Astron. Astrophys. 77, 121–127 (1979)

ASTRONOMY AND ASTROPHYSICS

Solar Activity Cycle during Classical Antiquity

R. Stothers

NASA Goddard Institute for Space Studies, Goddard Space Flight Center, 2880 Broadway, New York, NY 10025, USA

Received August 28, 1978

Summary. Early accounts of phenomena that may be identified as auroral displays have been abstracted from reports of unusual celestial prodigies in the classical literature. An extensive catalog of ancient aurorae and a new mathematical method of analyzing fragmentary time series of observations have been used to demonstrate, provisionally, that an auroral cycle actually existed in antiquity, at least during the 2nd century BC, and that it had an average length and amplitude comparable with those of the modern auroral cycle. On the reasonable supposition that solar activity has always been the factor responsible for aurorae, it can be concluded that the solar cycle two millennia ago was very similar to what it is today.

Key words: aurora borealis — solar cycle — statistical analysis — ancient observations

I. Introduction

In a now famous monograph on the aurora borealis, de Mairan (1733) first drew modern attention to reports of certain transient sky phenomena found in the works of some of the better known classical authors. One of these authors was Pliny the Elder, who wrote (AD 77):

> There are also stars that suddenly come to birth in the heaven itself; of these there are several kinds.... "beams", in Greek *dokoi*, for example one that appeared when the Spartans were defeated at sea and lost the empire of Greece. There also occurs a yawning of the actual sky, called *chasma*, and also something that looks like blood, and a fire that falls from it to the earth – the most alarming possible cause of terror to mankind.... A light from the sky by night, the phenomenon usually called "night-suns", was seen in the consulship of Gaius Caecilius and Gnaeus Papirius and often on other occasions causing apparent daylight in the night.

These remarkable phenomena de Mairan concluded were displays of the aurora borealis. The simple forms and colors that Pliny and other ancient authors described call to mind the rare modern aurorae that occur at low geomagnetic latitudes. But historical studies of the aurora have left largely unsurveyed the available classical record. It is true that four catalogs of aurorae from this record have been compiled and published (Frobesius, 1739; Fritz, 1873; Schove, 1948; Link, 1962), but because only token recourse to the original literature was made by the modern compilers, who chose to rely on a number of incomplete and heterogeneous 16th and 17th century collections of ancient prodigies, the catalogs abound in unfortunate errors of mistaken contexts, incorrect dates, and numerous omissions. A detailed criticism of these catalogs is given in a paper to appear soon in *Isis*, to which the reader is réferred (Stothers, 1979). In their place I have compiled a new, exhaustive catalog directly from the ancient literature itself.

With such a catalog it becomes possible, for the first time, to analyze the data statistically in order to determine whether or not an auroral cycle existed during antiquity, and, if so, what its length and amplitude were. These results then immediately disclose something about the solar activity cycle, if the reasonable assumption is made that solar activity was then, as now, the controlling factor in producing aurorae. A statistical treatment of the ancient data seems to be unavoidable if for no other reason than that a certain number of mistaken identifications of aurorae are bound to have been made by any compiler, no matter how conscientious. Previous authors, such as Schove (1955), Nicolini (1963, 1976), and Link (1964), did not use a statistical approach. They essentially assumed an auroral cycle approximately equal to the 11-year cycle of solar activity that exists today. Eddy (1976), among others, has recently questioned whether this solar cycle actually existed before the 17th century. I hope to show not only that it did, but also that it had virtually the same characteristics during the 2nd century BC as it has today.

II. The New Catalog

It is clear from Pliny's brief description that ancient aurorae are to be sought among accounts of unusual celestial portents and prodigies. A useful procedure, therefore, is to retain the ancient classifications of sky phenomena for the sake of uniformity, since these classifications seem to have kept their meanings unchanged throughout antiquity. Phenomena that I tentatively class as auroral are the following: "chasms" (X), "sky fire" (SF), "night suns" (NS), "blood rain" (BR), "milk rain" (MR), "beams" (B), "pillars" (P), aurora-like "torches" (T), and aurora-like "comets" (K). Objects in these categories have, of course, some possibility of being confused occasionally with ordinary comets, meteoric fireballs, zodiacal light, and so on. Therefore, the ultimate test of the correctness of the classifications must be based on an objective statistical analysis.

