THE ASTROPHYSICAL JOURNAL, 263:854–869, 1982 December 15

ON THE DETECTION OF OTHER PLANETARY SYSTEMS BY ASTROMETRIC TECHNIQUES

DAVID C. BLACK AND JEFFREY D. SCARGLE
Space Science Division, NASA Ames Research Center
Received 1982 February 8; accepted 1982 June 18

ABSTRACT

We examine the problem of detecting periodic signals in astrometric data, with specific emphasis on searching for a perturbation in a star's motion due to planetary companions. It is shown how spectral analysis of such data by means of the periodogram provides an objective statistical basis for evaluating the reality and characteristics of suspected perturbations. A brief description of the properties of the periodogram is followed by a demonstration of the technique in an analysis of 12 years of US Naval Observatory astrometric data for three stars with low-mass stellar companions. The periodogram is shown to provide a basis for both a comparison of the perturbation detection efficiencies of various astrometric systems and optimization of the efficiency of a given system.

A quantitative discussion is given of an error source generally overlooked or ignored in analysis of perturbations in astrometric data, namely, absorption of the linear component of the necessarily finite sample of a perturbation into the proper motion. This error decreases less rapidly with the number of periods observed than one might guess (with a full period observed, the error leads to underestimation of the amplitude and period by as much as 47% and 25%, respectively). In a quantitative comparison of two idealized astrometric systems, it is found that the space-based one is vastly superior to the ground-based one for the purpose of searching for and studying other planetary systems.

Subject headings: numerical methods — stars: stellar dynamics

I. INTRODUCTION

Observational evidence concerning the statistics and structural characteristics of other planetary systems is an important, perhaps essential, ingredient in understanding not only the processes by which such systems are formed but also the processes by which stars are formed. In spite of this importance *there is currently no unambiguous evidence for even the existence of another planetary system*.

There are, in principle, many observational techniques that could be used to search for other planetary systems. A review of those techniques and recent advances in related instrumentation is given by Black (1980). One of the more promising approaches is astrometry, whereby planetary companions to other stars would be detected indirectly by observation of their effect on the apparent motion of those stars. A star without companions will appear to move linearly across the sky, whereas a star with companions will have an additional component of motion arising from its orbital motion about the barycenter of the star-companion system. A detailed discussion of this technique is given by van de Kamp (1967).

The angular perturbation θ_* due to a single companion in a circular orbit is given by

$$\theta_* = 0.98 \frac{aM_c}{DM_*} \text{ milli-arcsec}, \quad (1)$$

where a is the semimajor axis of the companion's orbit (in AU), D is the distance (in pc) between the observer and the star under study, and M_c and M_* are, respectively, the masses of the companion (in Jovian masses) and of the star (in solar masses). Some appreciation for the difficulty of the detection problem can be gained by considering the magnitude of θ_* for two examples, viz., the Sun and Jupiter, and the Sun and the Earth. Taking the Sun-Jupiter pair alone, $\theta_* \approx 5D^{-1}$ mas (1 mas = 1 milli-arcsec), whereas for the Sun-Earth pair alone, $\theta_* \approx 3 \times 10^{-3} D^{-1}$ mas.

Equation (1) does not include projection and eccentricity effects. If the plane of a circular orbit lies in the plane of the sky, the x and y projections of the perturbation will be equal, with amplitudes given by equation (1). If the orbital plane is inclined with respect to the observer's line of sight, there will always be some

projection of that orbit on the plane of the sky which has the amplitude given by equation (1), but the amplitude at other position angles will be less. If the orbit is highly eccentric (say $e > 0.7$), there could be cases where the orientation of the orbital plane with respect to an observer's line of sight gives rise to significantly smaller apparent amplitudes than indicated by equation (1). In addition, if the orbit is eccentric, then the time dependence of the apparent displacement $[\theta_*(t)]$ of the star is no longer sinusoidal. Another potential complication concerns the possible presence of more than one (stellar or planetary) companion. The orbital motion of the star about the barycenter of a multicompanion system will contain components reflecting the orbital periods of each of the companions (e.g., eq. [2.12] in Black 1980). These complications will be discussed below (§ II*c*).

The majority of astrometric observations employ photographic plates as the detector, and the precision of a single night's observation ranges from about 8 to 30 mas, depending on the telescope being used and on observing conditions. Recent work by Gatewood and collaborators (Gatewood and Stein 1980) using a photoelectric detector has yielded nightly precision of about 4 mas. The anticipated precision of the Space Telescope for astrometric purposes is about 1 mas (Black and Brunk 1980), while that of the European Space Agency's proposed astrometric satellite (HIPPARCOS) is ~1–2 mas (Black and Brunk 1980). Comparing the amplitude of possible signals with these figures, it is clear that the search for other planetary systems by astrometric means will involve data with low signal-to-noise ratios. One would like to have an objective means of assessing the statistical significance of any claimed detection.

The purpose of this paper is to present an objective method for analyzing astrometric data for the presence of perturbations induced by a planetary companion to a star. The proposed method (1) minimizes subjectivity in the analysis and interpretation of astrometric data, (2) allows a quantitative comparison of the efficiencies of different astrometric systems for detecting planetary companions to stars, and (3) provides a guide to optimization of the parameters in a planetary detection program. The method referred to is that of the *periodogram*. The general mathematical basis for and statistical properties of the periodogram are discussed in a companion paper (Scargle 1982, hereafter S).

A brief review of salient aspects of the periodogram as used in analysis of astrometric data for the presence of periodic signals is given in § II. The technique is applied to real astrometric data in § III. A comparative analysis of the detection efficiencies of two astrometric systems is presented in § IV, and a summary is given in § V. Readers who are interested only in the general aspects of detecting other planetary systems using astrometric data could skip §§ II and III.

II. THE REDUCTION AND SPECTRAL ANALYSIS OF ASTROMETRIC TIME SERIES DATA FOR PERIODIC SIGNALS

Consider a star with a planetary system such that the observed coordinates in a rectangular system are given by

$$X_{\rm obs} = C_X + \mu_X t + \pi P_X + x(t) + n_X(t), \quad (2)$$

$$Y_{\rm obs} = C_Y + \mu_Y t + \pi P_Y + y(t) + n_Y(t). \quad (3)$$

The C's are constants which depend on the origin of coordinates, the μ's are constants giving the proper motion of the star, π is the parallax of the star, P_X and P_Y are the parallax factors (van de Kamp 1967), x and y are the displacements in the star's position due to its planetary companion(s) (i.e., the desired signals), and the n's are the observational errors, assumed to be random. These equations can be elaborated to include effects such as the aberration of starlight, perspective acceleration, etc., but the simple form in these equations is more than sufficiently general to illustrate the concepts of interest here. The data are taken to consist of a set of N_0 points $[X_{\rm obs}(t_i), Y_{\rm obs}(t_i)]$, $i = 1,2,\ldots,N_0$, representing the coordinate measurements at the observation times t_i. The constants C_X, C_Y, μ_X, μ_Y, and π are usually determined by fitting equations (2) and (3) to the data in a least-squares sense. Hence the problem of detecting the planetary-induced perturbations reduces to one of detecting the signals x and y in the presence of the noise n_X and n_Y, based on the reduced observations

$$X = X_{\rm obs} - (C_X + \mu_X t + \pi P_X) = x + n_X, \quad (4)$$

$$Y = Y_{\rm obs} - (C_Y + \mu_Y t + \pi P_Y) = y + n_Y. \quad (5)$$

a) Preprocessing: Spurious Proper Motion

There are several practical problems connected with this decoupling of the nonplanetary from the planetary-related stellar motions. For example, a planetary-induced perturbation with a period of approximately 1 yr will tend to be confused with the annual terms comprising the parallax and aberration (although the latter is, in principle, exactly known). Similarly, the linear component of the perturbation will be erroneously absorbed into the proper motion terms. This fact has been almost universally ignored, probably with the feeling that when the observations span more than a small fraction of an orbital period the resulting error is negligible. This feeling is incorrect; the effect can be large even if a full orbital period has been observed and becomes negligible only after several periods. To our knowledge the only published mention of this effect is a brief remark by Gatewood *et al.* (1980), although related

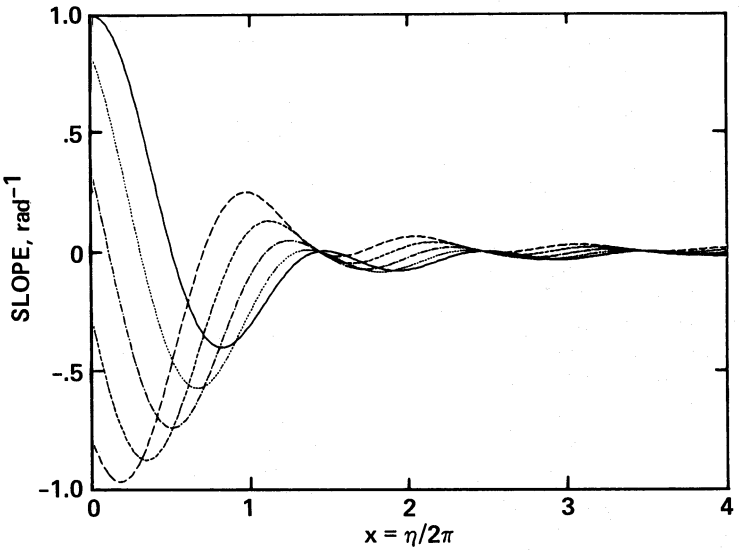

FIG. 1.—The least-squares slope (eq. [10]) of a sine wave of phase ϕ, as a function of the number of full periods observed, $\eta/2\pi$. The values of ϕ are as follows: 0, $\pi/5$, $2\pi/5$, $3\pi/5$, and $4\pi/5$.

general considerations have been discussed by Firneis and Firneis (1975).

