

The Failed Experiment That Failed to Fail

Mikel Aickin
Minneapolis, Minnesota, USA
mikel@aickin@gmail.com
Version 06 Jul 2025

Abstract. The Michelson-Morley experiment has become famous for disproving the existence of the luminiferous ether, and supposedly providing Einstein with the inspiration for special relativity. Although it has been much discussed in the literature, there has never been a comprehensive scientific and statistical analysis of the published results. The attempt to do so here uncovers statistical, logical, and scientific problems, sufficient to warrant the claim that the scientists actually did discover two effects, one of which they were seeking. This episode illustrates how a failure to apply sound philosophical principles of science can derail an otherwise excellent experiment, and create an enduring false history.

Introduction

The conventional interpretation of the Michelson-Morley experiment of 1887 is that they saw nothing, no evidence for the earth moving through the ether, while the minority contrary interpretation of those who promote the ether is that they merely saw what they had intended to see. Actual analyses of the data that the experimenters published are essentially non-existent, so that in practice both sides are reasoning on the basis of subjective impressions. The analysis presented here shows unquestionable evidence of the period π effect they sought, as well as a linear temporal effect and a period 2π effect. The null declaration caused Lorentz to propose an explanation that strained his own electrodynamic theory, and which Einstein passed along to relativists. There is still room for debate about exactly what Michelson and Morley saw, but there is no reason for further debate about whether they saw anything.

One measure of the importance of an event is the degree to which it is discussed, dissected, celebrated, vandalized, elevated and excoriated. By this criterion the Michelson-Morley experiment of 1887 is perhaps one of the most important events in all of physics. The attempt to measure the earth's speed through the luminiferous ether was reported by the experimenters as a failure, in the sense that no motion at all was detected. Most subsequent authors have accepted that judgment, and it has become the conventional physics

interpretation, but a smaller number have claimed that a positive speed was actually seen, although it was smaller than the experimenters had hoped. Given the importance of the Michelson-Morley experiment and its effect on the subsequent direction taken in electromagnetics, it is surprisingly how few of the judgments made about it are based on actual analysis. There does not appear to have ever been a complete and competent analysis of the data that Michelson and Morley published. Given the importance of their experiment, especially the modern emphasis on it as being determinative, this is more than just mildly surprising.

The purpose of the present note is to do as thorough an analysis as is possible based on the data that Michelson and Morley included in their article. The conclusion is that they unquestionably detected the pattern they were looking for, and that the method of analysis they chose was partly responsible for their incorrect conclusion. A close look at the data also raises a series of unanswered questions, however, which leaves a trail of doubts. One might argue, of course, that any such analysis is moot, in the light of subsequent more precise and carefully done experiments on the topic, but it is interesting that part of the legacy of the experiment was confusion about “length contraction” which went on to play a role in special relativity and continues as a misconception up to the current day.

I will also give some consideration to attempts that have been presented in the literature to re-analyze the Michelson-Morley data. To delimit my intentions here, I do not consider arguments that question the experimenters’ physical assumptions, which importune their mathematics (nearly always incorrectly), or which raise novel physical theories outside the bounds of conventional physics. This is not meant to disparage novel interpretations, but to restrict what I have tried to accomplish.

Design Concept

The basic foundation and physical apparatus have been well described in many places, including by Michelson and Morley themselves. In brief, the experimenters constructed a rotatable interferometer table, on which a beam of light was split in half so the two resulting paths went at a right angle to each other. They were reflected off mirrors and rejoined at an optical device that permitted visualization of interference fringes. Because the two half-beams travelled perpendicularly for part of their paths, they would have been speeded or slowed by movement through the luminiferous ether, to different degrees depending on the angle of their paths relative to the velocity vector of the earth heading into the ether. Since that direction was completely unknown, the table was rotated and the fringe shifts measured at each of 16 equally spaced points. If there were no ether (or improbably, the earth was stationary in the ether) there would be no pattern of variation of the shifts during a rotation, otherwise a sinusoidal pattern would be seen, the amplitude of which could be converted into a speed.

Michelson and Morley’s concept and analysis of their experiment was entirely Galilean, and so that is the viewpoint I will take here. Start by considering a single, one-directional light path represented by a vector $\mathbf{d}$ of length d . Suppose v is the speed of the earth through the ether, and let θ denote the angle between the light

path vector and the earth-velocity vector $\mathbf{v}$. As usual, c is the speed of light in any direction in the stationary ether. Then

$$c^2 t^2 = \|\mathbf{d} + \mathbf{v}t\|^2 = d^2 + v^2 t^2 + 2\mathbf{d} \cdot \mathbf{v} t \cos(\theta)$$

The left side is the squared distance a photon goes in the stationary ether in time t , and the right side embodies the Galilean assumptions about the effect of movement. Re-write as

$$(c^2 - v^2)t^2 - 2\mathbf{d} \cdot \mathbf{v} t \cos(\theta) - d^2 = 0$$

and solve for t

$$\begin{aligned} t &= \frac{\sqrt{d^2 v^2 \cos^2(\theta) + d^2(c^2 - v^2)} + \mathbf{d} \cdot \mathbf{v} \cos(\theta)}{(c^2 - v^2)} = \\ &= \frac{d}{c} \frac{\sqrt{c^2 - \sin^2(\theta)v^2} + v \cdot \cos(\theta)}{c(1 - v^2/c^2)} = \frac{d}{c} \frac{\sqrt{1 - \sin^2(\theta)v^2/c^2} + \cos(\theta)v/c}{(1 - v^2/c^2)} \end{aligned}$$

This represents the outbound light, and upon reflection the return photon traverses $\mathbf{d}$ in the opposite direction, which has the effect of replacing θ by $\theta + \pi$ and therefore changing the sign of the cosine term. When the two times are added, the round-trip time is

$$\frac{2d}{c} \frac{\sqrt{1 - \sin^2(\theta)v^2/c^2}}{1 - v^2/c^2} \cong \frac{2d}{c} \frac{1 - (1/2)\sin^2(\theta)v^2/c^2}{1 - v^2/c^2} \quad (\text{approximately})$$

This is based on the Maclaurin approximation

$$\sqrt{1 - u^2} \cong 1 - u^2/2$$

as was used by Michelson and Morley, which is quite accurate for the small earth speed they anticipated.