The previous auroral catalogs lacked a practical classification system. In fact, the last six of the nine categories listed above were largely ignored in the earlier work. Schove (1955), in his second paper, may have included some of these six, but he has published neither the classifications nor the sources of his reports – only the dates. In any case, there are only six auroral years between 223 and 91 BC that he regards as suitable for mathematical analysis. In this interval I find 36 suitable years.

An abridged version of the new catalog is given in Table 1 for convenience. Details that are not of direct scientific interest, such as source references, questionable reports, and dating methods, are to be found in my *Isis* article. Probably the individual dates listed are accurate to within ±1 yr. Continuity of the auroral record is poor in some centuries, but this can be explained adequately within the context of the fragmentary nature of the historical records. Climatic factors are almost certainly not responsible for the large gaps, as the ancient historians do not refer to whole decades of overcast skies; in fact, a stretch of even months of uninterrupted dimness of the sun was enough to cause surprised comment in ancient times. A high degree of geographical homogeneity characterizes the auroral catalog, since all the listed events refer to areas along the northern rim of the Mediterranean basin. In the almost continuously documented period 223–91 BC, the reporting area is narrowed down to central Italy and the source material is also very uniform, being taken, directly or indirectly, from Livy's great annalistic history of Rome (written between 27 BC and AD 17). In Livy's history the portents noted each year in public Roman territory are carefully enumerated, for the reason, he says (*Ab urbe condita* 43.13), that in those early days of Rome they were regarded as serious religious matters related to the welfare of the state. Their veracity was, it would seem, diligently checked into by the authorities because the rites necessary to expiate them were costly and time-consuming. I have no reason to doubt the trustworthiness of Livy's sources for these portents.

It is a happy accident of history that ancient civilization in Europe developed around low geomagnetic latitudes. For, if modern aurorae are admitted as a provisional guide to ancient ones, a small (but not negligible) number of aurorae are expected to have been easily visible in any decade down through the centuries at the low latitudes of the Mediterranean countries (Fritz 1881). In northern Europe, however, every year would have been an auroral year. In that case, unless the number of aurorae per year had been recorded (and it is known that such statistics were not kept during medieval times in the North), no reliable information about the ancient auroral cycle could have been derived today. On the other hand, at more southerly geomagnetic latitudes, only an extremely rare aurora every few decades would have been easily noticeable. It is, therefore, of critical importance that Table 1 is found to contain just a small number of auroral reports each decade during the best documented period 223–91 BC.

III. A Mathematical Method of Time Series Analysis

The data in Table 1 comprise a discrete time series of observations which could be seriously affected both by incompleteness and by a number of mistaken identifications. Periodicities (or mean cycle times) in poor, noisy records of this kind are usually looked for in one of two ways. The first way is to accumulate the observations into "bins", so that a power spectrum, or periodogram, analysis can be performed. In the present instance, the bin size necessary to obtain a statistically significant sample would be very large, probably larger than the anticipated period. The second way of locating the dominant period is to fit the observations to an equation of the form

$$t_{\max} = t_0 + nP, \quad (1)$$

where $t_{\max}$ is the time of the nth maximum in the observations, and P and t_0 are the period and epoch to be determined. In practice, it is necessary first to bin the observations so that the times of maximum can be determined, and then to assign an

Table 1. Catalog of ancient aurorae mentioned in classical literature

Year	Category	Year	Category	Year	Category
		BC 166	NS, BR	BC 95	MR
BC 467	SF, B	163	SF, NS, MR	94	SF
373	B, T	162	SF	93	X, SF
349	X, BR, SF	147	SF	92	MR
344	X, SF, T	134	NS, BR	91	X, BR
223	SF, NS	130	MR	63	SF, B, K
217	X, SF	128	BR	49	SF, BR
214	BR	125	MR	48	B, P
209	MR	124	MR	42	SF, NS
206	NS	118	MR	32	T
204	NS	117	MR	BC 17	T
200	SF	114	BR, MR	AD 9	SF, P, K
198	SF	113	SF, NS	14	SF, BR, K
197	NS	111	MR	50	SF
183	BR	108	MR	54	BR
181	BR	106	BR, MR	76	K
172	BR	104	BR, MR	196	SF
169	SF	102	NS, BR	333	SF