To demonstrate this proper motion effect, we evaluate the size of the error. Suppose that the motion of the star, after removal of the effects of parallax and aberration, is

$$X(t_i) = C_0 + \mu_0 t_i + x_0 \sin(t_i + \phi), \tag{6}$$

where $C_0 + \mu_0 t$ represents the actual motion, and time is measured in units of the period/2π. A least-squares estimate of the proper motion is obtained by minimizing

$$D(C,\mu) = \sum_i [X(t_i) - C - \mu t_i]^2 \tag{7}$$

with respect to C and μ. If this summation over the observations is replaced with a continuous integral over the interval $(0,\eta)$, where $(\eta/2\pi)$ represents the fraction or multiple of a period observed, an elementary analysis gives the following least-squares estimates:

$$\langle \mu \rangle = \mu_0 + x_0 a_\phi(\eta), \tag{8}$$

$$\langle c \rangle = C_0 + x_0 b_\phi(\eta), \tag{9}$$

where

$$a_\phi(\eta) = (6/\eta^2)\{(2/\eta)[\sin(\eta+\phi) - \sin\phi] - \cos(\eta+\phi) - \cos\phi\}, \tag{10}$$

$$b_\phi(\eta) = (6/\eta)\{\eta^{-1}[\sin\phi - \sin(\eta+\phi)] + \tfrac{1}{3}[\cos(\eta+\phi) + 2\cos\phi]\}. \tag{11}$$

The quantity $a_\phi(\eta)$ represents the *least-squares slope of a sine wave* (of phase ϕ) sampled over the range $(0,\eta)$ and is plotted in Figure 1 for a set of values of ϕ.

For $\eta < 1$ the sine wave is sampled over such a small range that it is effectively a linear segment, the slope of which is completely absorbed into the proper motion. But note that the slope error is still quite large for $\eta = 2\pi$, and not entirely negligible even for $\eta \approx 4\pi$. Figure 2 shows the rms phase average of a_ϕ, which is perhaps more relevant, since one does not know ahead of time what phase a perturbation will have.

The actual value of the slope error, $x_0 a_\phi$, may be small compared to the true proper motion, but what matters is the effect of this error on the resulting residuals. Figure 3 demonstrates the case $\eta = 2\pi$ (one full period observed) and $\phi = 0$. The line labeled "spurious proper motion" represents the perturbation which would erroneously be ascribed to proper motion. As the residuals would be referred to this line, the period and amplitude of the perturbation would both be underestimated. The inset in Figure 3 shows that the erroneous residuals closely mimic a sine curve characterized by the underestimated values of the period and amplitude, the major departure being that the residuals are of too large an amplitude near the ends of the observed interval. A similar analysis of the case $\eta = 4\pi$ (two full periods), $\phi = 0$, shows that the slope error is 4 times smaller, the apparent period is 0.96 times the true period, and the apparent amplitude is 0.83 times the true amplitude. For these two values of η, $\phi = 0$ yields the maximum slope error, but, as Figure 2 indicates, the phase-averaged error is only slightly less than for this worst case.

Not only is the slope wrong (eq. [8]), but so is the intercept (eq. [9]). It is more convenient to deal with the

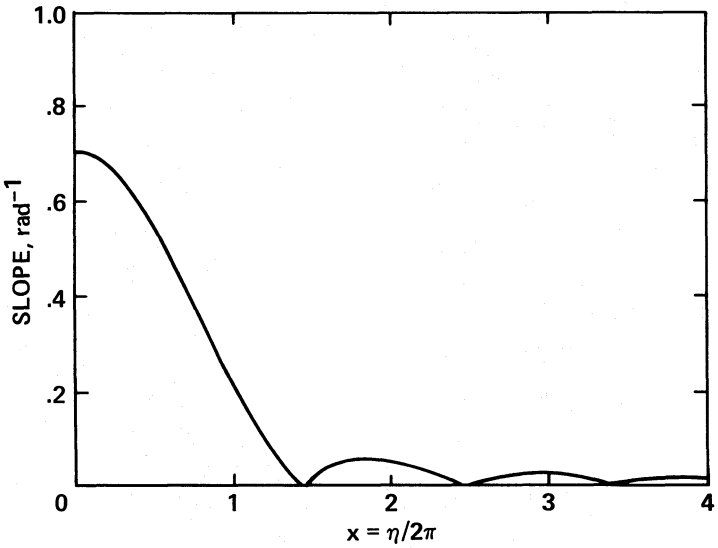

FIG. 2.—The rms phase average of the slope $\langle a \rangle = (6/\eta^2)[1+\cos\eta-(4\sin\eta)/\eta+4(1-\cos\eta)/\eta^2]^{1/2}$

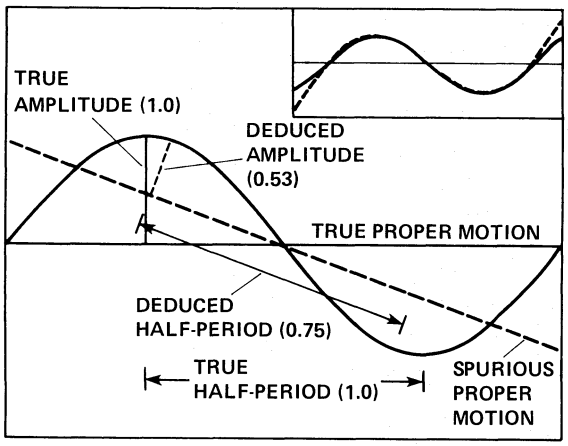

FIG. 3.—Example of how the spurious proper motion effect causes underestimation of the period and amplitude of a perturbation. The solid lines represent the actual proper motion and perturbation. The dashed line has the slope which is erroneously attributed to proper motion, so that the apparent residuals are smaller than the true ones. The inset shows these residuals (*dashed line*) plotted with a sinusoid of the incorrectly deduced amplitude and period, demonstrating how well the erroneous residuals mimic a sinusoidal perturbation.

error in the mean value of $X(t)$:

$$\delta\bar{X}_\phi = \eta^{-1}\cos\phi - \cos(\eta+\phi), \qquad (12)$$

which has a phase (rms) average of

$$\langle(\delta\bar{X}_\phi)^2\rangle^{1/2} = 2^{1/2}\eta^{-1}|\sin(\eta/2)|. \qquad (13)$$

This error drops even more slowly with η ($\sim\eta^{-1}$) than does the slope error ($\sim\eta^{-2}$). As expected, the mean-value error vanishes at integral values of $\eta/2\pi$ and is maximum near half-integral values.

To our knowledge, the slope error has not previously been corrected for in reductions of astrometric data. It is evident that such a correction can be accomplished if a model of the perturbation is incorporated into the determination of the proper motion. (For example, a model of the form in equation [6] could be fitted directly to the data.) A rigorous analysis would involve not only a correction for slope error but a *simultaneous* reestimation of all parameters once it became apparent that the original estimation model was inadequate.

There is an exception to this last statement: a run of historical astrometric data can be used to determine the proper motion over a long time baseline, and then the true proper motion can be eliminated from a shorter and more closely spaced set of observations, obtained specifically to detect perturbations. The point is that the value of η for the long run of data could be very large, even though η for the more concentrated (presumably more accurate and probably more closely spaced) run were relatively small.

b) The Periodogram

Consider now the problem of searching for the periodic signals x and y buried in noise (eqs. [4] and [5]). Because x and y are planetary-induced perturbations, they are, for the present purposes, *exactly periodic*. This is a particularly important point which derives from the fact that the time scales for changes in the parameters of

stable planetary orbits are very long compared with the periods of those orbits. We deal first with the simple case of a single planet in a nearly circular orbit. (In § II*c* we shall discuss more general cases.) This implies that the perturbations x and y will be purely sinusoidal—i.e., of the form suggested in equation (6).

Most published observational work presents graphs of the residuals as functions of time, which are then inspected for systematic quasi-sinusoidal trends. As shown in S, with sufficient data, perturbations which are small compared with the rms observational errors can be detected—whereas such signals will almost certainly be missed by visual inspection of the residuals. In addition, for a statistical analysis of detection efficiency it is necessary to introduce an objective and quantitative detection criterion. Some workers have used the increase in the variance of the residuals due to the perturbation for this purpose (e.g., Gatewood *et al.* 1980). Such a procedure discards the information contained in the correlation structure of the perturbation and, for example, makes no use of the sinusoidal nature of the perturbation and can therefore tell nothing about the period of disturbance.

