ω terminator instead, then the terminator partially absorbs the signal, so some reflects back into the spectrum analyzer. We will discuss this topic later in this Track when we look at transmission lines in more detail.

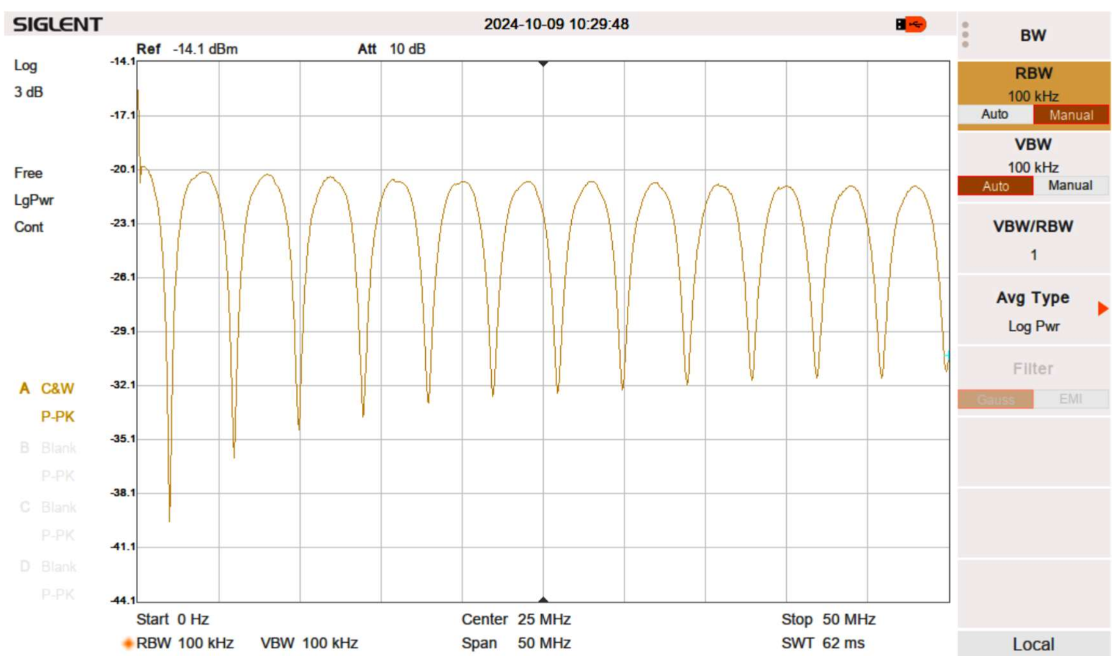


Figure 4. A measured transfer function for the setup shown in Figure 5. The periodic structure is caused by signal reflections in the long BNC cable.

Exercise 5. Reproduce the screenshot seen in Figure 4, but include a second trace on the same screen showing the signal with a 75Ω terminator at the end of the cable, and add a third with a 50Ω terminator. What happens when you replace the spool with a shorter open-ended cable (Hint: extend your spectrum to look at higher frequencies)? Add another screenshot describing what you observe.

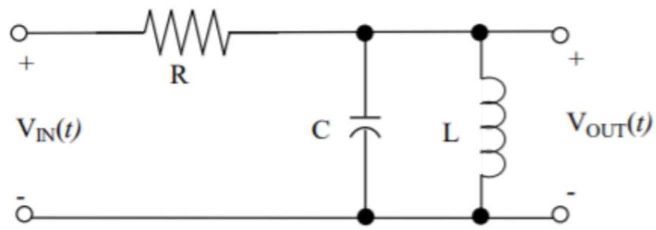


Figure 6. A schematic diagram of the 250kHz bandbass filter found in Box66 in the lab. The specified component values are 1kΩ±1%, 4.7nF±10%, and 100μH±10%.

Exercise 6. Set up a measurement of the transfer function of a 250kHz bandpass filter (Box 66) that has the circuit diagram shown in Figure 6. Note that this filter cannot drive the 50Ω input impedance of the 3021X – the 50Ω load would basically just ground the filter output, negating the intended operation of the filter. You can get around this by "buffering" the filter output using an SR560, as shown in the connection diagram Figure 7. The SR560 has a whopping 100MΩ input impedance, which does not disturb the filter, while the SR560's 50Ω output has enough power to drive the 3021X's 50Ω input. Adjust the various settings to obtain a nice-looking signal, then pause it (Trace/View), record a screenshot like that shown in Figure 8, and add it to your e-notebook. (For the analysis that follows, try to produce a screenshot that looks nearly identical to Figure 8.) Using the same paused trace, download the numbers to a .csv file and plot the data on your computer.

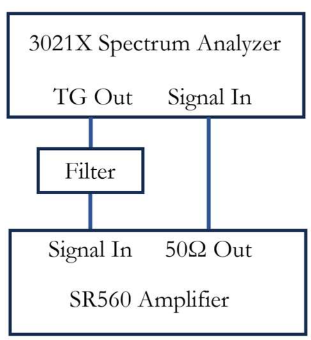


Figure 7. A wiring diagram showing how to "buffer" a high-impedance output (from the bandbass filter) so you can send it into a low-impedance input (in the 3021X).

Exercise 7. Show that the expected transfer function of this filter is

$$\left|\frac{V_{out}}{V_{in}}\right| = \sqrt{\frac{1}{\left(\frac{R}{\omega L}\right)^2 (1-\omega^2 LC)^2 + 1}} \tag{3}$$