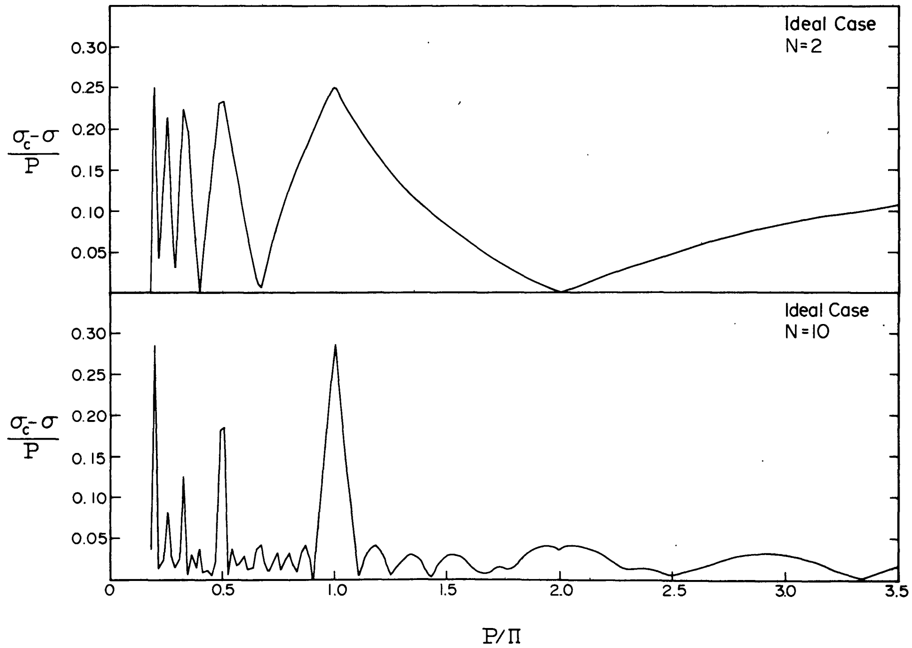

Fig. 1. Spectra of the residuals index $(\sigma_c - \sigma)/P$ for two ideal cases, where Π is the true period

integer n to each estimated $t_{\max}$ value on the basis of some assumed period P and assumed epoch t_0. Finally, a formal least-squares fit to Eq. (1) is made to find the definitive values of P and t_0. In suitable situations, graphical or tabular variants of this method can be devised. These procedures, however, do not generally determine the true period; they merely find the best-fitting period that lies in the immediate neighborhood of the assumed period, or, at most, they show that the assumed period is not grossly incompatible with the observations (Nicolini, 1976; Schove, 1955; Link, 1964).

In order to remove these inadequacies, I have used the following method of analysis, which should be generally applicable to any time series that is composed of sampled time intervals (not necessarily equally spaced) during which the only known information is whether an observation was or was not generated. First, a suitable array of trial periods is set up. For each trial period, P, a selection of trial epochs, running from t_0 to $t_0 + P$, is made. Then Eq. (1) is used to generate a continuous sequence of predicted times of maximum for each combination of trial period and trial epoch. A particular observation is assigned, within a given generated sequence, to the nearest predicted time of maximum. The difference between the observed time and the predicted time of maximum will be denoted $d_i (i = 1, 2, \ldots, N$, for N observations). Then, for each sequence, the following rms residual is computed.

$$\sigma = \left(\sum_{i=1}^{N} d_i^2/N\right)^{1/2}. \qquad (2)$$

A straightforward criterion of best fit is the minimum value of σ/P that is found when all the sequences have been analyzed in the foregoing fashion.