There are a variety of techniques for studying data as a function of the period of a possible periodic component. One example is folding (and then averaging) the data with respect to various periods and then examining the appearance of the resulting averaged curves. Another is the least-squares fitting of sine waves to the data and studying the variance of the residuals as a function of the period (Barning 1963; Lomb 1976). The technique advocated here is that of periodogram analysis, not because it is necessarily the best (e.g., most sensitive) technique, but because it has a well-developed and straightforward statistical theory. The practical importance of this attribute of the periodogram will be developed below.

Scargle (1982) introduced the following modified form of the periodogram:

$$P_*(\omega)=\frac{1}{2}\left\{\frac{\left[\sum_j X_j\cos\omega(t_j-\tau)\right]^2}{\sum_j\cos^2\omega(t_j-\tau)}+\frac{\left[\sum_j X_j\sin\omega(t_j-\tau)\right]^2}{\sum_j\sin^2\omega(t_j-\tau)}\right\}, \tag{14}$$

where τ is defined by

$$\tan(2\omega\tau)=\frac{\sum_j\sin 2\omega t_j}{\sum_j\cos 2\omega t_j}, \tag{15}$$

and where the t_j, $j=1,2,3,\ldots,N_0$, are the times of observation; N_0 is the number of data points; and $\omega=2\pi f$ is an arbitrary frequency. The motivation for modifying the usual form of the periodogram, as well as a full discussion of the properties of the periodogram, is given in the companion paper; here we only summarize those formulae and concepts which are germane to this discussion.

If the N_0 data points are evenly spaced in time, there exists a natural set of frequencies at which to evaluate the periodogram, viz.,

$$\omega_n=\frac{2\pi n}{T}, \tag{16}$$

where T is the total time interval of the observations, and the integer $n=0, 1,2,\ldots,N=N_0/2$. The fundamental frequency $\omega_1=2\pi/T$ corresponds to the lowest frequency about which there is information in the data. The so-called Nyquist frequency, viz., $\omega_N=\pi N_0/T$, corresponds to the highest frequency about which there is information in the data. Astronomical data are often not evenly spaced in time, in which case the fundamental frequency remains well defined, but the concept of the Nyquist frequency is altered (see Appendix D in S).

Using the statistical properties of the periodogram (see S) one can define a *detection threshold* z_0 as

$$z_0=-\ln\left[1-(1-p_0)^{1/N}\right]. \tag{17}$$

The meaning of z_0 is that, if the detection of a signal is claimed only when the observed power (signal to noise) exceeds z_0, then the probability that such a signal is due to a chance fluctuation in the noise power is p_0. In this sense p_0 is a *false-alarm probability*. Another useful statistical property of the periodogram is that of *detection efficiency* DE, given by

$$\mathrm{DE}=1-p^*(N,P), \tag{18}$$

where

$$p^*(N,P)=\{1-\exp\left[-(z_0+P)\right]\phi(z_0,P)\}\times(1-p_0)^{1-(1/N)}. \tag{19}$$

Here p^* is the *probability of missing a signal of power* P_s, P is the signal-to-noise power ratio

$$P=\frac{P_s}{P_N}=\frac{N_0}{4}\left(\frac{x_0}{\sigma_0}\right)^2, \tag{20}$$

with x_0 and σ_0 being, respectively, the signal amplitude and the noise variance, and

$$\phi(z_0,P)=\sum_{m=0}^{\infty}\sum_{k=0}^{m}\frac{z_0^k P^m}{k!m!}. \tag{21}$$

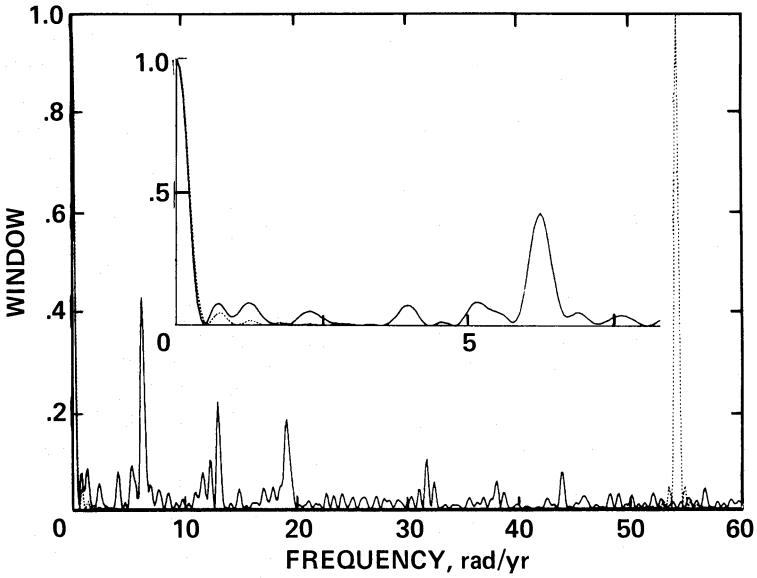

FIG. 4.—The classical spectral window for the 107 samples of the USNO study of G96-45. The dotted line is the (sinc^2) window which would obtain if the data were evenly spaced with the same average spacing as the actual observations. The strong alias peak at about 54 rad yr^{-1} is absent in the real window, but the pseudoalias peaks at multiples of 2π, due to the annual nature of the samples, are evident. In the inset, one can see that the near sidelobes of the real window are somewhat larger than those for even spacing, but the resolution (width of the main lobe) is essentially the same.

The number of frequencies sampled in the periodogram is $N \leq N_0/2$.

An important aspect of using any spectral analysis method, such as the periodogram, is understanding the way in which the analysis system responds to input data. (The term *system* is here taken to be the mathematical operation of transforming the data, as sampled in the time domain, into the frequency domain.) The normal way to calibrate this system is to determine its response to a known signal such as a DC signal or monochromatic sine wave. The system output in this calibration is called the *spectral window*. A detailed discussion of *spectral windows* is given in S, but it is useful to indicate the main features of this important concept.

Consider data which constitute a harmonic signal of frequency ω_0, and denote the Fourier transform and power spectral windows of the analysis system by $W(\omega)$ and $G(\omega)$, respectively. As a consequence of imperfections in the sampling (a finite-length data record and with finite intervals between samples), the spectral analysis of the monochromatic input will generally lead to nonzero power (sidelobes) appearing at frequencies other than ω_0. The spectral window gives a full characterization of this *spectral leakage* effect. If the data format is one of even spacing (in time), then $G(\omega)$ will also have large-amplitude peaks at frequencies well displaced from the input frequency. The displacement $\Delta\omega$ is quantized in units of the Nyquist frequency, $\Delta\omega_k = \pm 2k\omega_N$. This effect is referred to as *aliasing*, and it is due to periodicity in the data format (i.e., evenly spaced data). A type of pseudoaliasing occurs if there is quasiperiodicity in the data format. This often happens in astronomical work where one finds diurnal, monthly, or annual effects evident in the observational program. Such quasiperiodicity would lead to alias peaks displaced from the peak at $\omega = \omega_0$ by multiples of 2π radians per day, month, or year. Another important issue concerns the ability of the analysis system to *resolve* multiple signals at various frequencies. Here again, the spectral window indicates the resolution of the analysis system.

A graphic representation of the spectral window concept is shown in Figure 4. The window in this figure was generated using a DC signal that was sampled according to the format employed in a 12 yr study of the star G96-45 (data were kindly provided by Dr. R. Harrington). Various aspects of the spectral window are described in the figure legend. Due to the uneven sampling format there is no evidence of aliasing; however, there is clear evidence of pseudoaliasing due to the annual character of the observing program (i.e., parallax studies).

A further point to note about the periodogram is that it is *not* linear in the input data, $X(t)$ (see eq. [14]). As a consequence, the response to a sinusoidal signal of frequency ω_0 will, in general, contain peaks located at $\omega = \pm\omega_0$, as well as an overlap term (see Appendix D in S). The peak at $\omega = -\omega_0$ and the overlap can be negligible if ω_0 is large enough, but such is not always the case.

c) *Effects Due to Orbital Eccentricity and Multiple Companions*

The formulae given above describe the case of motion induced by a single companion in a circular orbit about

its star. The assumption of circular orbits assures the spectral purity of the signal one is trying to detect. That is, $P(\omega)$ will have a large value at a single frequency. The motion for noncircular orbits is still accurately periodic, but the projections (x and y) of the planet's motion are not sinusoidal functions of time—because of the ellipticity of the orbit and the fact that the planet does not move along the orbit at a constant speed. As is known from the theory of Fourier analysis, such nonsinusoidal, but periodic, motion can be represented as a series containing a sinusoidal term at the fundamental frequency (corresponding to the period of the motion) plus harmonics at multiples of this frequency. The effect for small eccentricities should be quite small, and, even if the eccentricity is of order unity, the motion is still largely harmonic and the fundamental should still dominate. Numerical results given by Jensen and Ulrych (1973) confirm these expectations. These authors find that even for e (the eccentricity) equal to 1, the ratio of the power in the first harmonic to that in the fundamental does not exceed 16% (even for the worst case orientation) and averages 12%. This ratio does not exceed 2.2% for $e < 0.3$. The corresponding values for the second and third harmonics, relative to the fundamental, are smaller by roughly one-half and a full order of magnitude, respectively. Although the effect of eccentricity is generally small, it could be detectable. Once a planetary perturbation is suspected, the first harmonic ($2\ \omega_s$) should be inspected for the presence of a possibly significant signal. (Note that the strict periodicity of the motion implies that any real signals at harmonics must occur at *exactly integral* multiples of the fundamental. There may be only the small shift due to the fact that ω_s may not lie exactly on one of the ω_n.) One could possibly maximize sensitivity by inspecting the quantity $P(\omega) + aP(2\omega) + bP(3\omega) + \cdots$ for signals, where $a, b, \ldots$ are coefficients whose values might be chosen based on a guess as to the likely range of eccentricities.

