Notes on Acoustical Measurement for Violin Makers and Researchers

Jim Woodhouse, Evan B. Davis

1: Philosophy

Making decent measurements is a skill of a similar kind to making a decent instrument. Doing it by recipe is like making a violin from Heron Allen — it will work, but it probably won't be great. Some people have more natural talent than others, but all beginners benefit from occasional hands-on guidance from someone with more experience. The issue is not so much "which buttons to press when", but how to recognise quickly that something looks wrong, then concentrate on fixing that rather than wasting time collecting a lot of misleading data.

Ask yourself the key questions "What do I need to know?" and "How well do I need to know it?" before starting any test project. Asking these questions can greatly simplify the selection of a test method. For example, a test which only requires tracking of resonant frequency changes or relative level changes is much simpler than one that involves absolute mobility levels, which requires a fully calibrated force and velocity measurement system (see section 3 below).

2: Reproducibility

All measurements have noise and accuracy limits. To avoid over-interpreting your measurements, and as a way of getting to know your particular measurement rig and its quirks, it is really important to go through something like the checklist given below. The important thing is the idea and the mindset, rather than the exact numbers of how many of each kind of test you do. But don't shortcut this learning and checking stage, the time it takes will be amply repaid in mistakes avoided later.

2.1: Repeat tests

Set your rig up for the measurement you intend. But don't just do the test once, repeat the measurement several times. Display the results on top of each other — on the screen if your software allows it, but if necessary print the plots out, stack them and hold them up to a window. Look carefully at the detailed comparison. How different are they? Close enough for what you wanted to know, or would you like the reproducibility to be better? Is the difference more or less uniform? Perhaps it is different in different frequency ranges, or the peaks are more reproducible than the dips, or there is some other pattern. Patterns are always trying to tell you something....so think about it. Can you guess what kind of thing might cause this pattern? Is it something you could tighten up in your procedure? Try it!

You don't have to go through this rigmarole every time you make a test, but you should do it occasionally — more often when you are learning, less often as you gain experience.

In this repeat-testing process identify the limits on the resolution of the measuring output devices. Assuming you are using some form of an FFT based analysis package the ultimate frequency resolution will be determined by the FFT size and frequency range. If the frequency step is 1.57 Hz then looking for changes smaller than about 6 Hz will be problematic. Set the analyzer so that it can clearly resolve the differences that are being measured, to the best of the analyzer's capabilities. This is no different than selecting a scale for weighing things: if the scale only reads out in tenths of a gram, weighing items that are on the order of a half of gram or less will have poor resolution.

2.2: Deliberate variations

Now do a deliberate experiment about variability. Keep the same test instrument (or wood sample, or whatever you are trying to measure), and repeat the test with "irrelevant" things deliberately varied. The first thing to do is to simply take the instrument (or whatever) out of the rig, put everything away, and then take it out again as if it was another day's testing, and try to put it all back together exactly as you had it before. The next thing to do is to move the rig around in the room, or to a different room. Move the furniture around a bit. Choose days with different temperature and humidity. Anything else you can think of. For all these tests, ask the same kind of questions as before. How repeatable are the results? Are there patterns to the variation?

Now you will have gained some experience about how far you can trust your measurement, and in the course of doing that you will also have got better at controlling what you are doing. Just keep thinking of the analogy with instrument making, and remember your first instrument: you were proud of it at the time, but many established makers would like to get that first instrument back and destroy it (or at least take their label out). Your first measurements may be similar. There is no shame in that, everyone is the same.

The caution can be extended: when you make your second instrument the "it really shouldn't take this long" voice is heard in the back of your mind. As it in many things in life, and for measurements in particular, 'haste make waste'.

A good idea for a final step in checking out your test rig and data analysis processes before you get down to serious measurement is to try to reproduce the results of a known test. This is best done as a non-destructive test such as adding lumps of mass to the instrument. Look around for a simple experiment someone else has done and see if you can get comparable results. The violin in your rig is different from the one tested by the previous author, so what you want to see is a similar *trend*, and broadly similar numbers. For a suggestion, see the "Homework assignment" at the end of this document.

