AN RR LYRAE STAR SURVEY WITH THE LICK 20-INCH ASTROGRAPH

II. THE CALCULATION OF RR LYRAE PERIODS BY ELECTRONIC COMPUTER*

J. LAFLER AND T. D. KINMAN
Survey Research Center and Lick Observatory, University of California
Received July 16, 1964

ABSTRACT

A program is described for finding the period of an RR Lyrae variable from magnitude observations by means of an electronic computer such as the IBM 1620. For a number of trial periods in the appropriate range, the computer forms the sum (θ) of the squares of the magnitude difference between observations of adjacent phase using a group of the observations which are most concentrated in time. A number (e.g., three) of these periods for which θ has the smallest values are selected as being the most likely to be correct. For each of these three periods, all the observations are used to compute θ for periods in decreasing ranges around the period for which θ had a minimum value in the previous calculation until the required accuracy is achieved. The finally selected period is that for which θ is a minimum. The ratio (Θ) of θ to the sum of the squares of the differences of all the magnitudes from their mean value may be used to estimate the reliability of the period obtained.

Expressions are given for the number of trial periods needed and the number of observations required as a function of amplitude of the variable and rms error of the observations at different levels of discrimination against spurious periods. The origin and probability of occurrence of spurious periods are briefly reviewed.

I. INTRODUCTION

The present survey, which has been described in Paper I, may well involve the period determination of a large number of variables, each requiring many hours with a desk computer. It therefore seemed worthwhile to consider using an electronic computer to increase the speed and, if possible, the reliability of these calculations.

Although analytical expressions may be found for variable-star light-curves (cf. Hagen 1913), a curve-fitting program would be difficult because of the loss of accuracy which would result from attempting to fit the observations to a single shape of light-curve. This difficulty might be avoided by a harmonic analysis of the observations. The light-curves of Bailey type *ab*, however, are quite steep at the time of rising light, and therefore many subharmonics of the fundamental period would be involved; it might therefore be difficult to obtain a solution from the rather small number of observations of limited accuracy that are usually available. It therefore seemed that the most promising method would be of the autocorrelation type in which the criterion for the correct period would be that the observations fall on a smooth curve when their phases are calculated with this period. A program of this type was originated at the Survey Research Center for use with their computer (IBM 1620 Mark 1 with 60000-position core memory storage in December, 1963). Since this program may well have applications to the period determination for other types of variables and possibly elsewhere, it is discussed here in some detail.

II. THE CRITERION FOR A SMOOTH LIGHT-CURVE

The commonly used criterion for correlation in the case of a non-linear regression is the correlation ratio. This is satisfactory, however, only when the number of observations is sufficiently large to form a grouped contingency table. It is found, in the present case, that the number of observations from which the initial selection of the period is

* *Contributions from the Lick Observatory*, No. 173.

made must be small if the number of trial periods is not to become excessive. On the other hand, since the phase interval between maximum and minimum light is $\sim$0.1, the phase interval for each phase array would need to be made at least as small as this in computing the correlation ratio and a relatively large number of observations would be needed to afford a satisfactory test.

A criterion which gives more satisfactory results in the case of a moderate number of observations and a sharply non-linear regression is obtained by requiring the sum of the squares of the magnitude differences between observations of adjacent phase to be a minimum. Our test parameter is then defined by

$$\Theta = \frac{\sum_i (m_i - m_{i+1})^2}{\sum_i (m_i - \bar{M})^2}, \tag{1}$$

where $\bar{M} = \Sigma_i m_i / N$ and N is the number of observations. Since the denominator is independent of the phase, only the numerator (θ) need be calculated for each trial period. Then, in principle, that period with the minimum value of Θ is the nearest to the correct period.

III. THE ARRANGEMENT OF THE PROGRAM

Most stars that were selected as variable in the present survey showed marked brightness variations in the course of a few hours. These stars are nearly all fainter than twelfth magnitude and are at galactic latitudes greater than 20° so that the majority are likely to be RR Lyrae stars. We shall therefore consider only periods in the range 0.20–1.00 day. The test statistic must therefore be computed for trial periods in this range, and from these some periods are selected that have the smallest value of θ. The calculation is then repeated for new trial periods close to these selected periods until a period is found with sufficient accuracy for which the value of Θ is considered satisfactorily small.

