Chapter 7

APPLICATIONS

Perhaps the greatest test of any theory is not so much how it was derived, but how it works. Here we apply the theory as developed in the preceding chapters to a number of specific examples including: NMR signals, economic time series, and Wolf's relative sunspot numbers. Also, we examine how multiple measurements affect the analysis.

7.1 NMR Time Series

NMR provides an excellent example of how the introduction of modern computers has revolutionized a branch of science. With the aid of computers more data can be taken and summarized into a useful form faster than has ever been possible before. The standard way to analyze an NMR experiment is to obtain a quadrature data set, with two separate measurements, 90° out of phase with each other, and to do a complex Fourier transform on these data [30]. The global phase of the discrete complex transform is adjusted until the real part (called an absorption spectrum) is as symmetric as possible. The frequencies and decay rates are then estimated from the absorption spectrum. There are, of course, good physical reasons why the absorption spectrum of the "true signal" is important to physicists. However, as we have emphasized repeatedly since Chapter 2, the discrete Fourier transform is an optimal frequency estimator only when a single simple harmonic frequency is present, and there are no conditions known to the author under which an absorption spectrum will give optimal frequency estimates.

We will apply the procedures developed in the previous sections to a time series

from a real NMR experiment, and contrast our analysis to the one done using the absorption spectrum. The NMR data used are of a free-induction decay [31], Fig. 7.1. The sample contained a mixture of 63% liquid Hydrogen-Deuterium (HD) and Deuterium (D_2) at 20.2°K. The sample was excited with a 55MHz pulse, and its response was observed using a standard mixer-modulation technique. The resulting signal is in the audio range where it has several oscillations at about 100Hz. The data were sampled at $\Delta t = 0.0005$ seconds, and $N = 2048$ data points were taken for each channel. The data therefore span a time interval of about one second. As was discussed earlier, we are using dimensionless units. The relation to physical units is given by

$$f = \frac{\omega}{2\pi\Delta t}\text{Hz}, \qquad \text{Period} = \frac{2\pi\Delta t}{\omega}\text{Seconds}$$

where f is the frequency in Hertz, ω is the frequency in radians per step, and Δt is the sampling time interval in seconds.

In these data there are a number of effects which we would like to investigate. First, the indirect J coupling [32] in the HD produces a doublet with a splitting of about 43Hz. The D_2 in the sample is also excited; its resonance is approximately in the middle of the HD doublet. One of the things we would like to determine is the shift of the D_2 singlet relative to the center of the HD doublet. In addition to the three frequencies there are two different characteristic decay times; the decay rate of the HD doublet is grossly different from that of D_2 [32]. However, an inhomogeneous magnetic field could mask the true decay: the decay could be magnet limited. We would like to know how strongly the inhomogeneous magnetic field has affected the decay.

The analysis we did in Chapter 3, although general, did not use a notation appropriate to two channels. We need to generalize the notation; there are two different measurements of this signal, (assumed to be independent), and we designate them as $d_1(t_i)$ and $d_2(t_i)$. The model functions will be abbreviated as $f_1(t)$ and $f_2(t)$ with the understanding that each measurement of the signal has different amplitudes and noise variance, but the same $\{\omega\}$ parameters.

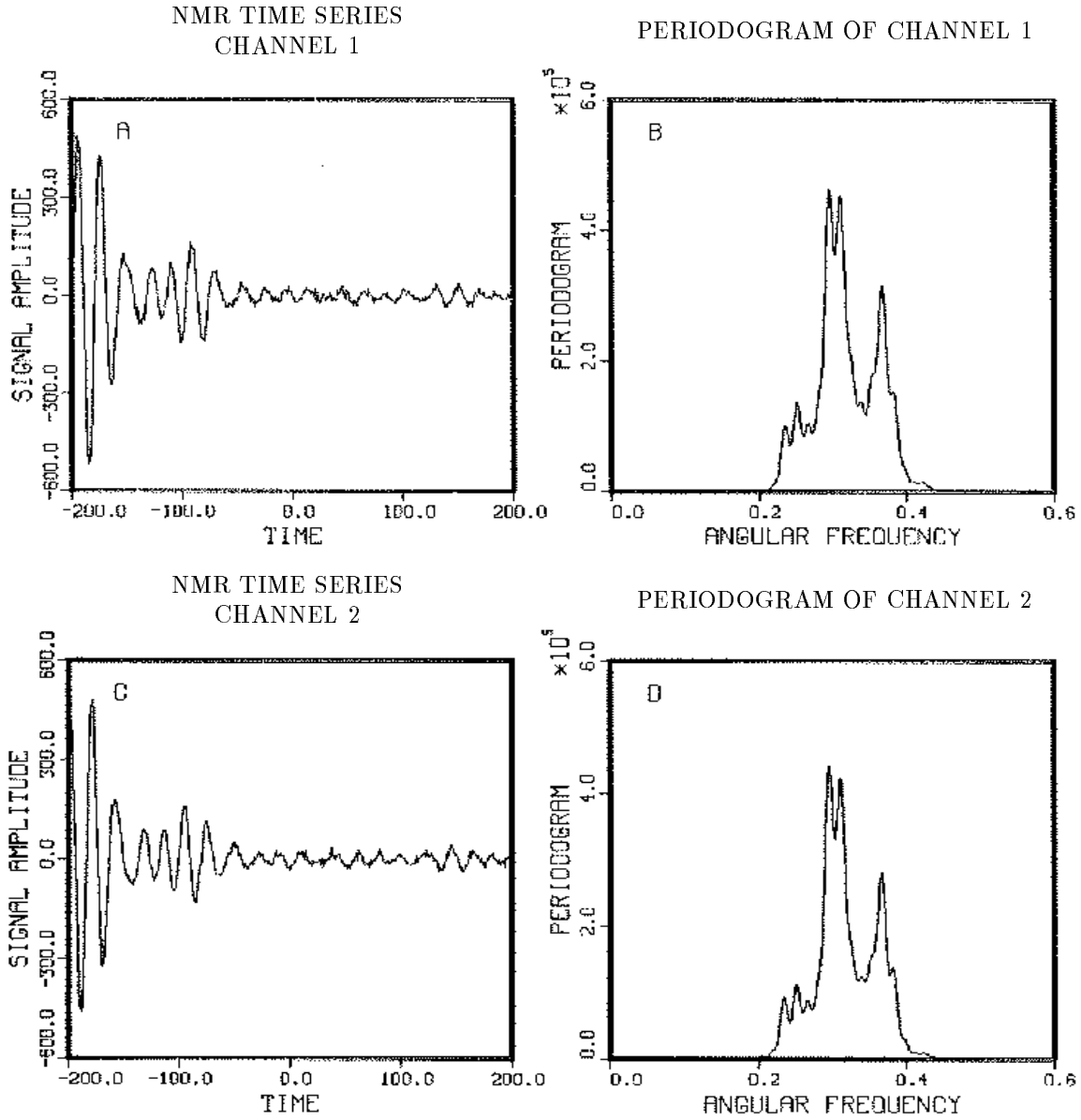
We can write the likelihood (3.2) immediately to obtain

$$L(f_1, f_2) \propto (\sigma_1\sigma_2)^{-N} \exp\left\{-\frac{X}{2\sigma_1^2} - \frac{Y}{2\sigma_2^2}\right\}$$

where

$$X \equiv \sum_{i=1}^{N}[d_1(t_i) - f_1(t_i)]^2$$

Figure 7.1: Analyzing NMR Spectra



The data are channel 1 (A) and 2 (C) from a quadrature detected NMR experiment. The time series or free-induction decay is of a sample containing a mixture of D_2 and HD in a liquid phase. Theory indicates there should be three frequencies in these data: A D_2 singlet, and an HD doublet with a 43Hz separation. The singlet should be approximately in the center of the doublet. In the discrete Fourier transform, (B channel 1) and (D channel 2), the singlet appears to be split.

$$Y \equiv \sum_{i=1}^{N}[d_2(t_i) - f_2(t_i)]^2.$$

Because the amplitudes and noise variance are assumed different in each channel, we may remove these using the same procedure developed in Chapter 3.

As in most of our examples, this procedure is conservative; if we had definite prior information linking the amplitude or variances in the two channels, we could exploit that information at this point to get still better estimates of the $\{\omega\}$ parameters. For example, if we knew that the noise was strongly correlated in the two channels, that would enable us to estimate the noise in each channel more accurately. After removing the nuisance parameters, the marginal posterior probability of the $\{\omega\}$ parameters is just the product of the "Student t-distributions" Eq. (3.17) for each channel separately:

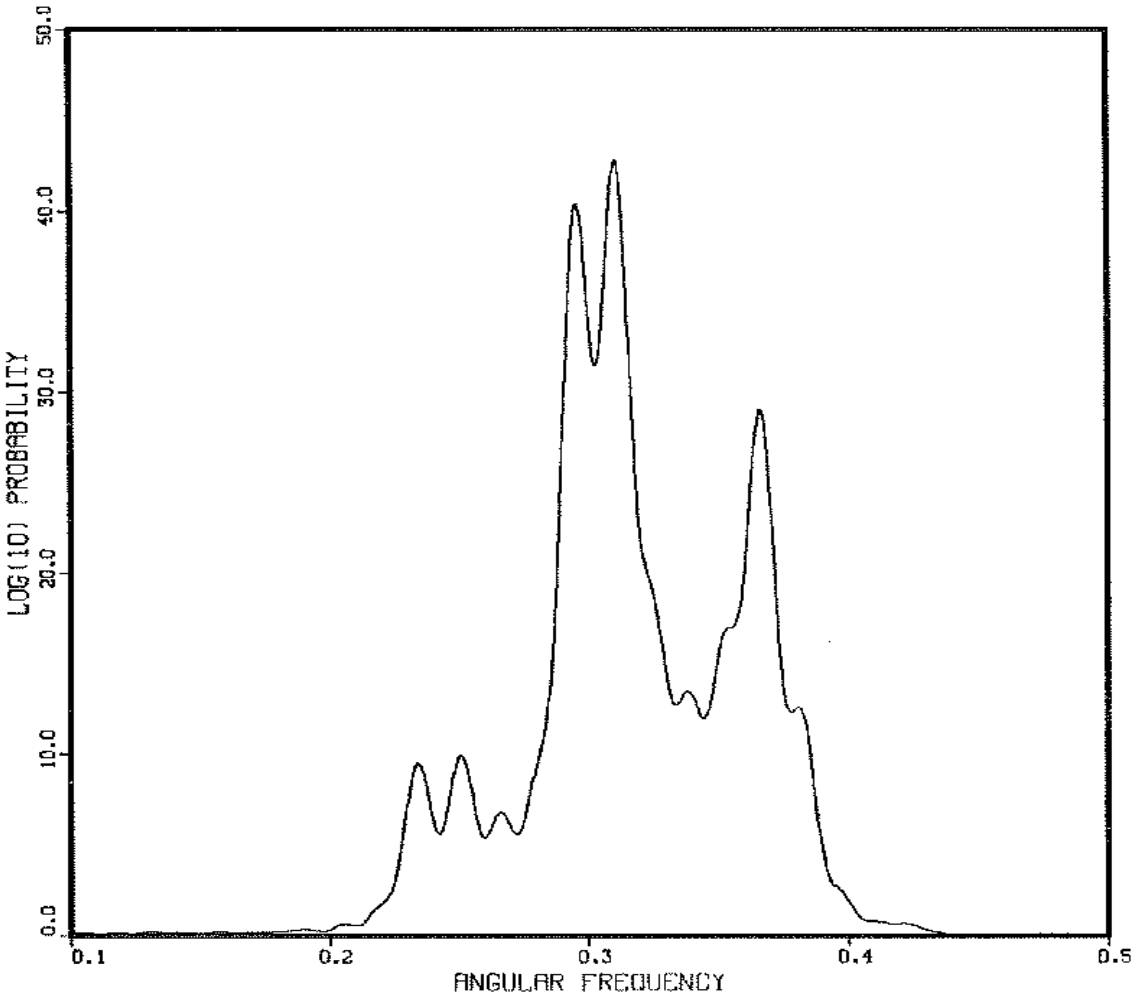
$$P(\{\omega\}|D,I) \propto \left[1 - \frac{m\overline{h^2}_1}{N\overline{d^2}_1}\right]^{\frac{m-N}{2}} \left[1 - \frac{m\overline{h^2}_2}{N\overline{d^2}_2}\right]^{\frac{m-N}{2}} \tag{7.1}$$

where the subscripts refer to the channel number. As explained previously, (7.1) in effect estimates the noise level independently in the two channels. This procedure is general and can be applied whenever two measurements of a signal are available; it is not restricted to NMR data. It is possible to specialize the estimation procedures to include this quadrature model, as well as the aforementioned phase and noise correlations. If all of this prior information is incorporated into the analysis (the author has, in fact, done this), we would expect to improve the results considerably. However, the present results will prove adequate for most purposes.

A procedure for dealing with the multiple frequency problem was outlined in Chapter VI, and we will apply that procedure here. The first step in any frequency estimation problem is to plot the data and the log of the probability of a single harmonic frequency. If there is only one data channel, this is essentially the periodogram of the data, Fig. 7.1(B) and Fig. 7.1(D). When more than one channel is present, the log probability of a single harmonic frequency is essentially the sum of the periodograms for each channel, weighted by the appropriate reciprocal variances. If the variances are unknown, then the appropriate statistic is the log of (7.1), shown in Fig. 7.2.

Now as was shown in Chapter 6, if the frequencies are well separated, a peak in the periodogram above the noise level is evidence – but not proof – of a frequency near that peak. From examining Fig. 7.2 we see there are nine resolved peaks in $0.2 < \omega < 0.4$ and suggestions of five more unresolved ones. This is many more peaks than theoretical physics indicates there should be. Is this evidence of more

Figure 7.2: The Log_{10} Probability of One Frequency in Both Channels



When more than one channel is present, the periodogram is not the proper statistic to be analyzed for indications of a simple harmonic frequency. The proper statistic (shown above) is log of the probability of a single harmonic frequency in both channels.

going on than theory predicts? To answer this question we will apply the general procedure outlined in the preceding chapter for determining multiple frequencies. We first fit the data with the single best frequency plus decay. We choose Lorentzian decay instead of Gaussian because physical theory indicates the decay is Lorentzian in a liquid phase. The model we used is

$$f_1(t) = [B_1 \cos(\omega_1 t) + B_2 \sin(\omega_1 t)]\mathrm{e}^{-\alpha t}.$$

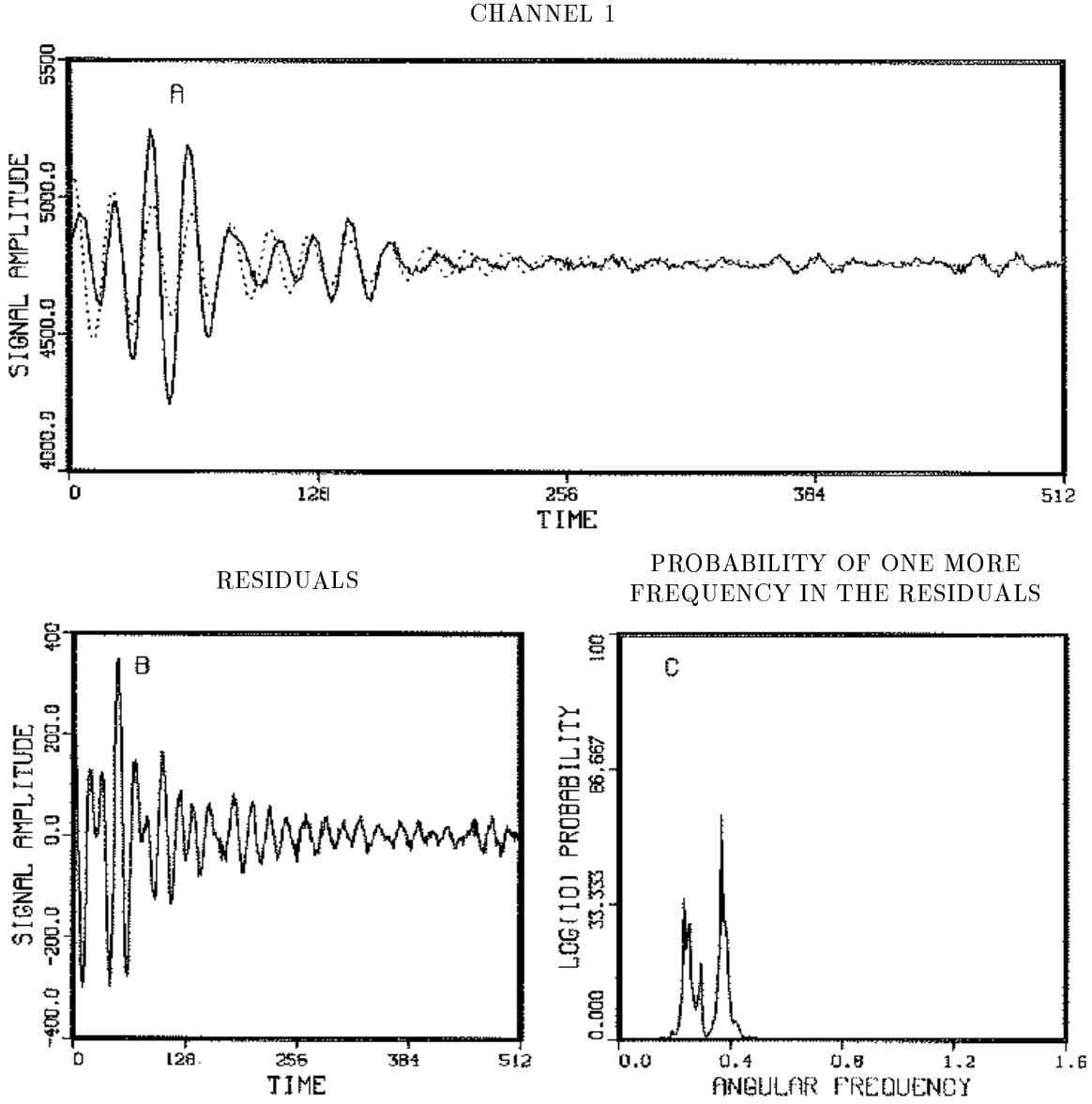
The computer code in Appendix E was used to evaluate the "Student t-distribution" Eq. (3.17) for each channel, and these were multiplied to obtain, Eq. (7.1). We searched in the two dimensional parameter space until we located the maximum of the distribution by the "pattern" searching procedure noted before. Next we computed the signal having the predicted parameters. The model (dotted line) and the data from the real channel are shown in, Fig. 7.3(A). It is clear from examining this figure as well as from examining the residuals in, Fig. 7.3(B), that there is at least a second frequency in this data. We see from the probability of a single harmonic frequency in the residuals, Fig. 7.3(C), that there is still strong evidence for additional frequencies near 0.3.

We then proceeded to a two-frequency-plus-decay model and repeated this procedure. That is, we estimated the second frequency plus decay from the residuals, and then used the results from the one-frequency model plus the estimates from the residuals as the initial estimates in a two-frequency model of the original data. We searched this four-parameter space until we located the maximum of the probability density. The results from the two-frequency model are displayed in Fig. 7.4. The model (dotted line) now takes on more characteristics of the signal (A), while the residuals (B) and the probability of a single harmonic frequency in the residuals (C) continue to show evidence for additional effects in the data. Notice the structure of the probability of a single frequency in the residuals. The addition of a second frequency removed one peak and essentially left the others unchanged. We demonstrated in Chapter 6 that when the frequencies are well separated the multiple-frequency estimation problem separates into a series of single-frequency problems, and this just confirms numerically that result.