The existence of more than one planet in orbit around a star would add further to the complexity of the power spectrum. In principle, with the assumption that the periods of the several planets are not commensurate, each planet would produce a simple, distinct peak in the power spectrum—each one at a value of ω_s corresponding to its orbital period. This would be realized in practice if the sampling interval is moderately long compared with the longest period present. In practice, the data for such a system would probably be analyzed before the observation interval became as long as the longest period present, so that there would be some confusion at the low-frequency end of the power spectrum. The existence of the shorter period planets would probably not be hidden—but the interpretation of the power spectrum would be somewhat confused, especially if the window function has large sidelobes.

III. EXAMPLES: THE USNO DATA FOR G96-45, G146-72, AND WOLF 1062

We shall now demonstrate the results discussed in § II by applying them to a specific set of data. We use the astrometric data for three stars obtained at the United States Naval Observatory (USNO) (Behall and Harrington 1976; Harrington 1977). Dr. Harrington kindly provided their reduced data, in the form of measured x and y coordinates for each plate, covering the interval 1965.8–1979.3. This represents approximately an additional 3 yr of data beyond what was presented in the referenced publications. The following discussion will center on G96-45, which was analyzed most thoroughly. Its reported perturbation has the longest period, and its detection is therefore the most challenging.

a) The Proper Motion Effect

It was noted in § II that, if the length of the data record involving a suspected sinusoidal perturbation is $\leq$ twice the period of that perturbation, there can be errors in estimating the proper motion of the star under study. The periods assigned by the USNO group to the perturbations for the three stars mentioned above are 7.2 yr (G96-45), 6.7 yr (G146-72), and 2.3 yr (Wolf 1062). Because its suspected period is the longest, the perturbation in the motion of G96-45 should lead to the largest error in estimating the proper motion. Using the equations in § II we find that the errors in estimating the proper motion of G96-45 using the original 9.28 yr of data are $\sim +0''.0025\ \mathrm{yr}^{-1}$ in both x and y coordinates. These errors are comparable to the actual difference between the USNO proper motion solutions for the longer (12.306 yr) and shorter (9.28 yr) data records on G96-45, suggesting that the difference in the estimate of proper motion can be attributed to this effect.

b) The Window

We next analyze the residuals from the parallax/proper motion fit for the presence of periodic perturbations. We do not attempt to correct for the proper motion error mentioned above because this can be done only in terms of a specific model of the perturbation, which for the purposes of this demonstration is as yet undiscovered.

The first step is to compute the window function $G(\omega)$ for the sampling times used in obtaining the 107 plates on G96-45. Nowhere in this discussion are the individual data points averaged into seasonal normal points. The window so calculated is shown in Figure 4. This window is similar to that which one obtains for evenly spaced data (also shown in the figure for comparison). The full widths at half-maximum are essentially identical for the two windows, but the sidelobe amplitudes of the window based on the actual sampling

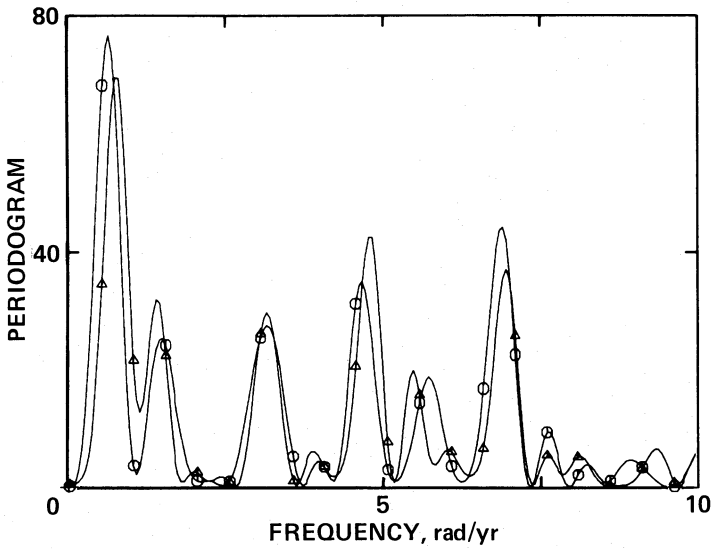

FIG. 5.—The periodogram of the x (*circles*) and y (*triangles*) residuals in the USNO astrometric study of G96-45. As detailed in the text, the peaks at frequencies corresponding to periods of about 9 yr are highly significant and verify the USNO conclusion that a periodic perturbation is present.

times are somewhat greater than those for evenly spaced data. Thus the resolution properties of the unevenly spaced power spectrum are only slightly inferior to those of the equivalent spectrum for evenly spaced data. Logarithmic plots of this window and the ones for the other two stars are shown in Figure 3 of S.

The window for the actual data is superior to that obtained for evenly spaced data in terms of its aliasing characteristics. The latter window has peaks of unit amplitude (i.e., identical to the amplitude of the main lobe located at zero frequency) at the set of alias frequencies $\omega_k = (2\pi k/\Delta t)$, whereas $G(\omega)$ has no trace of a peak at any of the higher frequencies (e.g., the Nyquist frequency, $2\pi/\langle\Delta t\rangle$). This suppression arises because aliasing depends on precise phasing deriving from equality of the time interval between data points, and uneven sampling destroys these coherent phase relationships. The sampling intervals in the USNO data possess a high degree of randomness and include many rather short intervals (e.g., plates taken on successive nights), so that aliasing is effectively eliminated. There remains, however, a strong peak at $\omega = 2\pi(f = 1\ \mathrm{yr}^{-1})$ which derives from the annual nature of the parallax program that led to these data. We refer to this type of peak as a *pseudoalias* peak and stress that care must be taken in interpreting any spectrum at frequencies that are displaced from a strong peak by integral multiples of 2π.

c) *The Periodogram*

Figure 5 shows the periodogram for the star G96-45, calculated with equation (14), using the residuals provided by Harrington for the USNO solution ($\pi = 0''.0625 \pm 0''.0034$, $\mu = 0''.3601 \pm 0''.0006$ toward position angle $172°.0 \pm 0°.1$). These residuals, x and y, are in microns on the plate (plate scale $= 13''.55\ \mathrm{mm}^{-1}$), so the power is in units of μ^2. As will be detailed below, this means that the ordinate in Figure 5 is numerically equal to the power signal-to-noise ratio. The points indicate the power as evaluated on an evenly spaced Nyquist grid. As discussed, these points contain essentially all of the information found in the periodogram. The lines drawn through the points are the interpolations obtained by evaluating the periodogram at intermediate frequencies.

The first thing to note is what is *not* there, namely, power at the frequency $\omega = 2\pi$ rad yr^{-1}. Because the parallactic displacement is exactly periodic at this frequency, the component of any perturbation at 1 cycle yr^{-1} is absorbed into the parallax and should not appear in the spectrum. This expectation is realized, because the observed power at this frequency is about 1 μ^2, consistent with the level of the observational noise power.

Next, a strong peak ($P \sim 70$) appears near the low-frequency end of the spectrum, at $\omega_s = 0.66$. It should be noted that the reduction procedure—specifically the proper motion determination—forces the power to be small (or zero) at zero frequency. [It can be seen from eq. (14) that $P_X(0)$ is $N\bar{x}^2$, and the fitting procedure forces $\bar{x} = 0$.] Hence, if the true spectrum has a continuous low-frequency component, for example, monotonically decreasing from a peak at zero frequency, the calculated spectrum can have an artificial peak at low frequency (Terrell and Olsen 1970). In this case it seems unlikely on astrophysical grounds that such a continuous component, which would indicate the presence of a

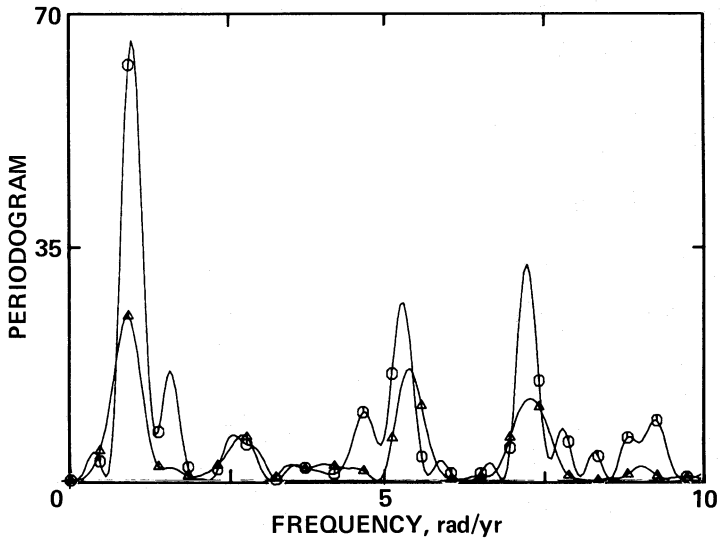

FIG. 6.—The same as Fig. 5, for G146-72

random perturbation with a strongly nonwhite spectrum, exists. An additional reason for believing that this peak corresponds to a monochromatic signal is that it mimics the overall shape of the window, including the pseudoalias peak located at 2π rad yr^{-1} above the signal frequency, as well as the negative spillover located at $\omega = -\omega_s + 2\pi$ rad yr^{-1}. The peaks in the range $2 \lesssim \omega \lesssim 5$ rad yr^{-1} are probably due to the cross-product term discussed in § II. These considerations can be made more quantitative (S), but that will not be done here.

