An Analysis of the Resolution of the Michelson−Morley Experiment

Source: http://sci.tech−archive.net/Archive/sci.physics.relativity/2005−01/4713.html

---

From: Tom Roberts (*tjroberts_at_lucent.com*)
Date: 01/23/05

```
Date: Sun, 23 Jan 2005 19:37:01 GMT
```

Title: An Analysis of the Resolution of the Michelson−Morley
Experiment
Author: Tom Roberts, tjroberts@lucent.com
Date: January 23, 2005

Introduction
−−−−−−−−−−−−

There have been several recent attempts to re−analyze the original Michelson Morley experiment [1][2]. Cahill[3] has interpreted these re−analyses as supporting his theory.

Unfortunately, none of these authors understand error analysis, and thus do not know how silly their analyses actually are. Their basic problem is that they, like the original authors, attempt to interpret this experiment as "measuring the velocity of the earth relative to the lumeniferous ether". While that was a reasonable approach in 1887, today it is completely ludicrous −− not because of the mention of "ether", but because today we use experiments like this to _test_theories_, not to try to make "measurements" on concepts contained in some particular theory.

In this case, this change in outlook of the scientific method is clearly required because of a simple observation:
In a hypothetical world in which:
a) a perfect MMX experiment would yield a truly null result
and
b) real measurements are subject to measurement errors
it is statistically highly unlikely that a real MMX measurement will yield a null result. In such a hypothetical world, of course, the non−null result is induced purely by the measurement errors. But with an error analysis of the measurement, it can be determined whether or not the measurement is consistent with a theory that predicts a null result.

So when Munera[1] repeatedly proclaims "this is a non−null result" for various experiments, he is repeating a fundamental error −− sure the measurements can be interpreted as a non−null result, but the important question is: are they _consistent_ with the predictions of a given theory? As we will see below, the actual MMX data are consistent with the predictions of SR, and with a wide range of theories in which the earth moves relative to the ether.

Michelson and Morley's data are given, in a reduced form, in their 1887 paper[4]. The above attempts at analysis are based on the data in the table on page 340 of [4]. Unfortunately these data are not the original readings, but each row is an average over 6 turns of the interferometer made over approximately 36 minutes. Note I am discussing only the six rows for their six runs, not any of the rows containing means.

In performing an analysis on an experiment performed long ago, with only limited access to the data and no access to the apparatus, we are limited in our ability to determine the experiment's actual resolution.
I have identified three approaches:

1. Look into a modern Michelson interferometer and estimate the measurement resolution.
2. Use the original authors' statements to infer their resolution.
3. Use the original authors' data in a statistical analysis of the resolution displayed by the actual data.

There are in increasing order of confidence and accuracy.

Note that it is important to refer to the actual measurements, and not to averages. Unfortunately, the available data are averages over 6 turns of the interferometer, not the original readings. So I will assume that the errors in the individual measurements are uncorrelated, and normally distributed. While such an assumption is undesirable, the available data essentially force it −− a competent modern repetition of this experiment would take pains to accurately measure the actual resolutions. Fortunately, the presence of a rather large systematic error in the data implies that this statistical independence is reasonably likely[#]. In keeping with the assumption of normal errors and with modern practice, when I discuss "resolution", I mean the sigma of the associated normal distribution for the original measurement (in this experiment the location of a fringe).

> [#] During each rotation the reading changed by 15−30 divisions. This forces the observer to reposition the micrometer for each reading. While statistical independence is not assured, it is clearly more likely for a system in which the micrometer is repositioned for each reading than for a system without the systematic error where the readings vary by so little that it would be easy for the observer to simply leave the micrometer untouched (thus inducing an enormous correlation among readings).

When you plot the data given in the table of [4] for each day, it is quite apparent that there is a large systematic error that dominates the measurements −− the measurements at mark 16 before and after the turn are not equal. In fact, for each of the six runs the difference in the two marker−16 values is larger than the variations among the other readings. The authors [4][1][2] all subtract off an assumed linear dependence of this systematic error, and the original authors [4] mention a "temperature effect". Given the limited availability of original data, this is the best one can do, and I will do likewise.

Note, however, that this analysis technique _forces_ the data to be cyclical. That is, the above subtraction ensures that at the beginning and end of each turn the value will be exactly zero; any non−zero measurement in between will naturally appear to be "cyclical". Given non−zero resolution and independent measurements, there will be non−zero measurements in between. So claims that somehow the "cyclical nature" of the results implies or supports the "motion of the earth relative to the ether" are bogus −− _any_ such data will be "cyclical".

Lets' look at the above three estimates of Michelson and Morley's actual measurement resolution:

1. Look into a modern Michelson interferometer

I believe that anyone who has ever done so will agree that
- a) it is fairly easy to note the location of a fringe to within about 1/5 of a fringe width
- b) it is unlikely to be able to locate fringes to better than 1/10 of a fringe width

Basically the fringes do not have sharp edges, and one must inherently guess where the center of a fringe is.

So this approach yields an estimate of resolution between 0.1 and 0.2 fringe widths.

2. Use the original author's statements to infer their resolution

Michelson and Morley[4] state "The width of the fringes varied between 40 to 60 divisions, the mean value being near 50[...]". In keeping with the assumption that the measurement errors are normally distributed, I'll assume that this means that 95% of measurements of fringe widths were contained in the interval from 40 to 60 divisions of their micrometer. That means their resolution for measuring fringe width is 5 divisions, or 0.1 fringe. As the measurement of a fringe width requires two measurements of the location of a fringe, their base resolution is sqrt(2) time this.