A simple mathematical example will suffice to demonstrate the method and to provide a useful paradigm by which more sophisticated physical results can be interpreted. Consider a sequence of N numbers with a constant difference Π between successive numbers. In this simple example, the true period is known in advance to be Π, but an unprejudiced period analysis will be performed for illustration. At each trial period the smallest value of σ/P is selected from all the values that have been generated with various trial epochs. Since it is more conventional to have an *increasing* index to measure probability of a good fit, I shall henceforth adopt the variable $(\sigma_c - \sigma)/P$, where the constant term is given by

$$\sigma_c/P = [(N^2 - 1)/12\, N^2]^{1/2}. \qquad (3)$$

This constant can be shown to represent the continuous part of the spectrum of residuals σ/P that lie between $P = 0$ and $P = N\Pi$. (For $P > N\Pi$, σ/P is proportional to P^{-1}.) Over the complete range of N, the continuum value varies only slightly: $0.250 \leqq \sigma_c/P \leqq 0.289$ for $2 \leqq N \leqq \infty$.

A period analysis for the present example is shown in Fig. 1. Here the trial periods run from 0.20 Π to 3.50 Π, in steps of 0.02 Π. The analysis has been performed for $N = 2$ and 10. In general, the spectrum of $(\sigma_c - \sigma)/P$ values takes the form of a zero-valued continuum broken by a series of high and low maxima. Wherever P equals the true period Π or any harmonic thereof ($\Pi/2$, $\Pi/3$, etc.) the spectrum has a high, sharp maximum. (On account of the limited spectral resolution in Fig. 1, not all the harmonics can be seen). Integer multiples of the true period (2Π, 3Π, etc.) are associated with wider, lower maxima. In general, as N is increased, all the maxima become narrower, and, because the noisiness of the spectrum is also reduced, they are more sharply defined.

From this idealized case certain general inferences can be drawn. First, since the continuous part of the spectrum of residuals σ/P has proven to be nearly independent of N, it is expected to be relatively independent of the pattern of observations in more complicated situations. Therefore $(\sigma_c - \sigma)/P$ remains a good index. Second, if there are any gaps in the series of observations (consider Fig. 1 with a harmonic of Π as the true period), these gaps are expected to be far less instrumental in lowering the spectral peak at the true period than are the accidental intrusions of spurious observations (consider Fig. 1 with an integer multiple of Π as the true period). Third, the

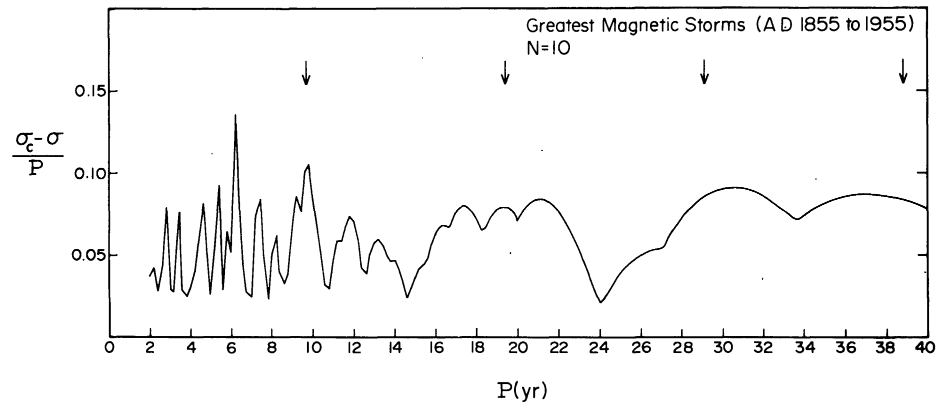

Fig. 2. Spectrum of the residuals index $(\sigma_c - \sigma)/P$ for the greatest magnetic storms on record between AD 1855 and 1955. Arrows mark the most probable period and its integer multiples

presence of gaps in the observations are expected to lead to a greater prominence of the spectral peaks that occur at integer multiples of the true period and hence to "alias" periods.

