

METHODS & DESIGNS

Fourier analysis of reaction time data

DAVID M. GREEN*

University of California at San Diego, La Jolla, California 92037

The popularity of the assumption of stages in models of the reaction time process and the availability of fast and efficient means of computing approximations to the Fourier transform makes the Fourier analysis of reaction time data attractive. This paper indicates some problems associated with such analyses and suggests convenient ways to overcome some of the difficulties.

Transforms, such as the Fourier transform, are useful whenever one thinks the reaction process is composed of stages. Consider the following simple model from Green and Luce (1971). The model assumed that the reaction time distribution reflects two separate and independent stages. One stage is the sensory decision process and represents the time taken to process the stimulus until a decision is reached. The second stage is a residual stage; it represents all delays, both fixed and variable, other than those associated with the sensory decision stage.

The critical assumption is that the resulting reaction times can be described as a convolution of the two stages, that is,

$$f(t) = \int_0^{\infty} s(\tau)r(t - \tau)d\tau \quad (1)$$

where $f(t)$ is the reaction time density, $s(t)$ the density of the sensory decision latencies, and $r(t)$ the density of the residual latencies. If one could obtain the Fourier transforms of $r(t)$, $s(t)$, and $f(t)$, then in the transform domain,

$$F(\omega) = S(\omega)R(\omega) \quad (2)$$

where capital letters are used to represent the transforms of the corresponding time processes. For example,

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t}dt.$$

Converting a convolution, Eq. 1, to a multiplication, Eq. 2, is the basic property of most classical transforms.

*This research was supported by a grant from National Science Foundation. The author is indebted to R. Duncan Luce for suggesting the original exploration of this technique, to James Patterson who helped interpret many of the convergence problems, to Virginia Maier who did the programming, and to Norman Anderson who read and criticized a number of earlier drafts.

Given a knowledge of one process, say $R(\omega)$, one could divide the transform of the obtained reaction time, $F(\omega)$, and thereby obtain $S(\omega)$. Taking the inverse transform of $S(\omega)$ would yield $s(t)$. In effect, this procedure removes $r(t)$ from $f(t)$ and allows one to isolate $s(t)$. As with most such schemes, there are lots of pitfalls, but the goal is certainly tantalizing.

Background

The stage assumption is a popular one in modern investigations of reaction time processes. Sternberg (1964, 1969) has written an extremely thorough and perceptive analysis of the various stage models. The assumption that the final distribution can be written as a convolution of the stages is strong, but not uncommon. This paper deals only with two stages,

but multistage models could be reduced to a two-stage process by isolating one stage and lumping all the rest together as a remainder. Once the isolated stage is removed, another stage is separated from the remainder, and so forth. The principles are the same for multistage as for two-stage models, although the estimation problems are worse, and they are bad enough with only two stages.

In the application discussed, the theory outlined in Green and Luce (1971) is used to specify that the sensory decision process, $s(t)$, is exponential in form. The form of the residual process is completely unknown. In this case, the information about $s(t)$ can be used to specify the Fourier transform, $S(\omega)$. An estimate of the Fourier transform $F(\omega)$ can be obtained from the reaction time histogram. By dividing the estimate of $F(\omega)$ by $S(\omega)$, one obtains $R(\omega)$, and by inverting, an estimate of $r(t)$, the residual process. Were $r(t)$ the same for many different experiments using different stimuli, one would gain some confidence both in the Fourier transform method itself and the assumptions about the form of $s(t)$.