So this approach yields an estimate of resolution of 0.14 fringe widths.

3. A statistical analysis of the resolution displayed in the data

The key to doing this is to find instances in the data where they measured the same value multiple times; then a histogram of the multiple measurements will give a distribution of the errors, and the resolution can be obtained from the distribution.

In an idealized Michelson interferometer, the interfering light rays travel both directions along each path, so there is exact 180 degree symmetry. In the actual apparatus, the ray paths are indeed out−and−back, so this symmetry should apply to the measurements. The original authors applied this symmetry in their analysis. Here we will use it to estimate their resolution.

> [In fact for perpendicular arms there is an additional 90−degree symmetry, unexploited by all authors including me.]

The idea is to first subtract the linear systematic from each of the six rows of data, thus forcing the two measurements at mark 16 to be equal for each run. Then histogram the eight differences for measurements 180 degrees apart, for all six rows, and determine the resolution of the measurement from the histogram. This was done in an Excel spreadsheet, but it is not feasible to display the details in this ASCII medium. The histogram does not look very Gaussian, but is rather flat between −5 and +7 divisions. The likely source of this non−Gaussian behavior is the systematic error that was _assumed_ to be linear, and nonlinearities due to either non−uniform behavior or non−uniform spacing of the measurements could cause this. The sigma of the histogram is 3.0 divisions, corresponding to 0.060 fringe widths; as each point is an average of 6 turns, the resolution of the original measurements is sqrt(6) times this value.

So this approach yields an estimate of resolution of 0.15 fringe widths.

Discussion

None of the above estimates are particularly compelling, mainly because the histogram of method 3 is not really Gaussian; this does not destroy that approach, but makes it less compelling that it would be with Gausssian errors. But their agreement indicates they are not crazy. Certainly there is no support for any resolution estimate much lower than 0.14−0.15 fringe width.

To compare to the original authors' data table, each row is an average of 6 turns, and so should have a resolution of 0.14/sqrt(6) = 0.057 fringe width. After subtracting the assumed−linear systematic from the 6 rows, the largest deviation from zero is 0.132 fringe widths, or 2.3 sigma; of the 96 data points, only 1 point exceeds 2 sigma, and 11 exceed 1 sigma. Clearly the readings are not Gaussian distributed. But equally clearly they are consistent with a null result, and provide only equivocal support for the notion that there is a non−null result.

Interestingly, when one histograms the data with the assumed−linear systematic subtracted, the deviations from zero are roughly Gaussian distributed with a mean of −0.01 fringe and a sigma for individual measurements of 0.1 fringe. While this is _not_ an error plot, when compared to the above resolution estimates it solidly demonstrates that the measurements are consistent with the hypothesis of a truly null result.

When Consoli and Costanzo [2] display a graph of the July 9 PM data, they drew error bars approximately 0.005 fringe −− more than a factor of ten too small. They give no indication whatsoever how they arrived at this value; certainly the original authors gave no error bars. The above estimate of 0.057 fringe is larger than their entire plot, and indicates their fit is meaningless. Their fit has 10 parameters for 16 data points, so it is not surprising that they can draw a line through most of the points, even with tiny error bars. They do not mention any chi−squared tests for goodness of fit, and without that and realistic error bars their estimates on the errors in their parameters are completely bogus. It is clear that with the above error estimate a zero−parameter flat line fits the data as well as their 10−parameter Fourier decomposition.

Munera[1] correctly points out that for a velocity relative to the ether the MMX only displays the projection of the velocity vector onto the plane of the interferometer, and this implies that it is unlikely that such a signal will be a pure cosine. He goes on to claim that even the intra−session average of 6 turns is invalid as during 36 minutes there is a change in this projection. While true, that is not important, as his values show it changes by at most ten percent −− this is wildly exceeded by the resolution of the measurement.

Cahill[3] has interpreted this as a positive observation of motion relative to his ether, with a value consistent with the CMBR dipole=0 frame. As mentioned above, he is performing an invalid comparison, and is basically imposing his hopes and dreams onto the data. A proper analysis would take his formulas with an unknown speed and direction of motion relative to the ether, and _predict_ the results of the measurement. Presumably this could then determine the speed and direction of that motion. Had he done so, it is clear that with the above resolution estimate his formula would fit the data for any speed between zero and several thousand km/s and any direction whatsoever.

Conclusion

The recent attempts to "re−analyze" the Michelson Morley experiment[1][2] are woefully incomplete, and do not include an accurate consideration of the experiment's actual resolution. If considered as a measurement of the motion of the earth relative to some ether, the value depends upon the details of the theory used to model such motion. For the ether theory used by the original authors, an upper limit of 5 km/s is appropriate, but might be reduced by a careful modern analysis. For Cahill's theory an upper limit of several thousand km/s is appropriate.

In any case, the experiment is indeed solidly consistent with the prediction of SR −− a null result.

[1] H.Munera, APEIRON _5_ (1998), p37.
[2] Consoli and Costanzo, http://arxiv.org/abs/astro−ph/0311576
[3] Cahill, http://arxiv.org/abs/physics/0501051
Cahill and Kitto, http://arxiv.org/abs/physics/0205070
[4] Michelson and Morley, Am. J. Sci., _XXXIV_ (1887), p333.
http://www.aip.org/history/gap/PDF/michelson.pdf