5061A) lasting several days to weeks gave normally distributed, zero-center closure times with a spread of about 60 nanoseconds per day of trip [G. M. R. Winkler, in *Proceedings of the Second Annual Precise Time and Time Interval Strategic Planning Meeting*, 10 to 11 December 1970, vol. 1 (available from Technical Officer, U.S. Naval Observatory, Washington, D.C. 20390)].
8. H. J. Schaefer, *Science* **173**, 780 (1971).
9. Equation 4 includes the effect discussed by W. J. Cocke [*Phys. Rev. Lett.* **16**, 662 (1966)]. Clocks at rest on the earth's surface (at average sea level) keep the same relativistic time independently of latitude differences. The effect of the difference in surface speed at different latitudes is canceled to lowest order by a corresponding effect from the difference in surface potential owing to the oblate figure of the earth.
10. Preliminary results of this work have been reported: J. C. Hafele, in *Proceedings of the Third Annual Precise Time and Time Interval Strategic Planning Meeting*, 16 to 18 November, 1971, vol. 1 (available from Technical Officer, U.S. Naval Observatory, Washington, D.C. 20390).
11. "Nanosecond" is abbreviated as "nsec" in this and the following report because this is the usage dictated by the editorial policy of *Science*. The U.S. Naval Observatory adheres to the "International System of Units" in which the abbreviation for "nanoseconds" is "ns."
12. We thank Pan Am, TWA, and AA for their cooperation, and particularly their flight crews for providing the necessary flight data. We also thank R. Agricola of Pan Am and J. Clay of TWA, who were sent as escorts at no extra cost. The Office of Naval Research provided financial support. H. N. Acrivos and J. McNeece of the Time Service Division of the U.S. Naval Observatory gave advice and assistance in scheduling the flight. We thank the Washington University Computer Center for computer time.
* Present address: Research Department, Caterpillar Tractor Co., Peoria, Illinois 61602.
24 March 1972; revised 8 June 1972

Around-the-World Atomic Clocks: Observed Relativistic Time Gains

Abstract. *Four cesium beam clocks flown around the world on commercial jet flights during October 1971, once eastward and once westward, recorded directionally dependent time differences which are in good agreement with predictions of conventional relativity theory. Relative to the atomic time scale of the U.S. Naval Observatory, the flying clocks lost 59 ± 10 nanoseconds during the eastward trip and gained 273 ± 7 nanoseconds during the westward trip, where the errors are the corresponding standard deviations. These results provide an unambiguous empirical resolution of the famous clock "paradox" with macroscopic clocks.*

In science, relevant experimental facts supersede theoretical arguments. In an attempt to throw some empirical light on the question of whether macroscopic clocks record time in accordance with the conventional interpretation of Einstein's relativity theory (*1*), we flew four cesium beam atomic clocks around the world on commercial jet flights, first eastward, then westward. Then we compared the time they recorded during each trip with the corresponding time recorded by the reference atomic time scale at the U.S. Naval Observatory, MEAN(USNO) (*2*). As was expected from theoretical predictions (*1*), the flying clocks lost time (aged slower) during the eastward trip and gained time (aged faster) during the westward trip. Furthermore, the magnitudes of the time differences agree reasonably well with predicted values, which were discussed in the preceding report (*1*). In this second report, we present the time difference data for the flying ensemble, and explain the methods by which the relativistic time differences were extracted.

The development of compact and portable cesium beam atomic clocks (*3*) permits a terrestrial test of relativity theory with flying clocks. The fundamental unit of time interval, the second, is now by definition equal to 9,192,631,770 accumulated periods of the frequency of the atomic transitions of an "ideal" cesium beam frequency standard (*2, 3*). Because these clocks are regulated by the frequency of a natural atomic transition, a particularly well defined hyperfine transition in the ground state of the ^{133}Cs atom, they approach the ideal standard clock of relativity theory.