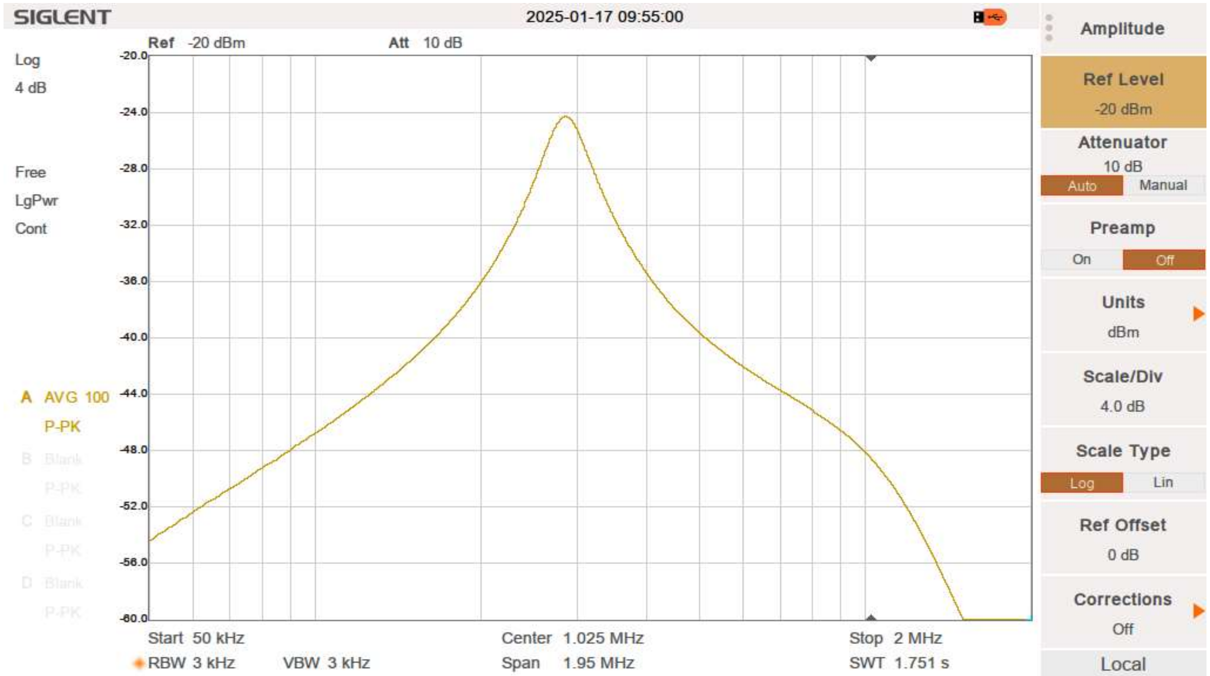
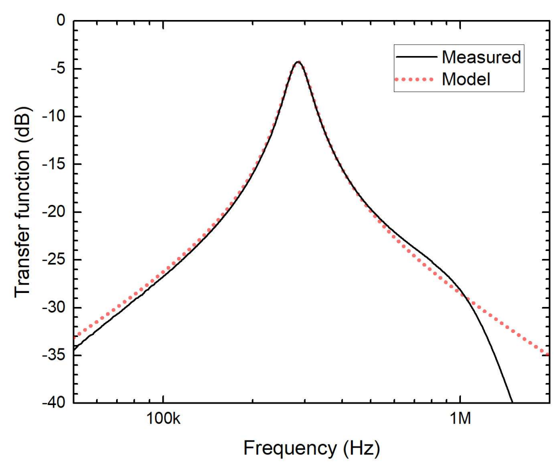
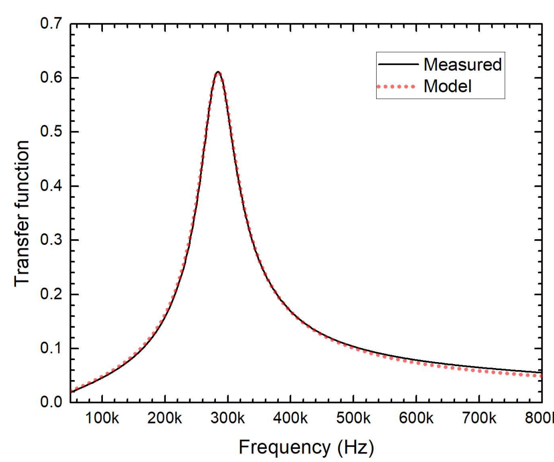


Figure 8. A measured transfer function for the 250kHz bandpass filter shown in Figure 7.

and add a theory line on your data plot. (Remember that your SA output is $x(dBm) = 20\log_{10}(|V_{out}/V_{in}|)$, plus a normalization constant.) Adjust the model parameters and see if you can produce a model that fits your measurements reasonably well. The image on the right shows what you might get, if you work at the model a bit. Document your work in your e-notebook, include your plots, and report on your best-fit values for R, C, and L.





If your model has a poor fit overall, that could be because you are using some kind of data-fitting algorithm built into Mathematica, Python, etc. Remember that these algorithms are designed to fit simple polynomial lines, not complex nonlinear functions. When faced with the latter, you may find that your fitting algorithm is dumb as a rock.

A better first approach is to try a "chi-by-eye" fit, where you simply define a fitting function and enter the parameters by hand. Then you look at the plot and adjust the parameters to get satisfactory results. That is what was done for the plots on the right. Because you are smarter than a rock, you can quickly home in on the correct parameters, or something close.

You can see that the data and model diverge badly above 1 MHz. This is because the SR560 is only designed to amplify signals below 1 MHz, effectively sending the signal through a low-pass filter with a 1MHz cutoff. In addition, the 3021X is only rated down to 9kHz, and this causes some rolloff at frequencies below 20kHz.

Even at lower frequencies, it is not possible to find a model that perfectly matches the data. This is because the model assumes ideal electrical components, which is unrealistic. Real inductors are especially problematic in this regard, as inductors invariably include some resistance and capacitive effects. The model, therefore, is simply not sophisticated enough to reproduce every nuance of the measurements. As you can see in the linear plot, however, the model can reproduce the behavior around the peak reasonably well.

In the world of experimental physics, you often find that real devices have a variety of quirks, limitations, and perhaps unspecified defects. And if you want to work at the cutting edge in some field, these problems become important. The struggle is real.

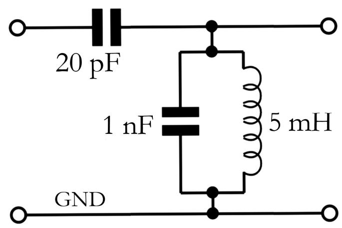


Figure 9. A schematic diagram of the 250kHz bandbass filter found in Box73 in the lab. The specified component values are all ±10%.

Exercise 8. Find Box73 in the lab, which is another bandpass filter using the circuit shown in Figure 9. Measure its transfer function, again buffering using the SR560, and add it to your e-notebook. Here you will find that the RBW is an important parameter in your measurement. Figure 10 shows a typical result, here with the vertical axis in linear (not dB) units. What do you measure for the Q ($= \nu_0/\Delta\nu_{FWHM}$) of this filter?

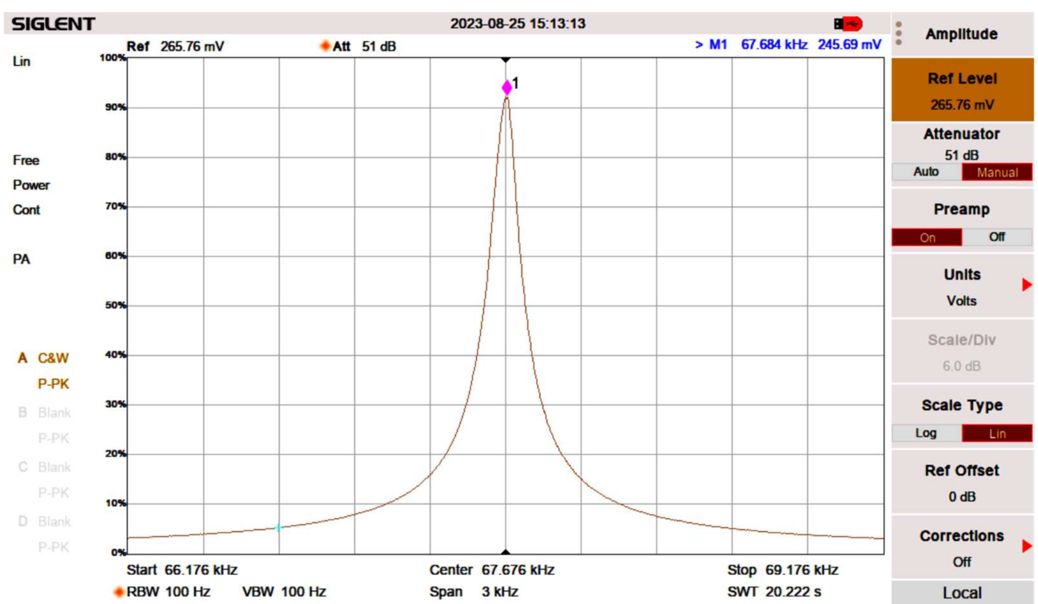


Figure 10. A measured transfer function for the high-Q bandpass filter shown in Figure 9.

Amusing tech note: It is difficult to build any kind of electronic oscillator with a Q above 1000, while Qs of a million or more are common in well-crafted mechanical systems. This is why portable stand-alone clocks (in wristwatches, smartphones, spectrum analyzers, oscilloscopes, computers, etc.) are usually based on quartz crystal oscillators. These are among the last vacuum tubes left in consumer electronic devices. The cavity magnetron powering your microwave oven is another vacuum tube. The vacuum envelopes are metal in these devices; glass vacuum tubes are uncommon.