The trial periods used in the initial selection must be chosen so that the maximum difference between the true period and the nearest trial period does not give rise to a phase error ($\Delta\phi$) between the first and last observations which produces a significant change in θ. If T is the time interval between the first and last observations and ΔP is the difference between the true period and the nearest trial period, then

$$\Delta\phi = \frac{T\Delta P}{P}. \tag{2}$$

Since the maximum value of ΔP is one-half the difference between adjacent trial periods, then the trial periods should be members of the series

$$P_n = P_0 \left(1 + \frac{2\Delta\phi}{T}\right)^n. \tag{3}$$

If the period range to be investigated is P_0 to GP_0, then the total number of trial periods required will be

$$\frac{T}{2\Delta\phi} \ln G \qquad \left(\text{if } \frac{\Delta\phi}{2T} \ll 1\right). \tag{4}$$

The maximum value of $\Delta\phi$ which is acceptable is not well defined. In practice it was found that with $\Delta\phi = 0.1$ (i.e., of the same order as the phase difference between minimum and maximum light for a variable of Bailey type *ab*), that period for which θ was the minimum was closest to the true period for the majority of the tested variables.

Smaller values of $\Delta\phi$ would give better discrimination but would also increase the number of trial periods so that some compromise was needed.

The calculation of θ for the first set of trial periods is that which is most time consuming so that the following procedures were adopted to keep the number of trial periods to a minimum:

a) Not all observations were used for the calculation of θ for the first set of trial periods. In this way T is kept small and hence also (by eq. [4]) the number of trial periods. The minimum number of observations needed is discussed in the next section.

b) In a survey of the present type, the majority of variables found have $P > 0^{d}40$. This is partly the result of the observational selection of variables of large amplitude. It is therefore profitable to make the original solution in the period range $0^{d}40$–$1^{d}00$ only. If no satisfactory period is found, then the range $0^{d}20$–$0^{d}40$ may be tried.

As an example, seventeen observations were selected out of a total of thirty-four. For these seventeen observations, T was 10^{d} so that the initial trial periods were

$$P_n = 0.40\ (1.02)^n\ . \tag{5}$$

From these trial periods, the *three* having the smallest values of θ were selected, and for these three periods the following procedure was carried out using all thirty-four observations:

1. Compute θ for ten equally spaced periods centered on the selected trial periods in the range $P \pm 0.01\ P$ and select the period for which θ is a minimum. Three periods are obtained, and the one for which θ is a minimum is taken to be correct if the value of Θ for this period is satisfactorily small. If this is not the case, a new period range must be investigated on a re-run of the program.

2. Compute θ for ten equally spaced periods centered on the trial period found in step 1 in the range $P \pm 0.001\ P$ and select the period for which θ is a minimum.

From equation (2) with $\Delta\phi = 0.1$, we see that the time interval between the first and last observations should not exceed 90 days in step 1 and 900 days in step 2. While step 2 gave a period of sufficient accuracy for the present observations, the process may be continued further to obtain a more accurate period if the observations permit.

The reason for selecting more than one period from the initial trial periods is to avoid spurious periods when the number of observations in the original selection is small. Such spurious periods commonly arise when there is an impressed periodicity in the times of observation (see Appendix).

IV. THE DISCRIMINATING POWER OF THE TEST STATISTIC

In the case when θ is calculated with the true period we may write

$$\theta_1 \sim n\sigma^2 + \sum_n \left(\frac{dm}{d\phi}\,\delta\phi\right)^2 , \tag{6}$$

where σ is the rms observational error in magnitude for a single observation, n is the number of observations, $dm/d\phi$ is the slope of the light-curve at phase ϕ, and $\delta\phi$ is the phase difference between two observations of adjacent phase. In general $dm/d\phi$ will be proportional to the amplitude (A) of the variable and $\delta\phi$ will be inversely proportional to n, and we therefore have

$$\theta_1 \sim n\sigma^2 + \frac{k_1 A^2}{n} . \tag{7}$$

Here k_1 was found to have a mean value of 30 from an analysis of twenty-five type *ab* variables for which $\sigma \sim 0.14$ mag.