Let us now replace d by d_1 for the first arm, and for the second arm at a right angle to the first assume length d_2 . Presumably it would have been impossible for Michelson and Morley to have made the two arms of equal lengths at a level of precision that would make it ignorable. For convenience let us use $\bar{d}$ for the average length and Δd for half their difference, so that $d_1 = \bar{d} + \Delta d$ and $d_2 = \bar{d} - \Delta d$. For the second round-trip time the sine and cosine are interchanged and the sign of one of them is changed. Thus the difference between the two round-trip times is

$$\begin{aligned} &\frac{\bar{d}}{c} \frac{(\cos^2(\theta) - \sin^2(\theta))v^2/c^2}{1 - v^2/c^2} + \frac{2 \cdot \Delta d \cdot (2 - v^2/c^2)}{c(1 - v^2/c^2)} \\ &= \frac{\bar{d}}{c} \frac{\cos(2\theta)v^2/c^2 + 2(\Delta d/\bar{d})(2 - v^2/c^2)}{1 - v^2/c^2} \end{aligned}$$

Therefore, the pattern of the time differences, and so anything that depends on it, would have a period of π , which would vanish if (and only if) the earth velocity v were zero.

Note that a difference between the arm lengths would at most contribute a nearly constant term, which would not interfere with the periodicity. If the second arm were not exactly perpendicular to the first, then from

$$\cos^2(\theta + \Delta\theta) - \sin^2(\theta) = \cos(2\theta) \cos^2(\Delta\theta) + \sin(2\theta) \cos(\Delta\theta) \sin(\Delta\theta)$$

the period would stay the same. Consequently, so long as one focuses on the period π pattern, small differences in arm lengths or departure of their angle from $\pi/2$ would have no effect.

It is worthwhile noting that if the returning light beam were affected either by reflection or due to one-way light speeds not being equal, then a period 2π effect proportional to v/c might be detected. Since the theoretical isotropy of light would pertain to a vacuum, but the experiment took place in air, one cannot automatically discount the longer period effect.

To be clear, I have presented these computations to clarify the analysis below, but I do not claim any originality. For example, they have essentially been published by Articulo (1969).

The Data and Analysis

The only data we have from the experiment was what was reported in the 1887 article. Six runs were given, each consisting of 17 fringe-shift measurements, the initial measurement being repeated at the end of the rotation. Each run was in fact the average of six identical runs made consecutively. Three of the summary runs were started at noon and three at 6 pm in the evening, over four days in July of 1887. A rotation took about six minutes, so that each reported summary was based on about 36 minutes of observation. The interferometer table floated on a pool of mercury, so it could be rotated at a constant rate with negligible vibration. One experimenter (Michelson) had to follow the optical eyepiece around as the table rotated, with measurements being taken rather quickly. The effort that was required evidently caused some physical strain.

The fringe-shift runs are shown in Fig. 1. All exhibit an obvious near-linear trend, upon which the experimenters made no comment. Since this effect dominated the data pattern, it is unclear why Michelson and Morley did not track down the cause. Although their instrument was of exceptional quality, it was assembled and tested over only about two months, and we have no evidence of what quality testing they did. The most frequently published surmise is that there was a temperature effect, but as shown in the analysis above, any uniform change in the length of the arms would have multiplied the effect, not added to it, and although a differential change might have appeared as an additive effect, it would have had to be very consistently different between the two arms as they were rotated, and it is not clear how that could have occurred. Measuring the temperature might have clarified its potential effect or lack of effect, but this was not done. More surprisingly, evidently no control runs were done without table rotation to assess the presence and magnitude of potential disturbances like temperature changes. Michelson and Morley removed the temporal trend as the first step of data manipulation, but they failed to report how. The suspicion has been that they took out a linear term in order to make the first and last

observations equal. For the analysis here I removed the linear trend as part of the model fitting.

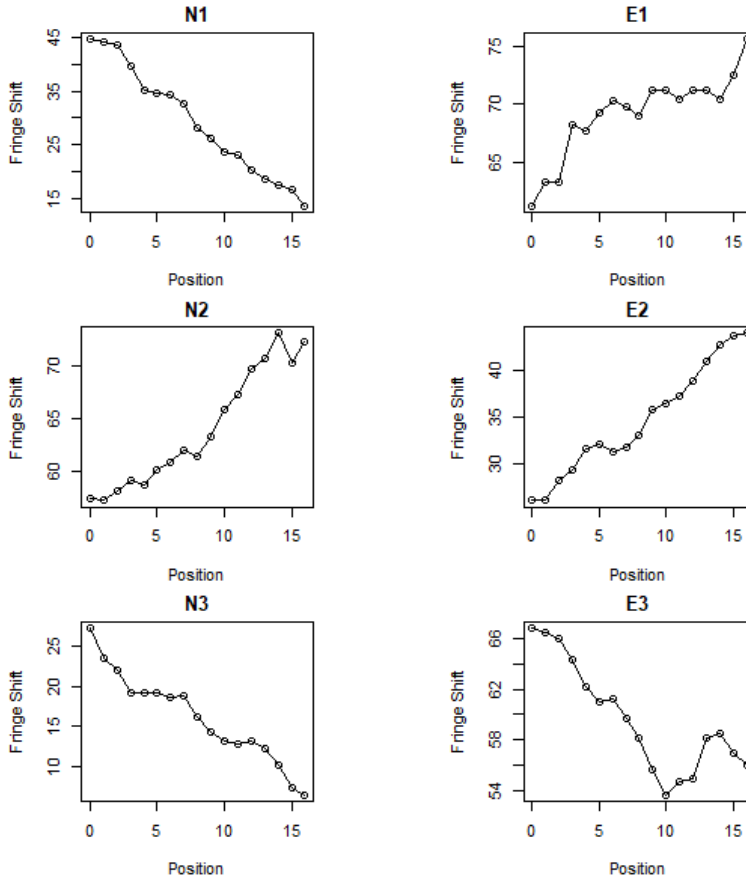


Fig. 1. Raw fringe-shift data for all six reported runs. N = noon, E = evening, days 1, 2, 3. (Day 3 was actually split over two days.)

A plausible alternative explanation for the linear trends was indirectly provided in the article, where Michelson and Morley remarked that the interferometer was rotated counter-clockwise for the noon runs and clockwise for the evening runs. This is suggestive because three runs show decreases and three show increases. If one imagines that the sense of the rotation produced consistent trends, then the pattern shown in Fig. 1. would be inconsistent with their description. Further, if instead it were the case that runs done at the same time of day would show the same temperature effect, then again their data would be inconsistent. This does raise the possibility, however, that the data were not entered into the published table correctly as to time and day. If true, this fact would invalidate the approach they took to analysis.

Fig 2. shows the raw data with the time trends removed, along with values fitted according to a model with two sinusoids, one of period π and the second of period 2π . Correlations between the observations and fitted values are displayed as rough indicators of the quality of the fit. As indicated in the analysis above, the period π pattern was of primary interest, but it is conventional in analysis of periodic data to include all harmonics below the one of interest, which in this case would just be the period 2π sinusoid. This decision was justified, in the light of subsequent analysis. It is to be noted that the N1 and E2 fits were of lower quality than the others, which could be caused either by excessive experimental error or a temporal pattern different from the other runs.