To compare the two-frequency model to the one-frequency we computed the posterior odds ratio, this is given by:

$$\text{posterior odds} = \frac{P(f_2|I)}{P(f_1|I)} \frac{P(D|f_2, I)}{P(D|f_1, I)}$$

Figure 7.3: The One-Frequency Model



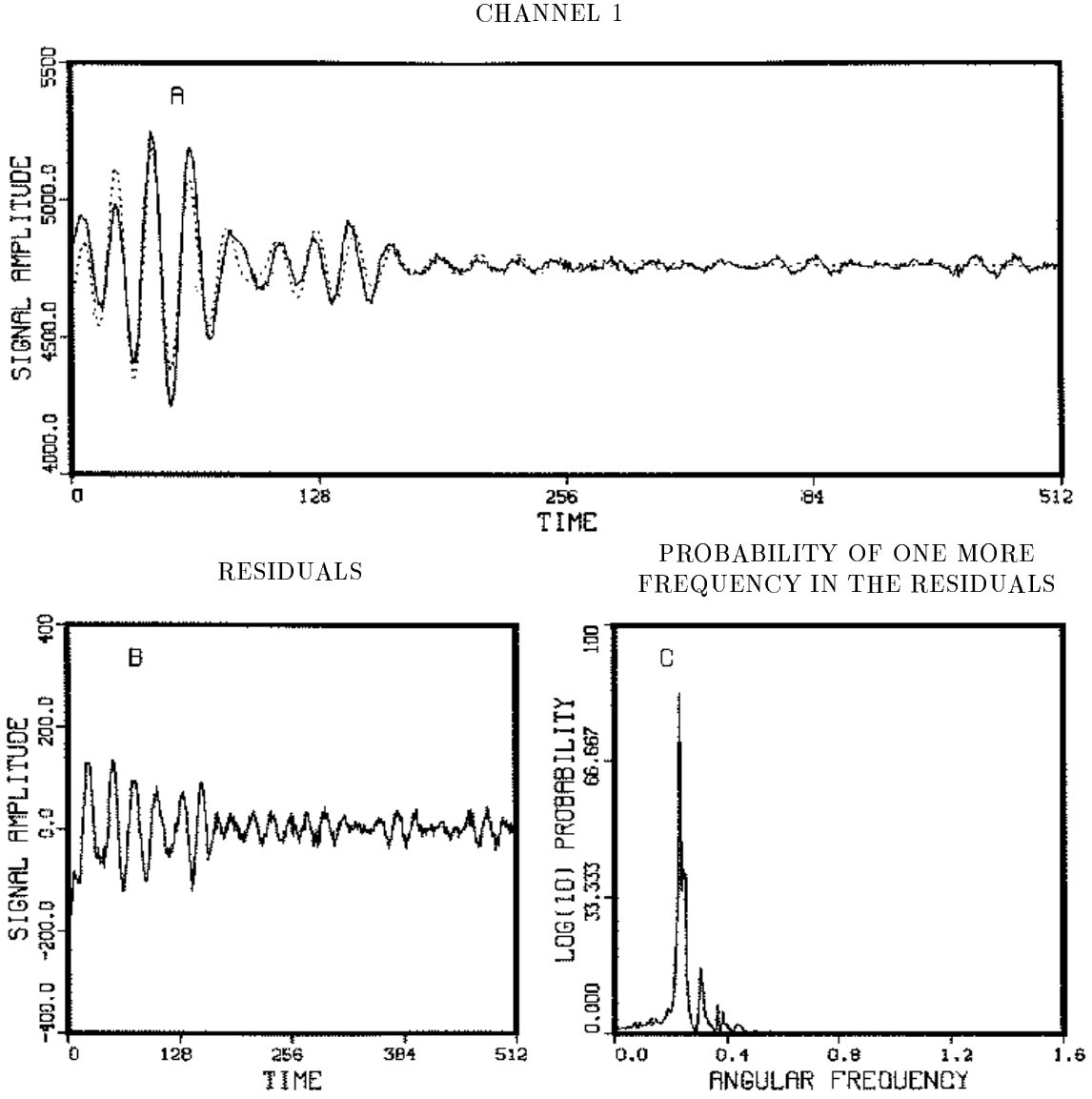
The data from one channel of the NMR experiment (solid line), and a one-frequency-plus-decay model with the predicted parameters (dotted line) are shown in (A). Next we computed the residuals: the differences between the data and the model (B). The residuals clearly indicate additional effects in the data. Last we computed the probability of a single harmonic frequency in the residuals (C). This clearly indicates there are additional effects in the data.

where, $P(f_1|I)$ and $P(f_2|I)$ are the prior probability of the one-frequency and two-frequency models, and $P(D|f_1, I)$ and $P(D|f_2, I)$ are the global likelihoods, Eq. 5.9, for the one-frequency and two-frequency models. We have some prior information about how many frequencies should be present: theoretical physics indicates there should be three frequencies, however, we will assume either of the models is equally probable and set $P(f_2|I)/P(f_1|I) = 1$. We then computed the likelihood ratio and find there is one chance in 10^{457} that the one-frequency model is a better description of the phenomenon than the two-frequency model. There is zero chance that the one-frequency model represents the data better! But it might be that very unusual noise is confusing us, and there is a tiny chance that the one-frequency model would represent the next data set better. To see how tiny that chance is, note that the number of microseconds in the estimated age of the universe is only about 10^{24}.

Then we proceeded to the three-frequency-plus-decay model. The most probable frequency is the low frequency peak in the vicinity of 0.3; so we ran the three-frequency-plus-decay model using this low frequency as the initial estimate for the third frequency, Fig. 7.5. The model has now taken on most of the dominant characteristics of the signal as in Fig. 7.5(A): indeed the triple is the largest effect in the data. However, fitting the triple does not account for the long time behavior of the system. Notice in the residuals, Fig. 7.5(B), that there is still more than enough signal left for the eye to make out the oscillations easily. We see from the probability of a single frequency in the residuals, Fig. 7.5(C), that there is still evidence for additional frequencies in the data. The posterior odds ratio for the two-frequency-plus-decay model compared to the three-frequency-plus-decay model indicates that there is one chance in 10^{703} that the two-frequency model is a better description than the three-frequency model.

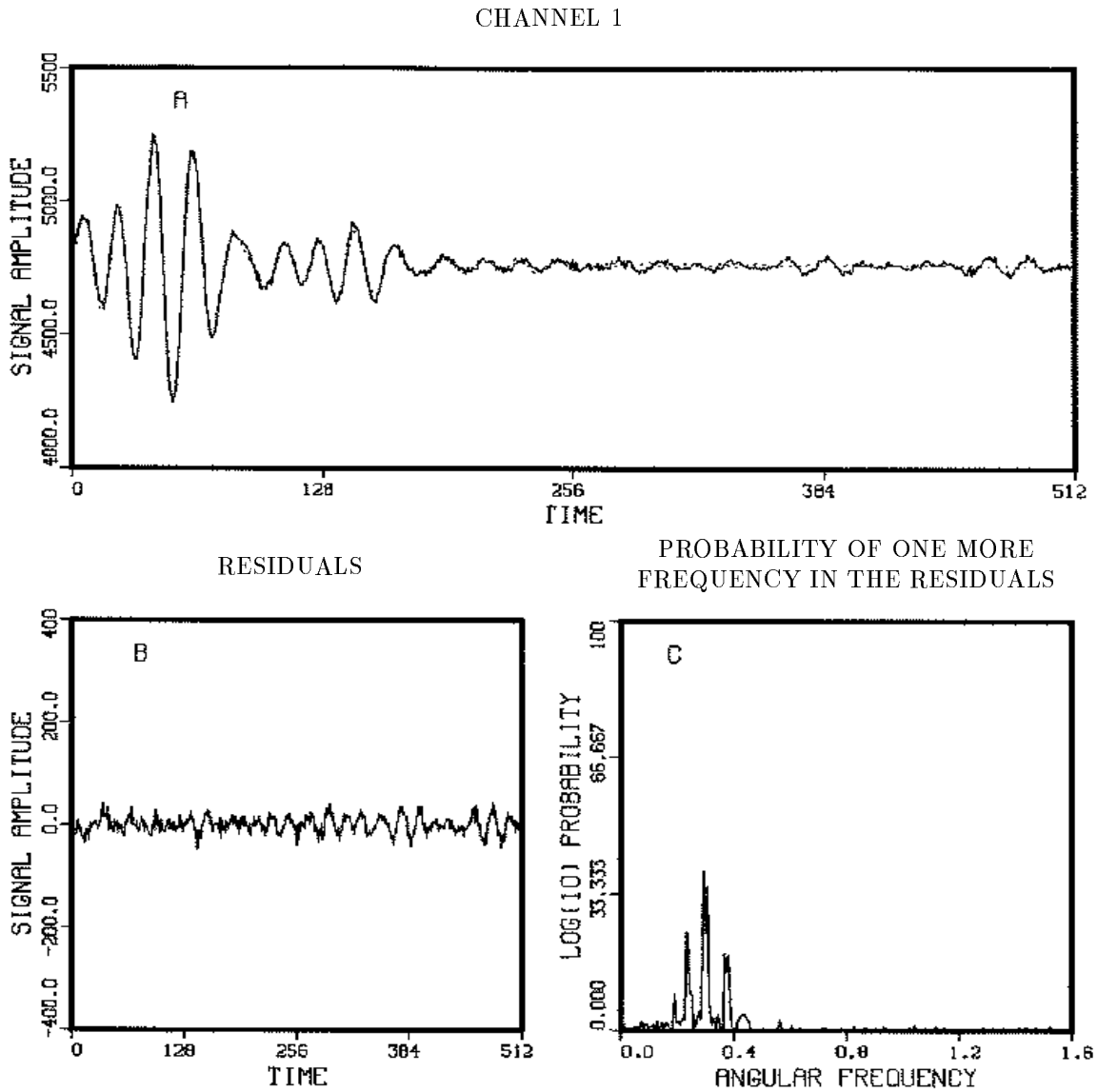
To see if there are additional effects in the data we proceeded to the four-frequency-plus-decay model. Figure 7.6(A) is a plot of the data and the model. Now the model is making a much better showing in the long time behavior of the system, but even here we have not accounted for all the effects in the data. Clearly in the residuals, Fig. 7.6(B), there is a small unaccounted for signal; the probability of a single frequency in the residuals, Fig. 7.6(C), verifies this, indicating it to be a high-frequency component (not shown in Fig. 7.6). The posterior odds ratio of the four-frequency-plus-decay model to the three-frequency-plus-decay model indicates there is one chance in 10^{80} that the three-frequency model is a better description than the four-frequency model.