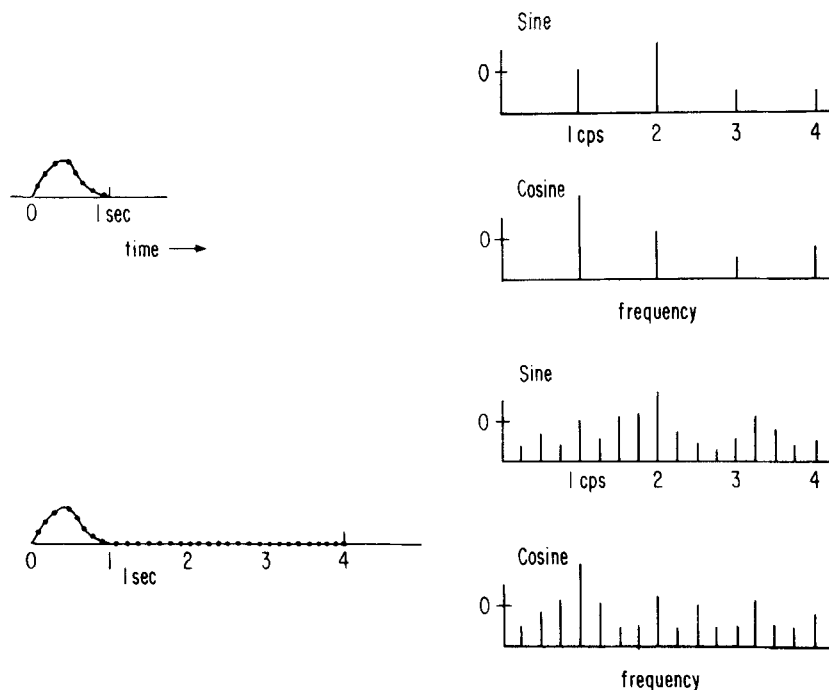


Fig. 1. Examples of Fourier representation of 8-point (top) and 32-point (bottom) time functions. The coefficients of the sine and cosine series are represented on the right. Note the coefficients for the four common frequencies are the same in both the 8-point and the 32-point examples.