Frequency Modulation

As soon as scientists and engineers realized they could transmit and receive radio signals over large distances, the concepts of amplitude modulation (AM) and frequency modulation (FM) became central in exploiting this new technology. In the modern world, AM and FM radio stations have largely been replaced by internet streaming, but the same concepts are now being applied in the optical regime, realizing a host of new laser applications. Modulating a 400 THz laser is certainly different compared to a 100 MHz FM radio station, but the underlying mathematical principles are generally frequency independent.

Of course, this subject requires a lot of study to fully understand, but it is quite useful to develop a practical intuition for the phenomenology simply by looking at some signals on the spectrum analyzer. To this end, set up a 100mVrms sine wave signal with a frequency of 10 MHz and display it on the 3021X. Set RBW = 1 kHz and you should see a nice peak on the screen.

Next press Mod on the signal generator and add some frequency modulation using: Type = FM, (Source = Internal), FMdeviation = 100 Hz, FMFreq = 10 kHz, and Shape = Sine. Center the 10MHz carrier on the SA and zoom in until you see the 10 kHz sidebands on the spectrum. We will just be looking at FM modulation for now (Type = FM). This is generally what happens when you slightly modulate an RF carrier frequency at 10 kHz … you get some 10 kHz sidebands.

We can write the signal with added frequency modulation as

$$V(t) = A\sin\left[2\pi f_0 t + \frac{\Delta f}{f_m}\sin(2\pi f_m t)\right]$$

where f_0 is the carrier frequency (10MHz), f_m is the modulation frequency (10kHz), and Δf is the maximum frequency deviation (100Hz). It is customary to also define the modulation index

$$h = \frac{\Delta f}{f_m}$$

When $h \ll 1$, as we have with the current settings, then the FM spectrum shows the carrier with two sidebands at $f_0 \pm f_m$. Press the Peak button and verify that these sidebands are where they should be. If you lower h by lowering Δf, then observe that the sideband amplitudes go down while the sideband frequencies remain unchanged.

Now turn Δf higher and see what happens. The sideband frequencies are still multiples of f_m from the carrier, but the number of sidebands goes up as Δf goes up. If you switch the SA to a linear scale (in amplitude), you will see that the number of significant sidebands on each side of the carrier is roughly equal to h. Try Δf equal to 100 kHz and you can count the number of sidebands. At 1 MHz the sidebands start to run together, so zoom out to see the envelope. Set Δf to 4 MHz $(h = 40)$ and you see a plateau that runs from (10-4)= 6 MHz to (10+4)= 14 MHz, as illustrated in Figure 11.

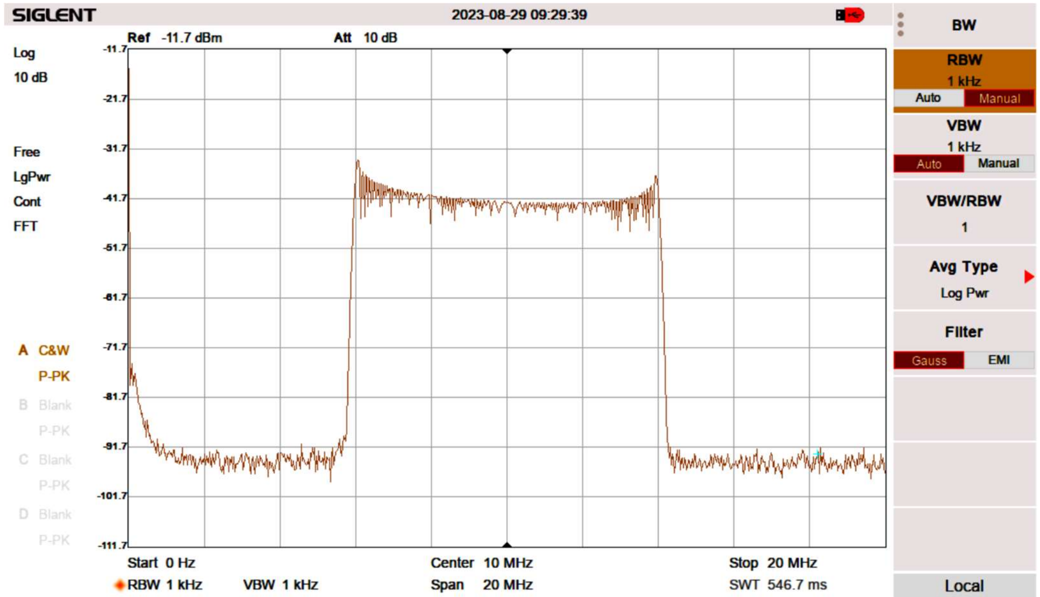


Figure 11. A spectrum resulting from frequency modulating a 10 MHz carrier with FMFreq = 10 kHz and Diviat = 4 MHz.

Exercise 9. As usual, get some practice working with the hardware by reproducing Figure 11 and adding it to your e-notebook. At this point h is very large, and what you see is mostly the carrier signal being essentially swept from $f_0 - \Delta f$ to $f_0 + \Delta f$. When you move a frequency back and forth sinusoidally like this, your swing spends more time at the endpoints than in the middle. Thus, the broad plateau is higher at the endpoints than at the middle.

Now zoom in on this signal, always staying centered at 10 MHz. Zoom in far enough and you will see the individual sidebands, still separated by f_m, as they should be. If you change Δf slowly while staying zoomed in, you can see the sideband amplitudes vary quite a bit. If you change f_m while still zoomed in, you will see the sideband spacing change accordingly. But if you zoom out again and now change f_m, you will see that the envelope does not change, as this depends only on Δf. Record some additional spectra with $h = 0.1$ and $h = 10$, and add these to your e-notebook (with suitable notes describing the various parameters, of course). For best results, adjust the SA parameters and average some traces to see the signal clearly.

If this all seems a bit confusing, that's because frequency modulation is a complicated phenomenon. Alas, doing all the math does not make it immediately clearer. Theory just gives you a lengthy series containing about h significant terms – one for each sideband. If you sum the series and plot the results, you will see what you see on the spectrum analyzer…but the spectrum analyzer does it a lot faster. Signal modulation is a staple in technology and finds many applications in experimental physics as well. If you see any lasers in a physics lab, chances are they are being modulated for a variety of reasons. Even if you take a different path and never delve deeper into this subject, it is useful to dabble into the basics here.

FM Radio

The 3021X is designed especially for RF analysis going up to 2.1 GHz, which covers a range of sub-microwave communication channels. To see what you might see permeating the air, connect a BNC-to-antenna device (located in the lab) to the SA input and look at the spectrum around 100 MHz. You can easily see a series of broad peaks separated by about 1 MHz. These are traditional over-the-air FM radio and TV channels. (Yes, OTA TV is still available, remaining popular because it requires no cable subscription.) The widths of the spectral peaks are determined by the frequency-modulated signal being transmitted. A pure carrier wave would display as a sharp peak, but the FM signal broadens the peak.

Exercise 10. Record a spectrum of the FM radio band around 100 MHz, and add it to your e-notebook. Be sure to zoom in so the signal fills the display, and lower the BW enough to see the individual peaks (each corresponding to a single radio station). Zoom in further to focus on KTWV-FM at 94.7 MHz, which is one of the stronger stations. Use a 1MHz span to show the carrier flanked by flat-topped sidebands. These sidebands contain the audio signal that you hear on the radio. Of course, FM radio stations are a bit old fashioned, but cell phones use essentially the same ideas, but with higher carrier frequencies (above the 2 GHz range of our spectrum analyzer).

