

A Direct Digital Method of Power Spectrum Estimation

Abstract: This paper discusses a method of digital power spectrum estimation involving the direct combination of the sample time function with sines and cosines. This treatment is in contrast to the normal indirect digital method which proceeds through the intermediary of the autocovariance function. All the practical design details necessary for the planning of a spectral estimation program are treated.

Introduction

Power spectra are generally estimated either digitally by estimating the autocovariance function and taking its Fourier transform or in an analog fashion using a bank of bandpass filters followed by square law and integrating devices. In this paper a digital method paralleling analog filter bank techniques will be described. It is a direct digital procedure (following the wording of Blackman and Tukey¹) in that it involves the direct combination of sines and cosines with the sample time function. It has been applied to the study of speech signals at the IBM Research Center, Yorktown Heights, N. Y.

The general form of the procedure is as follows. We will designate the sample time function by $x(t)$; from it we will generate a digital approximation, $P(t; f_0)$, to the following random function:

$$\left| \int_{-\infty}^{\infty} x(t-\tau) w(\tau) \exp(2\pi j f_0 \tau) d\tau \right|^2.$$

$P(t; f_0)$ will then be sampled at a uniform rate and the mean of these samples will constitute an estimate of the power spectrum in the region about the point f_0 .

An outline of the paper is given below. Most of the mathematical details are contained in the appendices. Section 2 and Appendix A describe the procedures for setting up the frequency location and resolution characteristics of an estimation program. Sections 3 and 4 and Appendices B, C, and D describe the accuracy characteristics of an estimation program; i.e., they give the consequences of taking the mean of n samples from $P(t; f_0)$ when the sampling period is ΔT . The problem of the quantization of amplitudes is discussed in Section 5 and Appendix E, and related to the available spectral dynamic range. Sections 1 and 6 contain some discussion and actual comparisons of the computational merit of the two procedures, direct and indirect. Section 7 discusses the

speed and basic characteristics of the existing set of IBM 704 computer programs which implement the direct procedure.

This paper is indebted to the background of spectral estimation procedures and concepts due to Blackman and Tukey.¹ In fact, this type of direct procedure is mentioned in Section 3 and Section B-4 of their paper but described there as not preferable to the indirect method in the digital case. As stated above, a discussion of this question of preference is contained in Sections 1 and 6 of this paper.

The initial thought to use this type of procedure in digital estimation is due to G. L. Shultz, who also wrote the computer programs embodying the procedure.

1. Situations in which direct approach holds computational advantages over indirect

The approach which we will discuss yields complete control over the resolution and accuracy variables involved in the estimation procedure. It makes possible estimates of the spectrum of arbitrary resolution and accuracy at arbitrary points throughout the available frequency range. Because of this, it has strong computational advantages over the indirect procedure for estimation requirements which deviate from the usual uniform resolution and accuracy requirements considered in connection with the indirect procedure.

Consider the following examples:

A. An experimenter has data sampled from a time series in such a way as to make possible frequency analysis up to 4000 cps. However, he is interested only in the spectrum from 2000 to 3000 cps. If he uses the indirect approach, there is no way he can calculate only that part of the autocovariance function which yields frequency estimates from 2000 to 3000 cps. He must calculate the

same autocovariance function he would need for the estimation of the entire spectrum.

B. An experimenter is studying a phenomenon such that his requirements for frequency resolution and accuracy differ from the low-frequency region to the high. Again if he uses the indirect approach he must calculate that autocovariance function which will meet his most strict resolution and accuracy requirements everywhere, and hence, which will probably (with the combinations of spectral estimates) exceed his accuracy requirements in most places.

C. An experimenter wishes to look "sequentially" at the spectrum, viewing certain regions with increased resolution and accuracy based upon what he has thus far seen. Again, this cannot be efficiently done indirectly.

We will return to this question with some specific calculations in Section 7.

2. Location and resolution characteristics of spectral estimation procedure

We are, in this and following discussions, considering a stationary random process with a sample function, $x(t)$, and a two-sided power (or variance) spectrum $P(f)$. This spectrum is two-sided in that it is defined over positive and negative frequencies with $P(-f) = P(f)$. In our analysis we will be dealing exclusively with this two-sided spectrum $P(f)$; however, we will normalize our estimates so that they properly estimate the one-sided power spectrum

$$P_+(f) = 2P(f) \quad \text{for } |f| > 0, \\ = 0 \quad \text{otherwise.}$$

Further we are assuming that $x(t)$ is such that

$$E\{x(t)\} = 0$$

and that $x(t)$ is limited to frequencies less than or equal to W ; i.e., that

$$P(f) = 0 \quad \text{for } |f| > W.$$

In Appendix A we develop, from $x(t)$, a random function

$$P(t; f_0) = u^2(t; f_0) + v^2(t; f_0) \\ = \{1/2W \sum_{m=-M}^M x(t-m/2W)w(m/2W) \\ \cos(2\pi f_0 m/2W)\}^2 \\ + \{1/2W \sum_{m=-M}^M x(t-m/2W)w(m/2W) \\ \sin(2\pi f_0 m/2W)\}^2.$$

The expected value of $P(t; f_0)$ represents a look at the spectrum $P(f)$ through a "spectral window" (in the sense of Blackman and Tukey¹) whose shape is determined by $w(t)$ and whose location is determined by f_0 . It is a set

of these quantities

$$\{E \ P(t; f_0)\}$$

which we are going to estimate. In the discussion below,

$$A(f) = \int_{-\infty}^{\infty} w(t) \exp(-2\pi jft) dt,$$

i.e., $A(f)$ is the Fourier transform of $w(t)$.

In general

$$E\{P(t; f_0)\} = \int_{-W}^W Q(f; f_0) P(f) df.$$

In Appendix A three cases are considered which are sufficient to enable us to construct a set of estimates whose expectations cover the frequency range of $x(t)$. These three cases are described below. These cases will be carried through Sections 3 and 4. An appropriate sketch of the resulting $Q(f; f_0)$ is included for each (see Figs. 1, 2 and 3).

• Case 1

$$A(f-f_0) = 0 \quad \text{for } f < 0 \text{ and } f > W.$$

This case will cover all estimators of $P(f)$ other than those at the end points. In this case

$$E\{P(t; f_0)\} = \frac{1}{2} \int_{-W}^W \{|A(f-f_0)|^2 + |A(f+f_0)|^2\} P(f) df.$$

• Case 2

$$f_0 = 0, \quad A(f) = 0 \quad \text{for } |f| > W.$$

This case covers the lower end of the spectrum. We have here

$$E\{P(t; 0)\} = \int_{-W}^W |A(f)|^2 P(f) df.$$

• Case 3

$$f_0 = W, \quad A(f) = 0 \quad \text{for } |f| > W.$$

This case covers the high end of the spectrum.

$$E\{P(t; W)\} = \int_{-W}^W \{|A(f-W)|^2 + |A(f+W)|^2\} P(f) df$$

The normalization of the spectral window such that

$$\int_{-\infty}^{\infty} Q(f; f_0) df = 2$$

is shown in Appendix A for all three cases to be equivalent to normalizing $w(t)$ such that

$$\int_{-\infty}^{\infty} w^2(t) dt = 2.$$

This normalization is proper for the estimation of the one-sided spectrum $P_+(f)$.

Using these results we are now in a position to design the location and resolution characteristics of a spectral

estimation procedure. With the restrictions of Cases 1, 2, and 3 in mind we select a set of center frequencies $\{f_{0i}\}$ and a set of weighting functions $\{w_i(t)\}$. These yield a set of random functions $\{P(t; f_{0i})\}$ whose expectations we are going to estimate. The functions $\{w_i(t)\}$ determine the type of average obtained at the points $\{f_{0i}\}$ through the resulting spectral windows $\{Q(f; f_{0i})\}$.