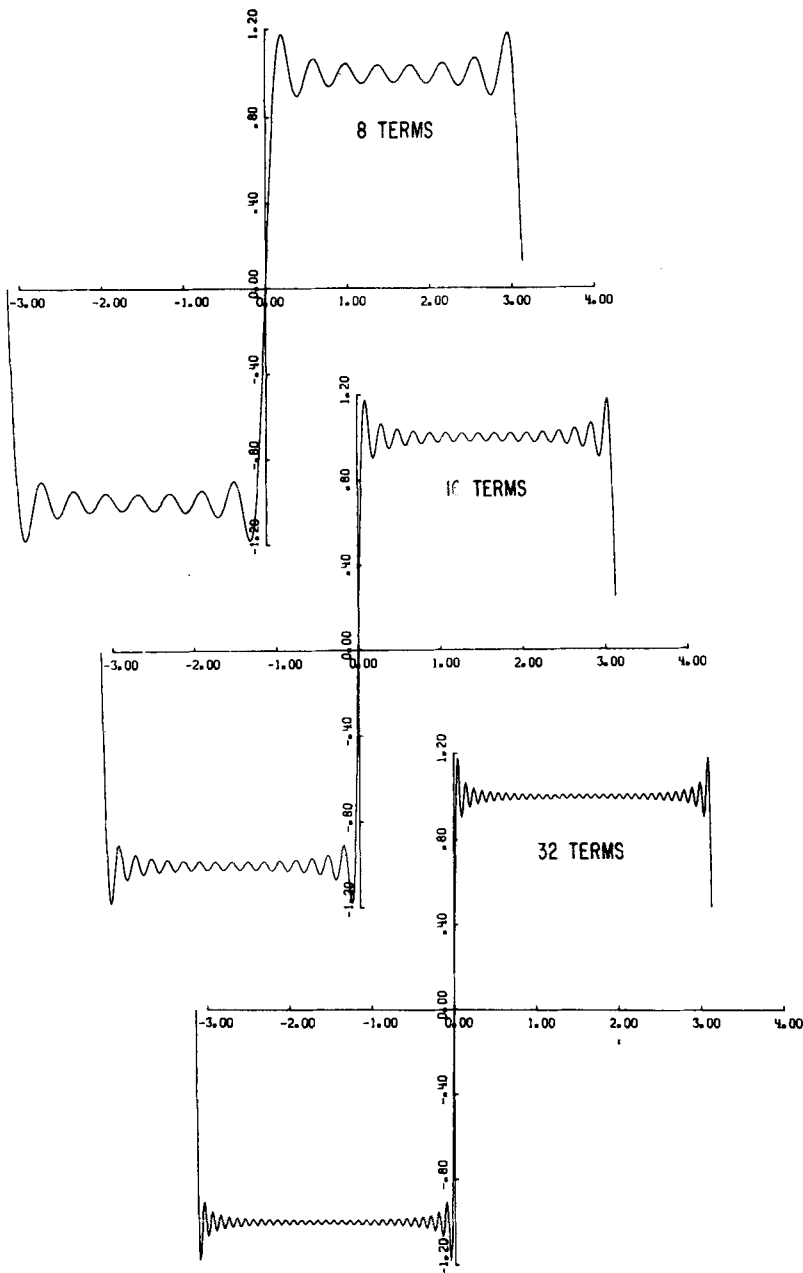


Fig. 2. The waveform associated with the Fourier series representation of a unit square wave. The number of terms in the series is indicated by the waveform. Note the overshoot of the value 1 at the point of discontinuity.

This paper is one of method and describes how a particularly efficient program, the fast Fourier transform (FFT), may be used to analyze the reaction time distribution. Although such an application may appear to be straightforward, there are, in fact, several rather subtle features, both in execution and in interpreting the results of such application. The main technical problem is that a reaction time distribution is most naturally estimated as a histogram, and the discontinuities of the histogram create rather severe problems in the convergence of a discrete approximation to Fourier transform.

Thus some form of smoothing must be used to obtain sensible results. A convenient way to accomplish this smoothing and some implications of this procedure is the main topic of this paper.

A few preliminary comments on the relation of this approach to the general problem of the analysis of reaction time data may be of interest. While latency data should provide very clear evidence about the nature of the processing mechanism, this promise is largely unfulfilled. In the case of simple reaction time, there appear to be two basic problems.

First, even simple reaction time data

are often contaminated with simple anticipations and time estimates as well as reaction to the primary stimulus (Snodgrass, Luce, & Galanter, 1967). The resulting distribution of reaction time judgments is therefore a composite of different distributions such as those suggested by the fast guess models of Ollman (1966) and Yellott (1967). The Fourier analysis suggested in this paper will not aid in separating composite distributions. Instead, various experimental controls must be exercised to avoid the contaminating influence of a mixture of processes. Only if the data are reasonably pure is the application of Fourier analysis likely to help in understanding the stages of the process.

Another general problem that hinders a more extensive exploitation of reaction time data is that the motor side is poorly understood. The exact physical details of the reaction button can contribute delays of 50 to 150 msec. Standardized apparatus has not been adopted, and it is well known, therefore, that one can expect constant differences in mean reaction times from one laboratory to the next. Thus, only relative differences between conditions run within the same experiment can be trusted, and this severely limits the usefulness of reaction time as an experimental technique. One possible contribution of the present technique might be to provide some information about the motor as opposed to the sensory decision processes. Thus it might be possible to correct for such mundane items as the travel time of the button and make possible the direct comparison of reaction times taken in different laboratories. In any case, for the purposes of this paper, we assume that $f(t)$ can only be estimated by a histogram and that one wishes to transform $f(t)$ to obtain $F(\omega)$.

TRANSFORMS OF HISTOGRAMS

Fast Fourier Analysis

Calculating the Fourier transform of reaction time data, except on a very limited set of data, would have been unthinkable a decade ago. Using the then available approximation techniques, the cost was proportional to the square of the number of data points, and this simply made calculation too expensive, even on the fastest machines. In 1965 a computational scheme suggested by Cooley and Tukey (1956) yielded computational time proportional to $n \log n$ rather than n^2 and made routine calculation of Fourier transforms possible. The programs employing this efficient computational algorithm are known as the fast Fourier transform (FFT) programs and are available at all computation