More specific inferences for the case where N is close to 10 and where gaps in the observations may occur are of interest in the applications below. According to the idealized case just discussed, four or five subsidiary maxima occur between $P = \Pi$ and $P = 2\Pi$; and between $P = 2\Pi$ and $P = 3\Pi$ one or two occur. Thus, counts of subsidiary maxima may also help in identifying the true period. That this pattern does depend on the existence of a true period and is not simply an artefact produced by the number of observations may be seen by considering the very different pattern that exists in the vicinity of a trial period which is N times the true period (compare the vicinity of $P/\Pi = 1$ with the vicinity of $P/\Pi = 2$ for $N = 2$ in Fig. 1).

IV. The Modern Auroral Cycle

Since the ancient auroral observations will be analyzed in the fashion just indicated, it is instructive to have also a comparable analysis of modern auroral data. A sample closely analogous to the ancient sample insofar as the geomagnetic latitude and total number of events are concerned is provided by the greatest "magnetic storms" and aurorae on record in modern times. During the period 1855–1955, they occurred in the following years: 1859, 1872, 1882, 1903, 1909, 1921, 1938, 1940, 1941, and 1946 (Chapman, 1957). A straightforward average interval between these events is 9.7 yr. But this type of average is very misleading because the observations are clustered in some cases (e.g., 1938, 1940, 1941) and show apparent gaps in other cases (e.g., between 1882 and 1903). Instead, a high-resolution spectral analysis of these observations has been performed in the manner indicated above, and is shown in Fig. 2. Obviously, very short periods such as 1 yr can exactly fit the data (since an undetermined number of gaps may be present), but common sense suggests examining trial periods in the vicinity of the average data interval, or less restrictively, trial periods longer than, say, 2 or 3 yr. Clearly, if no hint of a period exists in a preliminary scrutiny of the data, any formal period analysis is pointless. In the present case, the assumption of a very short period is manifestly unrealistic because this would imply that the majority of auroral maxima have been missed (either because of a lack of attention on the part of potential observers or because of bad weather or an unusual weakness of the missed maxima). On the other hand, an examination of the average data interval does suggest that some sort of periodicity is present in the modern auroral data.

Despite the paucity of observations and their spotty distribution in time, a best-fitting period of 9.7 yr (accidentally the same as the straightforward average interval) can be derived. Notice in Fig. 2 that both the relatively large size of the main maxima (corresponding to the period and its integer multiples) and the number of subsidiary maxima lying between successive integer multiples of the period are in accord with the spectrum expected for a periodic phenomenon that is affected by the presence of some gaps as well as some clustering or noise. Additionally, the large peak near 6 yr appears to be the first harmonic of another period of about 12 yr. The simple conclusion to be drawn is that the "true" period lies somewhere in the range 9–12 yr.

Over the same time interval, sunspot maxima (with which great magnetic storms and aurorae are known to be correlated) occurred in the years 1860, 1870, 1883, 1893, 1905, 1917, 1928, 1937, and 1947 (Abetti, 1957). These maxima are best fitted by a period of 11.1 yr (the straightforward average interval is 10.9 yr). Even with such a very small number of observations, the results for the modern auroral and sunspot cycles are in essential agreement with previously known results based on a much larger number of observations, and demonstrate the usefulness of the new method of analysis.

V. The Ancient Auroral Cycle

From the catalog of ancient aurorae in Table 1, sufficient data are at hand to consider separately the categories SF, NS, BR, and MR, in the well-documented time interval 223–91 BC. An analysis of each case is shown in Fig. 3. Despite the appreciable noise, the individual spectra yield the following periods:

SF, 8.7 yr;
NS, 10.2 yr;
BR, 13.3 (?) yr;
MR, 11.6 (?) yr.