Figure 7.4: The Two-Frequency Model



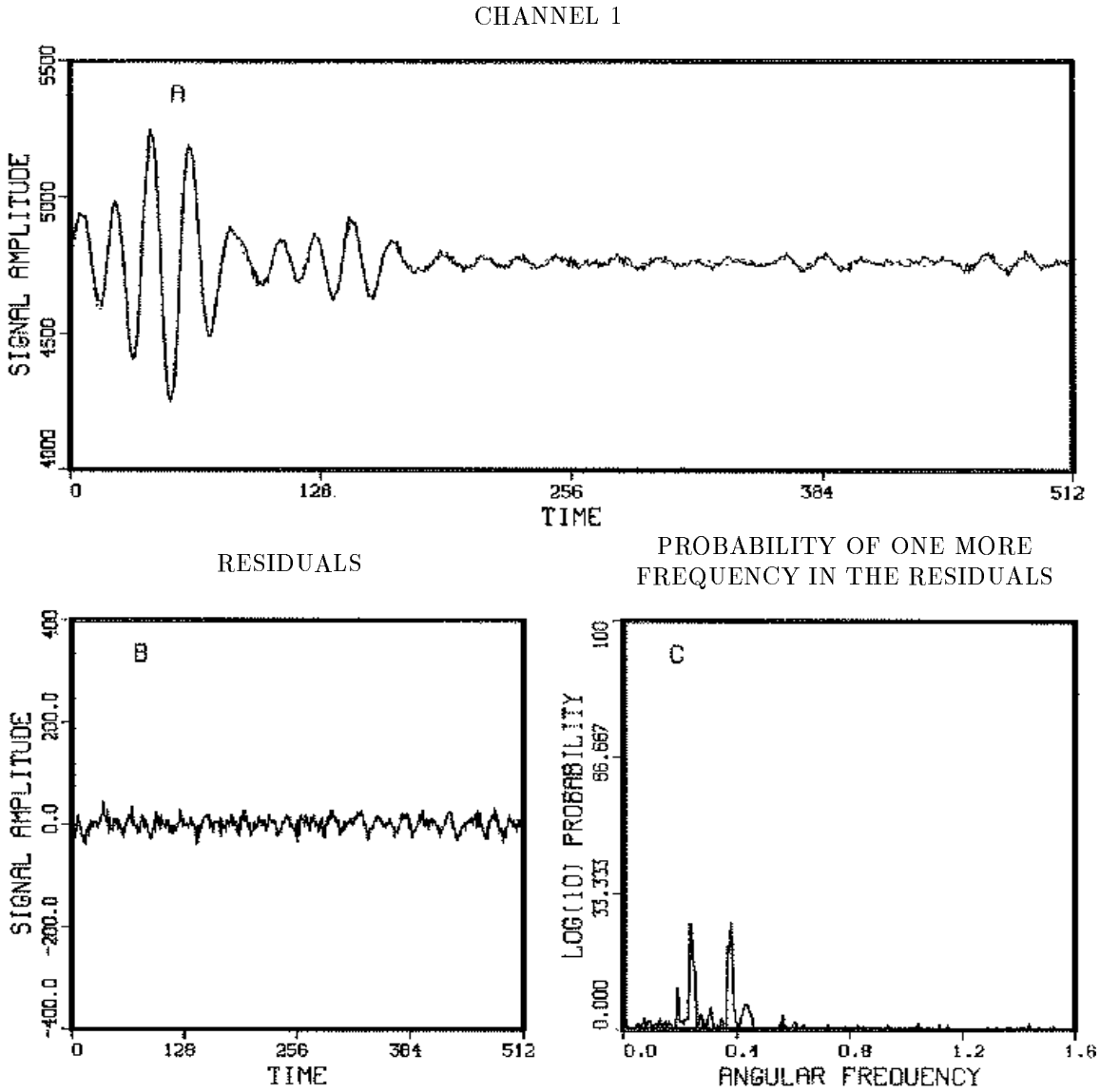
Next we computed the probability of two frequencies plus decay in both channels. The model (dotted line) and the data (solid line) are displayed in (A). The residuals (B) clearly indicate additional effects in the data. We then computed the probability of a single frequency in the residuals and displayed that in panel (C).

Figure 7.5: The Three-Frequency Model



Because the two-frequency-plus-decay model did not take up all of the signal, we proceeded to a three-frequency-plus-decay model, (A) dotted line. The residuals (B) clearly indicate additional effects in the data. We computed the probability of a single frequency in the residuals and displayed that in panel (C). Again we see there are additional effects in this data set.

Figure 7.6: The Four-Frequency Model



Because the three-frequency-plus-decay model did not account for all of the signal, we proceeded to a four-frequency-plus-decay model, (A) dotted line. The residuals (B) continue to indicate additional effects in the data. We computed the probability of a single frequency in the residuals and displayed that in panel (C). Again we see that there are additional effects in this data set.

We continued repeating this procedure until we accounted for all systematic components – see Fig. 7.7 through Fig. 7.9. Now in Fig. 7.9(C) the residuals are finally beginning to look like Gaussian white noise. However, there is some evidence for a very small additional frequency, see Fig. 7.9(C). We did not go further because this frequency, although present in the real channel, is not present to any significant degree in the quadrature data channel.

Our probability analysis indicates there are at least seven frequencies in these data, of which one is attributable to the instrumentation. That leaves six frequencies located near 0.3 in dimensionless units. The posterior probability of a single harmonic frequency in the combined data, Fig. 7.2, gives evidence of multiple complex phenomena around 0.3 but it could not sort out what is going on. This is not too surprising, given that there are six frequencies in this region. The one-frequency model has done surprisingly well. The absorption spectrum, Fig. 7.10(A), on the other hand, shows only three peaks in this region. This simple example illustrated that the discrete Fourier transform gives evidence of frequencies in the data that an absorption spectrum does not. Although the probability of a single harmonic frequency or the Schuster periodogram is not an exact estimator for multiple frequencies, it is adequate as long as the frequencies are well separated. The only time one must worry about this statistic being incorrect is when the frequencies are close together (as they were here). But by contrast, there are no conditions under which the absorption spectrum is an optimal frequency estimator, and the global phase adjustment on the absorption spectrum can suppress indications of frequencies in the data.

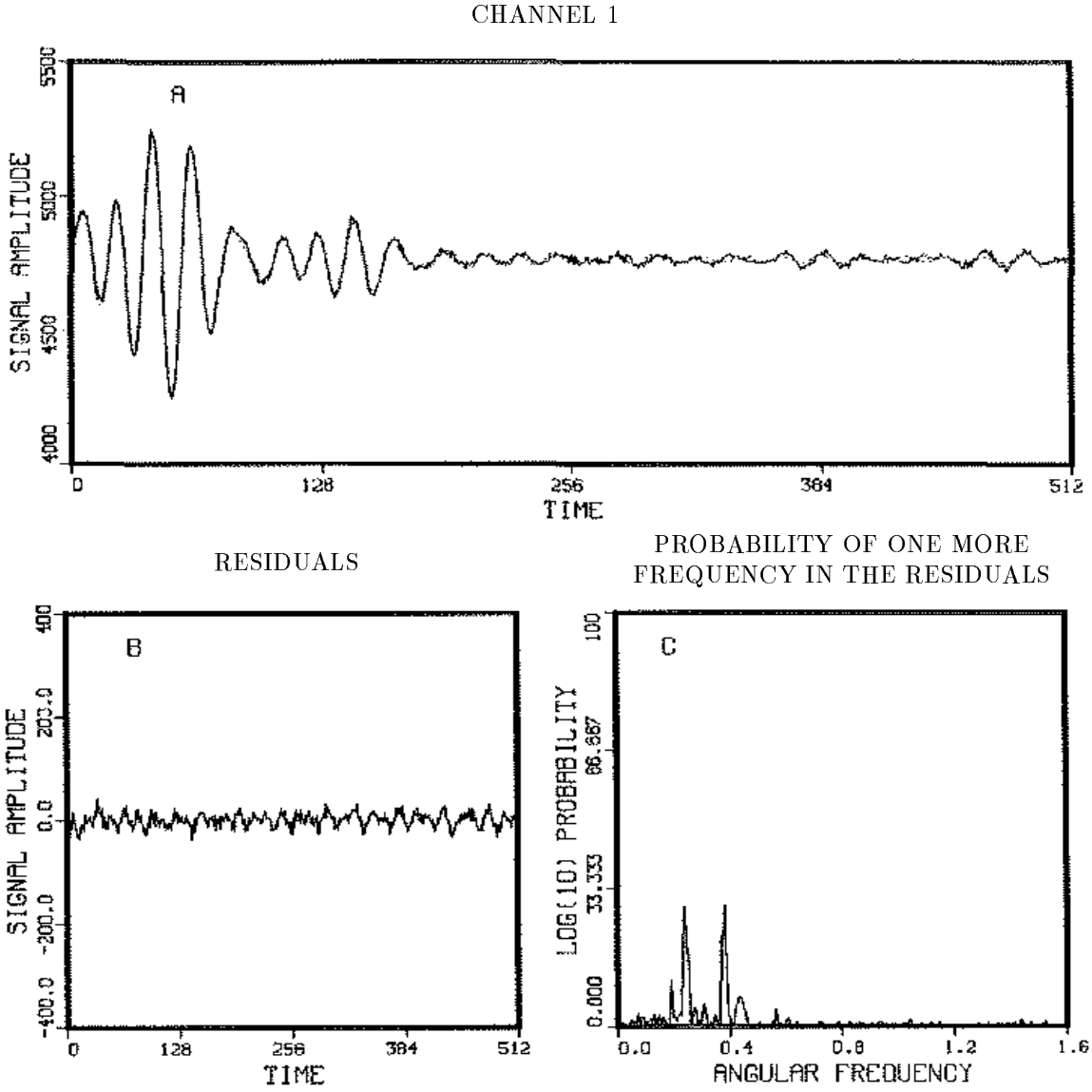
We developed the procedures for estimating the accuracy of the frequencies and the amplitudes and we have used those procedures here [to apply them we calculated the second derivatives numerically (4.13)]. The results of this calculation are:

Frequency Hertz	Decay Rate Hertz	Amplitude Real	Amplitude Imaginary
75.0695 ± 0.0005	7.294 ± 0.003	49	46
78.1231 ± 0.0002	19.613 ± 0.001	170	160
94.1207 ± 0.0008	8.569 ± 0.001	71	72
98.0187 ± 0.0001	23.211 ± 0.001	354	318
117.6052 ± 0.0001	16.336 ± 0.001	193	188
121.0824 ± 0.0002	11.270 ± 0.001	67	66

We also estimated the signal-to-noise ratio, Eq. (4.8), for each channel:

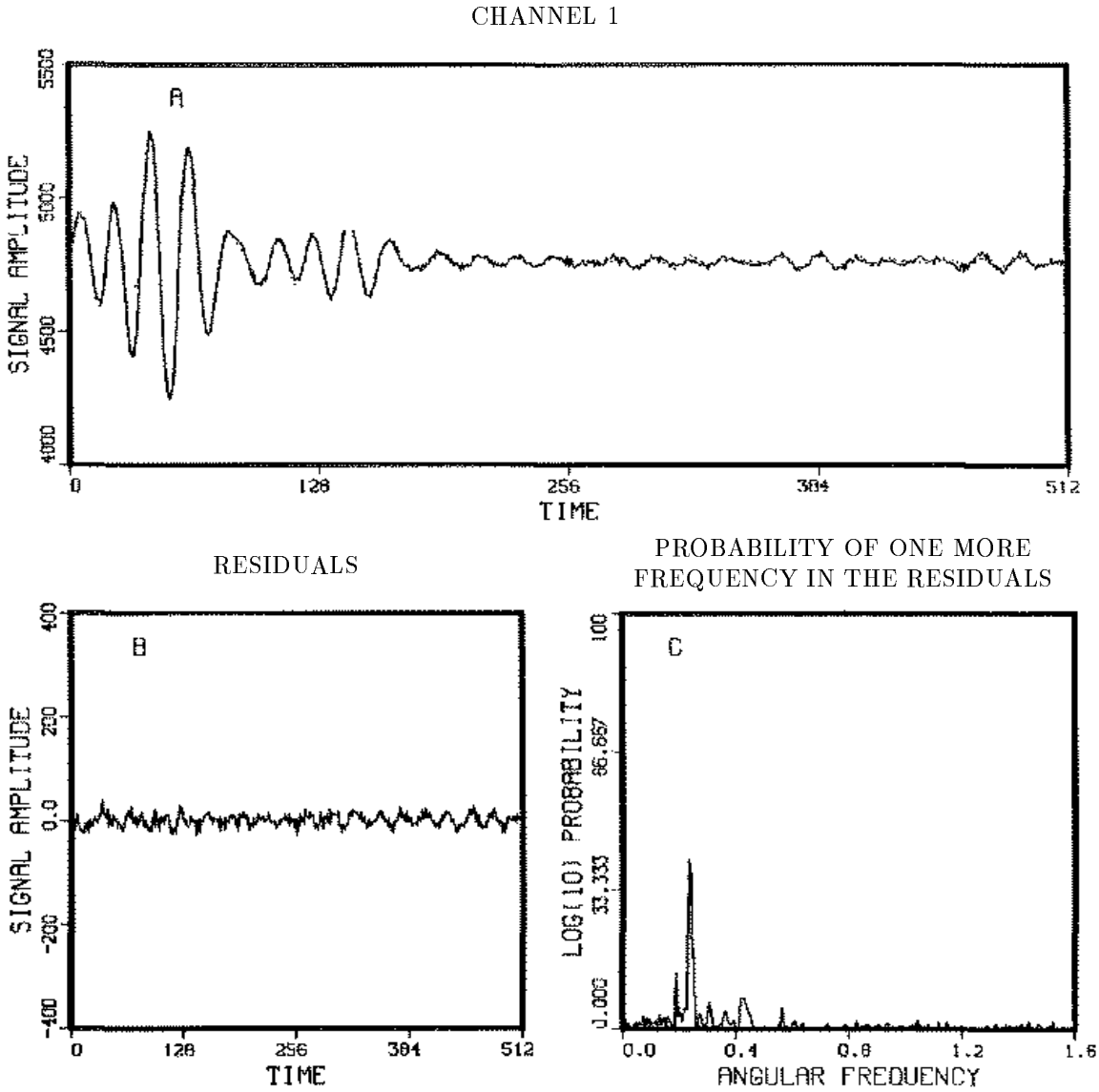
$$\frac{\text{Signal}}{\text{Noise}} = 1606 \text{ in channel 1},$$

Figure 7.7: The Five-Frequency Model



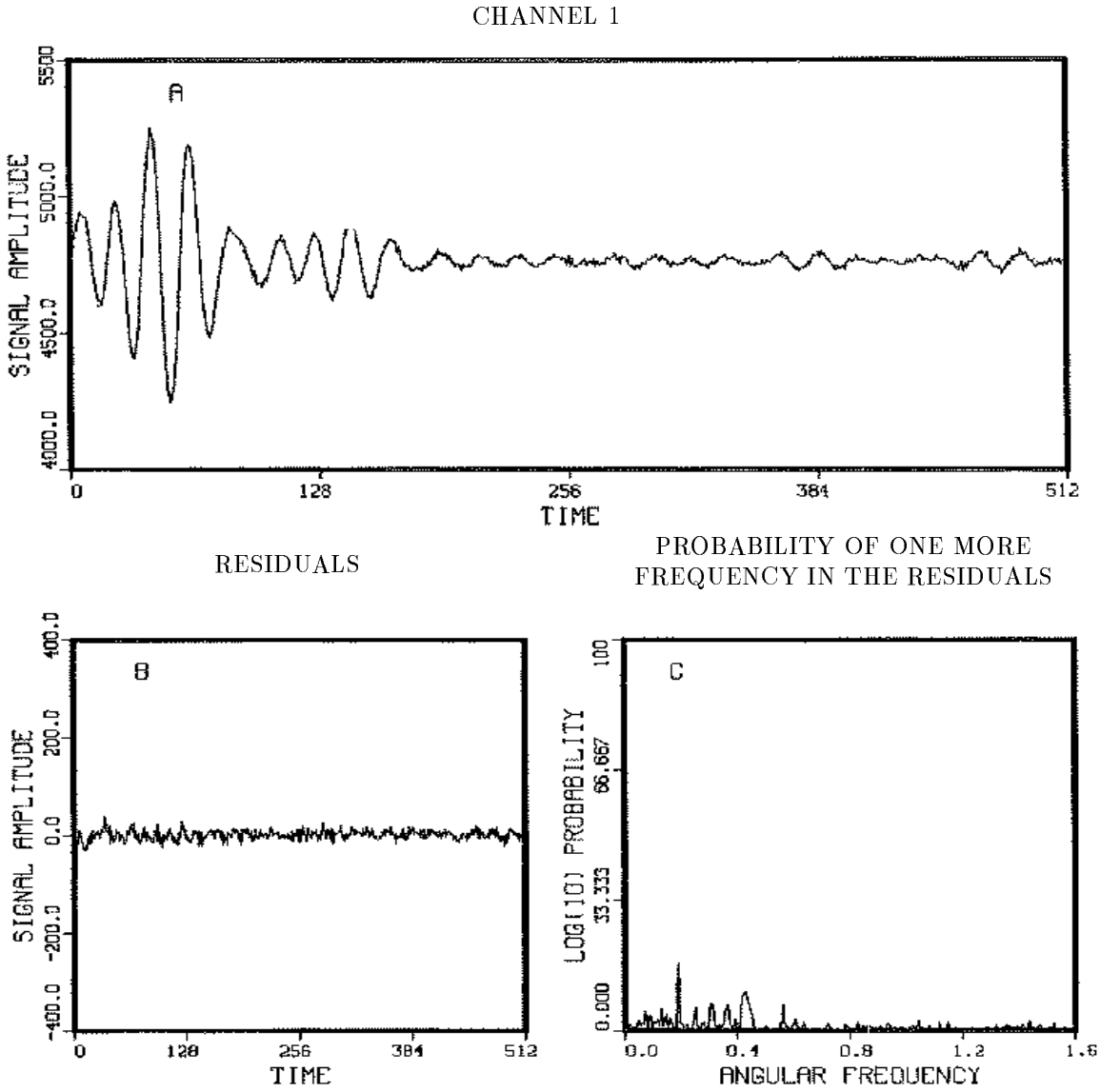
Because the four-frequency model did not account for all of the signal, we proceeded to a five-frequency model, (A) dotted line. The residuals (B) continue to indicate additional effects in the data. We computed the probability of a single frequency in the residuals and displayed that in panel (C). There is no apparent change in (C) because a very high frequency component was remove by the four-frequency model. Again we see that there are additional effects in this data set.