In the case when θ is calculated for an incorrect period, we assume that there is no correlation between m and ϕ and write

$$\theta_2 \sim n(\sigma^2 + k_2 A^2) , \tag{8}$$

where k_2 was found to average 0.45 from an analysis of twenty-five RR Lyrae stars.

An estimate of the discrimination afforded by θ is given by the ratio Λ between θ_2 and θ_1, which should be as large as possible. Both k_1 and k_2 depend upon the shape of the light-curve and on the fortuitous distribution of the observations with phase. In any particular case therefore, θ may depart significantly from the mean values given in

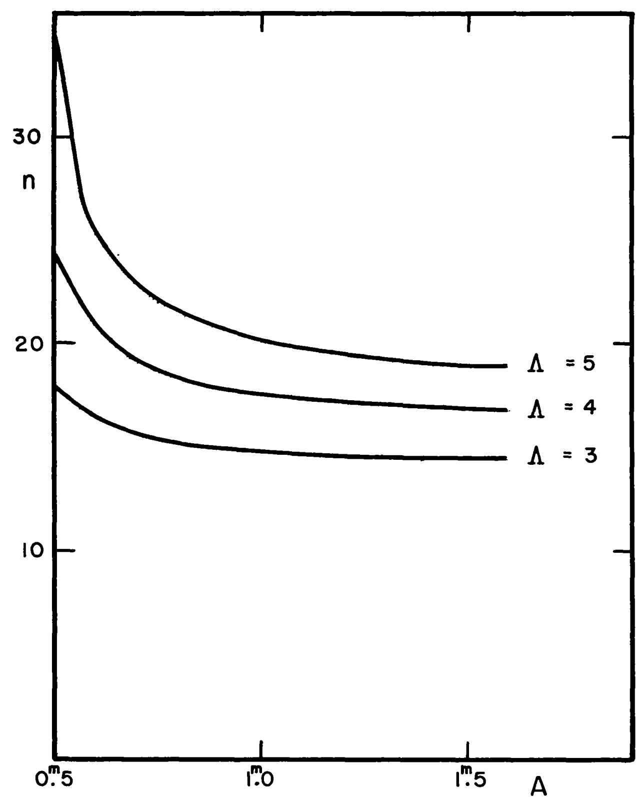

FIG. 1.—The number of observations (needed to give a discrimination Λ in the selection of the period) as a function of the amplitude (in magnitudes) of a variable of Bailey type *ab*. The range $3 < \Lambda < 5$ corresponds roughly to that of from poor to good discrimination.

equations (7) and (8). As a rough rule, we suggest that for a moderate number of observations ($10 < n < 40$), $\Lambda = 3$ affords rather poor discrimination, $\Lambda = 4$ is adequate in most cases, and $\Lambda = 5$ is rather good. Figure 1 shows the relation between amplitude of the variables and the number of observations needed to find its period at these levels of discrimination for $\sigma = 0.14$. It may be noted that the value of σ becomes important only when it is greater than about 0.2 of the amplitude of the variable.

These relations between n and A are for asymmetrical light-curves; in the case of symmetrical light-curves, k_1 may be expected to be significantly smaller and the number of observations needed for discrimination of their periods should be lower for a given amplitude than for the asymmetrical types.

It is seen from equation (4) that for a period range of $0^d\!.40$–$1^d\!.00$ ($G = 2.5$) and $\Delta\phi = 0.1$ the number of trial periods needed for the initial determination will be 4.6 times the number of days between first and last observations. This number of trial periods should not exceed about 100 if the total computation time with a relatively slow computer such as the IBM 1620 is not to exceed 1 hour. This limits T to about 2 weeks, or the observations during one dark-of-the-moon period. From Figure 1 we see that the present method requires a minimum concentration of about fourteen observations in such a period. It is clear also that, for the concentrations of observations which are likely to be available in practice, that the discrimination afforded by θ is not good enough to assure that the minimum value of θ always corresponds to the true period. The smaller the number of observations, the greater this uncertainty will be, and the greater the number of the lowest values of θ which should be tested (using all observations) in order to eliminate spurious periods.

