PHY340 Data Analysis Feedback:

Group P05 doing Problem P6

# Data Analysis

You determined the visibility values by fitting a sine curve to the data and reading off the am­pli­tude. This seems a reasonable procedure, although I found that the visibility data were not always very consistent with a simple sine wave (see for example figure 1), and this may intro­duce a systematic error into your determinations. In order to convince your readers that your procedure is valid, you should have included some plots showing the fits, for a range of ampli­tudes (the procedure is obviously most likely to fail at small amplitudes, and this may lead to a systematic error in your subsequent fit to the line shape). It appears from your report that you tried two different ways of determining the visibility for the RF data, presumably because some of these fringe sets are more obviously not good sine functions. This would have provided an excellent way to investigate the likely uncertainties of your results, but you do not seem to have taken advantage of this, making only a vague claim that the sine fits gave a “smoother” result.

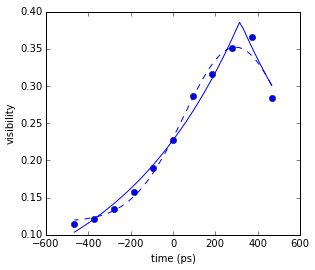
*Figure 1: PL data for 56 mm, with a sine wave fit.   
The data do not seem to be a pure sine wave (the 3rd, 4th and 9th peaks look slightly misaligned),   
and as a consequence the best fit amplitude is a bit smaller than seems optimal.*

Having obtained your visibilities, you then fitted the line shapes. The instructions state that the PL visibility is described by   
which is essentially a Gaussian with standard deviation and modified normalisation, while the RF data are described by a pure exponential,

In frequency space, these correspond to a Gaussian and a Lorentzian respectively. However, ***you are not working in frequency space***, so I have **no** idea why you tried to fit your RF data with a Lorentzian. Nor do I understand why you write your equation 1 as : as written (with the minus sign inside the square, and therefore removed by squaring), this is a function that would unphysically run off to +∞ as |*τ*| becomes large; assuming you meant to put the minus sign outside the square, it would be another Gaussian (you are just redefining *T*2 as instead of ). In your derivation, you sneakily remove this square before you Fourier trans­form, so you get the “right” answer—but not the answer that corresponds to your stated star­ting point.

Your RF fit therefore makes no sense, since it’s not fitting the correct function. Your PL fit may also be unreliable, because there is an obvious constant “tail” in your data extending out to large time delays: looking at your fit, it is possible that you restricted your fit range to avoid this tail (since if you didn’t, I would have expected it to increase the width of the Gaussian, which does not seem to have happened)—but if you did this, you have not said so. The tail is an inevitable consequence of most methods of fitting the visibility: the noise in the data gives a non-zero value of *V* even when there are really no visible fringes.

Your assertion that “it was found that…the resonance fluorescence graph *required* a Lorentzian fit” (emphasis added) is extremely unconvincing. There is nothing in your figure 2b that looks as though it could not be fitted with a Gaussian: in fact, I found that a Gaussian fit works quite well, see figure 2. The fact is that you fitted a Lorentzian because you thought for some reason that this was the expected form of the distribution, *not* because it is in any way required by the data. (It is, in fact, extremely unclear what you *did* fit: you say “a Lorentzian”, but you do not give the form of the function that you fitted, so it is not clear whether this is true.)



*Figure 2: Fits to the RF data. Solid line, fit to , yielding A = 0.387± 0.009, t*0 *= 317±11 ps and T2 = 594 ±33 ps, with an RMS of 0.013 and an ad­jus­ted coefficient of determination of 0.981. Dashed line, Gaussian fit (with con­s­tant background, which fits as 0.118±0.010), yielding t*0 *= 296±15 ps, σ = 244± 21 ps (hence T2 = 432±37 ps), with an RMS of 0.013 and an adjusted coefficient of deter­mination of 0.979. The two fits are es­sen­tially equally good.*

I did try fitting a Lorentzian to these data, even though there is no reason to do so. The fit is less good than either the Gaussian or the exponential (the coefficient of determination is lower at 0.965, and the RMS correspondingly higher, 0.017), demonstrating that there is no way that the data can be said to “require” such a fit.