Table 1. Observed relativistic time differences from application of the correlated rate-change method to the time intercomparison data for the flying ensemble. Predicted values are listed for comparison with the mean of the observed values; S.D., standard deviation.

Clock serial No.	$\Delta\tau$ (nsec)	
	Eastward*	Westward
120	− 57	277
361	− 74	284
408	− 55	266
447	− 51	266
Mean ± S.D.	− 59 ± 10	273 ± 7
Predicted ± Error est.	− 40 ± 23	275 ± 21

* Negative signs indicate that upon return the time indicated on the flying clocks was less than the time indicated on the MEAN(USNO) clock of the U.S. Naval Observatory.

However, no two "real" cesium beam clocks keep precisely the same time, even when located together in the laboratory, but generally show systematic rate (or frequency) differences which in extreme cases may amount to time differences as large as 1 μsec per day. Because the relativistic time offsets expected in our experiments are only of the order of 0.1 μsec per day (*1, 4*), any such time divergences (or rate differences) must be taken into account.

A much more serious complication is caused by the fact that the relative rates for cesium beam clocks do not remain precisely constant. In addition to short term fluctuations in rate caused mainly by shot noise in the beam tubes, cesium beam clocks exhibit small but more or less well defined quasi-permanent changes in rate. The times at which these rate changes occur typically are separated by at least 2 or 3 days for good clocks. Some clocks have been observed in the laboratory to go as long as several months without a rate change (*2, 5*).

These unpredictable changes in rate produce the major uncertainty in our results. Because of the nature of these changes, however, their effect on the observed time differences can be removed to a large extent in the data analysis. Under normal conditions changes in relative rates occur independently, that is, there are no known systematic correlations between rate changes of one clock and those of another. Consequently, the chance that two or more clocks will change rate by the same amount in the same direction at the same time is extremely remote. Because of the random and independent character of these rate changes, the long-term average rate of an ensemble of clocks is more stable than the rate of any individual member.

Starting at 0^h U.T. on 25 September 1971, we recorded more than 5000 time differences during the data period. Figure 1 shows the time difference data relative to MEAN(USNO) for the entire data period, which lasted 636 hours. The labels in Fig. 1 are the serial numbers of the corresponding clocks, and the traces give the measured differences in time between the corresponding clock and MEAN(USNO). Of course, no comparisons with MEAN(USNO) were possible during the trips. Exactly the same electronic arrangement was used for all time intercomparison measurements, and time differences were measured with an electronic time interval counter to the nearest nanosecond (*5*).

The axis of abscissas in Fig. 1 is the running time for the data period, and, because at any given instant all clocks involved read the same time to within much better than the nearest second, it is immaterial whether the running time is the time of MEAN(USNO) or any member of the flying ensemble. The times at which the time differences were measured are certain to within about $\pm$ 5 minutes.

The tangential slope of a trace in Fig. 1 equals the instantaneous relative rate for the corresponding clock. For example, a positive slope of 30° means that the clock was gaining on MEAN(USNO) by about 10 nsec per hour. Figure 1 shows that no member of the flying ensemble maintained a constant rate relative to MEAN(USNO) during the entire data period. Spontaneous rate changes between periods of virtually constant rate are apparent.

Figure 2 gives magnified views of the average time difference data of Fig. 1 immediately before and after each trip. These magnified views clearly show that, on the average, the flying clocks lost time during the eastward trip and gained time during the westward trip. Hence the predicted east-west directional asymmetry (*1*) is qualitatively confirmed.

We have extracted relativistic time differences from our data with two different methods, the "average rate" method and the "correlated rate-change" method (*6*). Although the average rate method is less reliable, it is the simpler of the two methods and we will describe it first.