2.3: Real tests

Now get back to what you were trying to do. Probably you wanted to measure differences between instruments, or changes after some procedure. Are the differences you measure bigger than the ones you saw in these reproducibility tests? If they are, you (probably) have meaningful data. If not, you may still have meaningful data, but it isn't obvious. Look for the patterns again — does the deliberate variation produce a different pattern from the ones you saw before, or does it look suspiciously similar to the effect of the weather changing or moving the furniture?

Hypothesis testing is the best way to explore the data. A hypothesis is based on your current understanding and expectations. A hypothesis is a guess, hopefully an educated guess. It maybe the null hypothesis "If I do this nothing will change" or based on conjecture "If I block one f-hole A0 will change by a fifth". The hypothesis may be right or wrong. If the hypothesis is correct you most likely understand how the system is working (great!). If the hypothesis turns out to be incorrect, what you learn from the experiment will help you pose a new hypothesis. After you have checked the hypothesis, look around — did something else change in a pattern? For example, in most cases changing something that moves a low frequency mode will also change the system at high frequencies.

3: Calibration

You have now learned to make measurements that you understand and trust, at least to an extent. Now you may want to compare with someone else's measurements of what is supposed to be the same thing. Maybe something published, or maybe you are exchanging results with a friend with a similar test rig. The first thing you need is some kind of *calibration* procedure, to fix the absolute level of your plots so that it makes sense to compare. The essence of any calibration procedure is to apply your measurement procedure *exactly as you were doing it, with no changes whatever* to measure something that you already know the answer to. You compare what you get with what you expect. If all is well, the pattern should look right, but you need to apply a scaling factor to your measurements to convert to the right answer. This is the calibration factor for your rig, write it down and apply exactly the same factor to all measurements made with the same rig and the same settings.

If you change a setting by twiddling a knob or switching a range switch, you should repeat the calibration experiment — don't trust the manufacturer to have marked the dial accurately enough, check it out yourself. Any instrumentation rig that has 'knobs' which vary the gain continuously (as opposed to switching between preset ranges) will be a constant source of problems with maintaining calibration. One approach is to arrange to work with the knob right at the top of the range, which is something you can reliably repeat every time. Unfortunately top of the range settings can have other consequences related to signal limiting or signal distortion. The best advice to work around the 'knob' problem is find a good safe setting, then tape over the knob securely, do your calibration measurement in that condition, and never touch the knob again!

3.1: Absolute and relative calibration

There are two levels of calibration you might do. The "gold standard" is an *absolute* calibration, which turns your results into internationally agreed units: for example admittance (or mobility) in N s/m (or s/kg, which is the same thing although it looks different). If you can easily do such a calibration, as you can for admittance, then it is obviously the best thing. For many measurements, especially for sound pressure measured by a microphone, it is not so easy to do an absolute calibration using reasonably cheap equipment. Then you have to think what you are really trying to do. Do you want to express your radiated sound power in absolute terms (in milliwatts, for instance)? Then you need to bite the bullet and find a way to get an absolute calibration. But perhaps you just want to be able to compare measured sound levels with someone else's results from a similar rig, being sure that *differences* are correctly captured. Then all you really need is the *relative* calibration factor which would convert your friend's results onto the same scale as yours.

3.2: How do you do it?

There is a choice of two approaches: calibrate each part of the system separately, or do an overall calibration. The second approach is usually much easier. With enough effort, you can calibrate each separate piece of equipment — for example accelerometer, hammer, conditioning amplifiers, computer card, and software FFT algorithm — then combine all these factors to give the overall factor for your frequency response function measurement. Some of these factors may be supplied by manufacturers, but the others may take some effort to determine. The chain is no stronger than its weakest link, so this approach is quite error-prone if you miss a step out or get one thing wrong.