Let us now evaluate the statistical significance of the peak at $\omega = \omega_s$. The maximum power level, at $n = 1.3$, is $P_s = 77\ \mu\text{m}^2$, but this peak falls between the ω_n—i.e., it is an interpolated value and was discovered only because Figure 5 is oversampled (we produced the figures by oversampling by a factor of 10). The statistical analysis only applies to the maximum over the ω_n, as discussed above. This maximum is $P(\omega_1) = 68.5\ \mu\text{m}^2$. This power is converted into a signal-to-noise ratio by dividing by an estimate of σ_0^2, the noise variance. We computed the variance in the x and y displacements for all time intervals less than about 0.1 yr. Since there is presumably no real variation on such short time scales, this computation should yield the variance associated with the observational errors. The result is

$$\sigma_0 = (1.03 \pm 0.11)\,\mu\text{m}, \qquad (22)$$

with no statistically significant evidence for variation from star to star or for a difference between the x and y variances. It is thus convenient and accurate to adopt $\sigma_0 = 1\ \mu$m, because then $P_N = \sigma_0^2 = 1\ \mu\text{m}^2$, and the signal-to-noise ratio $P = P_s/P_N$ and the signal power P_s are numerically equal. Hence, one can interpret the ordinate in Figures 5, 6, and 7 as either power levels or signal-to-noise ratios. (If the estimate of σ_0 should ever be modified, the signal-to-noise values would have to be rescaled accordingly.) The chance probability of finding a peak power greater than or equal to 68.5 with no signal actually present is given by equation (14) of S with $z = 68.5$ and $N = 107/2$, namely, 1.6×10^{-28}. Actually N should be much smaller, because we are effectively ignoring frequencies higher than 1 yr^{-1}, but this hardly matters. The maximum y power (for integer n) is $P(\omega_1) = 34.8$, corresponding to the chance probability of 5.6×10^{-14}. Thus in both coordinates we have highly significant peaks, well above the threshold for any reasonable value of p_0 and occurring at the same frequency (to the nearest integral n). The best estimates of the signal power are just one less than the observed power, which includes both the signal and the noise power. Another interesting quantity is the *range* of signal powers that could have yielded the observed power, at a specified confidence level. This range can be easily read off of the figures given by Groth (1975) by drawing a vertical line at the P value observed and reading off the P_s values at which this line intersects the contours of the appropriate value of the cumulative distribution function (f_1 in his notation). For example, with a confidence of 98%, the signal power lies within the intersections with the contours $f_1 = 0.01$ and $f_1 = 0.99$. Thus, reading approximate figures from Groth's Figure 1 and using the above observed power levels, we arrive at the following estimates of the true power:

$$P_s(x) = \left(67.5^{+35}_{-22}\right)\,\mu\text{m}^2,$$

$$P_s(y) = \left(34.8^{+25}_{-18}\right)\,\mu\text{m}^2.$$

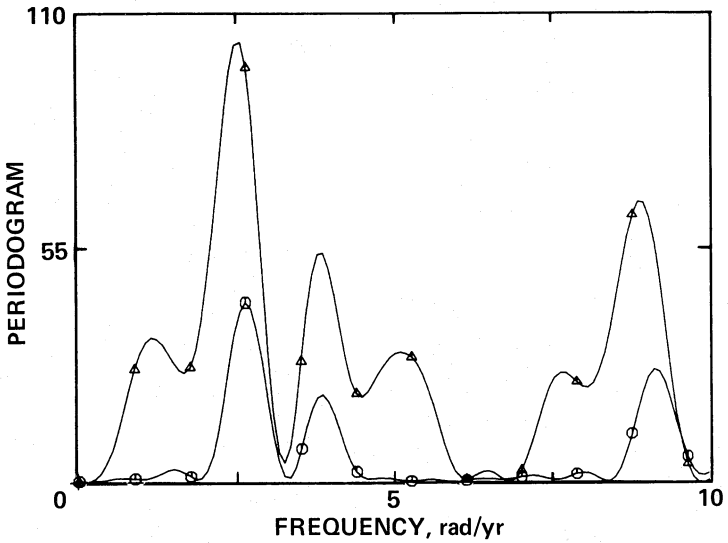

FIG. 7.—The same as Fig. 5, for Wolf 1062

TABLE 1

RESULTS OF PERIODOGRAM ANALYSIS OF USNO ASTROMETRIC DATA

STAR	RESULTS FROM THIS PAPER						USNO RESULTS		
	P_x (max)	x_0 (μm)	T_x (yr)	P_y (max)	y_0 (μm)	T_y (yr)	T (yr)	x_0 (μm)	y_0 (μm)
G96-45	68.5	1.60	9.6	34.8	1.14	8.3	7.2	1.2	2.6
(Interpolation)	69.1	1.61	10.4	56.8	1.46	8.3	...	...	...
G146-72	62.0	1.45	6.1	24.6	0.91	6.4	6.7	1.9	1.1
(Interpolation)	68.4	1.52	6.1	13.3	0.67	7.1	...	...	...
Wolf 1062	42.2	1.43	2.3	97.0	2.16	2.4	2.3	1.5	2.2
(Interpolation)	30.1	1.20	2.5	66.0	1.78	2.3	...	...	...

Shown in Table 1 are the results of the periodogram analysis on both the x and y residuals for the three stars in the USNO program. The USNO amplitudes were read off Figures 1 and 2 of Behall and Harrington (1976) and Figure 1 of Harrington (1977). It can be seen that (1) for these cases interpolation is not very different from periodogram analysis, presumably because the frequency of the signal is relatively low; (2) there is a tendency for the period to be less well determined (e.g., as shown by comparing the x and y results) when the period is of the same order as the data interval; and (3) the periods and amplitudes found by Harrington (based on several years' less data) are in good agreement with our results, except for G96-45, where the spurious proper motion effect may be important. The comparison is based on slightly different data, as our results were obtained with the full data set and the USNO results were based on the data through about 1976 (the parallax and proper motion solutions are slightly different for the two data sets).

In addition, there is a somewhat subtle point which arises in connection with the determination of the best estimates of the amplitude and frequency of the signal. As seen in Figure 5, the peak of the oversampled P_X curve comes close to one of the grid frequencies given by equation (16), whereas for P_Y the peak comes approximately midway between two of the grid frequencies. Thus the maximum over n of the $P(\omega_n)$ gives estimates of this peak power and the corresponding frequency which are not the best estimates. Indeed, the best procedure is to fit the window function to the data $P(\omega_n)$, allowing arbitrary shifts in frequency. The resulting power can be used to determine the signal amplitude (using eq. [7] of S). While this is a better estimate of the amplitude, the statistical formulae summarized in S refer only to the $P(\omega_n)$ with integer n. Because we are primar-

ily interested in statistical matters, the maximum powers (e.g., in Table 1) used here are for integer n. On the other hand, the periods quoted in Table 1 correspond to the maximum of the oversampled curves in Figures 5, 6, and 7—and thus give the best estimates of the period.

IV. APPLICATION TO THE DETECTION OF OTHER PLANETARY SYSTEMS

A basic goal of this paper is to provide means for evaluating the efficiency of various astrometric techniques for searching for other planetary systems. We next indicate how the statistical results summarized in § II can be used in such an evaluation. There are many ways in which such data can be obtained, and we cannot analyze every possible observational scheme. What is presented here is a general discussion followed by application to one specific situation. It is hoped that this discussion will provide and exemplify tools that can be used in a variety of situations.

a) The Basic Observational Systems (a Gedankenexperiment)

We now outline a kind of gedankenexperiment which is an idealized representation of a possible astrometric program. It will be used to set up the efficiency analysis in a precisely defined way and is not meant to represent in detail any real or practical astrometric survey.