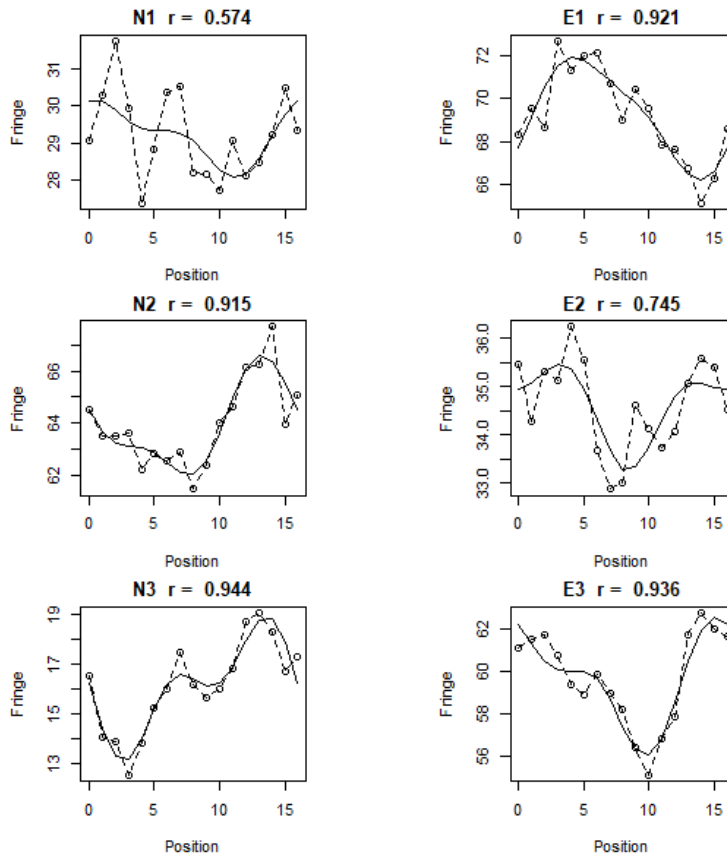


Fig. 2. Data with time trends removed (circles) and values fitted by least squares to a model with sinusoids of periods π and 2π . r = correlation between observed and fitted values.

The two sinusoidal components of fit are shown in Fig. 3. Unexpectedly, the longer period effects were if anything greater than the shorter period effects,

indicating that there was some influence related to the full rotation. Going back to the underlying mathematics, a difference between the outbound and return light velocities would contribute a period 2π effect proportional to v/c . If this were the source of the longer period component, then the experiment might have solved the long-standing problem of one-way as opposed to two-way light speeds, while at the same time demonstrating a potential ether effect on the one-way speeds.

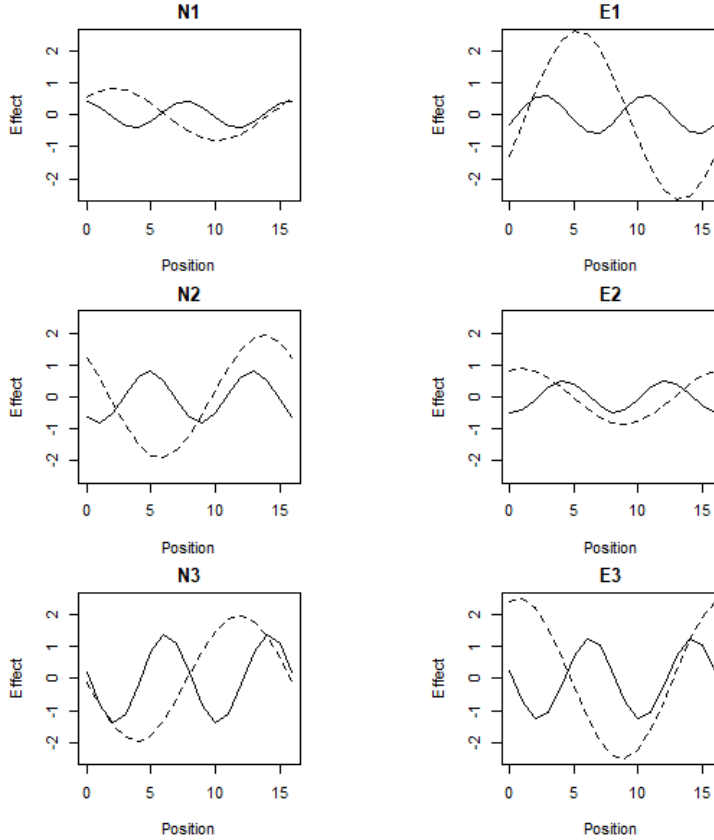


Fig. 3. Fitted values of the period π sinusoids (solid) and period 2π sinusoids (dashed). Position is in terms of $1/16$ of a rotation, effect is in fringe-shifts.

It is noteworthy that the phase offsets of the period π components appear relatively uniform across five of the six plots, whereas the phase offsets of the period 2π components vary considerably. In theory, the offsets would be influenced by the direction of the interferometry arms at the start of a run, and possibly on any adjustment of the mirrors. Michelson and Morley did not make any mention of any standardization of initial conditions, but they did label their graph with NS (north-south) and EW (east-west), suggesting that they might have

started each run with one of the arms pointing north. (They may have altered their data in their graph, however, because it does not correspond to their table.) In this case Fig. 4 would show the potential direction of a period π effect, with the plots aligned so that north is upward, but east-versus west depends on the sense of rotation. Four directions are necessary, since we can only say that the actual direction would correspond to one of the four extrema in the period π effects. The plots in Fig. 4. do not make any adjustment for the sense of the rotation, so that cases where the sense was reversed would need to be flipped horizontally. If we take the experimenters' words at face value, the counter-clockwise rotations at noon would need to be flipped to make the directions of the plot line up with the evening plots in terms of east-west direction.