V. TESTS FOR CORRECTNESS OF THE PERIODS

The final data may be readily plotted by the computer in the form of a light-curve, and inspection of this curve is the usual method of estimating whether the period is correct. According to Preston (1964), however, 30 per cent of RR Lyrae stars have variable

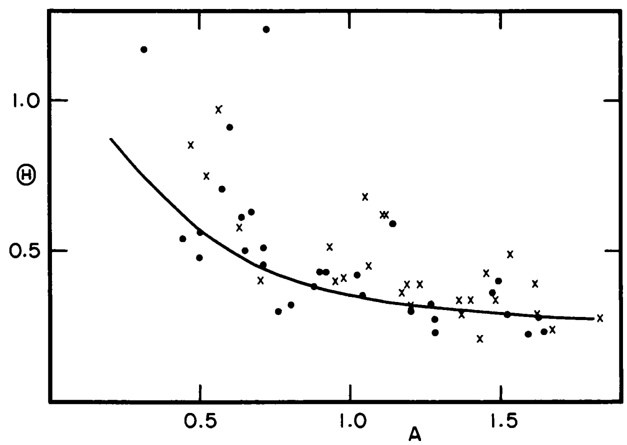

FIG. 2.—A plot of the value of the test parameter Θ as a function of the amplitude of the variable in magnitudes. Variables with mean magnitudes fainter than 16.0 mag. are shown by crosses; those brighter by filled circles. The curve is a relation derived for type *ab* light curves assuming an rms observational error (σ) of $\pm$ 0.15 mag.

light-curves, and consequently there are always likely to be some variables in any sample where the correctness of the period will be in doubt with a limited number of observations.

Figure 2 shows a plot of the values of Θ obtained for RR Lyrae variables in the field MWF 361A as a function of their amplitudes. Variables with mean photographic magnitudes greater and less than 16.0 mag. are shown by crosses and filled circles, respectively. The curve in this figure was computed using the expression for θ_1 given in equation (7) and assuming a value of σ of ± 0.15 mag. Since equation (7) applies only to type *ab* variables, the curve is only strictly applicable to amplitudes greater than about 0.75 mag. For amplitudes greater than 0.75 mag. most of the values of Θ are seen to cluster around the curve, and the exceptions are mostly fainter stars for which σ is larger than ± 0.15 mag. The large values of Θ found for many of the variables with small amplitudes indicate the difficulty of obtaining reliable periods for such stars from small numbers of observations. It may be noted that the values of Θ which may be expected for purely random light-curves (obtained by using θ_2 from eq. [8] instead of θ_1) vary from 1.4 to 3.7 as the amplitude increases from 0.2 to 1.5 mag.

The above analysis suggests that Θ may be used as a quantitative indicator of the reliability of a light-curve. While better criteria could doubtless be devised, that suggested here has the virtue of relative simplicity and ease of calculation which is essential when it must be applied to a large number of trial periods.

Thus, in addition to the necessary but subjective impression obtained from the light-curve, the present program provides a numerical estimate of the reliability of the period obtained. Considerable caution must clearly be used in interpreting periods obtained by any method from a limited number of observations of moderate photometric accuracy. The use of a program such as that described in this paper does, however, allow a more thorough analysis than is convenient with conventional trial and error methods. Whatever method of computation is used, both the reliability and the ease of computation will depend upon having a suitable temporal distribution of the observations. One substantial group of observations which is well concentrated in time is needed to make the initial period selection, and the remaining observations should be made over an extended period of time in order to distinguish between the selected trial periods.

This work was supported by a grant from the National Science Foundation (NSF GP-907).

APPENDIX: SPURIOUS PERIODS

It is well known that spurious periods may be derived because the times of observation are not random but conditioned by some external period such as the day, lunation period, or the year. The periods of RR Lyrae stars are commonly found by selecting observations at an identifiable phase such as maximum light, and by calculating the least common denominator which will satisfy the time intervals between these observations. Any such interval ΔT_i between maxima is related to the period P by

$$\Delta T_i P^{-1} = \phi_i ,$$

where the phase ϕ_i is integral. If the observations are made at intervals which are close multiples of some period p, then $\Delta T_i \sim \alpha_i p$, where α_i is an integer. We then have a number of relations to solve for the period of the form

$$P^{-1} = \phi_i / \alpha_i p .$$

Spurious periods (Π) which may also satisfy these relations will be related to the true period by