We assume the following. Positional measurements are made for a set of stars, M_0 in number, each one of which is observed with the same accuracy. The errors of observation are normally distributed (with a Gaussian distribution function of variance σ_0^2 about a mean of zero), and therefore this assumption is the same as saying that σ_0 is the same for all stars in the program. The observations extend over a period of T_0 yr, after which the data are to be analyzed via the periodogram for planetary-induced (i.e., periodic) signals. We examine the periodogram calculated for each star for the presence of strong signals and ask, How likely is it that these signals are due to one or more planets around this star? (Alternatively, we could analyze the data as they are taken and continuously ask how likely it is that the largest signal detected so far is real. This may be closer to procedures used in practice, but such a scheme is evidently biased toward the situation where the suspected planetary period is on the same order as the observational interval—and this rough equality has the important consequences outlined in § II*a*.) There are N_0 observations of each star, spaced at a constant time interval, Δt_0, so that $T_0 = N_0 \Delta t_0$. Let it further be assumed that M_* out of the M_0 stars do indeed have a single planet of period T_* which produces an astrometric signal of amplitude X_*. (To simplify the discussion we assume that the observations refer to a single coordinate x and ignore the orthogonal coordinate.) The phase of the planetary perturbation at the beginning of the observations will be denoted ϕ_*; then the perturbation can be written

$$\delta X_*(t) = X_* \sin\left(2\pi t/T_* + \phi_*\right) \qquad (0 \le t \le T_0). \tag{23}$$

(In general one would be concerned with averages over a random array of phases ϕ_*, and, for that matter, over the other parameters T_* and X_*. But for now we assume fixed constant values for these parameters.) It will further be assumed that it is undesirable to report the detection of a signal that is not present (although one could argue that such false alarms will be very soon weeded out by further, presumably intensified, observations). Therefore, the false-alarm probability p_0 will be made as small as is practical.

This gedankenexperiment then has the following *system parameters* (i.e., referring to the observational procedures):

1. The period of observation T_0,
2. The observation interval Δt_0,
3. The number of data points per star N_0,
4. The variance of the observational errors σ_0,
5. The number of stars surveyed M_0,
6. The false-alarm probability per star p_0;

and the following *signal parameters* (i.e., referring to the stars and planets):

7. The period of the planetary orbit T_*,
8. The phase of the planetary motion ϕ_*,
9. The amplitude of the planetary motion X_*,
10. The number of stars in the survey with such planets M_*.

b) The False-Alarm Rate and Detection Efficiency

We shall now discuss the false-alarm probability p_t for the *total gedankenexperiment* in terms of the false-alarm rate per star, p_0. Note that, because a false alarm is due to noise fluctuations, p_0 is independent of the signal parameters. If the data for each of the M_0 stars in the gedankenexperiment are treated in the manner outlined in § II, then the probability of M' false alarms is given by

$$p_t(M') = \left[1-(1-p_0)^{M_0}\right]^{M'}, \tag{24}$$

$$= \left\{1-\left[1-\exp\left(-z_0\right)\right]^{NM_0}\right\}^{M'}, \tag{25}$$

where equation (17) has been used to express p_0 in terms of the detection threshold (z_0) for a single star. As one would expect, the probability of a single false alarm ($M' = 1$) in a study of the M_0 stars can be quite large even if the false-alarm probability for a single star (p_0) is small. We will examine specific quantitative examples

later in this section. Equation (24) can be used to define a detection threshold in terms of $p_t(M')$, viz.,

$$z_0 = -\ln\left\{\left[1-(1-p_t)^{1/M'}\right]^{1/NM_0}\right\}. \quad (26)$$

We earlier defined a detection efficiency for observations of a single star. Generalizing that concept in the context of our gedankenexperiment is not as straightforward as is the case for the false-alarm probability because the detection efficiency does depend on the signal parameters. Rather than introduce additional formalism into this discussion in terms of assumed distributions of signal amplitudes X_*, we shall consider only the detection efficiency for observations of a single star (note that, *if* all of the M_* stars in our sample which have companions also have the same signal amplitude, then the detection efficiency for the total experiment is identical to the detection efficiency for the single-star case). The power level for the assumed planetary-induced signal is

$$P_* = \frac{N_0}{4}\left(\frac{X_*}{\sigma_0}\right)^2, \quad (27)$$

and the detection efficiency DE is given by

$$\mathrm{DE} \approx \exp\left[-(z_0 + P_*)\right]\phi(z_0, P_*). \quad (28)$$

The function ϕ is defined by equation (21), and we have ignored the term $(1-p_0)^{1-(1/N_0)}$, which for reasonably small values of p_0 is nearly constant with N_0 and is roughly equal to 1.

c) *Maximizing the Detection Efficiency*

It is desirable to pick the system parameters (§ IV*a*) to minimize the probability of missing a given real signal, while maintaining an acceptably low false-alarm rate p_0. This is a complicated problem in general because the detection efficiency depends on both the signal *and* system parameters. What follows is a simplified discussion of some of the general features of this problem. Specific quantitative examples are given in the following subsection.

The system parameters for the gedankenexperiment fall into three categories: (1) those on which DE explicitly depends (N_0, σ_0), (2) those on which DE does not depend but which are important in evaluating detection systems $(T_0, \Delta t_0)$, and (3) those which are fixed by other considerations (M_0, p_0). It should be noted that DE does depend on M_0 through the detection threshold z_0 (eqs. [26] and [28]); as more stars are observed the threshold must be raised to maintain a fixed false-alarm probability for the whole gedankenexperiment. However, the general features of maximizing DE can be seen by considering the single-star case $(M_0 = 1)$. We will also assume a fixed value of p_0. This means that the category 3 parameters are fixed and that z_0 is an explicit function of only one parameter, N_0.

Consider now the role of category 2 parameters—parameter in the case of evenly spaced data, as then $\Delta t_0 = T_0/N_0$. In this highly idealized gedankenexperiment, category 2 parameters do not enter explicitly. They do, however, serve to define the practical frequency range for a detection program in the sense that one would like the signal frequency ω_s to fall in the range $\omega_1 \ll \omega_s$, where $\omega_1 = 2\pi/T_0$. It is sometimes considered that one should also have $\omega_s \lesssim \omega_N = 2\pi/\Delta t_0$. However, this condition is necessary for correct determination of the period, *not* for detection of the signal. The signal can be detected even if $\omega_s \gg \omega_N$—the problem is that the signal frequency will be "aliased down" only to a frequency between 0 and ω_N. These requirements on ω_s clearly indicate that, all things being equal, the larger the value of ω_N and the smaller the value of ω_1, the better off one is. We discuss possible astrophysical constraints on ω_s below.

As one would guess, it is the category 1 parameters, N_0 and σ_0, which are most crucial in terms of maximizing DE. Imagine that one has two separate instruments which could be used to conduct our gedankenexperiment. In comparing the merits of these two instruments we will assume that they are both to detect the same signal (X_*), both are used for observations covering the same total interval of time (T_0), and that we assign the same false-alarm probability (p_0) to each data set. We have then the following:

$$z_{0,i} = -\ln\left[1-(1-p_0)^{1/N_i}\right], \qquad N_i \le 0.5N_{0,i},$$

$$P_{*,i} = \frac{N_{0,i}}{4}\left(\frac{X_*}{\sigma_{0,i}}\right)^2,$$

$$\mathrm{DE}_i = 1-\exp\left[-(z_{0,i} + P_{*,i})\right]\phi(z_{0,i}, P_{*,i}), \quad (29)$$

where the subscript i denotes the instrument (1 or 2) and $N_{0,i}$ and $\sigma_{0,i}$ are, respectively, the total number of observations with instrument i and the noise variance of instrument i (N_i is the number of frequencies investigated for the presence of a signal).

General remarks on possible ways to maximize DE are given in the companion paper (S), so we will only summarize the major points of that discussion here. If one assumes that the errors associated with a given instrument are normally distributed and independent, so that $\sigma_{0,i}^2 \propto N_{0,i}$, it then follows that DE is maximized *for a given instrument* by making $N_{0,i}$ as small as possible. This can be accomplished chiefly in two ways. One way is to average the available data set, obtaining a smaller number of higher accuracy data points. A danger is that this procedure increases the effective Δt_0, reducing ω_N.

Such averaging should not lead to $\omega_N \ll \omega_s$. Another possibility is to reduce N_i, the number of frequencies examined for the presence of a signal. For a given p_0, this reduces $z_{0,i}$ but leaves P_i unchanged and thereby increases DE. An obvious danger with this procedure is that one could miss the signal frequency.

d) Quantitative Comparison of Two Idealized Systems

The following discussion gives a quantitative comparison between two idealized astrometric systems. For the purpose of this comparison we assign the following values to the category 3 parameters: $M_0 = 1$, $p_0 = 0.01$. Values for the category 2 parameters, Δt_0 and T_0, are estimated as follows. If we assume that gas-giant planets (e.g., Jupiter) can only form if the temperature in a putative circumstellar nebula is less than ~ 200 K, a temperature appropriate to condensation of icy material, and, if we further suppose that nebular temperatures are set by radiative equilibrium with the luminosity of the central star in that nebula, one has the following constraint on T_0:

$$T_0 \gtrsim T_{0,\min} = 2.7 \frac{L_*^{3/4}}{M_*^{1/2}} \text{ yr}, \qquad (30)$$

where L_* and M_* are, respectively, the luminosity and mass *of the central star expressed in solar units*. (We are not suggesting that this approach is the correct way to estimate where or whether gas-giant planets will form. It is only intended to obtain a quantitative estimate and to indicate that orbital period data *can be used to test models of planet formation*.) Using equation (29) one obtains the T_0 values shown in Table 2 as a function of stellar spectral type. Also given in Table 2 are values of Δt_0 under the assumption of even spacing of data (i.e., $\Delta t_0 \leq t_{0,\max} = 1/2\ T_{0,\min}$).