Table 1
Fast Fourier Analysis of Unit Square Wave

$$g(x) = \begin{cases} +1 & 0 \leq x < \pi \\ -1 & \pi < x < 2\pi \end{cases}$$

$$g'(x) = \frac{4}{\pi} (\sin x + 1/3 \sin 3x + 1/5 \sin 5x \dots)$$

Entry is magnitude times $\pi/4$

Frequency	64 Point Analysis			256 Point Analysis		
	Sine	Cosine	Error (Sine)	Sine	Cosine	Error (Sine)
D.C.	0	0		0	0	
f_0	.9992	.0491	.01%	1.000	.0122	0%
$2f_0$	0	0		0	0	
$3f_0$.3309	.0491	0.7%	.3332	.0122	0.05%
$4f_0$	0	0		0	0	
$5f_0$.1960	.0491	2.0%	.1998	.0122	0.13%
$6f_0$	0	0		0	0	
$7f_0$.1372	.0491	4.0%	.1425	.0122	0.25%
$8f_0$	0	0		0	0	
$9f_0$.1038	.0491	6.6%	.1107	.0122	0.4%
$10f_0$	0	0		0	0	
$11f_0$.0819	.0491	9.9%	.0904	.0122	0.6%
etc.						

centers. Today, calculation of a transform involving 2,000 or 4,000 data points takes 10 sec or so rather than many minutes.

Since we are calculating a discrete approximation to the transform, we can view the program very simply as calculating the coefficients in a series approximation to the reaction time distribution using the sine and cosine functions. If we start with n points on the histogram, we end with $n/2$ coefficients for the sine and $n/2$ coefficients for the cosine terms, as shown in the top part of Fig. 1. Normally one adds a number of zeros at the end of the histogram both because the best estimate of the probability of a reaction in that region is zero and because such a procedure contributes to better frequency resolution by increasing the density of the frequency components. This is illustrated in the bottom portion of Fig. 1. The FFT programs require that the number of input points equal some power of 2; hence, enough zeros are added to increase n to some convenient number such as 4,096. For further information on FFT, an extensive discussion is contained in a special IEEE (1967) volume.

Convergence and Gibb's Phenomena

Let $f(t)$ be the original histogram and $f'(t,n)$ be the series approximation obtained by using n terms. The notation emphasizes that the approximation depends on a finite number of terms, n . The function $f'(t,n)$ is an approximation in the mean squared sense; if

$$f(t) - f'(t,n) = e(t)$$

then

$$\int_0^\infty e^2(t) dt \rightarrow 0 \text{ as } n \rightarrow \infty$$

Unfortunately this does not guarantee absolute convergence, that is, it does not guarantee that the error itself will approach zero. In fact, the value of $e(t)$ at a discontinuity in f approaches a definite nonzero number. This anomalous behavior is usually referred to as Gibb's phenomena. It is illustrated in a simple example, namely a unit square wave, in Fig. 2. Near the value of the discontinuity, $f'(t,n)$ overshoots or undershoots $f(t)$ by a considerable amount. But as n

increases, this overshoot does not diminish, even though the total area of the error approaches zero. Increasing the number of terms in the series improves the mean square error but the failure of convergence in the absolute sense, or Gibb's phenomena, is evident in Fig. 2. Normally, we are not using a series approximation of a time function but rather the converse. We generally have a time function, the reaction time distribution, specified at a set of evenly spaced values of time, and we wish to calculate the series that approximates that time waveform. The approximation error then occurs in the frequency rather than the time domain.

To illustrate this problem, we reverse the problem just considered and calculate the transform of a simple square wave specified at a finite number of points. The first half of the square wave has a value of +1, whereas the last half of the square wave has a value of -1. Table 1 gives the FFT analysis for such a square wave defined by either 64 or 256 input terms. The sine coefficients of the odd harmonics approximate the series, 1, 1/3, 1/5, 1/7, 1/9, with better accuracy as the number of input terms increases. Note, however, that there is also a constant magnitude in the odd harmonics of the cosine series. The cosine terms correspond to a constant amplitude pulse train alternating in sign at the discontinuities of the square wave (plus on a transition from -1 to +1, negative on a transition from +1 to -1).