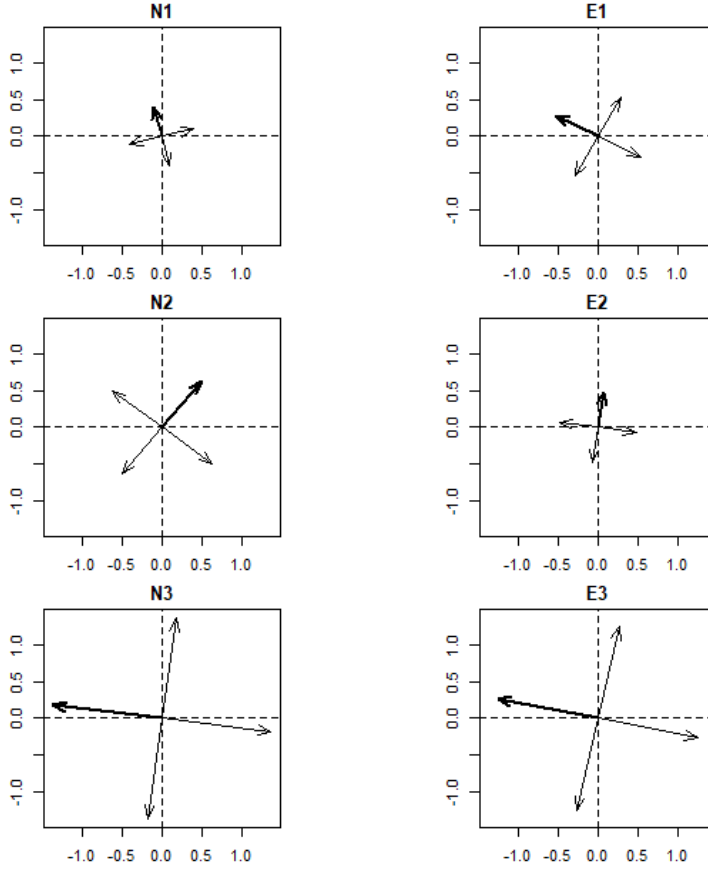


Fig. 4. Potential projections of the velocity vector onto the interferometry table surface for the period π component. Scales are in fringe-shifts. No adjustment is made for the sense of the table rotation. The bold arrow represents the arm initially pointing north, and on the plot north would be in the upward direction.

The phase offset issue is relevant to the question of whether the results indicate a terrestrial effect, something carried by the rotation of the planet, or an extra-terrestrial effect like a direction of motion through the ether. From the three-dimensional geometry, an extra-terrestrial effect would appear as a one-quarter rotation, which in Fig. 4 would pertain to the bold arrow, but this is unfortunately confounded with the sense of the rotation. Accepting Michelson and Morley's claim about this would amount to saying that a horizontal flip of the noon plots in Fig. 4 should give the same configuration as the evening plot, but with the bold arrow rotated one-quarter turn. Clearly this does not happen. Given the doubts about the data, it is probably the case that no speculation about the terrestrial/extra-terrestrial issue is warranted. The experiment would have been better designed if the times at which runs were carried out were not confounded with the issue of geometric rotation.

Taken all together, one might roughly group the six runs into three where the interesting period π pattern was more pronounced and three where it was less so. There is a sense, looking only at the period π patterns, that their quality might have improved over subsequent runs.

	N1	E1	N2	E2	N3	E3
	Period π					
Amplitudes	0.417	0.608	0.806	0.491	1.385	1.286
Speed	6.1	7.4	8.5	6.6	11.2	10.8
P-values	0.36837	0.12823	0.00664	0.06424	0.00011	0.00202
Offsets	0.038	0.171	-0.106	-0.025	0.229	0.217
	Period 2π					
Amplitudes	0.809	2.653	1.929	0.875	1.966	2.512
P-values	0.47964	0.00620	0.00173	0.04546	0.00002	0.00001
Offsets	-0.136	0.169	0.141	-0.053	-0.242	-0.048

Table 1. Inference from the fitted models. Amplitude is the square root of the sum of the squares of the sine and cosine components of the fringe-shift data. P-values are for the hypothesis of zero amplitude. Speed is corresponding earth speed through the ether. Offsets are the sinusoidal phase offsets, in units of 1 rotation (so 1 corresponds to 2π radians).

The statistical analysis from the model fits appears in Table 1. With regard to the pattern of interest, the first day amplitudes would not be counted as being detected, and we recall that the noon run was potentially less reliable. Similarly, the E2 amplitude is not technically significant by prevailing standards, although the low p-value does indicate lack of support for a zero value. The other three runs would be statistically significant by any standard. The period 2π amplitudes are

in each case larger than the period π amplitudes, in some cases considerably so, and five of the six would be considered to have been detected, by conventional standards. There is, therefore, virtually no question about the fact that the experiment detected effects of the sort that were sought.

Using the computations that Michelson and Morley gave in the article, the earth speed through the ether would be $3\sqrt{10A}$ km/sec, where A is the amplitude given in Table 1. The resulting values appear in the “Speed” row of Table 1, so that it is reasonable to summarize the results as estimates between 6 and 11 km/sec. By their reckoning, the experimenters concluded that with respect to their expectation of 30 km/sec, their result was “probably less than one-sixth and certainly less than one-fourth” of that value; that is, probably below 5 and certainly below 7.5 km/sec.

In the light of these graphical and statistical results, account must be given for why Michelson and Morley declared their experiment null. There were three mutually reinforcing errors. The first was that the experimenters based all of their calculations on the assumption that the earth’s velocity vector lay in the plane determined by the surface of the interferometry table. There was, of course, no evidence for this, and consequently the only speed they could have measured would have been the component of the earth’s velocity in the interferometry plane. In other words, their value of v was the magnitude of the orthogonal projection of the velocity vector onto that plane. If the velocity vector had happened to be perpendicular to the interferometry surface, they would have seen no effect no matter how high v was. To the extent that it was close to perpendicular, they would have seen a much reduced effect, as indeed they claimed to have done. This problem has been mentioned sporadically (Cartmel 1937, Miller 1933, Consoli 2018) and was especially re-emphasized by Múnera (1998).

The second error was to compare their results with what they had expected, as opposed to a speed of zero. They took the earth’s speed around the sun of approximately 30 km/sec as their expectation, despite the fact that any nonzero speed would have supported the existence of the ether. Even if one granted their expectation, the indeterminacy of their first error would have nullified their null assertion.

The third error was in the data manipulation, and in many ways this was the most surprising mistake. The whole basis of interferometry is the interference effect of super-imposed waves, which had been demonstrated for light by Thomas Young around 1800. The effect is computed in Fig. 5, showing how the amplitude of an average of $\sin(2\pi\omega t)$ and $\sin(2\pi\omega t + \delta)$ depends on δ . Exactly the same effect occurs for the type of sinusoids that Michelson and Morley recorded. By averaging six runs in order to produce single noon and evening runs each day, they risked reducing the magnitude of any effect due to potential phase offsets, which one should have expected due to random variations. Moreover, because they regarded the second half of each run as a replication of the first half, they averaged the two halves together to create their result for each noon and evening run (going now from 0 to π), and then finally they averaged over the three noon runs to produce an overall noon signal, and similarly for the evening runs. Consequently, they very likely reduced the effect they wanted to see by three successive averaging

processes. The results of the second two averagings are shown in Fig. 7. This does not correspond to the graph given in Michelson and Morley's article, which does not appear to be consistent with their own data. It shows amplitudes rather smaller than would be indicated by a correct analysis. The graph given in the article, which constitutes the only approximation of an analysis, is on the squared velocity scale (that is, in units of v^2/c^2), so that tiny differences are rendered infinitesimal. This was partially hidden by not providing an annotated labeling of the graph.