Optical frequency combs

The overarching topic of signal modulation is so broad that we cannot do much more than scratch the surface in Ph77. If you spend any time working in tech-heavy fields (including experimental physics), you will find that people are modulating signals for all sorts of reasons, in frequency ranges from kHz to THz.

To focus on just one modern example, we next consider *optical frequency combs*, for which John Hall and Theodor Hänsch received a Nobel Prize in 2005. The basic frequency-modulation concepts are the same as described above, except the carrier frequency is shifted from a few MHz or GHz to the optical regime around 400 THz. And, using this correspondence, we can demonstrate the essential ideas behind optical frequency combs using just the SSA 3021X and a signal generator.

Exercise 11. Once again, follow through the steps here and record your progress in your e-notebook. Start with a 30mVrms sine-wave signal with a frequency of $f_0 \approx 5$ MHz, created on *channel 2* of your spectrum analyzer. Make the actual frequency some random number, like 5.123123 MHz. Write this number down in your e-notebook and view the signal on the spectrum analyzer. Now you have to imagine that this electromagnetic signal is coming from a laser at around 400 THz, and you want to measure its frequency with great precision. There are no electronics technologies that can measure THz frequencies directly, so pretend for a moment that you don't actually know what f_0 is, and your goal is to measure it using a frequency comb.

To accomplish this, you first send your laser through a frequency-doubling crystal (look it up), and you can simulate this by generating a second signal on channel 1 of the signal generator at double your chosen frequency (e.g., 10.246246 MHz) and 100mVrms. Of course, again you do not know the frequency of this new laser beam, but you do know it is exactly twice the original laser frequency, because that's how frequency doubling works. Add a BNC Tee to the SA input and combine the two signals there. Verify that there are two peaks on the screen, at f_0 and $2f_0$.

Next, add a bunch of FM sidebands at $f_m = 20\ \text{kHz}$ to the channel 1 signal to get something that looks like you see in Figure 12. Note that the sidebands added to the $2f_0$ signal overlap the lone f_0 peak. In this context, those sidebands are called a frequency comb. Reproduce this screenshot for your e-notebook. Zoom in on the main peak at f_0 until you see the forest of sidebands around it. As you zoom in, decrease the RBW so the peaks get sharper. If you change f_m you can see the sidebands move in relation to f_0. Keep zooming in until you see the f_0 peak and just 1-2 sidebands around it, with RBW = 10Hz. Next adjust f_m until one of the sidebands aligns with the f_0 peak. Make this alignment as accurate as you can. This way you can find that the peaks have optimal overlap when f_0

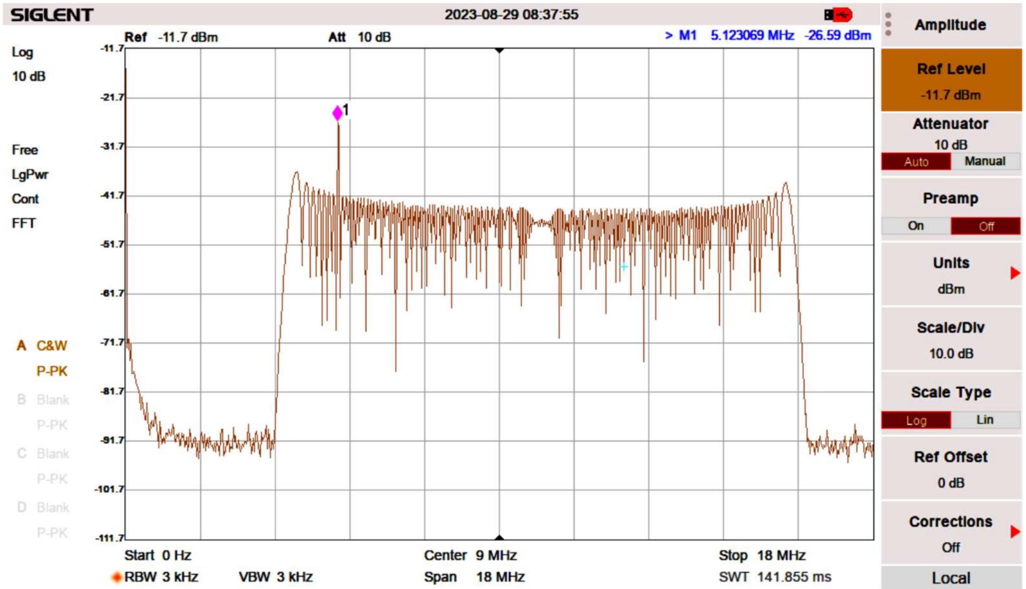


Figure 12. This spectrum shows two combined signals: a 5MHz sine wave signal (the tallest peak, marked) together with a strongly frequency-modulated 10MHz signal.

equals some number 20.abcd kHz, with all four .abcd digits being meaningful (the last digit being a bit of an estimate). Write this number down.

Your next step would be to count how many sidebands you have spanning the distance from f_0 to $2f_0$. In the optical regime, you might determine this number by dispersing the modulated laser using a diffraction grating, revealing a series of resolved optical peaks. You would not need any frequency information to make this count – it is just an integer number of sidebands. Figure 13 shows a real-world example of what this might look like, and you can see that making the sideband count is straightforward.

For this current exercise, you could work your way through the spectra on the 3021X and count the sidebands that way. But that would be tedious and not very educational. So, you can cheat a bit at this point and calculate the number of sidebands as $N \approx f_0/f_m$ (the cheat being that you know f_0 from your signal generator.) If you calculate this ratio, you should find that it indeed gives you a specific integer, to an accuracy of a few percent. (For example, if your calculated N is 251.03, then clearly the correct integer is 251.) If you had laboriously counted the sidebands by some other method (because you did not know f_0 *a priori*), then you would certainly have gotten the same N. Record that calculated N as well. Round it to the nearest integer, and this gives you an independent measurement of f_0, namely $f_0 = Nf_m$. If you did the measurement carefully, then your measured $f_0 = Nf_m$ should equal the preset value to better than 100ppm.

The main take-away message from this exercise is that you can use a frequency comb to measure high-frequencies (5 MHz in this example) using only low-frequency electronics (adding sidebands at 20 kHz in this example). And that is the magic of frequency combs – they allow you to connect high and low frequency signals, so you can use low-frequency signals to precisely measure frequencies that are N times higher. The only limitation is how many sidebands you can add to a carrier signal. (There are advanced techniques that allow you to precisely overlap the spectral peaks, but we will skip over that detail for now.)

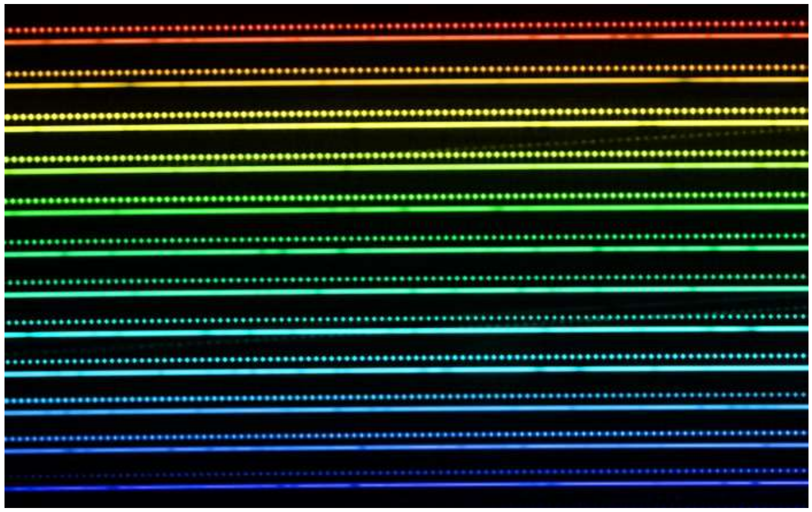


Figure 13. This image shows a grating-dispersed optical spectrum from the HARPS instrument on the ESO 3.6-metre telescope at the La Silla Observatory in Chile. The colors are real, as the spectrum goes from red to blue. The top spectrum shows a calibrated laser frequency comb used as a wavelength/frequency ruler. The lower spectrum is from a star with (weak) absorption lines. This is just one scientific application of optical frequency combs.