Other periods in the range 8–13 yr are also possible from these data. The resemblance of both the overall spectra and the derived periods to the results for modern aurorae tend to confirm the identification of "sky fire" and "night suns" as being in most cases auroral displays. However, the results for

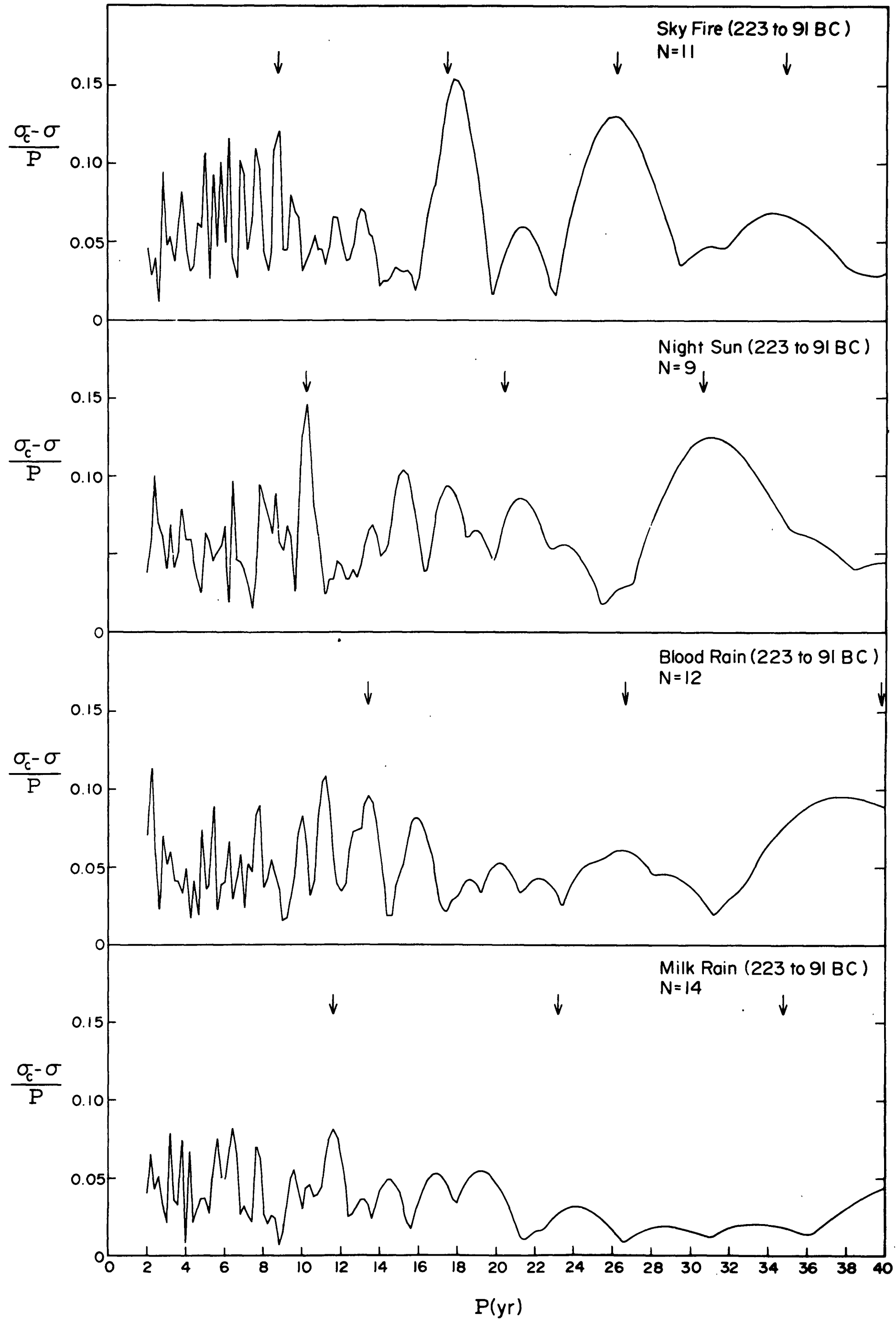

Fig. 3. Spectra of the residuals index $(\sigma_c - \sigma)/P$ for four types of ancient sky phenomena that may be auroral displays (223–91 BC). Arrows mark the most probable period and its integer multiples

"blood rain" and "milk rain" are rather marginal, and show what a nearly random spectral record might look like. If all four categories are analyzed together, the resulting period is 11.5 yr.