The set of functions $|A_i(f)|^2$ constitute the Blackman and Tukey¹ spectral windows. As concerns the choice of functions $w_i(t)$, in the operation of our programs, we used the cosine arch

$$w_i(t) = \frac{1}{2}(1 + \cos 2\pi t/T_i) \quad \text{for} \quad -T_i/2 \leq t \leq T_i/2, \\ = 0 \quad \text{otherwise,}$$

keeping the form of the function constant and varying only the parameter T to vary the resolution. Any of the acceptable "lag windows" of Blackman and Tukey would serve as good choices for $w_i(t)$. The spectral windows would not be the corresponding spectral windows of Blackman and Tukey, however, but their square. Thus the side lobes would be down twice as many decibels as in the indirect approach.² In the case of the cosine arch this means the first side lobe is down approximately 32 db from the main lobe peak. In this sense one can obtain a better estimate of the spectrum using this method than using the indirect method.

3. Confidence intervals for estimation of $E\{P(t; f_0)\}$ - the Gaussian case

Once the designer has settled the question of where and to what degree of resolution he desires estimates of the spectrum, that is, once he has settled on a set of random functions $\{P(t; f_{0i})\}$, there remains the problem of the estimation of the set $\{E\{P(t; f_{0i})\}\}$.

$E\{P(t; f_{0i})\}$ will be estimated by sampling $P(t; f_{0i})$ and averaging over the samples obtained. Thus the designer needs to know for each $P(t; f_{0i})$ the relationship between the sampling period, $\Delta_i T$, the number of samples taken, n_i , and the accuracy to which he can estimate $E\{P(t; f_{0i})\}$ with the estimator

$$\bar{P}(f_{0i}) = 1/n_i \sum_{k=1}^{n_i} P(k\Delta_i T; f_{0i}).$$

We will treat the Gaussian case (i.e., assume that $x(t)$ is a sample function from a Gaussian random process), and we will derive formulas for the variance of $\bar{P}(f_{0i})$ as a function of the sampling and averaging procedure. We will then follow Blackman and Tukey¹ and suggest that to establish confidence intervals, we assume $\bar{P}(f_{0i})$ to be a multiple of a χ^2 variable. Thus, to establish confidence intervals, the $\chi^2_{k_i}$ distribution can be used with the number of degrees of freedom, k_i , given by

$$k_i = 2E^2\{\bar{P}(f_{0i})\}/\text{Var}\{\bar{P}(f_{0i})\}, \quad (3.1)$$

that is, a χ^2 distribution with the same (mean)²-to-variance ratio as our estimator. These assumptions and the

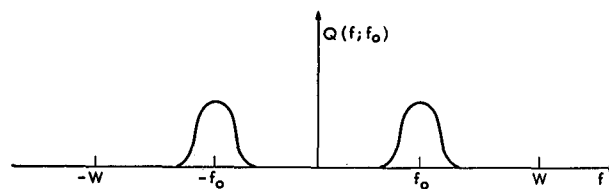


Figure 1

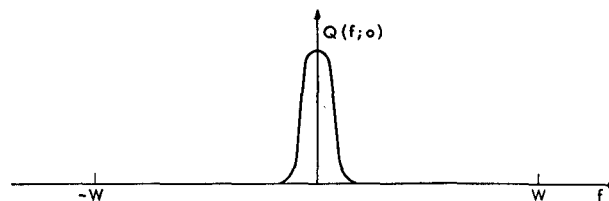


Figure 2

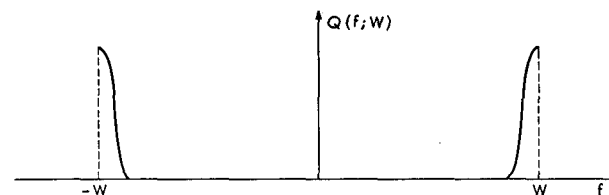


Figure 3

resulting confidence intervals are discussed in detail in Blackman and Tukey,¹ Sections 6, 7, 8, and 9.

The single-sample case

We will first derive formulas for the ratio of the (mean)² to the variance for the case of a single sample (that is, for $P(t; f_0)$) for the three cases considered in Section 2 and Appendix A. We will be using the results of Appendix B where the variance spectrum and total variance of $P(t; f_0)$ are derived.

• Case 1

From (A.2) we have

$$E\{P(t; f_0)\} = \frac{1}{2} \int_{-\infty}^{\infty} \{|A(f-f_0)|^2 + |A(f+f_0)|^2\} P(f) df \\ = \int_{-\infty}^{\infty} |A(f-f_0)|^2 P(f) df,$$

and from (B.4)

$$\text{Var}\{P(t; f_0)\} = \left\{ \int_{-\infty}^{\infty} |A(f-f_0)|^2 P(f) df \right\}^2,$$

and hence

$$E^2\{P(t; f_0)\}/\text{Var}\{P(t; f_0)\} = 1 \quad (3.2)$$

and

$$k = 2E^2\{P(t; f_0)\}/\text{Var}\{P(t; f_0)\} = 2.$$

Hence, regardless of the shape of $P(f)$ we get the equivalent of 2 degrees of freedom in the estimation of $E\{P(t; f_0)\}$ from one sample.

• Case 2

From (A.3) we have

$$E\{P(t; 0)\} = \int_{-\infty}^{\infty} |A(f)|^2 P(f) df$$

and from (B.5)

$$\text{Var}\{P(t; 0)\} = 2 \left\{ \int_{-\infty}^{\infty} |A(f)|^2 P(f) df \right\}^2.$$

Hence

$$E^2\{P(t; 0)\} / \text{Var}\{P(t; 0)\} = \frac{1}{2} \quad (3.3)$$

and $k=1$.

This could be derived directly since $P(t; 0) = u^2(t)$ and under the assumption of a Gaussian process this is dis-

tributed exactly as a constant times a χ_1^2 variable.

• Case 3

This is identical to Case 2 in that again $P(t; W) = u^2(t)$ and $u^2(t)$ is equal to a constant times a χ^2 distributed variable with one degree of freedom. Thus $k=1$ and (3.3) holds for $P(t; W)$.

The general case

We will now use the results of Appendix C, where an expression for $\text{Var}\{P(f_{0i})\}$ is derived, to obtain formulas for the ratio

$$E^2\{P(f_{0i})\} / \text{Var}\{P(f_{0i})\}$$

and the value of k_i which it yields. We recall that

$$P(f_{0i}) = 1/n_i \sum_{k=1}^{n_i} P(k\Delta_i T; f_{0i}).$$

Applying the expression for $\text{Var}\{P(f_{0i})\}$ given by (C.4) we obtain

$$\frac{E^2\{P(f_{0i})\}}{\text{Var}\{P(f_{0i})\}} = \frac{E^2\{P(t; f_{0i})\}}{\text{Var}\{P(t; f_{0i})\} \int_{-\infty}^{\infty} \frac{\sin^2(\pi n_i \Delta_i T f)}{(\pi n_i \Delta_i T f)^2} \sum_{j=-\infty}^{\infty} N. \text{Var}_{P(t; f_{0i})}\{f - (j/\Delta_i T)\} df}, \quad (3.4)$$

where

$$N. \text{Var}_{P(t; f_{0i})}(f) = \text{Var}_{P(t; f_{0i})}(f) / \text{Var}\{P(t; f_{0i})\}$$

is the normalized variance spectrum of $P(t; f_{0i})$. Applying (3.2) we get for Case 1

$$\frac{E^2\{P(f_{0i})\}}{\text{Var}\{P(f_{0i})\}} = \frac{1}{\int_{-\infty}^{\infty} \frac{\sin^2(\pi n_i \Delta_i T f)}{(\pi n_i \Delta_i T f)^2} \sum_{j=-\infty}^{\infty} N. \text{Var}_{P(t; f_{0i})}\{f - (j/\Delta_i T)\} df}, \quad (3.5)$$

where, using (B.3) and (B.4),

$$N. \text{Var}_{P(t; f_{0i})}(f) = \frac{\int_{-\infty}^{\infty} |A(g - f_0)|^2 P(g) |A(f - g + f_0)|^2 P(f - g) dg}{\left\{ \int_{-\infty}^{\infty} |A(f - f_0)|^2 P(f) df \right\}^2}. \quad (3.6)$$