In the optical regime, sidebands are added using electro-optic modulators, and these nonlinear crystals can only be driven so hard. If you can make 1000 sidebands, then you can connect a 400 THz laser to a 400 GHz microwave source. From there you add additional steps to build a ladder of measurements, eventually connecting the 400 THz laser to the 9.192,631,770 MHz Cesium hyperfine transition that we use to define the second. And this, dear reader, is how you can precisely measure optical laser frequencies.

Of course, once you have all this optical/microwave/RF metrology set up at a few labs around the world, you can use it to calibrate specific atomic transitions at many wavelengths, and these can be used as secondary standards to measure whatever optical frequencies you need to measure. Having good metrology is important in engineering, but it is especially so in physics, where we are obsessed with making ultra-precise measurements.

Optical-frequency-comb technology is still relatively new, and applications are appearing in LIDAR, spectroscopy, and many other fields. FM radio has mostly been replaced by internet streaming, but FM lasers are all the rage these days as diode-laser technology becomes cheaper and easier to use. Hall and Hänsch received their Nobel Prize for two reasons: 1) people had been trying for decades to figure out a way to accurately measure laser frequencies; and 2) frequency-comb techniques have quickly found all sorts of applications in other fields of science and technology.

Noise Spectra

Getting back to our core topic of general spectral analysis, one of the most important uses of spectrum analyzers is to measure and study the noise in electronic circuits and the physical systems attached to those circuits. This feature is especially important in experimental physics, as noise is a central issue whenever you set out to measure something well. In your smartphone, for example, you likely have a barometer that can measure pressure to about one part in 10^4, thus giving you a handy altimeter with a vertical resolution of about 30cm. Somebody had to think hard about fundamental noise limits to make that happen!

To the newcomer, noise analysis can be quite a boring subject, because you want to focus on the signal, not the noise. But consider this: laser-interferometer gravitational-wave detection (including the LIGO project) started to ramp up around 1980, when people started building prototype interferometers. There were no detections then, only measured noise floors, which were terrible. So people studied the noise, improved the noise floor, and making models of the physical processes causing the noise. For decades, the field mainly studied interferometer noise in all its forms, slowly improving the detection limits. After a great deal of effort by thousands of researchers… lo and behold, a small gravity-wave signal was spotted above the noise in 2015, revealing a pair of coalescing black holes. The field of observational gravity-wave astronomy is now thriving, but it took 35 years of careful noise analysis to make that happen.

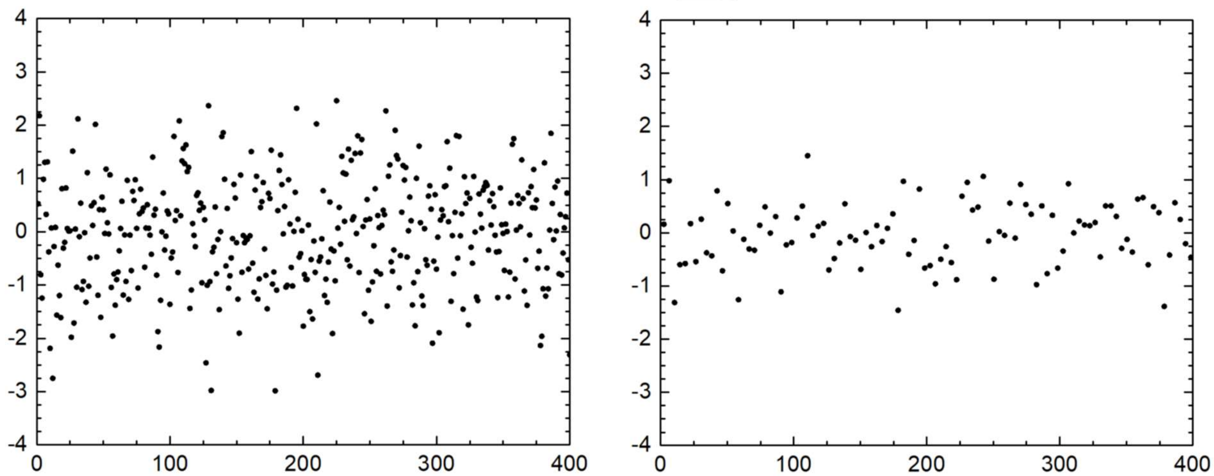


Figure 14. (Left) A series of data points showing Gaussian noise. (Right) The same data after averaging sets of four points. This illustrates why reducing the bandwidth by a factor F reduces the noise by a factor of $\sqrt{F}$. Because of this, voltage noise is often described in nV/rtHz, pronounced "nano-volts per root Hertz."

Noise units

The first thing you have to understand about noise is that what you measure depends on the *bandwidth* of your measurement. To see why this is, consider the example in Figure 14. The left panel shows a timeseries of points depicting normal Gaussian noise. The points are sampled once every time interval Δt, so the bandwidth of the measurement is simply $B = 1/\Delta t$. In this context, a bandwidth of B means that there cannot be any signal information at frequencies with $\nu > B$. Each measurement point averages the signal for a time Δt, so any high-frequency information with $\nu > B$ would be averaged to zero.

The second panel in Figure 14 shows the same data, except this time each individual point is the average of four adjacent points in the first panel. (Thus 400 points in the left panel becomes 100 points in the right panel.) In the second panel the bandwidth is $B_{ave} = 1/4\Delta t$ (because 4 points were averaged) and the noise level has been reduced by a factor of 2 (because averaging N points reduces the noise by a factor of $\sqrt{N}$).

Putting this another way, we can calculate the RMS noise in Figure 14 as the root-mean-square of the string of values, giving

$$\sigma_V = V_{RMS} = \sqrt{\frac{1}{N}\sum V_i^2}$$

If you calculate this for both panels in Figure 14, it quickly becomes apparent that the RMS noise σ_V is proportional to $\sqrt{B}$. This fact about noise may not be immediately intuitive, but you soon get used to the terminology. If you sample a voltage very rapidly, the noise is high. But if you sample slowly (running the signal through low-pass filter first, thus reducing the bandwidth), the noise is lower.

So the units for voltage noise are often nV/rtHz (or $nV/\sqrt{Hz}$) pronounced "nanovolts per root Hertz". However, if you look at the noise spectrum using a spectrum analyzer, then you have to remember that dBm is a unit of power, and power goes like V^2. Thus noise power is measured in dBm/Hz, and the noise floor will be proportional to the RBW. In either case, the noise spectrum is often called the noise *power spectral density* (PSD).