An additional analysis can be made by using all the data in Table 1, from 467 BC to AD 333. In analyzing these data, it must be remembered that there is no year "zero" in the historical system of dating, so that for computational purposes the BC dates must be decreased by 1 yr and then made negative in order to conform to the astronomical system. A period of 15.0 yr is found to provide a best fit to all the data. The next most probable periods are 8.7 and 11.1 yr. But since the large gaps in the record of eight centuries of observation have probably tended to distort the derived period, the previous value of 11.5 yr should be definitely preferred as the best period that can be extracted.

Since it is doubtlessly true that aurorae were caused in the past, as now, by the interaction of the solar wind with the earth's upper atmosphere, the ancient auroral cycle can be directly equated with the ancient solar cycle, apart from a possible difference of phase. In respect to the average cycle length, the agreement between ancient and modern solar cycles is good and implies a certain regularity of the sun's rhythm over an interval of over 2000 yr. In respect to the amplitude of the cycle, solar activity in the 2nd century BC was also probably very nearly the same as it is today, since the incidence of visible aurorae near Rome is found to be, in both periods, about 3 per decade. However, our knowledge that a prolonged minimum in solar and auroral activity occurred between AD 1645 and 1715 serves as a warning that the 2nd century BC, like the present century, may not be representative of all eras.

It is curious that ancient reports of "chasms" recur cyclically at much longer intervals of about 125 yr. The recorded dates of observation are 349, 344, 217, 93 and 91 BC. The great event of 467 BC also fits this sequence. Whether such a long period has any physical significance cannot be determined because the historical records are so incomplete. But during the well-documented time interval 223–91 BC a secondary cycle of 80 to 100 yr seems to fit the data well, as is indicated chiefly by the long auroral minimum from 162 to 134 BC. It is perhaps not merely coincidental that, in more recent centuries, an auroral and sunspot cycle of some 80 yr has been in steady operation (Schove, 1955; Link, 1964; Abetti, 1957).

A few words are necessary concerning the nature of direct observations of solar variability during classical antiquity. To begin with, there exist a number of reports in the classical literature of unusual changes in the sun's size, brightness, and color. These changes have not attracted much modern attention, perhaps because they are so obviously explained by purely optical effects produced in the earth's atmosphere and by unrecognized solar eclipses. However, minor solar changes have been the subject of frequent commentary over the years. These changes fall into three categories. First are the dark "spots" often mentioned by classical writers on weather lore, beginning at least with Theophrastus (*De signis tempestatum* 1.11, 2.27, 4.50) and Aratus (*Phaenomena* 819–839). However, these "spots" are almost certainly not what we mean today by "sunspots", which are rarely and with difficulty perceptible to the naked eye; rather, since they are stated to be readily visible at sunrise, to occur sometimes on the moon, and to prognosticate a rainy day, they would seem to be merely small terrestrial clouds seen projected on the sun's disk. A more interesting inference of a sunspot observation has been made by Bicknell (1968), who has noted that Anaxagoras (5th century BC) once predicted, at a time of probable auroral maximum, that a large stone would fall from the sun. A second category of solar change is the occurrence of a bright or dark "halo" around the uneclipsed sun. This atmospheric phenomenon is frequently reported in classical literature, and was explained, correctly, as early as the 4th century BC, by Aristotle (*Meteorologica* 371b–378a). Third, there are phenomena associated with the eclipsed sun. At the time of solar maximum, the corona is greatly enhanced and could be easily visible during a total solar eclipse. However, if the eclipse is annular rather than total, the sun's rim itself would appear as a fictitious "corona". In the absence of further details, an unavoidable ambiguity must attend the ancient reports of "the visible light about the rim of the eclipsed sun" (Cleomedes, *De motu circulari corporum caelestium* 2.105; Plutarch, *De facie quae in orbe lunae apparet* 932B) and "the comet that once was seen near the sun when the latter was eclipsed" (Posidonius in Seneca, *Naturales quaestiones* 7.20.4; Ptolemy, *Tetrabiblos* 2.9; Arrianus Meteorologicus in Stobaeus, *Eclogae* 1.28.2). In summary, there seems to be reasonable doubt that either sunspots or the solar corona was ever observed in the ancient West, although sunspots were more certainly reported in China during the same period (Kanda, 1933; Schove, 1951).