Using (3.4) and (3.3) we get for Cases 2 and 3

$$\frac{E^2\{P(0 \text{ or } W)\}}{\text{Var}\{P(0 \text{ or } W)\}} = \frac{1}{2 \int_{-\infty}^{\infty} \frac{\sin^2(\pi n_i \Delta_i T f)}{(\pi n_i \Delta_i T f)^2} \sum_{j=-\infty}^{\infty} N. \text{Var}_{P(t; 0 \text{ or } W)}\{f - (j/\Delta_i T)\} df}. \quad (3.7)$$

For Case 2 the normalized variance spectrum is, using (B.5) and (B.6), given by (3.6) with $f_0=0$. For Case 3, using (B.7) and (B.8), we can obtain a similar expression for the normalized variance spectrum.

Finally, the equivalent number of degrees of freedom, k_i is given in each case by (3.1).

4. Application of results of Section 3 to design of estimating procedures

Equations (3.5) and (3.7) in combination with equation (3.1) and the χ^2 tables yield confidence intervals for estimates of $E\{P(t; f_{0i})\}$. However, (3.5) and (3.7) involve the unknown spectrum, $P(f)$, and hence, any statement about the accuracy of an estimate must be based upon an assumption about $P(f)$.

• Sharp concentration of power in pass band

Let us assume first that, within the resolution band³ of the estimator, $P(f)$ is concentrated in a very narrow region, narrow in comparison to the width (approximately $1/n_i \Delta_i T$) of the averaging filter (see Appendix C) being used. Then in Case 1 the function

$$\sum_{j=-\infty}^{\infty} P_{P(t; f_{0i})}(f - j/\Delta_i T)$$

will be concentrated about 0, $\pm 1/\Delta_i T$, $\pm 2/\Delta_i T$, ..., and there will be no reduction of the variance through averaging. The ratio and degrees of freedom will be the same as for a single sample. In Cases 2 and 3 half the area

under

$$\sum_{j=-\infty}^{\infty} P_{P(t; 0 \text{ or } W)}(f - j/\Delta_i T)$$

will be concentrated at $0, \pm 1/\Delta_i T, \pm 2/\Delta_i T, \dots$ and the other half at $\pm 2f', \pm 1/\Delta_i T \pm 2f', \pm 2/\Delta_i T \pm 2f', \dots$ where f' is frequency location of the narrow region of concentration in $P(f)$. Hence, for these Cases and this assumption concerning $P(f)$, one-half the variance may be removed by sampling, meaning a maximum gain of one degree of freedom.

• $P(f)$ constant over pass band

In the second and most important case we will assume $P(f)$ to be constant over the resolution band of the estimator. This is a quite reasonable assumption since often the resolution characteristics of the estimators are designed in an attempt to achieve it.

With this assumption $P(f)$ cancels out of the expression for $N. \text{Var}_{P(t; f_0)}(f)$ for all three cases (see (3.6)) and consequently, the ratios (3.5) and (3.7) involve only the characteristics of the spectral windows. Thus, under these circumstances k_i , the equivalent number of degrees of freedom, can be calculated for any estimator $\hat{P}(f_{0i})$ from its corresponding $\Delta_i T$ and n_i .

For Cases 1 and 2 we get, from (3.6) and our assumption,

$$N. \text{Var}_{P(t; f_{0i})}(f) = \frac{\int_{-\infty}^{\infty} |A(g - f_{0i})|^2 |A(f - g + f_{0i})|^2 dg}{\left\{ \int_{-\infty}^{\infty} |A(f - f_{0i})|^2 df \right\}^2}.$$

Following the procedure discussed in Appendix C, we can define a *doubled-back* (or aliased) spectrum.

$$N. \text{Var}_{P(t; f_{0i})}^* = \sum_{j=-\infty}^{\infty} N. \text{Var}_{P(t; f_{0i})} \{f - (j/\Delta_i T)\} \quad \text{for } |f| \leq 1/2\Delta_i T$$

$$= 0 \quad \text{otherwise.}$$

Then using (3.2) and (C.5), we get for Case 1

$$\frac{E^2\{\hat{P}(f_{0i})\}}{\text{Var}\{\hat{P}(f_{0i})\}} = \frac{1}{\int_{-\infty}^{\infty} \frac{\sin^2(\pi n_i \Delta_i T f)}{(\pi n_i \Delta_i T f)^2} \left\{ \sum_{j=-\infty}^{\infty} N. \text{Var}_{P(t; f_{0i})}^* \{f - (j/\Delta_i T)\} \right\} df} \quad (4.1)$$

and using (3.3) and (C.5) we get the same expression for Cases 2 and 3 with the exception of an additional factor of 2 in the denominator. Now $N. \text{Var}_{P(t; f_0)}(f)$ is non-negligible for frequencies whose absolute value is less than the width of the resolution band of the estimator. Hence, with $1/2\Delta_i T$ less than this width we will have doubling back. $|A(f)|^2$ as a spectral window will be generally smooth with one major lobe and consequently,

$N. \text{Var}_{P(t; f_0)}(f)$ will have its maximum at zero and drop off smoothly to either side. Now the larger $\Delta_i T$ is chosen, the more severe the doubling back, redoubling back, et cetera, becomes and the closer $N. \text{Var}_{P(t; f_{0i})}^*(f)$ approaches the uniform spectrum from 0 to $1/2\Delta_i T$. In Appendix C we show that for this spectrum we get the reduction in variance expected for uncorrelated variables. This reduction yields using (4.1) and (3.1) for Case 1

$$k_i = 2n_i, \quad (4.2)$$

and using (3.1) and an appropriately modified (4.1) for Case 2,

$$k_i = n_i. \quad (4.3)$$

Hence, it would seem desirable to select a large $\Delta_i T$ so that $1/2\Delta_i T$ is very much smaller than the width of the resolution band of the estimator and thus assure this strong reduction of the variance. However, there is good reason for using a small $\Delta_i T$ since $n_i \Delta_i T$ determines the record length required, and generally there is a cost proportional to the record length.

The procedure followed in planning the speech signal estimation programs was to assume $P(f)$ to be constant over the spectral window (in this case a reasonable assumption), then choose a small $\Delta_i T$ such that $N. \text{Var}_{P(t; f_{0i})}^*(f)$ is approximately uniform and use the χ^2 confidence limits where k_i was determined from (4.2) and (4.3). For the cosine arch weighting function

$$w_i(t) = \frac{1}{2}(1 + \cos(2\pi t/T_i)) \quad \text{for } -T_i/2 \leq t \leq T_i/2$$

$$= 0 \quad \text{otherwise,}$$

we used

$$\Delta_i T = 2T_i/3. \quad (4.4)$$

Rough checks of the variations observed in the estimates of spectra of noise-like speech sounds confirmed the χ^2 confidence intervals.

Case 3 is special since in this case $N. \text{Var}_{P(t; W)}(f)$ is not concentrated around zero only but also around the points $\pm 2W$. Thus the doubling back of $N. \text{Var}_{P(t; W)}^*(f)$ probably will not be as uniform as the comparable quantity for Cases 1 and 2. This might mean a greater reduction in the variance or it might mean a smaller one.

• Periodic sample function

The estimation problem when $x(t)$ is a sum of sinusoids and particularly when $x(t)$ is periodic is discussed in Appendix D.