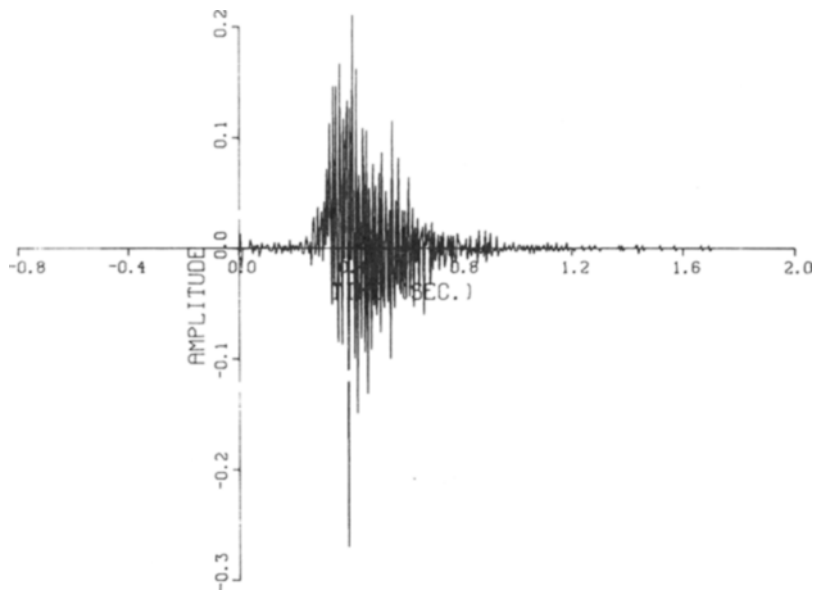


Fig. 3. Straightforward application of the Fourier analysis technique to data from single S_s ($N = 10,020$). See Green & Luce (1971) for details of experiment. Parameters are $\lambda = 0.5$, $v = 0.065$, $\mu = 8.2$, $R_{-\lambda} = 0.81$.

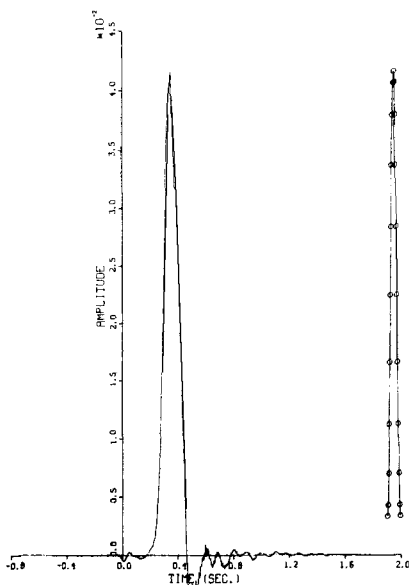


Fig. 4. Same data as that shown in Fig. 3, except 21-point Hamming window shown to right of graph used to smooth data.

Unfortunately, a histogram of reaction times may have as many discontinuities as class boundaries, and therefore some means of suppressing these discontinuities is needed. Fortunately, simple sliding averages provide a simple and effective way to avoid many problems. Suppose we simply replace our input with the average of adjacent points. We also average the first and last point since the waveform is, in effect, periodic for the purpose of Fourier series analysis. Thus an input of eight points would read 0, +1, +1, +1, 0, -1, -1, -1. The FFT analysis of such a smoothed square wave is the same as that shown in Table 1, except that the cosine terms vanish. The sine term remains the same. This simple experiment motivates us to consider more sophisticated forms of smoothing.

Smoothing to Avoid Discontinuities

The value of smoothing discontinuous functions such as square waves is obvious. In effect, smoothing, or sliding, averages are a convolution of the input data with some "averaging" function. A variety of smoothing functions, usually called "windows" in the literature, have been proposed. They have different strengths and weaknesses and some comparative analysis is available (Blackman & Tukey, 1958). The material that follows used a Hamming window. This window performs a weighted running average on the values at successive intervals in time. The weights are given by the following formula:

$$d(\tau) = 0.54 + 0.46 \cos \pi \tau / T_m \quad (3)$$

where $\tau = k\Delta t$, $k = 0, \pm 1, \pm 2, \dots, +m$, and $\Delta t = T_m/m$. The number of points ($2m + 1$) can be varied by the E to achieve the amount of smoothing desired. The more one smooths, the more one suppresses the convergence problems discussed earlier. But smoothing blurs time, for it renders similar two events in time that were formerly distinct. Let us now illustrate one application of the analysis of reaction time data.