You have redefined your time axis to place the maximum of the distribution at *t* = 0. This is a reasonable thing to do, *but you need to explain your procedure*: since there is no absolute defi­nition of “the maximum”, you need to explain and justify what you actually did. It is certainly not justifiable simply to take the highest data point: as can be seen from figure 2, and is also fairly clear from your figure 2b, the maximum in the RF data lies *between* two of the measured points, so fitting a curve with the maximum arbitrarily set at the highest point would not give the best fit. It is also not clear what you did with the data on the other side of the maximum: did you throw these points away (which would be a waste of useful information), or did you reflect them back on to the other side—and if so, how? What uncertainties have you introduced by this procedure? What alternatives could you have used? (The answer to the latter question is that you could have included the parameter *t*0 in your fit, which is what I did, and perhaps what you did too, before redefining the scale—but you don’t *say* so.)

Another serious problem is that you do not quote *any* *uncertainties at all*. There should be error bars on your visibility points, arising from the errors in your sine wave fit (or, better, from the comparison between the two methods that you used to determine the visibility for the RF data, since this would also give some idea of any systematic error introduced by the sine wave fit), and also errors in your fitted parameters, but you do not quote any of these. As a consequence, it would be very difficult to make a meaningful comparison with values from the literature, though as a matter of fact you make no attempt to do so. The absence of any attempt to quantify uncertainties means that this cannot be re­garded as a serious scientific report: assessing and reporting the uncer­tainties in your results is a key ingredient of experimental science.

I think your visibility data are quite plausible—they look very similar to mine, which is reas­suring since I did not use a sine wave fit—but you have not used them to best advantage. It would not have taken very much more work to produce a decent scientific analysis—I do not know Matlab’s fitting tools, but I am sure that they provide access to errors on the fitted para­meters[[1]](#footnote-1), and you already *had* the ideal information to estimate the errors in your visibility measurements, in that for one dataset you did it twice with two different methods.

Your final results also appear to be sensible, perhaps surprisingly so for the RF case given that you claim to have fitted the wrong function: I got *T*2 = 179±5 ps for the PL data, and 594±33 ps for the RF. The fact that your RF value is fairly consistent with this (about two of my error bars off) makes me wonder what you actually fitted: the width of the Lorentzian fit to my RF data was 463±35 ps, and your value seems to be midway between the two. Admittedly Lorentzian fits are notoriously unstable, so a minor difference in your data points might have made a big difference to the fitted width. But without a quoted uncertainty, and in the RF case without even a clear indication of what function you actually fitted, it is difficult to make rational com­parisons. Most of your conclusion does not follow from your data: there is simply no way that you can deduce from your data that “[u]nlike PL, RF caused homogenous [sic] broadening”: as we have seen, the RF data are perfectly well fitted by a Gaussian, so the statement that a Gaus­sian line profile “can not [sic] explain the line profile observed” is just not true.

Average mark for this section: 26.5/50

# Data Presentation

The quality of the data presentation in this report is not good. As noted above, there are no uncertainty estimates and no discussion of goodness of fit—the χ2 of the Gaussian fit in figure 3a must be *terrible*, because of its complete failure to account for the data from 0.2 ns onwards. The *x* axes of figures 2a and 2b are labelled as being in ns, but in fact they are in seconds (I categorically refuse to believe that the time delays are of order 2×10−10 ns!); in figures 3a and 3b they are also labelled as ns, but this time they seem to be in tenths of ns, since the 10−10 factor is not present. Figures 2a and 2b do not even show the data points, and in all four figures the axis labels are much too small to read comfortably (or at all, in the case of the legends on figure 3). There is inadequate detail at all stages: as noted above, you need to discuss how well your sine wave fits represent the data, and whether they could introduce systematic errors; to allow the reader to judge this him/herself, you also need to present some typical plots (of course, you do not want to show *all* the datasets—there are far too many—but a typical high-visibility PL plot, a typical low-visibility PL plot, and the same for RF, in each case with your sine wave fit inclu­ded, and in the RF case with your “turning point” values also shown, would have provided the necessary information) and any cases that were especially problematic, such as the 0 mm RF data, which you single out as being difficult to fit with a sine wave (I’m not at all surprised!).

You do not define all your terms: neither “coherence time” nor “radiative lifetime” is explained (these terms may be explained in your literature search, but it is not reasonable to assume that your reader has this to hand, when a couple of extra sentences is all that is needed).