5. Weighting and sample time function quantization and spectral dynamic range

So far we have discussed only the time sampling aspects of the digital calculation. We will now discuss the question of amplitude quantization of the quantities involved. We want to know how many levels are required for the quantization of the sample time function and for the

weighting functions of the estimators; that is, when we represent these quantities digitally what are the consequences of a choice of so many bits for the representation of each? We will see below that these choices determine the dynamic range available for spectral estimation. The question is an important practical one since many of the engineering and computational problems involve difficulties and costs strongly related to the number of levels of these representations. Hence it is frequently important that no more bits than necessary be used for their representation.

In Appendix A we define a random function

$$P(t; f_0) = \left\{ \frac{1}{2W} \sum_{m=-M}^M x(t-m/2W) w(m/2W) \cos(2\pi f_0 m/2W) \right\}^2 \\ + \left\{ \frac{1}{2W} \sum_{m=-M}^M x(t-m/2W) w(m/2W) \sin(2\pi f_0 m/2W) \right\}^2,$$

which is the basis of our estimation of $P_+(f)$ at the point f_0 . In actuality, one doesn't evaluate exactly this but rather something with further random contributions due to the quantization of the functions involved.

We will assume the weighting functions to be

$$w(m/2W) \cos(2\pi f_0 m/2W) + e_{1m}$$

and

$$w(m/2W) \sin(2\pi f_0 m/2W) + e_{2m},$$

where $e_{1m}, m=0, \pm 1, \dots, \pm M; e_{2m}, m=0, \pm 1, \dots, \pm M$ are all independently and identically distributed random variables with a uniform distribution from $-K/2N_1$ to $K/2N_1$; that is, that e_{1m} has a density

$$f_{1m}(x) = N_1/K \quad \text{for} \quad -K/2N_1 \leq x \leq K/2N_1, \\ = 0 \quad \text{otherwise},$$

where $K = \max |w(t)|$. These weighting functions are quantized into $2N_1 + 1$ levels; i.e., that they are represented by $\log_2 N_1$ bits plus a sign.

Further we will assume the sample time function to be

$$x(t-m/2W) + e_{3m},$$

where $e_{3m}; m=0, \pm 1, \dots, \pm M$ are independently and identically distributed with density

$$f_{3m}(x) = 1 \quad \text{for} \quad -\frac{1}{2} \leq x \leq \frac{1}{2} \\ = 0 \quad \text{otherwise}.$$

$x(t)$ is quantized into $2N_2 + 1$ integer values $0, \pm 1, \pm 2, \dots, \pm N_2$ and we assume the variance of $x(t)$ is controlled so the clipping which results from placing the sample function in the range $(-N_2 - \frac{1}{2}, N_2 + \frac{1}{2})$ is negligible. We also assume that e_{3k} is independent of e_{1i} , and e_{2j} for all i, j, k and also that e_{1i}, e_{2j} , and e_{3k} are independent of $x(t-m/2W)$ for all i, j, k and m .

Hence, $P(t; f_0)$ becomes

$$P^*(t; f_0) = \left\{ \frac{1}{2W} \sum_{m=-M}^M [x(t-m/2W) + e_{3m}] \right. \\ \left. [w(m/2W) \cos(2\pi f_0 m/2W) + e_{1m}] \right\}^2 \\ + \left\{ \frac{1}{2W} \sum_{m=-M}^M [x(t-m/2W) + e_{3m}] \right. \\ \left. [w(m/2W) \sin(2\pi f_0 m/2W) + e_{2m}] \right\}^2 \\ = \left\{ \frac{1}{2W} \sum_{m=-M}^M x(t-m/2W) w(m/2W) \cos(2\pi f_0 m/2W) + x(t-m/2W) e_{1m} \right. \\ \left. + e_{3m} w(m/2W) \cos(2\pi f_0 m/2W) + e_{3m} e_{1m} \right\}^2 \\ + \left\{ \frac{1}{2W} \sum_{m=-M}^M x(t-m/2W) w(m/2W) \sin(2\pi f_0 m/2W) + x(t-m/2W) e_{2m} \right. \\ \left. + e_{3m} w(m/2W) \sin(2\pi f_0 m/2W) + e_{3m} e_{2m} \right\}^2 \quad (5.1)$$

Now if the power within the resolution band of the estimator is large, then the first terms within each of the brackets in (5.1) will dominate and the discussion of Sections 3 and 4 on confidence intervals for the estimators, $P(f_{0i})$, will be valid. However, as $E\{P(t; f_{0i})\}$ becomes small these confidence intervals become more and more inaccurate (too narrow) because the effects of the quantizing errors become non-negligible. Eventually the variation caused by this quantizing will place a lower limit on spectral averages which can be measured and thus determine the dynamic range available for spectral estimation.

To solve this problem completely one would have to study the small-signal distribution of the estimator and obtain exactly the pattern in which the confidence intervals change as the signal power within the resolution band decreases. We have not done this.⁴ Instead, we took the expectation of $P^*(t; f_0)$ and obtained an approximate lower limit for spectral estimates at the point where they equal the expected value of the contributions due to the quantization; that is, at the point the expected "signal power" out of the estimator is equal to the expected quantization "noise power." This was felt adequate for planning purposes—a feeling which was confirmed empirically through actual operation of the program.

If we expand (5.1) and make use of the independence properties of the quantization error random variables, we get

$$E\{P^*(t; f_0)\} = E\{P(t; f_0)\} \\ + (1/2W)^2 (2M+1) \text{Var}(x) \text{Var}(e_1) \\ + (1/2W)^2 (2M+1) \text{Var}(x) \text{Var}(e_2) \\ + (1/2W)^2 \text{Var}(e_3) \sum_{m=-M}^M w^2(m/2W) \cos^2(2\pi f_0 m/2W)$$

$$\begin{aligned}
& + (1/2W)^2 \text{Var}(e_3) \sum_{m=-M}^M w^2(m/2W) \\
& \quad \sin^2(2\pi f_0 m/2W) \\
& + (1/2W)^2 (2M+1) [\text{Var}(e_1) \text{Var}(e_3) \\
& \quad + \text{Var}(e_2) \text{Var}(e_3)] . \quad (5.2)
\end{aligned}$$

The last term of (5.2) we will neglect. If we combine the fourth and fifth terms using the approximation

$$1/2W \sum_{m=-M}^M w^2(m/2W) \approx \int_{-\infty}^{\infty} w^2(t) dt = 2$$

and combine the second and third terms, we get

$$\begin{aligned}
E\{P^*(t; f_0)\} &= E\{P(t; f_0)\} + 2(1/2W)^2 (2M+1) \\
& \quad \text{Var}(x) \text{Var}(e_1) + \text{Var}(e_3)/W . \quad (5.3)
\end{aligned}$$

The first term of (5.3) is the quantity we are interested in measuring. The second and third terms represent the contributions due to the quantization.