Analysis of Reaction Time Data

For illustration, consider some data collected by Green and Luce (1971) and the analysis of that data in terms of the theory mentioned earlier. Briefly, the theory asserts that the reaction time distribution $f(t)$ is the result of the convolution of $r(t)$, an unknown distribution, with $s(t)$. The form of $s(t)$ is essentially exponential, and the parameter of the exponential distribution is presumably related to signal intensity. Thus, if we transform $f(t)$ by numerical techniques, we can solve for $R(\omega)$. The transform of $R(\omega)$ then gives us an estimate of $r(t)$.

Going through the steps outlined above produces Fig. 3. This irregularity reflects various problems of convergence. We can, as we might expect, improve the appearance of the result by smoothing $f(t)$ before we transform it. The problem with this approach is that for each smoothing window we have to repeat the entire

transform process, and this is both time-consuming and expensive. Ideally, we would prefer to average the final result, $r(t)$, rather than the input, $f(t)$. Then the transform programs are used once. Any change in smoothing windows would involve the final averaging used to estimate $r(t)$. The next section outlines how this trick can be accomplished.

Impulse Analysis

We wish to smooth $r(t)$ and obtain the same results as if we had smoothed $f(t)$. Since $f(t)$ is a linear function of $r(t)$, this is possible as long as no spurious elements are introduced in calculating r from f . However, this is exactly what happens when we transform a finite approximation of $S(\omega)$. The time waveform, $s(t)$, will contain errors at the discontinuities in $s(t)$ just as the square wave did (see Fig. 2). No matter how many terms we employ in approximating the transform $S(\omega)$, the transform of this approximation would not exactly equal $s(t)$.

However, since the theory specifies both $s(t)$ and $S(\omega)$, our goal can be achieved by the following trick. We calculate $S(\omega)$ by asking the computer to transform $s(t)$. Then $s(t)$ will be exactly correct at all the sample points and $S(\omega)$ will contain all the errors associated with convergence problems, just as our transform of the square wave did (Table 1). Call the approximation calculated in this way $S^f(\omega)$. The procedure then is to

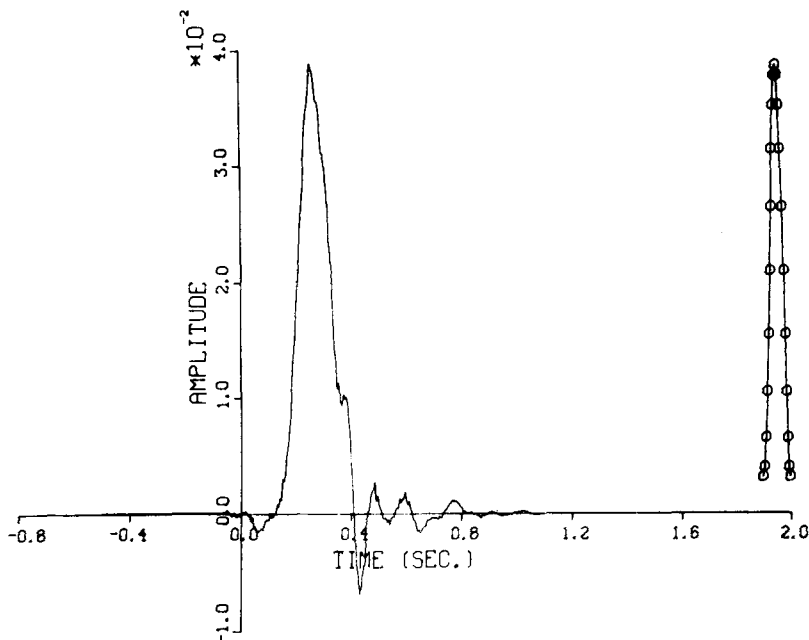


Fig. 5. Estimated $r(t)$ for same S and similar condition as used for Fig. 3. The original reaction time distribution based on 2,145 observations. Parameters are $\lambda = 2.0$, $\nu = 0.96$, $\mu = 10.4$, $R_{-\lambda} = 0.59$.

transform $f(t)$, solve for $R(\omega)$ using $S'(\omega)$, and transform the result to obtain $r(t)$. Smoothing $r(t)$ now produces exactly the same results as smoothing $f(t)$. In effect, we have swept all our worst convergence problems into the frequency domain.