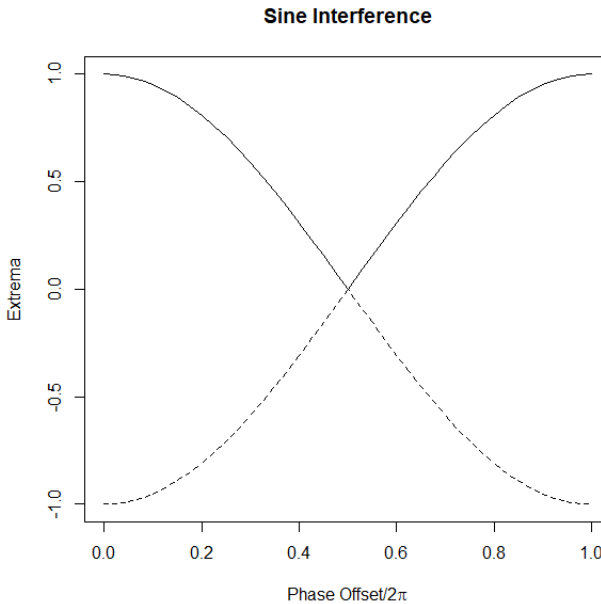


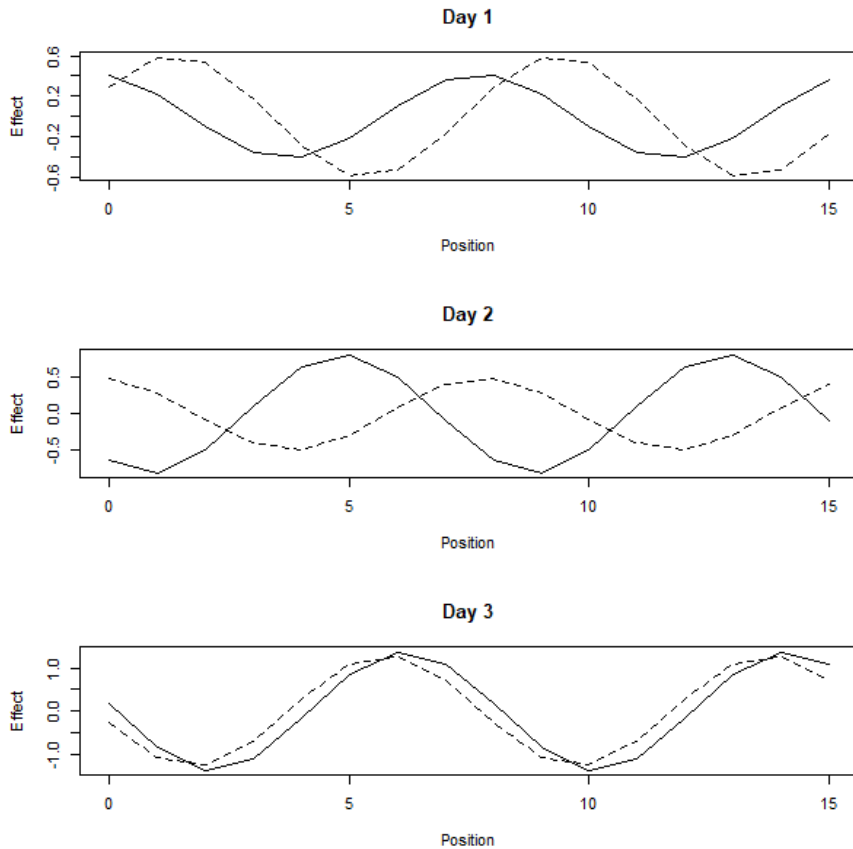
Fig. 5. Positive and negative amplitudes of a sinusoid composed of two sines of the same frequency but different phase offsets. The effect is independent of the frequency.

There is an additional issue with phase shifts. As stated above, Michelson and Morley wrote that the noon runs were done counter-clockwise and the evening runs were done clockwise. This implies that to make the noon and evening runs homogeneous, the time-order of one of them would have to be reversed. Moreover, the one-quarter earth rotation between daily runs implies a one-quarter rotation of the period π signal. Thus consistency of the pattern of results would imply that when these adjustments are made the two signals on each day should be in phase. The results in Fig. 6. suggest that the third day gave highly consistent results, but the previous two days did not.

We can make several remarks about the experiment, some that we can justify with the data, and some that must remain speculative. There is essentially no doubt that Michelson and Morley succeeded in detecting the effect suggested by

The Failed Experiment

a Galilean analysis of their design. What they saw was an unknown fraction of any actual speed, due to the unknown inclination of the earth's velocity vector to the surface of the interferometry table. There are two anomalous effects in the data, a linear time trend and a full period pattern. Both were at least as strong and convincing as the signal of interest. The method of data manipulation and presentation almost certainly reduced the effect they were seeking, although we will never know the magnitude of this reduction because the data from the underlying 36 runs were not preserved.



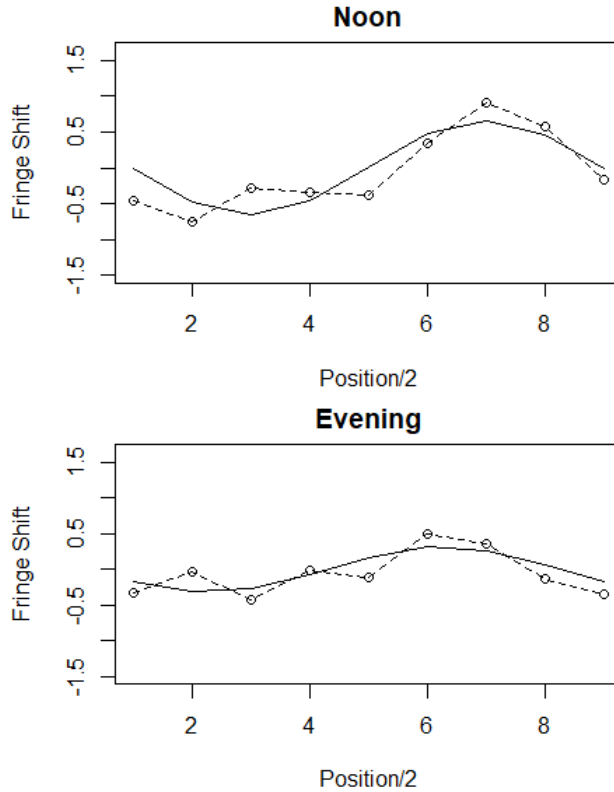


Fig. 7. Averaged data over all noon runs and all evening runs, folded so that the second half of each rotation is averaged with the first half. These were the basis of Michelson and Morley's conclusion, in which they used an amplitude of 10 as their expectation.

Having made some positive assertions about the data, there remain difficult problems. One would want physical explanations for the anomalous effects. Specifically, if the linear trend were related to the sense of the interferometer rotation, then the data do not correspond to Michelson and Morley's description. This in turn clouds the extent to which one might cite a terrestrial rather than ethereal cause for what they clearly did detect. We do not have the original 36 runs, and so we cannot subject them to a competent modern analysis. No control runs were reported, nor were there any daily or yearly repetitions.