Figure 2 shows least-squares straight-line fits to the average data of Fig. 1 for an interval of 25 hours immediately before and after each trip. The slopes for these lines give the mean rates for this interval (electronic noise causes the short-term fluctuations about the local mean rate). If we assume that only one rate change occurred in the "average" clock during each trip, and that it occurred at the midpoint, relativistic time differences follow from extrapolation with the average of the initial and final rates. The slopes of the dashed extrapolation lines in Fig. 2 are equal to the average of the slopes of the linear fits, and with this fitting interval the average rate method gives a time loss of 66 nsec for the eastward trip and a time gain of 205 nsec for the westward trip.

Reliability of results with the average rate method, however, depends on the unlikely chance that only one rate change occurred during each trip and that it occurred at the midpoints. Furthermore, there is no obvious method for estimating the experimental error. Nevertheless, the average rate method does produce convincing qualitative results.

The correlated rate-change method produces more reliable quantitative results because, with this method, rate changes for individual members of the flying ensemble during each trip are taken into account. Because rate changes relative to MEAN(USNO) are uncorrelated whether or not there is an actual comparison with MEAN(USNO), records of the time differences between each member of the flying ensemble permit reconstruction of the rate-change history for each clock during the trips.

The correlated rate-change method is basically an old technique (*7*). It was first used with atomic clocks to supplement a frequency averaging method that employed very low frequency radio transmissions (*8*), and then became a main element in the generation of the Naval Observatory atomic time scale. The computer algorithm (*2*) currently used to construct MEAN(USNO) evolved from studies of time scales constructed with the correlated rate-change method. Because the stability of a single cesium beam clock would probably be inadequate to permit an absolutely unambiguous detection of the expected relativistic effects, we used four flying clocks so that the correlated rate-change or other similar methods could be ap-

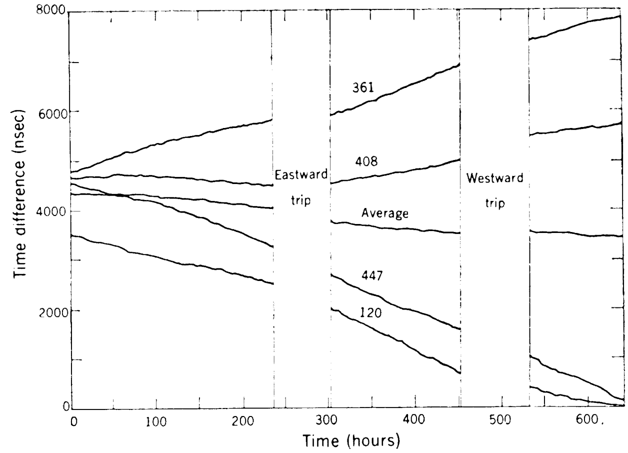

Fig. 1. Time differences between each clock in the flying ensemble and the Naval Observatory clock MEAN(USNO) at hourly intervals throughout the data period. The hour count along the abscissa begins at 0^h U.T., 25 September 1971. The trace of each clock is labeled with the clock's serial number: and the trace labeled *Average* is the average of the four time differences.

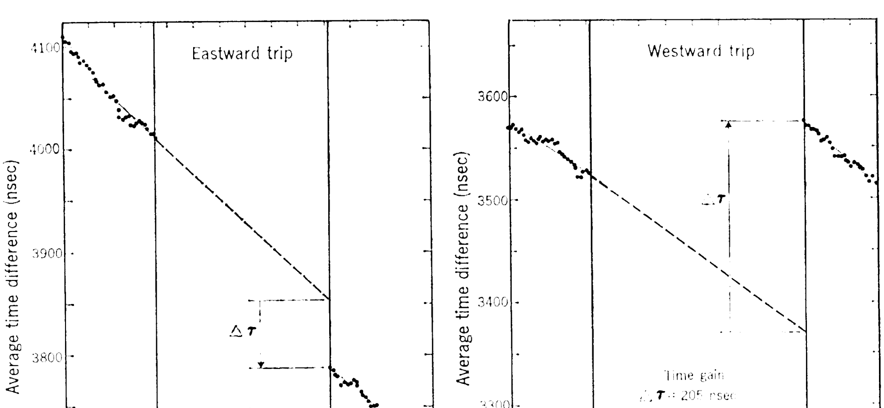

Fig. 2. Magnified views of the average time difference data of Fig. 1 in the vicinity of (a) the eastward trip and (b) the westward trip. The small step down for the eastward trip indicates an average time loss, and the somewhat larger step up for the westward trip indicates an average time gain. The indicated time differences follow from application of the average rate method with a fitting interval of 25 hours.