Exercise 12. To get you acclimated to thinking about noise, measure the noise floor of the SR560 amplifier as a function of amplifier gain. Start by connecting a 50Ω (or 75Ω) terminator to the SR560 input, set the gain to unity, and set up the low-pass filter with a 1MHz bandwidth. Then send the output to the oscilloscope. If you turn up the oscilloscope gain, you should see some noise that is above the 'scope noise floor. Set up the 'scope further with 1) A horizontal scale of 100 μsec/div, so you will be measuring noise near 10kHz; 2) Acquire = normal; and 3) Measure the AC RMS of the signal, which is equal to σ_V described above.

When your setup looks good, measure σ_V on the 'scope as a function of the SR560 gain. Convert σ_V to nV/rtHz by dividing by $\sqrt{B}$ (where B is set by the filter bandwidth). Remember also that amplifier noise is *referred to the SR560 input*, so the input noise is just the measured output noise divided by the gain.

If you put everything together correctly, you should get results that look something like you see in Figure 15. If you get something radically different from this, check all your settings again. The exact noise levels may differ substantially (by as much as 2x) between instruments, so your results may not look exactly like Figure 15. But the overall amplitudes and trends should be there. As always, if you are having trouble, ask.

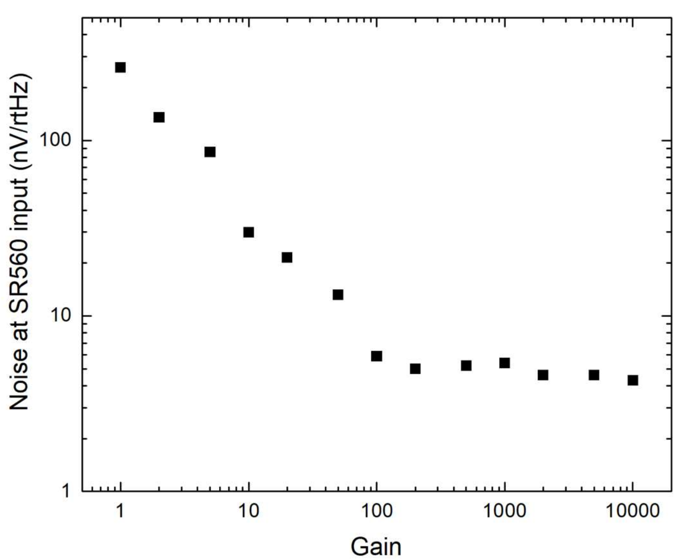


Figure 15. A measurement of the SR560 noise (referred to the input) as a function of the amplifier gain.

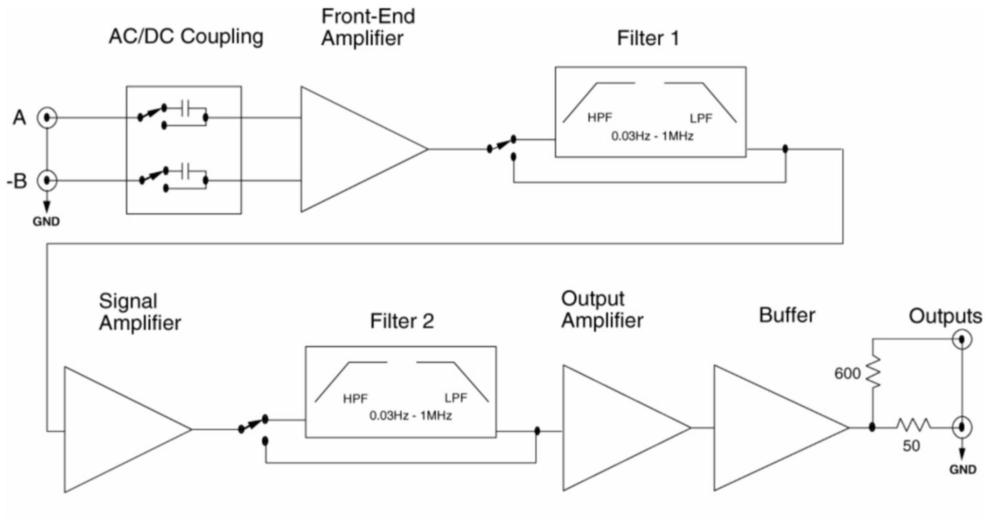


Figure 16. A simplified internal schematic of the SR560 amplifier.

Your next challenge is to interpret these results. Starting at G=1, you see that the noise floor is quite high, especially when you consider that the SR560 spec sheet says the input noise is as low as 5 nV/rtHz. To understand this, consider the SR560 internal schematic shown in Figure 16. At G=1, the instrument assumes that you have a large input signal, so the front-end amplifier will have a low gain. After going through the filters, the noise level is probably quite low. But the output amplifier and buffer add noise, and these give you a fairly high noise at the output. Thus, all you are seeing at G=1 is the noise from these last amplifier stages.

At higher G, the front-end amplifier has a corresponding higher gain, so the front-end amplifier noise matters more while the output-amplifier noise matters less (relatively speaking). When you get up to G=100, you see lots of front-end amplifier noise, so the output-amplifier noise basically doesn't matter anymore. Thus, only at gains of G=100 and above do you realize the low input noise described in the instrument specifications.

This kind of behavior is typical for general-purpose electronics test equipment. You cannot do everything for everybody all the time, so compromises are made. This is why it helps to know a bit about what's inside the box, so you understand enough to use the instrument to maximum effect. Fortunately, you can get pretty far just being an ignorant user, because the SR560 was designed with you in mind. When you have a small signal, you naturally turn up the gain, and the input noise goes down without you thinking about it. Nevertheless, if you want to get the best results when using a precision instrument, it is often beneficial to know what you are doing.

Exercise 13. Moving on, let's look at the noise floor of the 3021X. First connect a 10,000x attenuator to the 3021X input, because we want to look at small signals just above the noise. Then create a 10 MHz signal with an amplitude of 1 Vrms and connect this to the 10,000x attenuator input. Push the Preset button on the 3021X to give you the default instrument settings, then look at the spectrum from 0-20MHz with RBW=10 kHz. With these settings, you should see a noise floor of about -80 dBm at 10MHz, and the attenuated signal will not yet be visible above the noise.

The 3021X has a specified noise floor of -150 dBm/Hz at 10 MHz, which is called the Displayed Average Noise Level (DANL) in the trades. The "/Hz" means you divide by the RBW to give dBm/Hz. (In log notation, subtract 10dBm for each factor of 10 in RBW.) When you express your measurement in dBm/Hz, -80 dBm divided by 10 kHz gives -120 dBm/Hz, which is 30 dB higher than the DANL.

The reason for the discrepancy is because the instrument is not yet giving you its best performance. To get this, go to the Amplitude menu and set the input attenuator to 0dB, which removes this layer of safety from the instrument. Now you should see the 10MHz signal peak, although the noise level is still not below the DANL. If you do not see the signal, something is amiss, so try again.

For additional improvement in the noise level, turn on the Preamp (also in the Amplitude menu), which inserts a special low-noise preamp at the front end. Try averaging 100 traces to see the noise floor more clearly. Now measure the noise floor, convert to dBm/Hz, and document that the noise floor is indeed close to the specified DANL. Now the 10 MHz signal peak is far above the noise, so add a screen shot to your e-notebook.

Remember that all these measures make the instrument more susceptible to damage, so do not touch the input while the instrument is in this state. In particular, ***do not disconnect the 10,000x attenuator in this low-noise mode***, as this helps protect the instrument. Amusingly, the 3021X will never hand you its best performance by itself. You have to manually remove the input attenuation and manually turn on the preamp. Hiding these features is just the instrument protecting itself from inexperienced users.