There is evidence that Michelson and Morley did not carry out a full testing program before doing the experiment, due to the short time from the decision to go ahead and the actual experiment itself. This is further supported by the anomalies, and as well by the fact that there seemed to be an improvement of the results with time, indicative perhaps of a learning process over the four days.

Taken in total, the data on the last two runs are of higher quality, suggesting that one might regard the first two days as preliminary runs.

Existing Re-evaluations

There are no satisfying analyses of the data that Michelson and Morley published. Nearly all mentions of the experiment restrict themselves either to the verbal conclusion, or to a rehearsal of Michelson and Morley's approach to their data. Mis-statements are common, including assertions that the experiment was repeated many times, throughout the year, and at different altitudes. A number of authors appear to go into more detail, but in effect they simply give elaborated accounts of the Galilean transformation. There are various attempts to argue that the conventional Galilean analysis is mistaken, but I have made no attempt to collect these. Here I cite the small collection of articles I have found that are in some way relevant to the analysis I have presented. I have not included claimed analyses where there is credible evidence of misrepresentation or fraud.

Hicks (1902) carried out a forbiddingly complex analysis of the equations underlying the experiment, focusing on a putative shift in wavelength caused by reflection off moving mirrors. He claimed that an effect of the desired sort had been detected, but he provided no estimates of actual effects, relying on the data tables and an unlabeled graph to make his points. The linear effects were attributed to temperature change, with no evidence presented. He criticized the averaging over runs that was performed by Michelson and Morley. Hicks' article was later retracted due to a serious error, but it nevertheless stimulated Morley and Dayton Miller to try further experiments with the Michelson interferometer.

The most readable and accurate account of ether-drift experiments was by Miller (1933). Typical data from Michelson-Morley were plotted, using the two- and then three-times averaged figures. Miller noted that the effects were small but not zero, without any formal analysis. He included an exhaustive 3-dimensional geometric analysis of the consequences of the earth's velocity vector changing its orientation to the surface of the interferometer, primarily a diurnal effect, but also having a yearly effect. This formed the basis for his estimates of the velocity and direction of the earth through the ether from his own data. He noted that the strong diurnal effect would cause results to vary by time of day as well as terrestrial latitude, making it perhaps rather difficult to draw any conclusions from the Michelson-Morley experiment in this regard. The final conclusion from his data about the direction of the earth's movement placed it about 20° - 30° from being perpendicular to the interferometry tables of all experiments, which would imply that the detected fringe shift would be between $1/3$ and $1/2$ of the actual shift. Miller made many thousands of careful observations in varying locations and altitudes over a period of more than two decades. Interestingly, Miller noted that there were observable linear and period 2π components in all of his experiments, as well as in the Michelson-Morley data. He accounted for this as a result of a tiny rotation necessary for one of the mirrors in order to bring the fringe-shifts into view, an argument he copied from Hicks (1902), elaborated on in 1928, and which has been disputed by Valle (1934) who claims that mirror adjustments could only affect period π effects.

Notably with regard to the speeds estimated in Table 1, if Miller's directionality results are accurate then Michelson and Morley had actually detected a total speed somewhere between 12-22 km/sec or 18-33 km/sec. It is, of course, remarkable that Michelson and Morley would have missed the implications of three-dimensional geometry for their design, but there is independent evidence that they had also missed other things. In 1886 they had repeated Fizeau's experiment and confirmed his result, which was that the ether was partially dragged along by matter that it penetrated. They then interpreted this to mean that the ether was unaffected by matter moving through it, contradicting their own result while solidifying the argument to justify their eventual interferometry experiment. Even more curiously, Michelson evidently believed in the ether entrainment hypothesis for the rest of his life, without noting that its effect might have been to diminish the apparent interferometry results below his expectation. Even the somewhat vague published "analyses" of the interferometry results fail to note these various disturbing points.

Perhaps the earliest actual analysis was by done by Handschy (1982). Although he used a statistic that was said to be appropriate for detecting a periodic pattern, he applied it to the multiply-averaged data published by Michelson and Morley in which the effect had been artificially suppressed. He mentioned in passing the problem that would result from the earth's velocity vector not lying in the plane of the interferometer table, but he drew no conclusions from it, and gave it no particular attention. His conclusion, that Michelson and Morley were right, tends to ring rather hollowly.

Statistical methodologists occasionally use the data to make pedagogical points; for instance, Carver (1993) used them to make some points about hypothesis testing. From an "ANOVA" analysis, otherwise unspecified, he reported a $p\text{-value} < 0.001$. He was of the opinion that it would have been a mistake to report this, because

the eta squared value associated with the "significant" effect that direction had on the speed of light was .005. This means that if Michelson and Morley had been forced by journal editors and referees to do a test of statistical significance, they could have minimized its influence by reporting this effect size measure indicating that less than 1 % of the variance in the speed of light was associated with its direction. There seems to be no good excuse for ignoring effect size whenever a p value is reported.

He goes on; "Fortunately, Michelson and Morley did not corrupt the scientific method by testing the null hypothesis before they interpreted their data with respect to their research hypothesis." While it is true that Carver's purpose was didactic rather than an attempt at real analysis, and while it is also true that his point about hypothesis tests without effect estimates is right, he reports the wrong effect estimate (the eta squared statistic compares the squared proportion of variation in the signal Michelson and Morley were seeking to the total variation in the data, including the anomalous linear trend and period 2π effects). He mistakenly cited his estimate in terms of the speed of light, when it is in fact with respect to fringe shifts. He further missed the point that the purpose of the experiment actually was to detect *any fringe-shift effect*, regardless of its

magnitude, which is almost the definition of a situation in which a hypothesis test should lead the analysis. In this case, a profoundly misguided opinion in a statistics-oriented journal was injected into the literature, to be quoted by the unwary. The separation of statistical methodology from science in the last half of the previous century has done neither activity any good.

Cahill's program of research (2005) questioned both the null effect and the interpretation of the experiment. He showed the patterns of the individual runs, but carried out no statistical analysis. Cahill's interpretational argument accepted the FitzGerald-Lorentz contraction but argued that the index of refraction of the gas in which the experiment is done is a critical component of the design, and he showed the consequences of taking it into consideration.

A nearer approximation to a complete analysis was reported by Consoli et al. (2013, and summarized in Consoli & Pluchino 2019). Although they actually fitted the data with the appropriate sinusoids, instead of using the implied standard deviations of the estimates from those analyses, they computed the six amplitudes and then apparently reasoned on the basis of their average. In other words, the variation in the data that they used for precision was the between-runs component rather than the appropriate within-run components. Although their method was erroneous, they reported more analysis than any other ether-drift experimenters, mentioning the unexplained period 2π effect as well as the variation in phase shifts in the period π effects. Their overall conclusion was that Michelson and Morley had detected something, but in the end their reasoning was essentially based on the same argument about detectable limits as Michelson and Morley used.