$$\Pi^{-1} = P^{-1} \pm \epsilon / \alpha_i p \qquad (\epsilon \text{ is an integer}),$$

which is a generalization of relations previously given by Hagen (1913), Payne-Gaposchkin and Gaposchkin (1938), and others. An infinity of spurious periods is possible but only some such periods are a practical nuisance. When $\epsilon \ll \alpha_i$, Π is nearly equal to P so that the derived period is merely slightly inaccurate. If $\epsilon \gg \alpha_i$, then Π becomes very small and the solutions would be rejected on physical grounds. The spurious periods which are most troublesome in practice arise when $\epsilon \sim \alpha_i$. It is found by inspection, although no formal proof is attempted, that the most common values of ϵ / α_i will be

$$1; \text{ with relative frequency } 1;$$
$$\tfrac{1}{2}, 2; \text{ with relative frequency } \tfrac{1}{2};$$
$$\tfrac{1}{3}, \tfrac{2}{3}, \tfrac{3}{2}, 3; \text{ with relative frequency } \tfrac{1}{3}, \text{ etc.}$$

This leads to the well-known result that the most common spurious period is given by

$$\Pi^{-1} = P^{-1} \pm \frac{1}{p} \sim P^{-1} \pm 1 ,$$

where $p \sim 1$ day. In the case of RR Lyrae star observations, $p \sim 1$ is the only external period which is likely to be effective, and to minimize this effect, observations were made at hour angles (both east and west) as large as possible. Even so, the effect of the external period of 1 day was far from eliminated. This is shown in Figure 3, which is a plot of the test statistic θ for periods in the range $0^d.40$–$1^d.00$ for a variable of period $0^d.4587$. In the lower part of the figure, the most likely spurious periods in this range have been plotted with ordinates to indicate relative probability. The most common spurious period ($\epsilon/\alpha_i = 1$) is very apparent and gives a minimum nearly as deep as that corresponding to the true period. Cases were found where the spurious period gave a deeper minimum than the true period. Therefore, when the number of observations and their concentration in time is small enough for the machine to identify the period from a wide range of possible periods, then the minimum value of θ may occasionally correspond to a

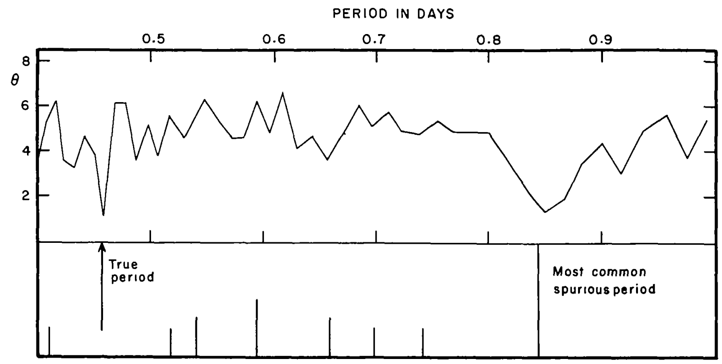

FIG. 3.—(*Top*) The variation of θ among the trial periods (in days) for a variable with period $0^d.4587$. (*Bottom*) The location in this period range of the most common spurious periods. Each ordinate is proportional to the probability of occurrence of the period.

spurious period. It is therefore necessary to investigate several of the deepest minima in order to see which of these, when the calculations are repeated with all the observations, yields the smallest value of the test statistic.

I am grateful to Dr. George Preston for pointing out that in the past a number of stars (e.g., V453 Oph, BE Mon, PW Cas, V342 Cas) have been assigned periods less than 1 day when in fact their true periods are much longer. The present program assumes that the true period is less than 1 day. It is therefore most important to make a preliminary inspection of the observed magnitudes to see that a variable shows significant variations in the course of a few hours and has no obvious longer periodicities.

REFERENCES

Hagen, J. G. 1913, *Die Veränderlichen Sterne*, **1** (Freiburg im Br.: Herdersche Verlagshandlung), 430.
Payne-Gaposchkin, C., and Gaposchkin, S. 1938, *Variable Stars* (Cambridge, Mass.: Harvard Observatory), p. 359.
Preston, G. W. 1964, *Ann. Rev. Astr. and Ap.* (Stanford, Calif.: Annual Reviews, Inc.), **2**, 23.