Figure 4 shows the result of Hamming Fig. 3 with a 21-point window. The actual values used in the window are displayed on the right side of Fig. 4, so that one can judge what variation is possible and significant.

The general shape of $r(t)$, as estimated in Fig. 4, has been replicated. In the course of another experiment, the same S was observed under similar circumstances for 2,000 observations. Figure 5 shows the $r(t)$ estimated in that experiment. Both figures show negative loops, which are theoretically impossible. One must not expect too much stability in $r(t)$, however, since it is a linear transform of a histogram and the estimates of the various cells of the histogram are not very stable. The average, of course, contributes both to the stability and the smoothness of $r(t)$, since it makes successive points highly correlated.

CONCLUSIONS AND SUMMARY

The obtained estimate $r(t)$ is not unreasonable, although several features of the distribution tell us that the

theory is wrong in detail. First, the negative loops between 500 and 700 msec are impossible. They are a consistent feature of other analyses of similar data performed on this S and other S s in similar experiments. As yet, we have not been able to determine exactly what feature of the model is at fault. Second, according to our theory, if the signal level were increased, the exponential wait for the signal should be negligible. Thus at large signal-to-noise ratios, we would expect that $f(t)$ should become $r(t)$. This does not happen. Although $f(t)$ resembles the distribution shown in Fig. 4, the entire distribution of responses is about 150 msec quicker.

Despite this failure, the Fourier technique appears to be useful in the analysis of reaction time data and will undoubtedly be used in future investigations. Although other, more fundamental, solutions to the convergence problem may be developed, the present technique is convenient and efficient and does aid in a more detailed analysis of the reaction time process.

REFERENCES

BLACKMAN, R. B., & TUKEY, J. W. *The measurement of power spectra*. New York: Dover, 1958.
 COOLEY, J. W., & TUKEY, J. W. An

algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation*, 1965, 19, 297-301.
 GREEN, D. M., & LUCE, R. D. Detection of auditory signals presented at random times, III. *Perception & Psychophysics*, 1971, 9, 257-268.
 I E E E *Transaction of audio and electroacoustics*. Vol. AU-15. IEEE, 1967. Pp. 43-113.
 LINK, S. W., & TINDALL, A. D. Reaction times to comparative judgement of line length. Technical Report No. 32 (not dated). Department of Psychology, McMaster University, Hamilton, Ont., Canada.
 NICKERSON, R. S. Expectancy waiting time and the psychological refractory period. *Acta Psychologica*, 1967, 27, 23-34.
 NICKERSON, R. S., & BURNHAM, D. W. Response times with nonaging foreperiods. *Journal of Experimental Psychology*, 1969, 79, 452-475.
 OLLMAN, R. Fast guesses in choice reaction time. *Psychonomic Science*, 1966, 6, 155-156.
 SNODGRASS, J. G., LUCE, R. D., & GALANTER, G. Some experiments on simple and choice reaction time. *Journal of Experimental Psychology*, 1967, 75, 1-17.
 STERNBERG, S. Estimating the distribution of additive reaction-time components. Paper presented at the meeting of the Psychometric Society, Niagara Falls, Ont., October 1964.
 STERNBERG, S. The discovery of processing stages: Extensions of Donder's method. In W. G. Koster (Ed.), *Attention and performance II*. *Acta Psychologica*, 1969, 30, 276-315.
 YELLOTT, J. Correction for guessing in choice reaction time. *Psychonomic Science*, 1967, 8, 321-322.