The most recent analysis is from Correa et al. (2020), who conclude that no ether-wind was detected. While they spend some time considering the individual curves, they commit the same (threefold) averaging error that Michelson and Morley did. They distinguish between the curves for which the raw data show a rising or falling temporal pattern, although they propose no reason for doing so. As Figs. 1 and 2 above show, there is rather little evident homogeneity within the rising and falling subsets that would suggest a rationale for combining them. A sense of the analytic thinking can be gotten from a quote:

Any measurement falling within 2 standard deviations from the mean would likely fall within the random distribution...The similar shape of all three curves suggests a possible common pattern involving at least one harmonic and subharmonic oscillations. However, all points fall within 2 standard deviations of the mean for each curve type.

This is followed by an irrelevant t-test for a mean. Although the language and graphics suggest that a statistical analysis has been done, in fact it is little more than an elaboration of the impressionistic analysis done by the experimenters 133 years earlier.

Correa et al.'s analysis portrays an important misconception, which was perhaps more popular in Michelson and Morley's day than it is currently. When a lower detection limit is established for a measurement process, and then an average effect falls below it, the results are claimed to be due to random error. This ignores the fact that the detection limit for an average of uncorrelated

measurements is generally lower, and often quite a bit lower than the detection limit for a single observation. It is not uncommon to read proponents of the null interpretation of the Michelson-Morley experiment falling into this “detection limit” trap. In addition, as the quote above shows, Correa et al. in effect use the temporal variation in the data, which was the point of the experiment, as an argument for a “random” interpretation. That is, they use the pattern that was to be detected as an argument for saying nothing was detected, conveniently ignoring whether there was any pattern at all.

The distinction between the importance of the question that Michelson and Morley were asking and the quality of the subsequent analyses, is difficult to explain. Conventional methods of statistical inference have been ignored, and the tendency is to repeat the over-averaging performed by the original authors, despite the existence of better approaches. For perhaps the first half of the 20th century one might forgive analyses that fall below the standards of the 21st century, but the nearer the analyses come to the present day, the more the argument for forgiveness diminishes.

Why Re-analyze?

The Michelson-Morley experiment left a legacy of results. It is true that it may seem less than relevant today, due to further technical issues about whether its underlying theory was correct or not, as well as better experiments that are focused on the same question. But its short-term effects were considerable, and to some extent we are still affected by them.

Despite the modern view that the results were theory-shattering, reaction at the time was both slow and muted. Electromagnetics researchers had spent two generations building up a theory that depended fundamentally on the concept of the ether. Michael Faraday used the concept constantly in his descriptions of electrical and magnetic phenomena, and James Clerk Maxwell felt obliged to supply a mechanistic theory for his equations that absolutely required an ether. True, the longer the ether theory developed without any experiments on its properties, apart from its definitional characteristics with respect to light, the more tenuous and ambiguous the whole concept became. Before Michelson and Morley, the belief was that eventually a sufficiently precise apparatus would be developed to detect and study the ether. The Michelson interferometer was that device.

Together with the null interpretation, however, the very accuracy of the interferometer had a deleterious effect. If the best device for detecting the ether had failed, then what was the point of repeating the experiment? Evidently the answer at the time was that there was no point. Michelson drifted off into using his apparatus for establishing a light-based definition of the meter, for which he received the 1907 Nobel Prize, abandoning Morley in the process. Although various electromagnetics luminaries urged further experimentation, only Morley and Miller made any further attempts, and when Miller devoted the rest of his career to ether detection, he met hostile denunciation from the physics community. Perhaps the lesson is that false positive results tend to promote the

further repetitions that ultimately show them to be false, while false negative results discourage repetitive work, thus solidifying their apparent veracity.

Just after 1887, the quickest to mount a response against the null result were two researchers who profoundly believed in the ether, Hendrik Lorentz and George Francis FitzGerald. Both clearly understood the mathematics behind the measurements, and so around 1890 they were independently able to propose an alternative explanation. In terms of the description given above, they claimed that the two arms of the interferometer (lengths d_1 and d_2 in my notation) should be multiplied by

$$\sqrt{1 - \cos^2(\theta)v^2/c^2} \quad \text{and} \quad \sqrt{1 - \sin^2(\theta)v^2/c^2}$$

respectively, thereby canceling out any difference between the round-trip times, irrespective of the value of v . Since this amounted to saying that lengths contract in the direction of motion by a multiplier $\sqrt{1 - v^2/c^2}$, it has been called the FitzGerald-Lorentz contraction factor.

As he proposed the contraction explanation, Lorentz was attempting to develop a method of applying Maxwell's equations, which he supposed to be valid in the stationary ether, to an electron moving through the ether. He employed several slightly different space-time transformations for the purpose, the details of which are not important here. These transformations were in effect conversions between the coordinate system of the stationary ether and a coordinate system in which the moving electron was at rest. The dilemma faced by Lorentz was that while his discounting of the Michelson-Morley result depended on material objects contracting in the direction of motion, his transformations appeared to make objects expand in the direction of motion. In his final pre-relativity publication on the topic in 1904, Lorentz rehearsed his arguments for his transformation (which was essentially the modern Lorentz transformation) and then made the subtle implication that it involved a material contraction. It would not be fair to say that Lorentz misrepresented the situation, but it would be reasonable to say that he muddled it.

Within a year, Albert Einstein published his article introducing special relativity, in which he featured the Lorentz transformation as the correct coordinate conversion for uniformly moving systems. In a very short part of that article, Einstein provided an interpretation of the physical consequences of Lorentz transformations. He gave an example of a sphere, which he claimed squashed into an ellipsoid when it moved. He achieved this by the artifice of arbitrarily fixing time in one of the two systems, or alternatively by redefining the meaning of the spatial coordinates in one system, in either case making the supposed phenomenon irrelevant to the Michelson-Morley experiment, while at the same time promoting the mistaken concept of "length contraction".

It seems that some maneuver like this is the only way to make a contraction appear consistent with Lorentz transformation. To see this, let p_m denote the component in the direction of motion of a rod at rest in the moving system. Let p_s be the same component measured by the stationary observer, who takes time t_s to determine the two endpoints. Then Lorentz transformation is

$$p_m = \gamma(p_s - vt_s) \quad \text{where} \quad \gamma = 1/\sqrt{1 - v^2/c^2}$$

and so

$$p_s = p_m/\gamma + vt_s$$

In the Michelson-Morley setting, p_m is the component of the length of an interferometry arm in the direction of the earth's motion through the ether. p_s is the same component, as measured in the coordinate system of the ether. It is inherent in Lorentz transformations that the numerical value of any length depends in part on how long the measurement takes, which in this case is t_s in the ether's coordinates. Consequently, in order for Lorentz to have the ether-length of the arm shorter than the interferometry length by the FitzGerald-Lorentz contraction factor, the ether-measurement would have to be instantaneous.