Figure 7.8: The Six-Frequency Model



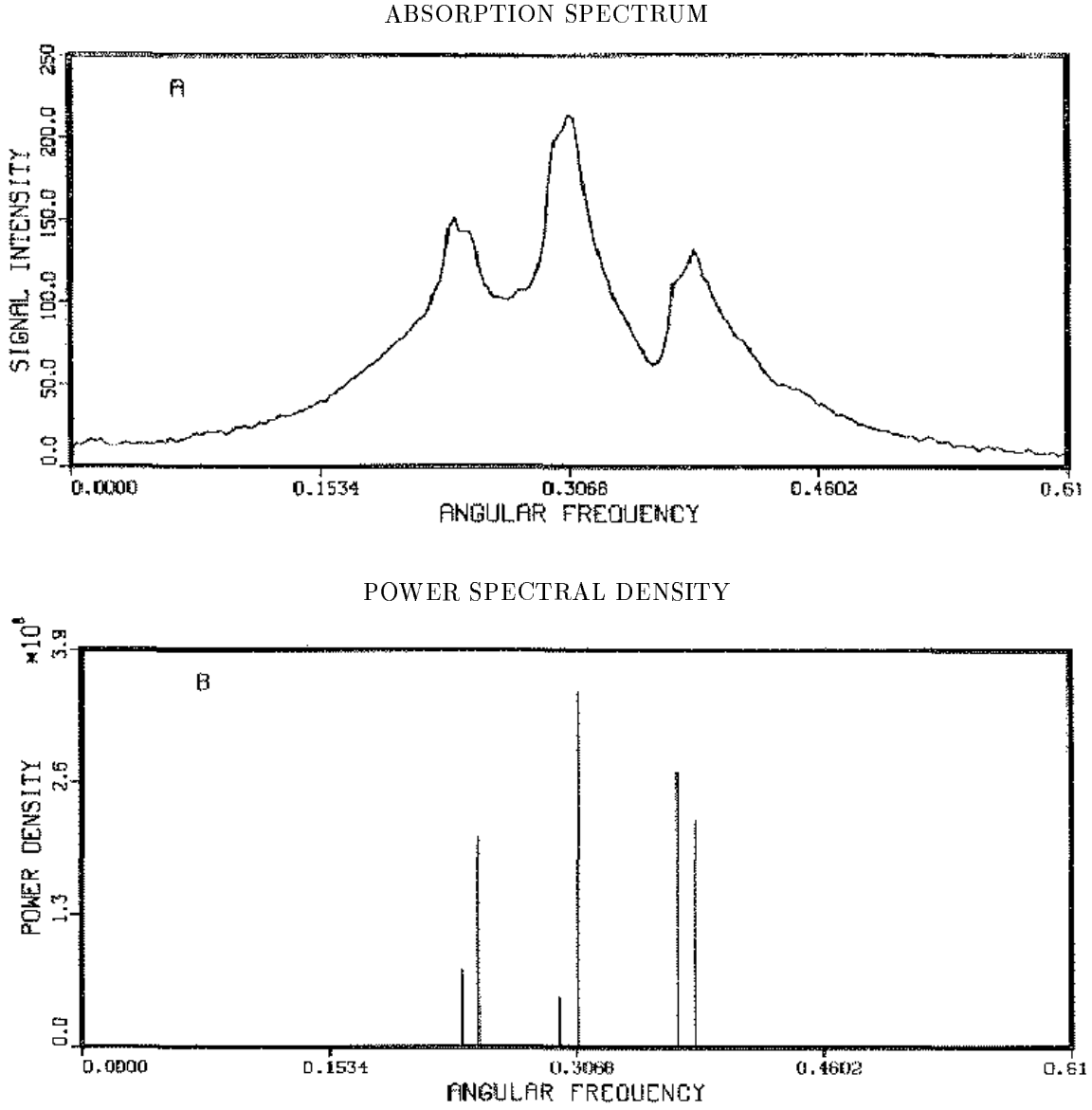
Because the six-frequency model did not account for all of the signal, we proceeded to a seven-frequency model, (A) dotted line. The residuals (B) clearly indicate additional effects in the data. We computed the probability of a single frequency in the residuals and displayed that in panel (C). Again we see there are additional effects in this data set.

Figure 7.9: The Seven-Frequency Model



At last, with the seven-frequency model we reached a point where the model and the signal look essentially identical (A). The residuals (B), now look much more like white noise. We computed the probability of a single frequency in the residuals and displayed that in panel (C). Again we see there are additional very small effects in this data set. However, these effects are not repeated in both channels: we interpret these effects be an instrumental artifact.

Figure 7.10: Comparison to an Absorption Spectrum



The absorption spectrum (described in the text, see page 117) gives a clear indication of three frequencies and hints at three others (A). Using the full width at half maximum of the absorption spectrum to determine the accuracy estimate and converting to physical units, it determines the frequencies to within ± 15Hz. The probability analysis (B) used a seven-frequency model with decay. The estimated accuracy is approximately ± 0.001Hz.

$$\frac{\text{Signal}}{\text{Noise}} = 1478 \text{ in channel 2,}$$

and the estimated standard deviation (4.6):

$$(\sigma)_{\text{est}} = 9 \text{ in channel 1,}$$

$$(\sigma)_{\text{est}} = 9 \text{ in channel 2.}$$

The amplitudes were estimated separately in each channel, and if the spectrometer is working correctly we expect the amplitude of each sinusoid to be approximately the same. This serves as an additional check on the model; if we were fitting an appreciable amount of noise, the estimated amplitudes would be different in the two channels.

The quantities of interest are the splitting between the two components of the HD doublet as well as the shift in the center frequency. But physical theory indicated there should be only three frequencies in the region of the main resonance: we find six. The calculation indicates there is clearly more going on here than physical theory indicates there should be. One of the major assumptions made in NMR is that the magnetic field is uniform over the sample. If it is not, the resonances will be spread out, corresponding to different intensities of the local field, and false structure may appear. Here we may be seeing this effect. However, the sharpness of the peaks suggests that the effect is real, conceivably arising from impurities in the sample or from association effects (such as H_4O_2 molecules) not considered in the theory.

However, we have derived a model of the process as if there were two major regions in the sample where the field was approximately uniform. If we wish to derive the splittings we must use the frequencies corresponding to uniform field. In each of the regions where the field is a uniform, the frequency shifts should be according to theory. Thus for the set of frequencies shifted to lower values (75, 94, and 117) the HD doublet separation is

$$\text{High - Low} = 42.536 \pm 0.001\text{Hz}$$

and the center frequency (94 Hz) is displaced from the center of the doublet by 2.217 ± 0.001 Hz. For the set of frequencies (78, 98, and 121) shifted to higher values we have

$$\text{High - Low} = 42.956 \pm 0.001\text{Hz}$$

and the center frequency (98 Hz) is displaced from the center of the doublet by 1.521 ± 0.001 Hz. Both of these tentative answers are in good agreement with the

simple theory; unfortunately, until the field shimming problems are cleared up we do not know which to believe, if either. The center frequency is displaced from the center of the doublet in the correct direction, and in reasonable agreement with prior measurements of this quantity [33]. In order to answer these questions it would be necessary to rerun the experiment with better shimming. Additionally, the estimates could be improved somewhat by sampling the data faster.

If one attempts to analyze these data using the standard absorption spectrum Fig. 7.10(A) only three peaks are found, with hints of three other frequencies. The splitting of the HD doublet is approximately correct, but the center peak is shifted in the wrong direction. We can compare these estimates directly to the absorption spectrum. The reason the analysis of this experiment is so difficult with the absorption spectrum is that the full-width at half maximum for the D_2 peak, Fig. 7.10(A), is 15Hz. But this width is indicative only of the decay rates; not the accuracy with which the oscillations frequency is determined. Probability theory has enabled us to separate these entirely different quantities. Figure 7.10(B) gives the estimates from Eq. (7.1). We have plotted these estimates as normalized Gaussians, each centered at the estimated frequency and having the same standard deviation as the estimated frequency. Clearly, the resolution of these frequencies is much improved compared to an absorption spectrum or a discrete Fourier transform. With separately normalized distributions, the heights in Fig. 7.10(B) are indications of the accuracy of the estimates, not of the power carried by the signal.

The accuracy of this procedure may be a little disturbing. To understand it, look at the estimated signal-to-noise ratio in these data. It is on the order of 1500 for each channel. There is essentially nothing in these data sets that can be ignored. Every little bump and wiggle in the discrete Fourier transform is indicative of some effect in the data, and must be accounted for. Because the accuracy of the estimates is inversely proportional to the signal-to-noise of the data, the estimates are very precise. It is rather the inaccuracy of the conventional method that should be disturbing to one.

7.2 Corn Crop Yields

Economic data are hard to analyze, in part because the data are frequently contaminated by large spurious effects, which one does not know how to capture in a

model, and the time series are often very short. Here we will examine one example of economic data to demonstrate how to remove some unknown and spurious effects. In particular, we will analyze one hundred year's worth of corn crop data from three states (Kansas, South Dakota, and Nebraska), Fig. 7.11(A) through Fig. 7.11(C) [34]. We would like to know if there is any indication of periodic behavior in these data.