We will now work on the second and third terms of (5.3) to put them in a form more suitable for discussion of the dynamic range. Consider first the second term

$$\begin{aligned}
2(1/2W)^2 (2M+1) \text{Var}(x) \text{Var}(e_1) &= \\
& (\text{Var}(x)/W) (1/2W) (2M+1) \frac{(K_1/N_1)^2}{12} \\
&= \overline{P_+(f)} (1/2W) (2M+1) \frac{\{\max |w(t)|\}^2}{12N_1^2} ,
\end{aligned}$$

where

$$\begin{aligned}
\overline{P_+(f)} &= 1/W \int_0^{\infty} P_+(f) df \\
&= \text{average power density} .
\end{aligned}$$

To a good approximation

$$\begin{aligned}
(1/2W) (2M+1) \{\max |w(t)|\}^2 &= k_0 \{w(t)\} \int_{-\infty}^{\infty} w^2(t) dt \\
&= 2k_0 \{w(t)\} ,
\end{aligned}$$

where

$$k_0 \{w(t)\} = \frac{\{\text{length of } w(t)\} \max \{w^2(t)\}}{\int_{-\infty}^{\infty} w^2(t) dt} .$$

Hence

$$2(1/2W)^2 (2M+1) \text{Var}(x) \text{Var}(e_1) = \overline{P_+(f)} k_0 \{w(t)\} / 6N_1^2 .$$

Now consider the third term. For any sample function $x(t)$ to be quantized within the region $(-N_2 - \frac{1}{2}, N_2 + \frac{1}{2})$ its variance must be such that $\text{Var}(x) \leq (N_2 + \frac{1}{2})^2 \approx N_2^2$. Hence depending upon the distribution of $x(t)$ there exists a constant k^* such that whenever $\text{Var}(x) = k_1 N_2^2$ and $k_1 \leq k^*$ then the clipping which results from limiting $x(t)$ to the range $(-N_2 - \frac{1}{2}, N_2 + \frac{1}{2})$ is negligible. Since

we are assuming this to be the case, the third term of (5.3) becomes

$$\begin{aligned}
\text{Var}(e_3)/W &= 1/12W = (1/12W) \text{Var}(x)/k_1 N_2^2 \\
&= \frac{1}{12k_1 N_2^2} \frac{\text{Var}(x)}{W} = \overline{P_+(f)} / (12k_1 N_2^2) .
\end{aligned}$$

Combining these two reductions we get

$$\begin{aligned}
E\{P^*(t; f_0)\} &= E\{P(t; f_0)\} + \overline{P_+(f)} \\
& \quad \left[\frac{k_0 \{w(t)\}}{6N_1^2} + \frac{1}{12k_1 N_2^2} \right] .
\end{aligned}$$

The first term is the expected *signal power*, the second the expected *quantization noise power*.

As was stated above, we will consider that the lowest meaningful estimate is given approximately by

$$E\{P(t; f_0)\} = \overline{P_+(f)} \left[\frac{k_0 \{w(t)\}}{6N_1^2} + \frac{1}{12k_1 N_2^2} \right] . \quad (5.4)$$

From this expression it is evident that, if possible, one should spread the bits for representation among N_1 and N_2 so as to make the terms equal inside the bracket on the right. We will consider dynamic range in terms of the lowest output available relative to the average power level, $\overline{P_+(f)}$. The usual definition involves the comparison of lowest to highest spectral estimates but it necessitates an assumption concerning the shape of $P_+(f)$. The consideration in terms of lowest to average power level is conservative. From (5.4) we can define

$$f(N_1, N_2; k_0, k_1) = -10 \log_{10} \left[\frac{k_0 \{w(t)\}}{6N_1^2} + \frac{1}{12k_1 N_2^2} \right] .$$

The expression $f(N_1, N_2; k_0, k_1)$ is the db dynamic range (lowest estimate to mean power level).

For purposes of illustration consider $N_1 = N_2 = 2^n$, then

$$\begin{aligned}
f(2^n, 2^n; k_0, k_1) &= 10 [\log_{10} \frac{1}{\left[\frac{k_0 \{w(t)\}}{3} + \frac{1}{6k_1} \right]} \\
& \quad + (2n+1) \log_{10} 2]
\end{aligned}$$

is the db dynamic range as a function of the number of bits, plus a sign, used in the representation of both the weighting functions and the sample time function. $f(2^n, 2^n, k_0, k_1)$ is plotted in Fig. 4 for

$$k_0 \{w(t)\} = k_0 \{\text{cosine arch}\} = 8/3$$

and

$$k_1 = 1/9 ,$$

which is a correct k_1 for a normally distributed time function with the cutoff point out three standard deviations.

If the general shape of the spectrum is known, then the ratio between the peak of the spectrum and its average can be determined and this estimate of the dynamic range can be increased by that amount. As it stands, it is con-

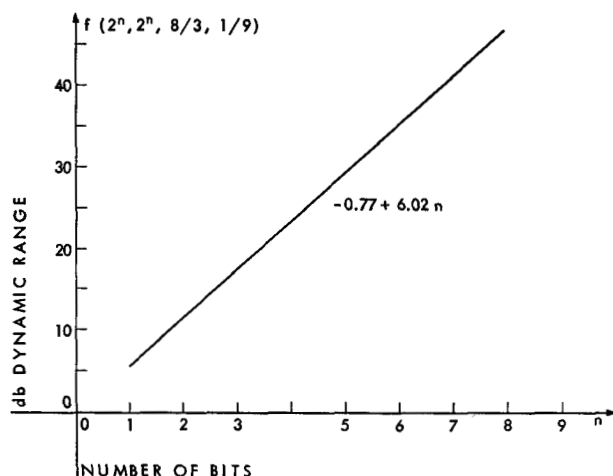


Figure 4

servative, a good characteristic for a planning procedure to possess.

This average level, due to quantization, was checked empirically on sinusoidal and speech inputs and found to describe accurately the lowest levels of the spectra. Also the fluctuations of these lowest levels, with a fixed input signal, convinced us that spectral estimates were essentially irretrievable if they were significantly smaller.

6. Some comparisons of number of multiplications required for direct and indirect methods

In Section 2 we stated that the direct approach in certain situations had considerable computational advantages over the indirect approach. In this section we will draw a comparison of the number of multiplications required in both cases for two particular sets of estimation requirements. In the discussion below we are assuming $P(f)$ is constant over the spectral window.

First, however, it should be pointed out that the sample function duration requirements for the two cases are approximately the same. If we let Δf be the half-power width of the spectral window and k be the number of degrees of freedom then (from Blackman and Tukey¹)

$$\text{duration (indirect case)} = (k/2 + 1/3) / \Delta f.$$

If for the direct case we consider the cosine arch weighting function of length T and let ΔT be the sampling period for $P(t; f_0)$ we get

$$\begin{aligned} \text{duration (direct case)} &= T + (k/2 - 1) \Delta T \\ &= T + (k/2 - 1) 2T/3 = T \left(\frac{k+1}{3} \right) \\ &= 2(k/2 + 1/2) / \Delta f. \end{aligned}$$

We are, in this second derivation, assuming that ΔT is selected as suggested by Eq. (4.4) and using the fact that the half width of the spectral window is given by

$$\Delta f(\text{direct case, cosine arch}) = 3/2T. \quad (6.1)$$

In the indirect case letting $W = 1/2\Delta t$ be the frequency range compatible with the sampling period, Δt , for $x(t)$, one obtains (Blackman and Tukey¹)

number of multiplications (indirect case)

$$\begin{aligned} &= \frac{4W^2}{(\Delta f)^2} (k/2 + 1/3) \\ &= 2W^2 k / (\Delta f)^2. \end{aligned} \quad (6.2)$$

In the direct case with estimators $P(f_{0i})$ with corresponding degrees of freedom k_i and spectral window half-widths Δf_i and assuming cosine arch weighting functions, we get using (6.1),

$$\begin{aligned} \text{number of multiplications (direct case)} &= \sum_i k_i T_i / \Delta t \\ &= 3W \sum k_i / \Delta f_i. \end{aligned} \quad (6.3)$$

The first set of estimation requirements we will consider is the one for which the indirect method is ideally suited, namely, constant resolution and accuracy requirements throughout the available frequency range. For this set and the direct case we get

$$\text{number of multiplications (direct case)} = 3W^2 k / (\Delta f)^2.$$

Comparison of this with (6.2) reveals that in this situation the direct case requires half again as many multiplications as the indirect case.

As a second set, for which the direct method is more efficient, in our study of speech the following requirements for spectral estimation were among several specified:

- a) resolution of 50 cps from 75 to 575 cps.
- b) approximate 1/6 octave resolution from 575 cps to 16,000 cps.
- c) constant accuracy over the frequency range.