When discussing your fits, you need to explain *what you actually did*, not just name the tool you used to do it. What does “Matlabs [sic] curve fitting tool” *do*: least-squares, minimum χ2, or maximum like­lihood? What algorithm does it use? What do you mean by “The data was fitted as close [sic] as possible before the autofit repeated the process”?

Average mark for this section: 15.75/30

# Style

It is good that you used LaTeX for the report, and the overall structure makes sense. However, the report is seriously lacking in necessary detail, as noted above, and is also not well written. Better proof-reading was needed to fix errors such as unwanted spaces before punctuation marks, missing or superfluous capital letters (“Michelson” is a name and should be capitalised; “voltage” and “intensity” are common nouns and should not be), spelling errors (“homogenous” for homogeneous) inverted word order (“value parameters” for “parameter values”) and more importantly an un­finished sentence in section 2 (“This was due to the nature of the beam­splitter changing the phase of the beam by.”). You appear to have mislaid the apostrophe on your keyboard alto­gether.

The final paragraph of the conclusion does not appear to have anything whatsoever to do with anything in the rest of the report, and is also backwards: what you have *said* is that a delta func­tion could never produce a perfectly monochromatic radiation (which is the exact opposite of the truth: a delta function is by its nature perfectly monochromatic); what I assume you *meant* is that atomic transitions never produce totally monochromatic radiation, and therefore cannot be accurately represented by a delta function. This should also have been caught by a good proof-reading. Likewise, the statement that “there is no thermal motion” of atoms in solids is clearly untrue, unless the solid in question is at absolute zero: the atoms in a solid are re­stricted to vibrating about the equilibrium position, rather than moving around at random like particles in a gas, but their temperature is still defined by their mean kinetic energy.

Your “derivations” in the appendix are totally useless, because *you do not explain what you are doing*. In particular, in section A.1 you state that if *ω* = 0, and then on the fol­low­ing line state that , which would imply that at *ω* = 0 *V*max = 4*T*2, not 2*T*2. What you *mean* is that the full width at half maximum is defined by the values *ω*1, *ω*2 for which , but **without the words it is impossible to deduce this from the equations that you quote**. In fact, this so-called derivation is wholly unsatisfac­tory: it starts from an incorrect equation, which is modified without comment in the following line, and then presents the expression without any proof whatsoever (so, as a derivation, it’s a complete failure).

Section A.2 does at least start from the right expression and make some attempt to do the inte­gral, but the omission of a number of minus signs and integration variables renders some of the expressions nonsensical:   
literally does not make any sense whatsoever (it needs an = sign and a d*t*). As with section A.1, the complete absence of words makes it impossible to follow, with the same lack of distinction between *V*max(*ω* = 0) and the half-maximum values *V*(*ω*1,2) = *V*max/2.

There are some missing references: I do not believe you drew figure 1 yourself, so it needs a reference; Matlab needs a reference; the standard integral needs a reference, since you clearly did not derive it from scratch. The references that are given are all incorrect or incomplete: reference 1 is not “M. Fox et al.”, but simply M. Fox; reference 2 should have Léonard Monniello’s initial and should include the article ID (041303), not just the volume and issue number; reference 3 should have the volume number for *Nano Letters* (14), and, to be consistent with reference 2, also the issue number (12), though really issue numbers are not needed in cases where the pagination doesn’t start afresh with each issue. If you quote DOI numbers (some journals expect this, others do not), quote them for *all* your references, not just for one of them.

The lack of detail is a structural fault, and there’s some lack of understanding apparent too (the statement in the abstract that “the PL linewidth was shown to be broader than that of RF, indicating that PL preferred a Gaussian fit while RF is more suited to a Lorentzian fit” is *utter rubbish*: the width of a Gaussian can be anything from 0 to ∞, and so can a Lorentzian’s; it’s the *shape* that distinguishes the two), but most of the problems with the report could have been fixed by a thorough and critical proof-reading, ideally by someone not directly involved in the analysis.

Average mark for this section: 11.5/20

Overall average mark: 53.75%

1. After a bit of Googling: yes, they do: ci = confint(fitresult, 0.682) will give the 1σ upper and lower bounds. See <http://uk.mathworks.com/help/curvefit/confint.html>. [↑](#footnote-ref-1)