Usually, it is much simpler to do a single calibration for your entire frequency response function in one go. For an absolute calibration, you need to find a system for which the particular property you are measuring is already known — either because of some fundamental law of nature, or because someone else has already calibrated it using some standard test. For admittance, you can use Newton's law of motion "force equals mass times acceleration". If you measure a system which is essentially just a mass moving in a straight line, then the acceleration per unit force is simply the inverse of the mass, which you can determine with a weighing scale (but remember that the standard unit of mass is the kilogram, not the gram, so express the mass in kilograms!). In practice your mass might be suspended on strings as a pendulum, or supported on soft foam. You have to be careful that your hammer tap really does just make it move in a more-or-less straight line, without any significant rotation or cross-direction motion. You must also remember to choose a mass which is convenient to measure using the same settings of gain etc. that you use for your actual measurements. But then with care you can obtain an absolute calibration and compare your results with, for example, published ones by Erik Jansson or George Bissinger. (To compare with Jansson measurements you need to read the small print — his plots are not given in calibrated units, but he gives the required calibration correction in the text.)

For a relative calibration of a hammer-to-microphone measurement, for example, you need to choose some kind of stable and easily reproducible system which makes a noise when tapped. By agreeing to tap in a given position and place your microphone at certain place, you should get the same answer every time (but check it out by the approach of section 2). Once that is verified, then the same structure or a careful copy of it can be tested by your friend. You should obtain matching frequency response plots, except for a scale factor (i.e. a shift on a dB scale). That scale factor is what you need to convert your measurements ready to be compared.

Pay careful attention in two-channel measurements where one spectrum is being divided by another — it is easy to get the input and output cables swapped. Mobility (velocity over force, admittance) is quite different from impedance (force over velocity). Whereas mobility peaks at resonances, impedance peaks at anti-resonances. If the signature modes have suddenly jumped when you do a measurement on your standard test article, check the channel assignments: another reason to add that standard test article to your test rig kit bag.

A final note: with all these calibration factors, of any kind, it is fatally easy to multiply by the factor when you meant to divide, or vice versa. You need to be awake when you think about how to do this! It always sounds obvious when someone explains, but this is a very common mistake.

4: Spotting patterns

4.1: Patterns and surprises

So how do you recognise a “significant pattern” and decide what it means? That is the million-dollar question, and the extent to which you learn to do that is the extent to which you become a good scientist. The crucial first stage is to ask the question, and to be on the lookout for patterns all the time — patterns that might be some interesting feature of the results of your measurements, or might be indicating that there is something wrong. How do you recognise “wrong”? Quite often from a “gut feeling” based on experience. Data not meeting an expectation is not necessarily “wrong” but it may be the germ which will lead to new understanding.

There are some kinds of thing for which there is believed to be a good theoretical understanding, at least in general terms. If you see something which contradicts previous experimental and/or theoretical expectation, it either means you are on track for your Nobel prize, or more likely that there is something wrong with your measurement or your interpretation of it. Even if you are in fact on track for the Nobel prize with some remarkable discovery which flies in the face of what everyone believes, you still have to convince people that your results are right. So in either case, seriously unexpected results have to be checked, checked and checked again. Your new material shows up with ten times the stiffness-to-weight ratio of any other known material? Better check your rig by measuring something familiar, like steel or aluminium. You see nonlinear response when everyone expects linear response? Better check that the nonlinearity is really coming from the test object, not from some artifact of your rig and the way it is put together.

4.2: Data manipulation

Data manipulation can sometimes help to see patterns. For example, data smoothing: but be aware that this is a double-edged sword. Smoothed “trend lines” make looking at a lot of data easier, while obscuring some of the spectral details. If your software will take the raw narrow band data and plot into 1/12 octave bands (i.e. in semitones), that will give you a relatively similar curve without a lot of the hash. Going to 1/3 octave bands will smooth even further. The 1/3 octave band is a fair approximation to the critical bandwidth of the ear so they can be interpreted as meaningful tonal change bands. The work by the psycho-acoustic researchers suggests we are looking for roughly 3 dB changes to be heard clearly by trained ears. Band filtered data is better for looking at differences between spectra; subtracting two “hashy” spectra give more hash.

The curve fitting tools in Excel are not very good for looking at spectral data and should be avoided if possible: use the filtering tools in the spectrum analyzer instead. A poor curve fit can lead to very strange conclusions. Be careful how many generations of smoothing have been used in the data analysis. Data smoothing is an art form best done using physics-based curve fitting processes. Eye-balling a line through a hash of data by hand has been shown to be a remarkably good way, so don’t worry about taking a pencil out and sketching in an average line through the data to help your eye find patterns.