For this set of requirements we get from the application of Eq. (6.2) with $W = 16,000$ and $\Delta f = 50$

$$\text{no. of multiplications (indirect case)} = 102,000k.$$

Applying equation (6.3) we get

$$\text{no. of multiplications (direct case)} = 16,000k.$$

Thus in this situation the direct case requires less than one-sixth as many multiplications as the indirect. It is true that we get more resolution and accuracy in the indirect case but it is unneeded resolution and accuracy.

One might argue that by processing the sample time series more than once (first at a low sampling rate to obtain estimates in a low frequency range, then at a higher sampling rate to obtain estimates in a higher frequency range, and so on) one could greatly reduce the number of multiplications required in the indirect case. This is true, but by this same technique one could also greatly reduce the number of multiplications required in the direct case.

7. The 704 computer program embodying procedure and its application to speech studies

This direct digital procedure has been embodied in an IBM 704 EDPM program and applied to the study of speech at the IBM Research Center. The input equipment and the application are described in an IBM Research Report by G. L. Shultz.⁵ This report also contains sample spectra produced by the method. The IBM 704 programs were written by G. L. Shultz. The program presently used was written for a 32,000 word core memory machine. It stores the weighting functions within the core memory. An earlier program written for an 8,000 word core memory machine stores the weighting functions on tape. Both programs perform the multiplications using a table look-up procedure at approximately ten times the speed of ordinary machine multiplication.

The following formula gives the time for the 32,000 word core memory machine to obtain a set of spectral estimates:

$$\text{computing time} = \frac{3W}{2000} \sum_i \frac{n_i}{\Delta_i f} \text{ seconds,}$$

where

$\Delta_i(f)$ = half width of i^{th} estimator

n_i = no. of samples used in i^{th} estimator.

We are assuming that W is related to the sampling period, Δt , for $x(t)$ by $W = (1/2\Delta t)$; and that cosine arch weighting functions are used.

Appendix A: Construction of an estimator with given location and resolution

Let us consider a stationary random process with a sample function $x(t)$ and a two-sided power (or variance) spectrum $P(f)$. We want to estimate the one-sided power spectrum

$$P_+(f) = 2P(f) \quad \text{for } f \geq 0 \\ = 0 \quad \text{otherwise,}$$

of the process in the region about f_0 through a digital approximation to the following random function

$$\left| \int_{-\infty}^{\infty} x(t-\tau) w(\tau) e^{2\pi j f_0 \tau} d\tau \right|^2 = \\ \left\{ \int_{-\infty}^{\infty} x(t-\tau) w(\tau) \cos 2\pi f_0 \tau d\tau \right\}^2 \\ + \left\{ \int_{-\infty}^{\infty} x(t-\tau) w(\tau) \sin 2\pi f_0 \tau d\tau \right\}^2. \quad (\text{A.1})$$

This digital approximation we will call $P(t; f_0)$. In this Appendix we will determine $E\{P(t; f_0)\}$ under conditions of interest for spectral estimation.

Let $w(t)$ be an even function and let it have a Fourier transform $A(f)$. Then $A(f)$ will be real and even. We

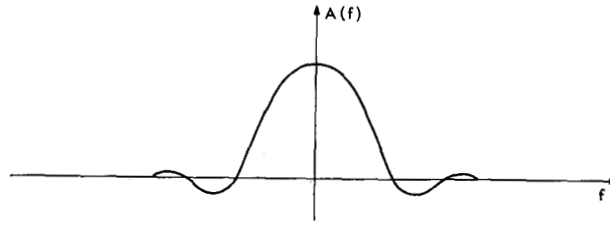


Figure A.1

want to consider functions $w(t)$ such that $A(f)$ is a typical spectral window (in the sense of Blackman and Tukey¹); that is, we want $A(f)$ concentrated with a predominant main lobe about $f=0$ as indicated in Fig. A.1.

We will discuss the first term on the right-hand side of (A.1) and carry this term to the sampled case before considering the complete expression. Hence, consider $w(t) \cos 2\pi f_0 t$. It has a Fourier transform $A(f-f_0)/2 + A(f+f_0)/2$; i.e., it has a frequency composition centered about f_0 whose shape is determined by $A(f)$. Now consider

$$\bar{u}(t) = \int_{-\infty}^{\infty} x(t-\tau) w(\tau) \cos 2\pi f_0 \tau d\tau.$$

The power spectrum of $\bar{u}(t)$ is

$$P_{\bar{u}}(f) = \frac{1}{4} |A(f-f_0) + A(f+f_0)|^2 P(f)$$

and

$$E\{\bar{u}^2(t)\} = \int_{-\infty}^{\infty} P_{\bar{u}}(f) df \\ = \int_{-\infty}^{\infty} \frac{1}{4} |A(f-f_0) + A(f+f_0)|^2 P(f) df.$$

Now we will assume $x(t)$ to be limited to frequencies in the band from $-W$ to W , then $x(t)$ is determined by its samples taken every $1/2W$ in time and⁶

$$x(t) = \sum_{n=-\infty}^{\infty} x(n/2W) \text{sinc } 2W(t-n/2W).$$

Consider

$$b(t) = \sum_{n=-\infty}^{\infty} w(n/2W) \cos 2\pi f_0 n/2W \text{sinc } 2W(t-n/2W).$$

It has a transform⁷

$$B(f) = \frac{1}{2} \sum_{n=-\infty}^{\infty} \{A(f-f_0-n2W) \\ + A(f+f_0-n2W)\} \quad \text{for } |f| \leq W \\ = 0 \quad \text{otherwise.}$$

Finally consider

$$u(t) = \int_{-\infty}^{\infty} x(t-\tau) b(\tau) d\tau.$$

Substituting the above expansions for $x(t-\tau)$ and $b(\tau)$

and using the following property of the sinc function

$$\int_{-\infty}^{\infty} \text{sinc } 2W \left(t - \frac{n}{2W} \right) \text{sinc } 2W \left(t - \frac{m}{2W} \right) dt = \frac{1}{2W} \delta_{mn}$$

we get

$$u(t) = 1/2W \sum_{n=-\infty}^{\infty} x(t-n/2W) b(n/2W) \\ = 1/2W \sum_{n=-\infty}^{\infty} x(t-n/2W) w(n/2W) \cos 2\pi f_0 \frac{n}{2W},$$

which is the sampled approximation to $\tilde{u}(t)$ used in the digital calculations. The spectral density of $u(t)$ is

$$P_u(f) = \frac{1}{2} \left| \sum_{n=-\infty}^{\infty} A(f-f_0-n2W) \right. \\ \left. + A(f+f_0-n2W) \right|^2 P(f) \quad \text{for } |f| \leq W \\ = 0 \quad \text{otherwise,}$$

and

$$E\{u^2(t)\} = \int_{-\infty}^{\infty} P_u(f) df \\ = \int_{-W}^W \frac{1}{2} \left| \sum_{n=-\infty}^{\infty} A(f-f_0-n2W) \right. \\ \left. + A(f+f_0-n2W) \right|^2 P(f) df.$$

By a similar argument

$$v(t) = 1/2W \sum_{n=-\infty}^{\infty} x(t-n/2W) w(n/2W) \sin 2\pi f_0 n/2W$$

has the spectral density⁸

$$P_v(f) = \frac{1}{2} \left| \sum_{n=-\infty}^{\infty} -jA(f-f_0-n2W) \right. \\ \left. + jA(f+f_0-n2W) \right|^2 P(f) \\ = 0 \quad \text{otherwise,}$$

and

$$E\{v^2(t)\} = \int_{-\infty}^{\infty} P_v(f) df \\ = \int_{-W}^W \frac{1}{2} \left| \sum_{n=-\infty}^{\infty} -jA(f-f_0-n2W) \right. \\ \left. + jA(f+f_0-n2W) \right|^2 P(f) df.$$

We now define the random function

$$P(t; f_0) = u^2(t) + v^2(t) \\ = \{1/2W \sum_{m=-M}^M x(t-m/2W) w(m/2W) \\ \cos 2\pi f_0 m/2W\}^2$$

$$+ \{1/2W \sum_{m=-M}^M x(t-m/2W) w(m/2W) \\ \sin 2\pi f_0 m/2W\}^2$$

whose expectation we are going to estimate.