plied to the data. In retrospect, it is clear that the use of only one or two flying clocks would have substantially decreased the feasibility of our experiments.

A hypothetical example may help to follow the logic of this method. Of two clocks, C_1 and C_2, suppose clock C_1 is chosen as the reference clock and that at successive readings of C_1 the time differences, $\tau_2-\tau_1$, are plotted on a graph. A change in slope in the resulting time trace represents a change in rate, but with only two clocks there is no way of telling whether C_1 or C_2 actually changed. But suppose the time differences for a third clock, $\tau_3-\tau_1$, are added to the plot. If at a certain instant there is a correlated change in slope, that is, there occurs in both traces a change in slope by the same amount in the same direction at the same time, then by induction one concludes that it is the clock common to both comparisons which actually changed rate (C_1 in this hypothetical case). Therefore, time intercomparison data for an ensemble of three or more clocks permit reconstruction of their rate change behavior relative to MEAN(USNO) without actual comparisons with MEAN(USNO), and the greater the number of clocks in the ensemble the more redundant and self-consistent the procedure becomes. Relativistic effects are not directly evident in the internal time intercomparison data because such changes apply equally to all members of the flying ensemble.

We recorded the differences in the times indicated by each member of the flying ensemble at regular intervals before, during and after each trip, that is, throughout the entire data period. An analysis of these data revealed the times and magnitudes for correlated rate changes during each trip. Thus significant rate changes were identified and ascribed to each clock (*9*). A piecewise extrapolation of the time trace for each clock relative to MEAN(USNO), with proper accounting for these identified rate changes, then produced the relativistic time differences listed in Table 1.

The agreement between the mean of the measured values and the predicted values in Table 1 is very satisfactory. In addition, the consistency among the measured values is striking. For the westward trip, the standard deviation is less than 5 percent of the mean. According to the statistical theory of error, the standard deviation is a valid indication of the precision of the measurement. However, the number of measured values is too small for a good statistical analysis.

Another estimate of the probable error in the final time differences follows from consideration of the effects of possible errors in the rates and rate changes used in the piecewise extrapolations. A preliminary study of the effect of adjustments in the calculated rates on the residuals between the calculated and measured time traces showed obvious distortions when deviations greater than 0.4 nsec per hour were arbitrarily introduced. This value provides an estimate of the maximum possible error in the rates actually used. The product of this rate uncertainty with the duration of the trips gives an estimated final error of less than $\pm$ 30 nsec, which is about three times the standard deviation and probably represents an overestimate.

An estimate of probable systematic errors would require specific knowledge of (nonrelativistic) systematic environmental effects on the clock rates, such as a systematic effect of temperature or pressure differences. Of course, any known or unknown systematic effect would apply equally to each of the flying clocks, for they all experienced the same ambient conditions, so such effects would not be revealed by the correlated rate-change analysis. Previous studies have shown that moderate variations in temperature and pressure, such as those experienced during the trips, do not produce systematic rate changes. Although temperature or pressure changes sometimes cause individual random and unpredictable changes (*10*), such random and uncorrelated changes in rate do not contribute to systematic errors. The clocks are highly resistant to impulse accelerations (*10*). They are triply shielded against the earth's magnetic field, and we found no effect of different orientations in this field in the laboratory (*5*). In fact we have been unable to discover, either by theoretical considerations or by independent experimental tests, any sources of significant systematic errors.