These data have been analyzed before. Currie [35] used a high pass filter and then applied the Burg algorithm [36] to the filtered data. Currie finds one frequency near 20 years which is attributed to the lunar 18.6 year cycle, and another at 11 years, which is attributed to the solar cycle.

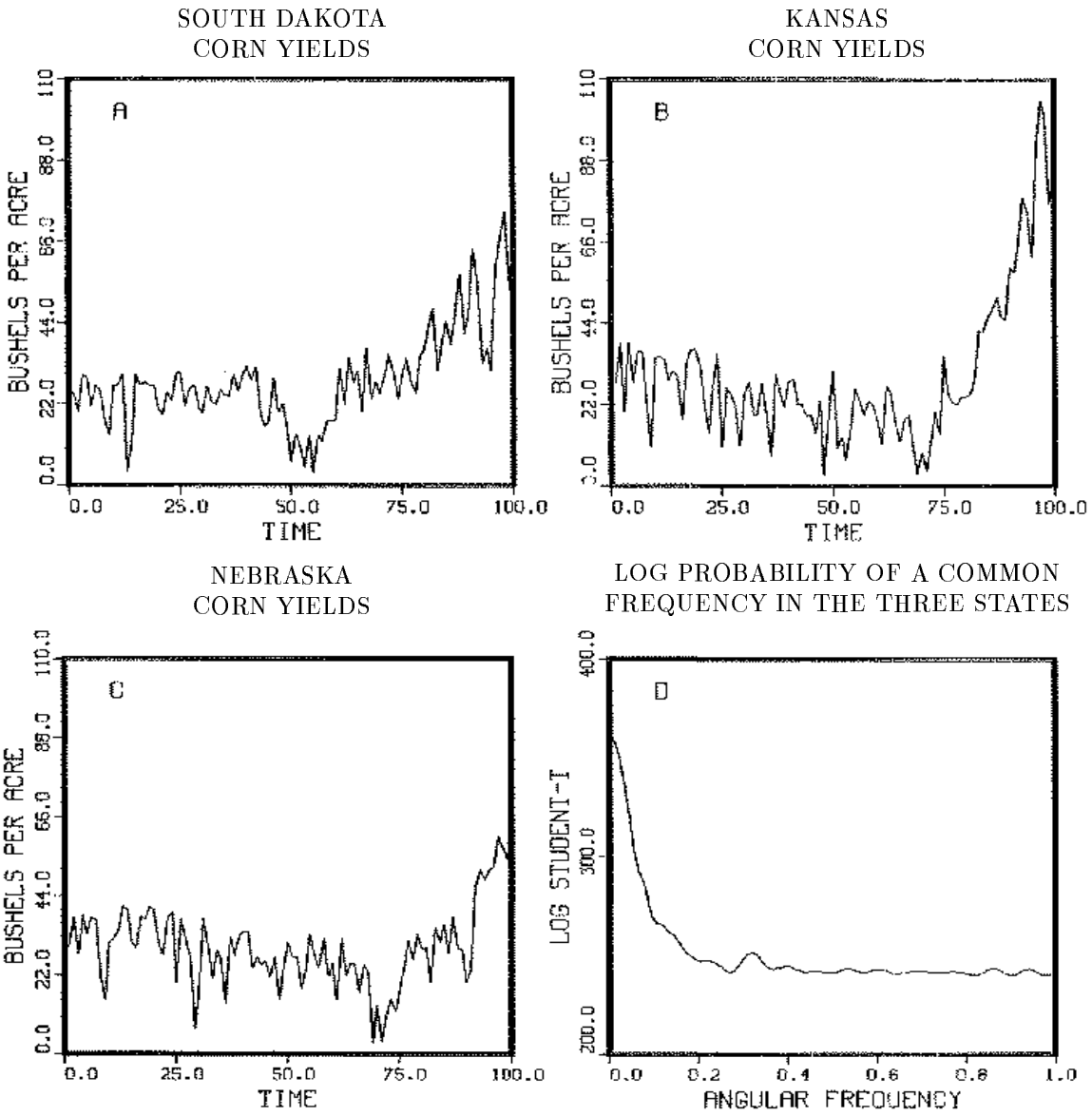
There are three steps in Currie's analysis that are troublesome. First, the Burg algorithm is not optimal in the presence of noise (although it is for the problem it was formulated to solve). The fact that it continues to work means that the procedure is reasonably robust; that does not change the fact that it is fundamentally not appropriate to this problem [36]. Second, one has doubts about the filter: could it suppress the effect one is looking for or introduce other spurious effects? Third, to apply the Burg algorithm when the data consist of the actual values of a time series, the autoregression order (maximum lag to be used) must be chosen and there is no theoretical principle to determine this choice. We do not mean to imply that Currie's result is incorrect; only that it is provisional. We would like to apply probability theory as developed in this work to check these results.

The first step in a harmonic analysis is simply to plot the data, Fig. 7.11(A) through Fig. 7.11(C) and the log of the posterior probability of a single harmonic frequency. In the previous example we generalized the analysis for two channels. The generalization to an arbitrary number of channels is just a repeat of the previous arguments:

$$P(\{\omega\}|D,I) \propto \prod_{j=1}^{r}\left[1-\frac{m_j\overline{h^2}_j}{N_j\overline{d^2}_j}\right]^{\frac{m_j-N_j}{2}} \tag{7.2}$$

where the subscripts refer to the jth channels: each of the models has m_j amplitudes, and each data set contains N_j data values. Additionally it was assumed that the noise variance σ_j was unknown and possibly different for each channel. The "Student t-distributions" Eq. (3.17) for each channel should be computed separately, thus estimating and eliminating the nuisance parameters particular to that channel, and then multiplied to obtain the posterior probability for the common effects, Eq. (7.2). Again if we had prior knowledge of correlations in the "noise" for different channels, we could exploit that information to get better final results, at the cost of more

Figure 7.11: Corn Crop Yields for Three Selected States



The three data sets analyzed were corn yields in bushels per acre for South Dakota (A), Kansas (B), and Nebraska (C). The log probability of a single common frequency plus a constant is plotted in (D). The question we would like to answer is "Is that small bump located at approximately 0.3, corresponding to a 20 year period, a real indication of a frequency or is it an artifact of the trend?"

computation.

For this harmonic analysis we take the model to be a single sinusoid which oscillates about a constant. The model for the jth channel may be written

$$f_j(t) = B_{j,1} + B_{j,2}\sin(\omega t) + B_{j,3}\cos(\omega t). \tag{7.3}$$

Here we have three channels, named "South Dakota", "Kansas", and "Nebraska". We allow $B_{j,1}$, $B_{j,2}$, and $B_{j,3}$ to be different for each channel; thus there are a total of nine amplitudes, one frequency, and three noise variances. To compute the posterior probability for each measurement, we used the computer code in Appendix E. The log of each "Student t-distribution" Eq. (3.17) was computed and added to obtain the log of the posterior probability of a single common harmonic frequency, Fig. 7.11(D).

What we would like to know is, "Are those small bumps in Fig. 7.11(D) indications of periodic behavior, or are they artifacts of the noise or trend?" To attempt to answer this, consider the following model function

$$f_j(t) = T_j(t) + B_{j,1}\cos(\omega t) + B_{j,2}\sin(\omega t)$$

where we have augmented the standard frequency model by a trend $T_j(t)$. The only parameter of interest is the frequency ω. The trend $T_j(t)$ is a nuisance function; to eliminate it we expand the trend in orthonormal polynomials $L_j(t)$. These orthonormal polynomials could be any complete set. We use the Legendre polynomials with an appropriate scaling of the independent variable to make them orthonormal on the region $(-49.5 \le t \le 49.5)$. This is the range of values used for the time index in the sine and cosine terms. After expanding the trend, the model function for the jth measurement can be written

$$f_j(t) = \sum_{k=0}^{E} B_{j,k+1}L_k(t) + B_{j,E+2}\cos(\omega t) + B_{j,E+3}\sin(\omega t).$$

Notice that if the expansion order E is zero the problem is reduced to the previous problem (7.3).

The expansion order E must be set to some appropriate value. From looking at these data one sees that it will take at least a second order expansion to remove the trend. The actual expansion order for the trend is unknown. However, it will turn out that the estimated frequencies are insensitive to the expansion order, as long as the expansion is sufficient to represent the trend without representing the signal of interest. Of course, different orders could have very different implications about other

questions than the ones we are asking; for example, predicting the future trend. That is an altogether more difficult problem than the one we are solving.

The effects of increasing the expansion order E can be demonstrated by plotting the posterior probability for several expansion orders – see Fig. 7.12(A)–7.12(H). For expansion order zero, Fig. 7.12(A), through expansion order 2, Fig. 7.12(C) the trend has not been removed: the posterior probability continues to pick out the low frequency trend. When a third order trend is used, Fig. 7.12(D), a sudden change in the behavior is seen. The frequency near $\omega \approx 0.31$ suddenly shows up, along with a spurious low-frequency effect due to the trend. In expansion orders four through seven, Fig. 7.12(E) through Fig. 7.12(H), the trend has been effectively removed and the posterior probability indicates there is a frequency near 0.31 corresponding to a 20.4 year period.