We have restricted the summation from $-M$ to M since in practice we will select functions $w(t)$ of finite nonzero extent.

Since we are interested in spectral estimation, we are interested in functions $A(f)$ whose width is small relative to W . Hence, we can use the approximations

$$\sum_{n=-\infty}^{\infty} A(f-f_0-n2W) = A(f-f_0) + A(f-f_0+2W)$$

and

$$\sum_{n=-\infty}^{\infty} A(f+f_0-n2W) = A(f+f_0) + A(f+f_0-2W).$$

Thus

$$E\{u^2(t)\} = \int_{-W}^W \frac{1}{2} |A(f+f_0-2W) + A(f-f_0) \\ + A(f+f_0) + A(f-f_0+2W)|^2 P(f) df$$

and

$$E\{v^2(t)\} = \int_{-W}^W \frac{1}{2} |jA(f+f_0-2W) - jA(f-f_0) \\ + jA(f+f_0) - jA(f-f_0+2W)|^2 P(f) df.$$

In the sampled case we have a doubling back of $A(f)$ at W and $-W$ in addition to the usual unsampled doubling back at zero.

We have in mind estimating the one-sided spectrum in the region about the point f_0 by estimating

$$E\{P(t; f_0)\} = E\{u^2(t) + v^2(t)\} = E\{u^2(t)\} + E\{v^2(t)\}.$$

We will examine three cases which are sufficient to enable us to design a set of region estimates with expectations covering the entire spectrum.

• Case 1

We first consider the case where f_0 is such that the following approximations can be made.

$$A(f+f_0-2W) = 0 \quad |f| \leq W$$

$$A(f-f_0+2W) = 0 \quad |f| \leq W$$

$$A(f-f_0) = 0 \quad f \leq 0$$

$$A(f+f_0) = 0 \quad f \geq 0$$

Under these conditions

$$E\{P(t; f_0)\} = E\{u^2(t) + v^2(t)\} \\ = \frac{1}{2} \int_{-W}^W \{|A(f-f_0)|^2 + |A(f+f_0)|^2\} P(f) df. \quad (\text{A.2})$$

• Case 2

Secondly consider the case where $f_0=0$, then with the approximations

$$A(f-2W)=A(f+2W)=0 \quad \text{for} \quad |f| \leq W$$

we get

$$E\{P(t; 0)\} = E\{u^2(t)\} = \int_{-W}^W |A(f)|^2 P(f) df. \quad (\text{A.3})$$

• Case 3

Lastly consider the case where $f_0=W$, then

$$\begin{aligned} E\{P(t; W)\} &= E\{u^2(t)\} \\ &= \int_{-W}^W \{|A(f-W)|^2 + |A(f+W)|^2\} P(f) df. \end{aligned}$$

Now in each of these we have an expression of the form

$$\int_{-W}^W Q(f; f_0) P(f) df$$

and to estimate $P_+(f)$ at the point f_0 we want $Q(f; f_0)$ to be such that

$$\int_{-W}^W Q(f; f_0) df = 2. \quad (\text{A.4})$$

This is 2 rather than 1 because $P_+(f) = 2P(f)$. If we apply the restriction (A.4) to the above three cases, and use Parseval's theorem, we get the following common requirement on $w(t)$,

$$\int_{-\infty}^{\infty} w^2(t) dt = 2.$$

Appendix B: The variance spectrum of $P(t; f_0)$, the Gaussian case

In this Appendix we will derive the variance or power spectrum of $P(t; f_0)$ for the three cases considered in Appendix A and Section 2 assuming our random process to be Gaussian. We have two functions

$$u(t) = \int_{-\infty}^{\infty} x(t-\tau) b(\tau) d\tau$$

$$v(t) = \int_{-\infty}^{\infty} x(t-\tau) c(\tau) d\tau,$$

where $b(t)$ has the Fourier transform

$$\begin{aligned} \Phi_u(f) &= \frac{1}{2} [A(f-f_0+2W) + A(f+f_0) \\ &\quad + A(f-f_0) + A(f-2W+f_0)] \quad \text{for} \quad |f| \leq W \\ &= 0 \quad \text{otherwise,} \end{aligned}$$

and $c(t)$ the transform

$$\begin{aligned} \Phi_v(f) &= \frac{1}{2} [-jA(f-f_0+2W) + jA(f+f_0) \\ &\quad - jA(f-f_0) + jA(f+f_0-2W)] \quad \text{for} \quad |f| \leq W \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

Now the variance spectrum $P(t; f_0) = u^2(t) + v^2(t)$ is the Fourier transform of

$$\begin{aligned} E\{(u^2(t) + v^2(t))(u^2(t+\tau) + v^2(t+\tau))\} \\ - E^2\{u^2(t) + v^2(t)\} \\ = E\{u^2(t)\}E\{u^2(t+\tau)\} + 2E^2\{u(t)u(t+\tau)\} \\ + E\{u^2(t)\}E\{v^2(t+\tau)\} + 2E^2\{u(t)v(t+\tau)\} \\ + E\{v^2(t)\}E\{u^2(t+\tau)\} + 2E^2\{v(t)u(t+\tau)\} \\ + E\{v^2(t)\}E\{v^2(t+\tau)\} + 2E^2\{v(t)v(t+\tau)\} \\ - E^2\{u^2(t) + v^2(t)\} \\ = 2[E^2\{u(t)u(t+\tau)\} + E^2\{u(t)v(t+\tau)\} \\ + E^2\{v(t)u(t+\tau)\} + E^2\{v(t)v(t+\tau)\}]. \end{aligned} \quad (\text{B.1})$$

This equality is established using

$$\begin{aligned} E\{wxyz\} &= E\{wx\}E\{yz\} + E\{wy\}E\{xz\} \\ &\quad + E\{wz\}E\{xy\}, \end{aligned}$$

a relation holding for normally distributed variables.

Taking the transform of (B.1) we obtain

$$\begin{aligned} \text{Var}_{P(t; f_0)}(f) &= 2[|\Phi_u(f)|^2 P(f) * |\Phi_u(f)|^2 P(f) \\ &\quad + \Phi_u(f) \widetilde{\Phi_v(f)} P(f) * \Phi_u(f) \widetilde{\Phi_v(f)} P(f) \\ &\quad + \Phi_v(f) \widetilde{\Phi_u(f)} P(f) * \Phi_v(f) \widetilde{\Phi_u(f)} P(f) \\ &\quad + |\Phi_v(f)|^2 P(f) * |\Phi_v(f)|^2 P(f)]. \end{aligned}$$

Here * indicates convolution and $\widetilde{\Phi(f)}$ is the complex conjugate of $\Phi(f)$. For a method of obtaining these convolved transforms see Davenport and Root,⁹ page 182. We will below consider the three cases treated in Section 2.