To see how this noise floor compares with the SR560, convert your dBm/Hz to nV/rtHz using

$$nV/\sqrt{Hz} = \sqrt{10^{(dBm/Hz)/10} \cdot Z \cdot 0.001} \cdot 10^9$$

where Z is the load impedance of 50Ω. (You can derive this for yourself, but the exercise is not so interesting and soon forgotten.) How does the 3021X noise floor compare with the SR560, which has a specified input noise as low as 4 nV/rtHz? (As you saw in Figure 15, the measured SR560 input noise is perhaps a bit worse than specified.)

In the electronics world, 1 nV/rtHz is the gold standard, as it is exceedingly difficult to get below this value at room temperature. The LT1028 op-amp you saw previously is specified at slightly below 1 nV/rtHz. Electronics test equipment cannot achieve this level, however, because the input would soon be destroyed by inexperienced users. Adding some input protection increases the noise.

Exercise 14. For your next challenge, reduce the 10MHz input signal (still going through the 10,000x attenuator) to measure the MOOS, now defined as the smallest peak that can be clearly seen above the noise floor on the 3021X screen, with an averaging time of about a second. Starting with the configuration you had above, reduce the signal Vrms while zooming in on the peak, changing the **BW**, **Span**, and **Scale** to zoom in on the peak to make it clearer. If the trace averaging time is greater than a few seconds, average fewer traces. Intermediate level is achieving a MOOS of 1μVrms at the input (10mVrms at the signal generator); Expert level is 100nVrms. Document your efforts, and include a screen shot of your best MOOS signal.

When you are finished with this exercise, hit the Preset button to restore all the input safety features in the 3021X.

Resistor Thermal Noise

The reason that 1 nV/rtHz is difficult to beat is because resistors exhibit thermal noise, which is also called Johnson noise. If you hook up an ideal voltmeter to a resistor, you will see that the average voltage is zero (because the average current through the resistor is zero), but the voltage fluctuates with a PSD (at room temperature) of

$$PSD_R \approx 0.9 \sqrt{\frac{R}{50\ Ohms}}\ \ nV/\sqrt{Hz}$$

And if you look at this noise signal on an ideal oscilloscope, you will see

$$V_{RMS} = PSD * \sqrt{B}$$

where B is the measurement bandwidth.

Resistor thermal noise often sets the noise floor for any instrument operating at room temperature. Looking at op-amps from 50 years ago, the lowest achievable input noise was about 1 nV/rtHz. In modern times, the number is still about 1 nV/rtHz. You cannot build op-amps without resistors, and it is hard to go much below 50Ω for practical reasons. Thus, 1 nV/rtHz is tough to beat for low-noise op-amps at room temperature. Thermal noise is what we call a "fundamental" noise source, meaning you cannot easily avoid it.

Using the SR770 to measure resistor thermal noise

Measuring resistor noise is difficult, but the SR770 spectrum analyzer can do the job. This instrument is designed specifically for low frequencies, covering a range from zero to 100kHz. Also, it has a high input impedance (1MΩ, similar to an oscilloscope) with an exceptionally low input noise floor. These qualities make the SR770 well suited for viewing small periodic signals and examining low noise levels in precision electronic circuits.

To turn on the SR770 and reset to its default settings (a good starting point), press and hold the "←" button while powering up. You can release this button when you see *Calibrating Offset* appear on the screen. You will also see this calibration screen occasionally during normal operation; all you can do is wait until it goes away.

Have a look at the SR770 features by setting your signal generator to produce a 5 kHz square wave signal at 100 mVrms, and then send this into a 20dB attenuator (10x amplitude reduction) attached to ChA on the SR770. Play around with the settings until you are comfortable using the instrument, especially the Freq, Scale, Average, and Meas buttons. Make sure you know how to auto-scale and average traces. The controls are vaguely similar to the 3021X, although the user interface is a bit clumsier. Also, there is no way to save a screenshot (although you can always take a picture using your phone).

When you are ready to proceed, replace the square-wave signal with a 50Ω terminator from the box of BNC terminators going from 50Ω to 250kΩ. Then select these settings that will allow you to measure resistor thermal noise:

1) Press Freq to set the frequency range to 0-12.5 kHz
2) Pres Meas to measure PSD with units of Vrms.
3) Press Marker and place the marker at around 10kHz. (The exact value doesn't matter, as the noise spectrum is quite flat.)
4) Make sure you have a spectrum displayed, and note the Y measurement on the screen, showing noise PSD in nV/rtHz. The actual noise of a 50Ω resistor is below 1nV/rtHz, so the noise floor you see is much too high.
5) Press Input and set the Input Range to -60dBV. This removes the normal input attenuation and thus reduces the noise floor. Remember that this attenuation protects the instrument, so be careful what you plug into the input now.

6) Average some traces, and you should see a noise floor of about 6 nV/rtHz. This is not low enough to measure the noise of a 50Ω resistor, but no commercially available test equipment has a noise floor low enough to accomplish that.

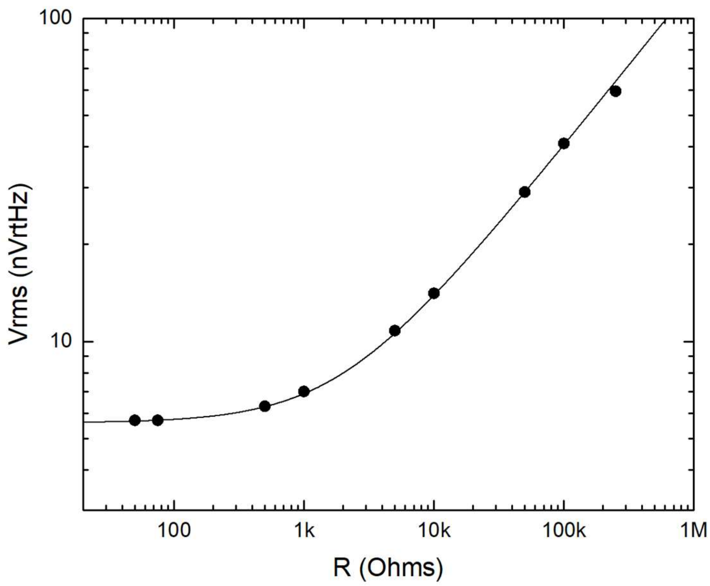


Figure 17. A measurement of resistor thermal noise using the SR770. The line shows a one-parameter fit of resistor noise along with an instrument noise floor.

Exercise 15. Now that you are set up, repeat the noise measurement with all the BNC terminators in the box, up to 250kΩ. Plot your data and you should see something like what you see in Figure 17. Add a theory curve, remembering that the resistor noise and the SR770 will add in quadrature. If you choose a suitable SR770 noise floor, your one-parameter theory curve should fit the data quite well. Once again, if your theory curve does not go through the data like in Figure 17, it may be because you hit the "fit" button and expected this to produce a good fit. Curve fitting routines often do not work well with log-log plots like Figure 17. Better to use a "chi-by-eye" fit (or at least start with that).

Noise Tells a Story

When you are new to the experimental physics game, you are probably not all that excited about instrument noise; you would rather look at some signal! Over time, however, you find that observing your desired phenomenon is hard, and measurements are limited in many ways. So, you naturally start to think more about instrument noise and how it can be reduced. And this is when it can be most beneficial to look at noise spectra using a spectrum analyzer.

For example, Figure 18 shows a noise spectrum of the LT1028 amplifier mounted on the printed circuit board in the lab. Clearly this is not the spectrum of simple white noise, but rather a mixture of several different noise sources that contribute to this complex spectrum. In this case, the noise spectrum can be largely accounted for with:

1) The sharp spikes come from the inexpensive power-supply "brick". This is called a "switching" power supply because it regulates the voltage by switching the power on/off rapidly. Figure 18 shows that the switching speed is a bit over 100kHz. These spikes can be reduced by using a more expensive power supply.