In conclusion, we have shown that the effects of travel on the time recording behavior of macroscopic clocks are in reasonable accord with predictions of the conventional theory of relativity, and that they can be observed in a straightforward and unambiguous manner with relatively inexpensive commercial jet flights and commercially available cesium beam clocks. In fact, the experiments were so successful that it is not unrealistic to consider improved versions designed to investigate aspects of the theory that were ignored in the predicted relativistic time differences (*1*). In any event, there seems to be little basis for further arguments about whether clocks will indicate the same time after a round trip, for we find that they do not.

J. C. HAFELE*

Department of Physics,
Washington University,
St. Louis, Missouri 63130

RICHARD E. KEATING

Time Service Division,
U.S. Naval Observatory,
Washington, D.C. 20390

References and Notes

1. J. C. Hafele and R. E. Keating, *Science* **177**, 166 (1972).
2. G. M. R. Winkler, R. G. Hall, D. B. Percival, *Metrologia* **6**, 126 (1970); A. G. Mungall, *ibid.* **7**, 146 (1971). We refer herein to the local atomic time scale AT(A.1 MEAN USNO) as simply MEAN(USNO).
3. R. E. Beehler, R. C. Mockler, J. M. Richardson, *Metrologia* **1** (No. 3), 114 (1965); L. N. Bodily and R. C. Hyatt, *Hewlett-Packard J.* **19**, 12 (December 1967).
4. J. C. Hafele, *Amer. J. Phys.* **40**, 81 (1972).
5. ———, in *Proceedings of the Third Annual Precise Time and Time Interval Strategic Planning Meeting*, 16 to 18 November 1971, vol. 1 (available from Technical Officer, U.S. Naval Observatory, Washington, D.C. 20390).
6. The term "correlated rate-change" is meant to imply that the presence of cross correlations in the data permits identification of the clock which changed rate, and not that the rate changes are intrinsically correlated.
7. The idea that a rate change in the time reference is indicated by an apparent correlation among changes in relative rates for a number of otherwise uncorrelated periodic phenomena was used long ago to suggest variability in the earth's rotational motion: S. Newcomb, *Amer. J. Sci. Arts* **8**, 161 (1874); W. de Sitter, *Bull. Astron. Inst. Neth.* **4**, 21 (1927).
8. W. Markowitz, in *Proceedings of the 13th Annual Symposium on Frequency Control* (Doc. No. PP146982, Library of Congress, Washington, D.C., 1959), p. 316; ——— and R. G. Hall, in *Proceedings 15th Annual Symposium on Frequency Control* (Doc. No. AD285086, National Technical Clearinghouse, Port Royal, Va., 1961), p. 168.
9. The time intercomparison data showed that clock 120 changed rate three times, 361 changed three times, 408 changed twice, and that 447 changed rate once during the eastward trip. For the westward trip, clock 120 changed once and 361 changed four times. No significant changes in rate were found for clocks 408 and 447 during the westward trip.
10. E. Hafner, in *Proceedings of the Third Annual Precise Time and Time Interval Strategic Planning Meeting*, 16 to 18 November 1971, vol. 1 (available from Technical Officer, U.S. Naval Observatory, Washington, D.C. 20390).
11. We thank the Office of Naval Research for providing funds for evaluating the performance and behavior of atomic clocks under flight conditions. We appreciate the help of G. M. R. Winkler, R. G. Hall, and D. B. Percival. We especially thank R. Stoneburner and A. Kummer for helping with the clock assemblies. We thank the U.S. Naval Observatory for providing the atomic clocks. We also thank the Hewlett-Packard Co., the manufacturer of the clocks, particularly L. S. Cutler and A. Walcek, for their cooperation.
* Present address: Research Department, Caterpillar Tractor Co., Peoria, Illinois 61602.

24 March 1972; revised 8 June 1972