To justify the FitzGerald-Lorentz contraction on the basis of Lorentz transformation, it seems necessary to say that the isotropy of light speed in the ether requires that the relevant lengths be measured in the ether's coordinate system, and that they be measured instantaneously. If this is what Lorentz meant, then he was not only retaining the ether as the medium in which light speed is isotropic, but he was promoting the ether-coordinates for space and time to be the relevant ones for physical measurements in electrodynamics, and in addition requiring that ether-lengths be determined instantaneously. While this is the only explanation I can see for Lorentz's argument, I admit that it is less than obvious in his writings. It would have been simpler to apply the Lorentz transformation directly to the initial Galilean equation and show that it made the round-trip times equal, but Lorentz did not take that path.

After 1900 there were several unsuccessful experimental attempts to verify the length contraction hypothesis. Their collective failure left physicists perplexed at the Michelson-Morley result. As we have seen, it is the positive result of that experiment that furnishes a denial of length contraction.

Recall that Lorentz was making his argument to claim that the Michelson-Morley experiment could not detect movement through the ether, as a consequence of its design. When Einstein indirectly repeated Lorentz's argument, he was in effect agreeing that Michelson and Morley's result provided no support for special relativity. It is conceivable that he saw this, which would account for why he failed to mention the experiment in his 1905 article, when he must have known about it from Lorentz's 1895 book.

Generations of relativists have repeated the length contraction story as a fundamental aspect of special relativity. In fact, when two uniformly moving observers measure a rod, the one who determines the positions of the two ends more quickly will find the shortest length. This is a consequence of the fundamental invariant of Lorentz transformations, and has nothing to do with the movement or stationarity of the rod.

The length contraction meme is one reason for continuing to debate Michelson-Morley in the current era. Mistakes made in an otherwise superb experiment caused the leading electromagnetics researcher of the age to create a subtle and probably incorrect explanatory theory, which was picked up by the creator of special relativity and passed along to generations of physicists.

If the observed results of the Michelson-Morley experiment are taken to have been positive, as argued in the analysis here, then there could be several results. First, the null result is supposed to have disproved the existence of the ether, paving the way for Einsteinian relativity. Whether or not this is historically accurate, it would be physically inaccurate. The constant speed of light required by Einstein occurs in a vacuum, but the Michelson-Morley experiment was conducted in air. If the experiment were null, it would imply that light speed is constant (and isotropic) in air. But a physically consistent set of Lorentz transformations can preserve only one speed, and if it were the air speed of light then it could not be the vacuum speed of light. In this sense a null Michelson-Morley experiment would disprove Einsteinian relativity, although it would have nothing to say on the obvious question about what would be so special about air. Secondly, a positive Michelson-Morley experiment would invalidate the counter-effect of the FitzGerald-Lorentz contraction in air, but this would not necessarily invalidate it in a vacuum, where Einsteinian relativity claimed it to exist.

Thirdly, one can only speculate whether a proper analysis of the Michelson-Morley experiment at the time might have inspired broader and faster development of subsequent interferometry experiments, especially ones that could have been oriented to a fuller understanding of what exactly was being detected.

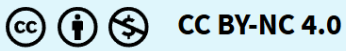
Conclusions

There do not seem to have been any proper analyses of the data produced by Michelson and Morley. This is partly due to the fact that they gave only averaged summaries of unpublished individual runs, and partly a failure to take a detailed look at the data they did publish. The analysis here shows unquestionable evidence of the period π effect they were seeking. It also shows unquestionable linear and period 2π effects that Michelson and Morley ignored. Since there are apparent anomalies in the data, it is unfortunate that the experimenters did not follow up their four days of data collection with further runs made under different controlled conditions, or in other locations and at various altitudes. Michelson and Morley's declaration of a null result was premature, and led to the proposal of a problematic length contraction hypothesis that has survived to become a part of special relativity.

References

- Articolo GA. The Michelson-Morley experiment and the phase shift upon a 90° rotation. *American Journal of Physics* 1969;37(2):215-216
- Cahill RT. The Michelson-Morley 1887 experiment and the discovery of absolute motion. *Progress in Physics* 2005;3:25-29 and arXiv:physics/0508174v1
- Cassini A, Levinas ML. Einstein's reinterpretation of the Fizeau experiment: how it turned out to be crucial for special relativity. *Studies in History and Philosophy of Modern Physics* 2018; 23 August:1-18

- Conference on the Michelson-Morley Experiment. *The Astrophysical Journal* 1928;68(5): at p. 397-398
- Consoli M, Pluchina A. The classical Michelson-Morley experiments: a new solution to an old problem. *Atti della Accademia Peloritana dei Pericoranti* 2018;96(S1,A2)
- Cartmel WB. A simple means of checking the Michelson-Morley experiment (letter). *Nature* 1937; 139:110
- Carver R. The case against statistical significance testing, revisited. *The Journal of Experimental Education* 1993;61(4): 287-292
- Consoli M, Matheson C, Pluchino A. The classical ether-drift experiments: a modern re-interpretation. arXiv:1302.3508v2, 2013.
- Consoli M, Pluchino A. *Michelson-Morley Experiments: An Enigma for Physics and the History of Science*. World Scientific: Hackensack NJ, 2019
- Correa PN, Correa AN, Pratt D, Askanas M. Re-examination of the experimental evidence for a nonzero aether drift. Part 1: Michelson-Morley-type experiments 1881-1930. *Journal of Aetherometric Research* 2020;3(1):1-41
- Handschy MA. Re-examination of the 1887 Michelson-Morley experiment. *American Journal of Physics* 1982;50(11);987-990
- Hicks WM. On the Michelson-Morley experiment relating to the drift of the Aether. , *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 1902;3(13): 9-42
- Lorentz HA. Electromagnetic phenomena in a system moving with any velocity smaller than that of light. *Proceedings of the Royal Netherlands Academy of Arts and Sciences* 1904;6:809-831
- Lorentz HA. *Versuch einer Theorie der electrischen und optischen Erscheinungen in bewegter Körpern*. E.J. Brill: Leiden, Netherlands, 1895 (Attempt at a theory of electrical and optical observations for moving bodies.)
- Michelson AA., Morley EW. On the relative motion of the earth and the luminiferous ether. *American Journal of Science* 1887;203:333-345.
- Miller DC. The ether-drift experiment and the determination of the absolute motion of the earth. *Review of Modern Physics* 1933;5:203-254
- Múnera HA. Michelson-Morley experiments revisited: systematic errors, consistency among different experiments, and compatibility with absolute space. *Apeiron* 1998;5(1-2):37-54.
- Valle G. Full period effect in Miller's ether drift experiment (letter). *Nature* 1934; 133(3368): 758-759.



Creative Commons Attribution-NonCommercial 4.0 International

This license requires that reusers give credit to the creator. It allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, for noncommercial purposes only.