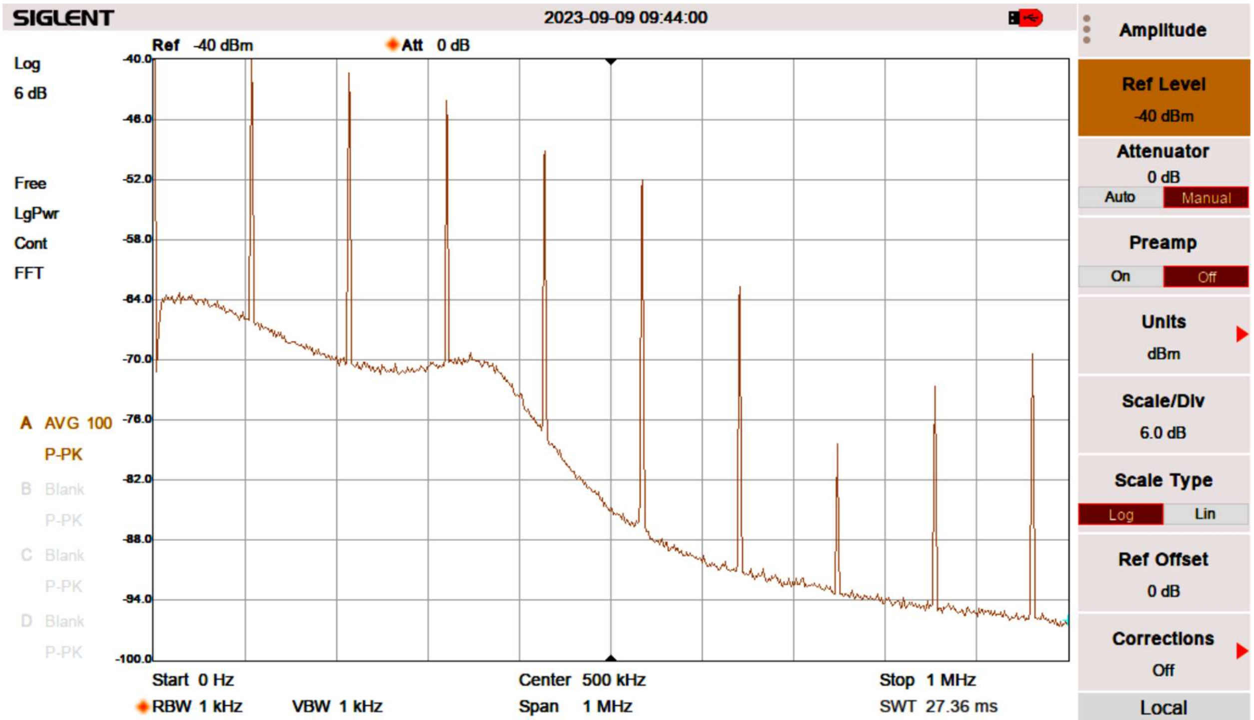


Figure 18. A noise spectrum of the LT1028 amplifier provided in the lab (on a printed-circuit board), measured using the Siglent 3021X spectrum analyzer. The series of sharp spikes come from the "switching" power supply provided, and these go away if a quieter power supply is used. The broad bump around 350 kHz is an intrinsic feature of the LT1028, described in its spec sheet. Noise spectra like this are especially useful for designing experiments and predicting their sensitivity to desired small signals.

2) The bump around 350kHz is an intrinsic feature of the LT1028, described in its spec sheet. This op amp is optimized to reduce the noise at low frequencies (to slightly below 1nV/rtHz), and it appears that achieving this comes at the expense of a 350kHz noise bump.
3) The high-frequency rolloff comes from the 1000x gain together with the LT1028's gain-bandwidth-product of 75MHz. Those together mean the gain starts to roll off at around 75kHz. Thus, both the signal and noise decrease together above 75kHz.

Exercise 16. Reproduce the results in Figure 18 and put a copy in your e-notebook. To do this, send the pcb output first into the SR560, as this is needed to buffer the pcb output to drive the 3021X's 50Ω input impedance. You can use a fairly low SR560 gain (say G=10), because the LT1028 already amplifies the noise signal by 1000x. But G=1 is a bit too low, given the results in Figure 15. Also, AC couple the SR560 because the LT1028 has a high DC offset.

To put all this in context, consider the LIGO noise spectrum shown in Figure 19. Because the underlying gravity-wave signals are so small, it took a great deal of effort to build an instrument capable of measuring them. And that meant studying the noise with great care. If your career path takes you into cutting-edge experimental physics or innovative new technologies, then you too may find yourself thinking about noise spectra from time to time.

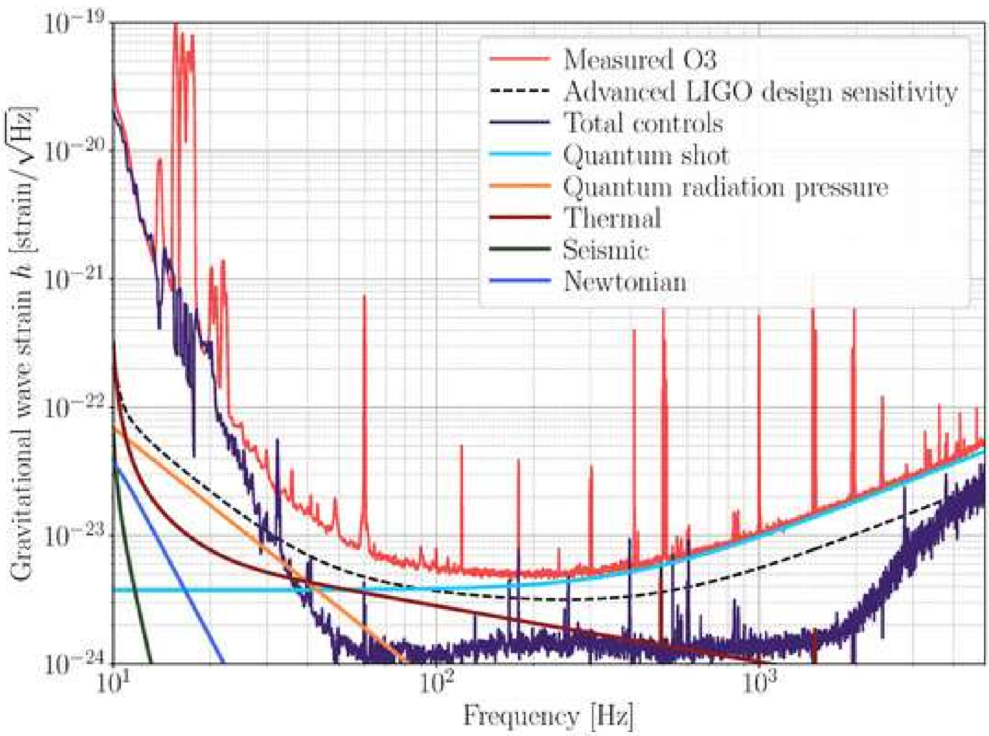


Figure 19. A sample noise spectrum from the LIGO (Laser Interferometer Gravitational-wave Interferometer) project. This shows the noise in measured strain (ΔL/L, where L=4km is the interferometer arm length). Thousands of people worked for several decades to measure, model, and reduce the noise, and this effort was essential for observing gravitational waves.