Recalling the caveat expressed above, one can draw the following *assumption-dependent* conclusions from Table 2. First, it appears to be impractical to use indirect techniques such as astrometry to search for massive planetary companions to stars of spectral type earlier than about B5; the orbital periods are too long. This is not a serious restriction because most stars are of later spectral type. (Note that if Jupiter-mass planets could exist at temperatures of 1000 K, the associated orbital period around a B5 star would be ~ 6.25 yr.) Second, if one wishes to search for planetary companions to the numerous late spectral type stars (i.e., K and M spectral types), one must take data *at least* 4 times per year. We will return to this point below. For convenience we will take $T_0 = 5$ yr and require that $\Delta t_0 \lesssim 0.25$ yr (i.e., $N_0 \gtrsim 20$ for evenly spaced data).

TABLE 2

ESTIMATES OF MINIMUM ORBITAL PERIOD AND MAXIMUM OBSERVING INTERVAL AS A FUNCTION OF STELLAR SPECTRAL TYPE

Spectral Type	$T_{0,\min}$(yr)	$t_{0,\max}$(yr)
B0	780	390
A0	38	19
F0	8	4
G0	3	1.5
K0	1.5	0.75
N0	0.5	0.25

Finally, we specify the category 1 parameters, N_0 and σ_0. System 1 we characterize by $\sigma_0 = 10^{-3}$ arcsec and $N_0 = 200$. This value of σ_0 is better (i.e., smaller) than that attainable with existing ground-based telescopes and is comparable to that expected for both the Space Telescope and the European Space Agency's HIPPARCOS satellite. However, current developments in astrometric detectors indicate that $\sigma_0 = 10^{-3}$ arcsec is attainable for ground-based astrometric systems. The parameters for System 2 are $\sigma_0 = 10^{-5}$ arcsec and $N_0 = 20$. This value of σ_0 is attainable, in principle, with a specially designed space-based astrometric telescope.

We compare these two gedanken systems on the basis of their detection efficiency as a function of signal amplitude. The DE of each system can also be determined for signals expected from specific sample planetary systems. Shown in Figure 8 are the DE characteristics of these two systems as a function of signal amplitude (expressed in arcsec). The solid curve is that for System 2 and the dotted curve is that for System 1. The dashed and dot-dashed curves pertain to the characteristics of current ground-based *photographic* observations. The vertical lines on Figure 8 indicate the signal amplitude expected from a planet with the mass of Jupiter (J), Saturn (S), or Neptune (N) revolving around a K5 dwarf with a 5 yr orbital period. The numerical subscripts designate the distance (in pc) that the K5 star is from the Sun. System 1 has a DE ~ 1 for signals similar to those expected from a Jupiter-mass planet revolving around a K5 star located 5 pc from the Sun, but it would have a very low (~ 0.03) DE for detecting Saturn-mass planetary companions to such a star. In contrast, System 2 has a DE $\gtrsim 0.8$ for detecting either a Neptune-mass planet around any K5 star within 30 pc of the Sun or a Jupiter-mass planet around any K5 star within 600 pc of the Sun. (Perhaps a more exciting measure of performance is that System 2 would have a DE ~ 0.8 for detecting Earth-mass planets around M dwarfs within ~ 7 pc of the Sun!)

System 2 is clearly more effective than System 1 on the basis discussed above. A crucial aspect of the DE is the signal power to noise power ratio (eq. [28]); the higher the value of P_i, the higher the DE for a given signal strength X_*. Thus, while 200 data points over a 5 yr interval with System 1 are inadequate to detect a

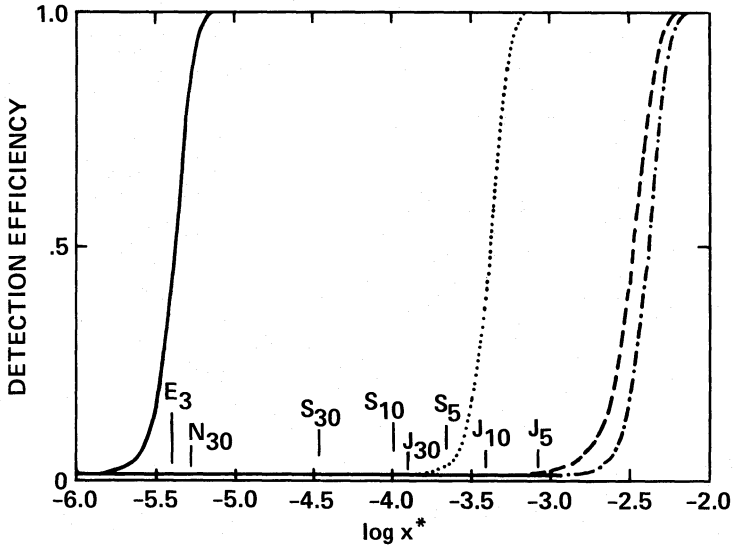

FIG. 8.—Detection efficiencies for two specific astrometric systems (*solid and dotted curves*), as described in the text, as functions of the amplitude of the unknown perturbation. Also shown are the efficiencies of currently available ground-based photographic observations (*dashed and dot-dashed curves*).

signal of 10^{-4} arcsec, increasing the number of data points will increase the signal-to-noise ratio. Using the system parameters assumed above, we find that one must obtain $20\times(10^{-3}/10^{-5})^2$ data points with System 1 to obtain a signal-to-noise ratio equal to that obtained in 20 data points with System 2. This number of data points, if taken at a rate of 40 per year, corresponds to 5000 yr of observing! If we relax the requirement on System 1 to that of requiring a DE ~ 0.8 at $X_*=10^{-4}$ arcsec, one still needs ~ 150 yr of data (again assuming 40 points per year).

It is clear that, in terms of detecting other planetary systems, System 2 is far superior to System 1. This superiority goes beyond the obvious gains attendant upon two orders of magnitude more accuracy; the essential point is that *the major scientific returns from a search for other planetary systems lie in the signal-strength realm of* $X_*\leq 10^{-4}$ *arcsec*. Given the rather sharp decrease in DE as a function of X_* once $X_*<\sigma_0$, a search program should employ a system with $\sigma_0<10^{-4}$ arcsec.

It was remarked earlier that one would need a Δt_0 of 3 months or less to be assured of adequate data sampling. Such an observation frequency would seem to require a free-flying space system, rather than a shuttle-based instrument package.

V. SUMMARY

The principal goal of this paper is to present an objective means of analyzing astrometric data for the presence of a periodic signal (or signals) such as a perturbation in a star's motion due to a planetary companion (or companions) to that star. The analysis method presented here is a type of power-spectral analysis, namely, the periodogram. In addition, this analysis technique is used to obtain a quantitative comparison of the efficiencies of various astrometric systems for detecting planetary companions to stars, as well as to establish a guide to optimization of parameters in a planetary detection program.

One finding of this study concerns an effect which has been overlooked or ignored in most published analyses of astrometric data, namely, that the linear component of a periodic perturbation in a star's motion is erroneously absorbed into the proper motion of that star, leading to underestimates of both the period and amplitude of the perturbation. It might be thought that if the available data span more than a reasonable fraction of the period of the perturbation then this error will be small. However, as shown with numerical examples in § II*a*, this is not the case. In particular, one obtains estimates for the amplitude and period of the perturbation which are, respectively, 4% and 13% smaller than their true values even for data covering two orbital periods. (The underestimates for a full period of data are 47% in amplitude and 25% in period.) This error can be avoided *only* if a model of the perturbation is incorporated into the determination of the proper motion. Evidence for this error can be found in a variety of sequential published reports on suspected perturbations in the proper motions of various stars, taking the form of suspected periods being comparable to the total time that an object has been studied (i.e., the apparent period increases with additional data). The apparent period asymptotically approaches the true period when data are available for more than about two orbital periods.