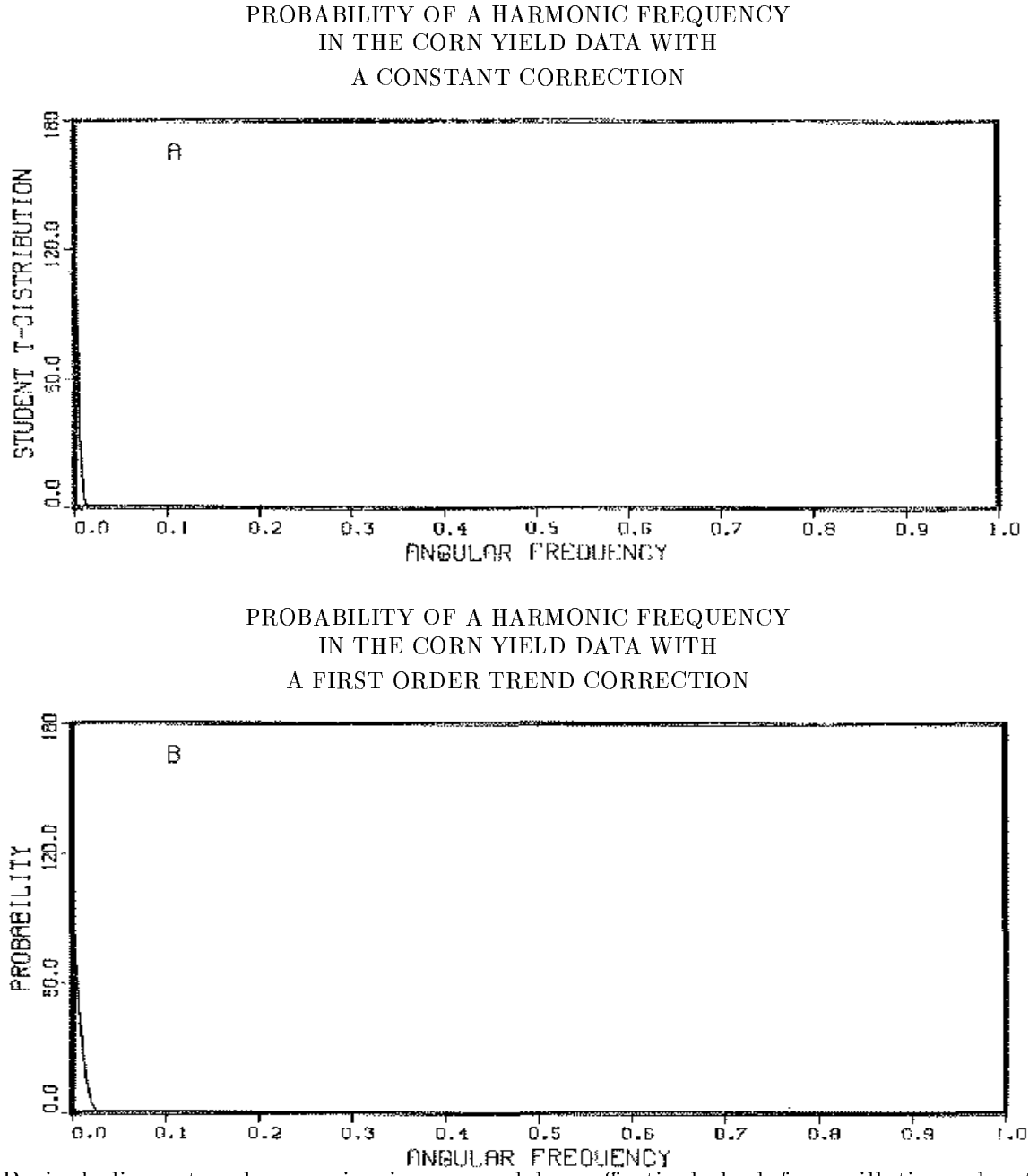
The amount of variability in the frequency estimates as a function of the expansion order will show how strongly the trend expansion is affecting the estimated frequency. The frequency estimates for the fourth through seventh order expansions are

$$\begin{aligned}
(f_4)_{\text{est}} &= 20.60 \quad \pm 0.16 \text{ years} \\
(f_5)_{\text{est}} &= 20.47 \quad \pm 0.18 \text{ years} \\
(f_6)_{\text{est}} &= 20.20 \quad \pm 0.14 \text{ years} \\
(f_7)_{\text{est}} &= 20.47 \quad \pm 0.18 \text{ years.}
\end{aligned}$$

Here the estimated errors represent two standard deviations. Thus, given the spread in the estimates it appears there is indeed evidence for a frequency of a period 20.4 $\pm$ 0.2 years.

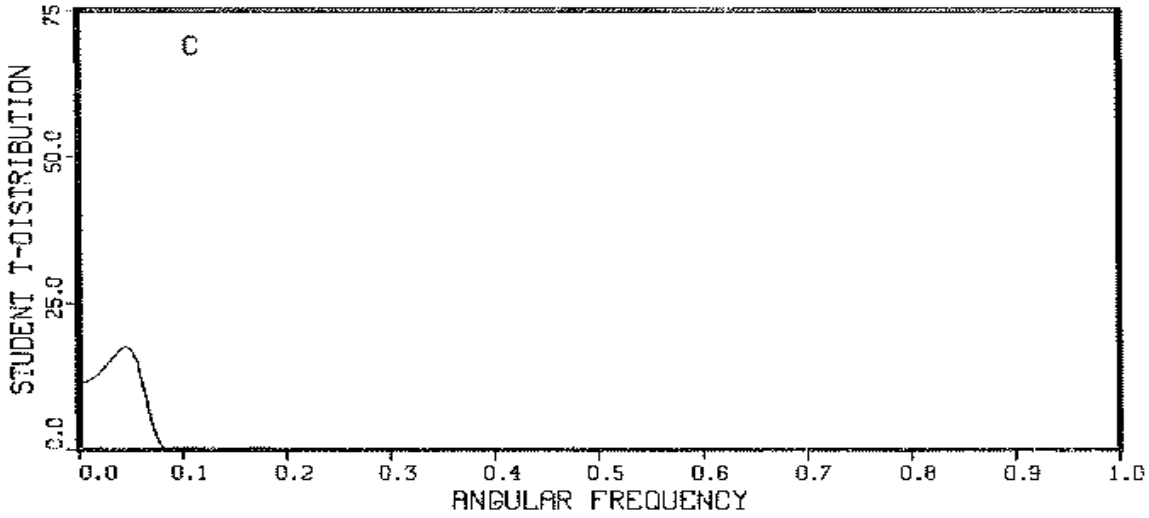
Now that the effects of removing a trend are better understood, we can proceed to a two-frequency model plus a trend to see if we can verify Currie's two frequency results. Figure 7.13 is a plot of the log of this probability distribution after removing a fifth order trend. The behavior of this plot is the type one would expect when a two-frequency model is applied to a data set that contains only one frequency. From this we cannot verify Currie's results. That is, for the three states taken as a whole these data show evidence for an oscillation near 20.4 years as he reports, but we do not find evidence for an 11 year cycle. This does not say that Currie's result is incorrect; he incorporated much more data into his calculation, and to check it we would need to include data from at least a dozen more states. While this is a worthy project, it is beyond the scope of this simple demonstration, whose main purpose is to show the good performance of the "theoretically correct" method of trend removal.

Figure 7.12: The Joint Probability of a Frequency Plus a Trend

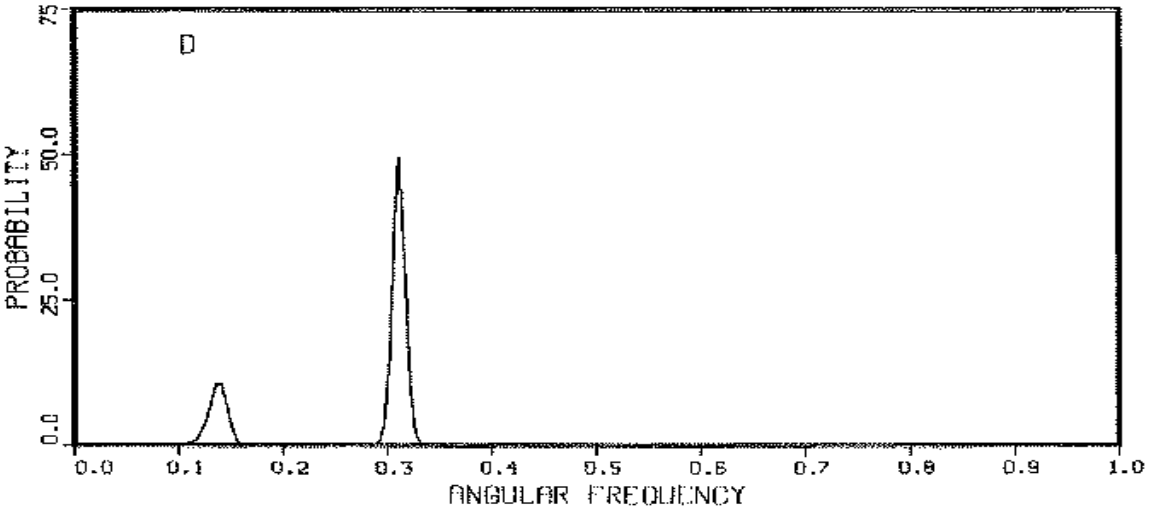


By including a trend expansion in our model we effectively look for oscillations about a trend. This is not the same as detrending, because the trend functions and the sine and cosine functions are never orthogonal. The zero order trend (or constant) plus a simple-harmonic-frequency model (A) is dominated by the trend. When we included a linear trend the height of the trend is decreased some, however the trend is still the dominant effect in the analysis.

PROBABILITY OF A HARMONIC FREQUENCY
IN THE CORN YIELD DATA WITH
A SECOND ORDER TREND CORRECTION



PROBABILITY OF A HARMONIC FREQUENCY
IN THE CORN YIELD DATA WITH
A THIRD ORDER TREND CORRECTION



The probability of a single harmonic frequency plus a second-order trend (C) continues to pick out the low frequency trend. However, the level and spread of the marginal posterior probability density is such that the trend has almost been removed. When the probability of a single harmonic frequency plus a third-order trend is computed, the probability density suddenly changes behavior. The frequency near 0.3 is now the dominant feature (D). The trend has not been completely removed; a small artifact persists at low frequencies.