VI. Conclusions

The ancient prodigies of "sky fire" and "night suns" (and, more uncertainly, of "blood rain" and "milk rain") appear to be phenomena closely allied to each other, since they show virtually the same cyclical variation during the well-documented time interval 223–91 BC. The period that best fits this cyclical variation is 11.5 years, with a scattering of other possible periods ranging from 8 to 13 years. It is remarkable how closely these periods resemble the ones found for the modern auroral cycle. Thus, they tend to confirm an auroral identification of the four classes of phenomena listed above, although the identification of the very rare "chasms" and "beams" necessarily rests on descriptive evidence alone. Unfortunately, it is not possible to construct a timetable for ancient aurorae that is trustworthy enough to identify individual cases as genuine or false. Great aurorae follow the unpredictably variable solar cycle and, moreover, can appear at almost any phase during this cycle (at least they have in modern times). But this lack of precise predictability does not, of course, vitiate the mean period derived from the accumulation of many cycles. Auroral statistics also suggest that the frequency of aurorae visible near Rome in the 2nd century BC was comparable with the frequency existing now. From these bits of evidence it may be fair to conclude that solar activity two millenia ago was not markedly different from what it is today.

If the available statistics of the ancient auroral record seem rather paltry to the reader, and the results correspondingly uncertain, he should reflect that no significant increase in the literary evidence is likely to be forthcoming (unless the ancient Far Eastern annals prove to be more fruitful than they have been in investigations to date – e.g., Kanda, 1933; Schove, 1951). Therefore, I have considered it worthwhile to present the evidence as it stands today. Perhaps in the future, entirely different methods will be able to provide data concerning the solar cycle in the ancient historical past.

Acknowledgements. The resources of the Columbia University Libraries and of the New York Public Library were indispensable to performing this work. Professor Harold Stolov kindly commented on the manuscript.

References

Abetti, G.: 1957, The Sun, Faber and Faber, London
Bicknell, P. J.: 1968, *Isis* **59**, 87
Chapman, S.: 1957, *Ann. IGY* **4**, (Part 2), 25
Eddy, J. A.: 1976, *Science* **192**, 1189
Fritz, H.: 1873, Verzeichniss beobachteter Polarlichter, Kaiserlichen Akademie der Wissenschaften, Vienna
Fritz, H.: 1881, Das Polarlicht, F. A. Brockhaus, Leipzig
Frobesius, J. N.: 1739, Nova et antiqua luminis atque aurora borealis spectacula, C. F. Weygandus, Helmstadt
Kanda, S.: 1933, *Proc. Imp. Acad. Japan* **9**, 293
Link, F.: 1962, *Geofysikální Sborník* No. 173

Link, F.: 1964, *Planetary and Space Sci.* **12**, 333
de Mairan, J. J. D.: 1733, Traité physique et historique de l'aurore boréale, Imprimerie Royale, Paris
Nicolini, T.: 1963, *Atti Convegno Astron. Milano, Soc. Astron. Ital.*, 219
Nicolini, T.: 1976, *Rendiconto dell'Accad. Sci. Fis. Math. in Napoli* (Ser. 4) **43**, 1
Pliny the Elder: 77, Natural History 2.89–100 *passim*, translated by H. Rackham, Harvard University Press, Cambridge, Mass., 1967
Schove, D. J.: 1948, *J. Brit. Astron. Assoc.* **58**, 178
Schove, D. J.: 1951, *Isis*, **42**, 133
Schove, D. J.: 1955, *J. Geophys. R.* **60**, 127
Stothers, R.: 1979, *Isis* (in press)