There are other kinds of data manipulation: for example one could make a series of measurements of the frequency response functions a large number of instruments and scale them so all the B1+ modes have the same level, to look for patterns of relative levels of other modes and features to the B1+ level. The selection of a normalizing parameter or feature like this is going to influence the patterns, so it is important to have a clear definition of what this reference point is and how it will be defined in the data (and to think carefully whether you are really sure you are doing something sensible). For example in level shifting you lose all claims to a study of the relative total output of the instruments. Scaling to a ‘known’ feature is a way to ‘semi-calibrate’ across data sets, but when this is done it should be well documented.

4.3: Turning patterns into numbers
Once you have seen a pattern try to describe in it words first, then into numbers. In English: when a lump of clay as added to the bridge the high frequency response was reduced. In numbers: the bridge hill peak moved on average 2 dB lower for each 1 gram of mass added to the G string notch.

5: Conclusions

The take away message here is that for measurements to be meaningful they must be reproducible and well understood. False data leads to erroneous conclusions that may steer you into a big project that is doomed to failure because the pattern gleaned from the experiment is false. In engineering the goal is to make safe, efficient products and the history of engineering is littered with miss-steps. Correcting the miss-steps in a timely manner by careful measurement, responding to your gut feelings when things don’t seem quite right, and confirmation of unexpected trends is the target. So get into your lab and get measuring!

A homework assignment: A0-B0 matching

A simple study which is of interest for violin adjustment is to see what happens to the modes A0 (the air resonance) and B0 (a fingerboard bending resonance) as mass is added to the end of the fingerboard to adjust the relative tuning. Whether or not you like the sound and feel of "matched" instruments, this makes a good topic for honing your measurement skills.

First, go to the web link: http://www2.eng.cam.ac.uk/~jw12/JW_publication_list.html and find the paper number 57 on the list. You can download the PDF. Figures 5 and 6 show some measurements (in the top panels) of a particular violin as mass is progressively added to the end of the fingerboard. Figure 5 shows the result from driving on the bridge and sensing with a small microphone inside the box, while Fig 6 shows results measured with a small accelerometer on the end of the fingerboard. You can try to repeat whichever of these matches the kit you have available.

Adding mass can only shift resonance frequencies downwards, so if you want to be able to go through the "matched" state during the test the violin you start with should have the fingerboard resonance B0 ABOVE the air resonance A0 (i.e. projecting fingerboard not too thin). You may be able to find the pitches of these two resonances by ear: twang the free end of the fingerboard and listen to the noise with your ear near the twanging point, then seek out the air resonance by blowing over the f-hole or by humming into it (or indeed by playing the violin: it is usually around C# in first position on the G string). For your added mass, get some oil clay (e.g. "plasticine") or poster-sticking putty (e.g. "Blu-tack"). To make a series of equal masses, roll a lump out with your hand on the bench into a uniform "sausage" then cut that into equal lengths. If you can weigh them, that will make your results more quantitative.

Now set up the violin for your measurement. Support it in a "suitable" way, which might mean hanging on rubber bands, or resting on soft foam blocks on the bench, or having it held by a long-suffering violinist prepared to hold the pose for as long as it takes. Attach your sensor, tap on the bridge corner (with your force hammer if you have one, or with your pencil if not). Record the signals into your FFT package. Now do some repeatability testing. Vary the support details, collect several different signals, and compare them carefully as explained in section 2. Zoom in to the frequency range somewhere around 280 Hz (for a typical violin) and look for the two peaks. Verify that you have really found A0 by repeating the measurement while one f-hole is lightly covered by a soft cloth: the peak should be strongly affected, while the rest of the vibration spectrum should not change much.

Now you are ready for the actual test. Add your masses one at a time, and repeat the measurement each time as carefully as you can. Compare the results with the different masses and look for a trend like the plots in the paper. The results will allow you to put that violin very readily into a "matched" condition, or deliberately mis-matched in one direction or the other. Try playing it in each state, then you can reach an informed judgement about whether you like the effects of "matching" or not.