TABLE 3

VALUES OF THE DETECTION EFFICIENCY AS A FUNCTION OF SIGNAL POWER AND NUMBER OF FREQUENCIES

P0=0.1

N= P	1	5	10	20	40	60	80	100	200	500	1000	2000	4000	8000
0	.100	.100	.100	.100	.100	.100	.100	.100	.100	.100	.100	.100	.100	.100
1	.334	.204	.170	.146	.130	.123	.119	.116	.110	.105	.103	.102	.101	.101
2	.542	.344	.280	.230	.192	.175	.164	.157	.139	.123	.116	.110	.107	.104
3	.702	.491	.410	.340	.283	.255	.237	.224	.191	.159	.142	.130	.121	.114
4	.814	.623	.539	.461	.390	.353	.329	.312	.264	.215	.186	.164	.147	.134
5	.887	.733	.655	.577	.502	.461	.433	.412	.353	.287	.247	.214	.187	.166
6	.934	.817	.751	.680	.608	.566	.537	.515	.449	.372	.322	.279	.242	.212
7	.962	.879	.826	.766	.701	.662	.634	.612	.546	.463	.406	.355	.310	.270
8	.979	.921	.882	.834	.779	.744	.719	.699	.636	.553	.494	.437	.386	.339
9	.988	.950	.922	.885	.841	.812	.790	.773	.716	.638	.579	.521	.466	.415
10	.994	.969	.949	.922	.888	.865	.847	.833	.784	.714	.659	.602	.547	.493
11	.997	.981	.968	.948	.923	.905	.891	.879	.840	.780	.730	.677	.624	.570
12	.998	.989	.980	.966	.948	.935	.924	.915	.883	.833	.790	.744	.694	.643
13	.999	.993	.988	.979	.966	.956	.948	.941	.917	.877	.841	.800	.756	.709
14	.999	.996	.992	.987	.978	.971	.965	.960	.942	.910	.881	.848	.810	.768
15	.999	.998	.996	.992	.986	.981	.977	.973	.960	.936	.913	.886	.854	.818
16	.999	.999	.997	.995	.991	.988	.985	.982	.973	.955	.937	.916	.890	.860
17	.999	.999	.998	.997	.994	.992	.990	.989	.982	.969	.956	.939	.918	.894
18	.999	1.00	.999	.998	.997	.995	.994	.993	.988	.979	.969	.956	.940	.921
19	.999	1.00	.999	.999	.998	.997	.996	.995	.992	.986	.979	.969	.957	.942
20	.999	1.00	1.00	.999	.999	.998	.998	.997	.995	.991	.985	.979	.969	.958
21	.999	1.00	1.00	1.00	.999	.999	.999	.998	.997	.994	.990	.985	.979	.970
22	.999	1.00	1.00	1.00	1.00	.999	.999	.999	.998	.996	.994	.990	.985	.979
23	.999	1.00	1.00	1.00	1.00	1.00	.999	.999	.999	.997	.996	.993	.990	.985
24	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.998	.997	.996	.993	.990
25	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.998	.997	.995	.993
26	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.999	.998	.997	.995
27	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.999	.998	.997
28	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.999	.998
29	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.999
30	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.999
31	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999
32	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
33	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
34	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
35	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
36	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

The periodogram was used to analyze astrometric data for three stars (G96-45, G146-72, and Wolf 1062) studied by Harrington and co-workers at the USNO. These data, which were provided to us by Dr. Harrington, include three additional years of observing beyond that available at the time of previous publications concerning these stars (Behall and Harrington 1976; Harrington 1977). Our analysis focused on the star G96-45 because the period of the suspected perturbation in its motion was longer than those for the other two stars and therefore provided a challenge to the detection technique as well as an example where the proper motion effect could be detected.

The periodogram confirms the presence of periodic perturbations in the motions of each of these three stars (see Table 1) and gives good agreement with the findings of previous analyses in terms of perturbation amplitudes and periods. A mild exception concerns the amplitude of the *Y*-component of the perturbation to G96-45 (Table 1); our results are a factor of 2 lower than those of Behall and Harrington (1976).

A parametric means, based on the statistical properties of the periodogram, was developed for intercomparing the detection efficiencies of astrometric systems. It was found that the detection efficiency of a system is determined, not unexpectedly, by its measurement accuracy and the number of observations. One interesting result is that, if the observational errors are proportional to the square root of the number of data points (as in averaging of normally distributed errors), then one maximizes the detection efficiency by minimizing the number of data points (§ IV*c*).

A comparison was made of the detection efficiency for an idealized ground-based astrometric system with that for an idealized space-based astrometric system (§ IV*d*). The measurement accuracy of the former was taken to be 1 mas, while that of the latter was taken to be 10^{-2} mas. If one requires a detection efficiency of 0.8 (20% chance of a missed signal) for a signal amplitude of 0.1 mas and a signal period of 5 yr, some 150 yr of data (assuming 40 observations per year) are required for the ground-based system. In contrast, the space-based system would have 100% detection efficiency for such a signal after only 5 yr (assuming 4 observations per year). In view of the fact that the major scientific returns from a search for other planetary systems involve signals of amplitude $\lesssim 0.1$ mas, these results argue strongly for a space-based astrometric telescope.

TABLE 3—*Continued*

P0=0.01

N= P	1	5	10	20	40	60	80	100	200	500	1000	2000	4000	8000
0	.010	.010	.010	.010	.010	.010	.010	.010	.010	.010	.010	.010	.010	.010
1	.084	.037	.027	.021	.017	.015	.014	.014	.012	.011	.011	.010	.010	.010
2	.204	.098	.071	.052	.038	.032	.029	.027	.021	.016	.014	.013	.012	.011
3	.345	.190	.144	.108	.081	.068	.061	.055	.042	.030	.024	.019	.016	.014
4	.487	.302	.239	.187	.145	.125	.112	.103	.079	.056	.043	.034	.027	.022
5	.615	.421	.348	.284	.229	.200	.182	.169	.133	.097	.075	.059	.046	.037
6	.721	.538	.461	.389	.325	.290	.268	.251	.204	.153	.122	.097	.077	.062
7	.804	.642	.568	.495	.426	.388	.362	.343	.287	.223	.182	.148	.120	.098
8	.867	.731	.664	.595	.526	.487	.459	.439	.377	.304	.255	.212	.175	.145
9	.911	.804	.746	.684	.619	.581	.554	.533	.469	.390	.335	.285	.241	.204
10	.942	.860	.812	.759	.701	.666	.640	.620	.559	.478	.420	.365	.315	.272
11	.963	.902	.865	.820	.771	.740	.717	.699	.641	.563	.504	.448	.394	.347
12	.977	.933	.904	.869	.828	.802	.782	.766	.714	.642	.585	.529	.474	.424
14	.991	.970	.955	.934	.909	.891	.878	.867	.830	.773	.726	.677	.625	.577
15	.995	.981	.970	.955	.935	.922	.911	.902	.872	.825	.784	.739	.692	.647
16	.997	.988	.980	.969	.955	.945	.936	.930	.905	.867	.832	.794	.752	.710
17	.998	.992	.987	.979	.969	.961	.955	.950	.931	.900	.872	.839	.803	.766
18	.999	.995	.992	.986	.979	.973	.969	.965	.951	.926	.903	.877	.846	.814
19	.999	.997	.995	.991	.986	.982	.979	.976	.965	.947	.928	.907	.881	.854
20	1.00	.998	.997	.994	.991	.988	.986	.984	.976	.962	.948	.930	.910	.887
21	1.00	.999	.998	.996	.994	.992	.990	.989	.983	.973	.962	.949	.932	.914
22	1.00	.999	.999	.998	.996	.995	.994	.993	.989	.981	.973	.963	.950	.936
23	1.00	1.00	.999	.999	.998	.997	.996	.995	.992	.987	.981	.973	.964	.952
24	1.00	1.00	1.00	.999	.998	.998	.997	.997	.995	.991	.987	.981	.974	.965
25	1.00	1.00	1.00	.999	.999	.999	.998	.998	.997	.994	.991	.987	.981	.975
26	1.00	1.00	1.00	1.00	.999	.999	.999	.999	.998	.996	.994	.991	.987	.982
27	1.00	1.00	1.00	1.00	1.00	.999	.999	.999	.999	.997	.996	.994	.991	.987
28	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.999	.998	.997	.996	.994	.991
29	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.999	.998	.997	.996	.994
30	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.999	.998	.997	.996
31	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.999	.998	.997
32	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.999	.999	.998
33	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.999	.999
34	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.999
35	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999
36	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

APPENDIX

TABLE OF THE DETECTION EFFICIENCY

Table 3 gives the values of the detection efficiency p^* as a function of the signal power P and the number of frequencies N, calculated from equation (19), using equation (17) for the threshold z_0 and equation (21) to evaluate the Bessel function integral. Separate tabulations are given for two values of p_0, 0.1 and 0.01.

REFERENCES

Barning, F. J. M. 1963, *Bull. Astr. Inst. Netherlands*, **17**, 22.
Behall, A. L., and Harrington, R. S. 1976, *Pub. A.S.P.*, **88**, 204.
Black, D. C. 1980, *Space Sci. Rev.*, **25**, 35.
Black, D. C., and Brunk, W. E. 1980, *An Assessment of Ground-based Techniques for Detecting Other Planetary Systems* (NASA CP-2124).
Firneis, M. G., and Firneis, F. J. 1975, *Astr. Nach.*, **296**, 95.
Gatewood, G., Breakiron, L., Goebel, R., Kipp, S., Russell, J., and Stein, J. 1980, in *An Assessment of Ground-based Techniques for Detecting Other Planetary Systems*, ed. D. C. Black and W. E. Brunk (NASA CP-2124).
Gatewood, G. D., and Stein, J. W. 1980, *Bull. AAS*, **12**, 455.
Groth, E. J. 1975, *Ap. J. Suppl.*, **29**, 285.
Harrington, R. S. 1977, *Pub. A.S.P.*, **89**, 214.
Jensen, O. G., and Ulrych, T. J. 1973, *A.J.*, **78**, 1104.
Lomb, N. R. 1976, *Ap. Space Sci.*, **39**, 447.
Scargle, J. D. 1982, *Ap. J.*, **263**, 835 (S).
Terrell, N. J., and Olsen, K. H. 1970, *Ap. J.*, **161**, 399.
van de Kamp, P. 1967, *Principles of Astrometry* (San Francisco: Freeman).

DAVID C. BLACK and JEFFREY D. SCARGLE: Space Science Division, NASA Ames Research Center, Moffett Field, CA 94035