• Case 1

For this case we have

$$\Phi_u(f) = \frac{A(f-f_0)}{2} + \frac{A(f+f_0)}{2}$$

and

$$\Phi_v(f) = -j \frac{A(f-f_0)}{2} + j \frac{A(f+f_0)}{2}$$

where we are assuming

$$\begin{aligned} A(f-f_0) &= 0 \quad \text{for} \quad f < 0 \quad \text{and} \quad f > W \\ A(f+f_0) &= 0 \quad \text{for} \quad f > 0 \quad \text{and} \quad f < -W. \end{aligned}$$

Hence we get

$$|\Phi_u(f)|^2 = |\Phi_v(f)|^2 = \frac{|A(f-f_0)|^2}{4} + \frac{|A(f+f_0)|^2}{4}$$

and

$$\begin{aligned}\Phi_u(f)\Phi_v(f) &= -\Phi_v(f)\Phi_u(f) \\ &= -j \frac{|A(f-f_0)|^2}{4} + j \frac{|A(f+f_0)|^2}{4}.\end{aligned}$$

Substituting these expressions in (B.2) after expansion and cancellation we obtain for the variance spectrum of $P(t; f_0)$

$$\begin{aligned}\text{Var}_{P(t; f_0)}(f) &= \frac{1}{2} \int_{-\infty}^{\infty} \{ |A(g-f_0)|^2 |A(f-g+f_0)|^2 \\ &\quad + |A(g+f_0)|^2 |A(f-g-f_0)|^2 \} \\ &\quad P(g)P(f-g) dg \\ &= \int_{-\infty}^{\infty} |A(g-f_0)|^2 P(g) |A(f-g+f_0)|^2 P(f-g) dg.\end{aligned}\quad (\text{B.3})$$

This latter equality follows from the commutivity of the convolution. The variance spectrum of $P(t; f_0)$ is the envelope spectrum (that is, the low frequency portion) of the spectrum after square law detection of the output of a filter with a power transfer function $\frac{|A(f-f_0)|^2}{2}$.

The upper frequency portions, those due to the sum terms, have been removed.¹⁰ This is the reason for using $u^2(t) + v^2(t)$ rather than just $u^2(t)$. With one additional computation we insure the reduction of the variance by a factor of two.

The total variance is given by

$$\begin{aligned}\text{Var}\{P(t; f_0)\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |A(g-f_0)|^2 P(g) \\ &\quad \times |A(f-g+f_0)|^2 P(f-g) dg df.\end{aligned}$$

Letting $f_1 = g$ and $f_2 = f - g$,

we get

$$\begin{aligned}\text{Var}\{P(t; f_0)\} &= \left\{ \int_{-\infty}^{\infty} |A(f_1-f_0)|^2 P(f_1) df_1 \right\} \\ &\quad \times \left\{ \int_{-\infty}^{\infty} |A(f_2+f_0)|^2 P(f_2) df_2 \right\} \\ &= \left\{ \int_{-\infty}^{\infty} |A(f-f_0)|^2 P(f) df \right\}^2.\end{aligned}\quad (\text{B.4})$$

• Case 2

We are considering $f_0 = 0$ and assuming

$$A(f-2W) = A(f+2W) = 0 \quad \text{for } |f| \leq W.$$

We thus have

$$\Phi_u(f) = A(f)$$

$$\Phi_v(f) = 0.$$

Hence substituting in (B.2) we get for the variance spectrum of $P(t; 0)$

$$\text{Var}_{P(t; 0)}(f) = 2 \int_{-\infty}^{\infty} |A(g)|^2 P(g) |A(f-g)|^2 P(f-g) dg \quad (\text{B.5})$$

and for the total variance

$$\text{Var}\{P(t; 0)\} = 2 \left\{ \int_{-\infty}^{\infty} |A(f)|^2 P(f) df \right\}^2. \quad (\text{B.6})$$

• Case 3

Now let $f_0 = W$, hence

$$\Phi_u^W(f) = A(f-W) + A(f+W) \quad \text{for } |f| < W$$

$$= 0 \quad \text{otherwise}$$

and $\Phi_v^W(f) = 0$.

Substituting in (B.2) we get for the variance spectrum of $P(t; W)$

$$\text{Var}_{P(t; W)}(f) = 2 \int_{-\infty}^{\infty} |\Phi_u^W(g)|^2 P(g) |\Phi_u^W(f-g)|^2 P(f-g) dg \quad (\text{B.7})$$

and integrating we get the total variance as

$$\text{Var}\{P(t; W)\} = 2 \int_{-W}^W \{ |A(f-W)|^2 + |A(f+W)|^2 \} P(f) df. \quad (\text{B.8})$$

In this last case the high frequency content remains and $\text{Var}_{P(t; W)}(f)$ contains content around $\pm 2W$ as well as around zero.

Appendix C: The variance of $\bar{P}(f_{0i})$

In this Appendix, we are going to derive the variance of

$$\bar{P}(f_{0i}) = 1/n_i \sum_{k=1}^{n_i} P(k\Delta_i T; f_{0i}),$$

our estimator of $E\{P(t; f_{0i})\}$. We will first consider the problem in general terms and then obtain the required particular expressions.

Suppose we have a stationary random process with a variance spectrum $P_s(f)$ and a sample function $s(t)$. Further suppose that $s(t)$ is sampled with a sampling period ΔT . We are going to derive the variance of

$$\bar{s} = 1/n \sum_{i=1}^n s(i\Delta T)$$

in terms of n , ΔT and $P_s(f)$.

A random process with the same total variance as $\bar{s}$ can be generated from $s(t)$ by the following two operations:

1) A sampler produces the function $\bar{s}(t)$ defined by

$$\bar{s}(t) = \sum_{i=-\infty}^{\infty} s(t) \delta(t + i\Delta T + \theta) \quad \text{where } \theta \text{ has the uniform}$$

distribution¹¹ from 0 to ΔT .

as, of course, it should since a constant spectrum means uncorrelated samples.

Applying (C.1) to

$$\bar{P}(f_{0i}) = 1/n_i \sum_{k=1}^{n_i} P(k\Delta_i T; f_{0i})$$

we get

$$\text{Var}\{P(f_{0i})\} = \int_{-\infty}^{\infty} \frac{\sin^2(\pi n_i \Delta_i T f)}{(\pi n_i \Delta_i T f)^2} \cdot \sum_{j=-\infty}^{\infty} \text{Var}_{P(t; f_{0i})}(f - j/\Delta_i T) df. \quad (\text{C.3})$$

Further if we define a normalized variance spectrum as

$$N. \text{Var}_{P(t; f_{0i})}(f) = \text{Var}_{P(t; f_{0i})}(f) / \text{Var}\{P(t; f_{0i})\}$$

then (C.3) becomes

$$\text{Var}\{P(f_{0i})\} = \text{Var}\{P(t; f_{0i})\} \int_{-\infty}^{\infty} \frac{\sin^2(\pi n_i \Delta_i T f)}{(\pi n_i \Delta_i T f)^2} \cdot \sum_{j=-\infty}^{\infty} N. \text{Var}_{P(t; f_{0i})}(f - j/\Delta_i T) df. \quad (\text{C.4})$$

Finally defining a doubled back $N. \text{Var}_{P(t; f_{0i})}^*(f)$, as we did previously a doubled back $P_i^*(f)$, (C.4) becomes

$$\text{Var}\{P(f_{0i})\} = \text{Var}\{P(t; f_{0i})\} \int_{-\infty}^{\infty} \frac{\sin^2(\pi n_i \Delta_i T f)}{(\pi n_i \Delta_i T f)^2} \cdot \sum_{j=-\infty}^{\infty} N. \text{Var}_{P(t; f_{0i})}^*(f - j/\Delta_i T) df. \quad (\text{C.5})$$

Appendix D: The sampling and estimation problem for a periodic input signal

We will in this Appendix analyze the sampling problem for Case 1 of Section 2 assuming the sample function, $x(t)$, to be a sum of sinusoids. The work will parallel Appendices B and C. Consider

$$x(t) = \sum_{i=-n}^n (D_i/2) e^{j\phi_i} e^{2\pi j f_i t}$$

where

$$D_i = D_{-i}, \quad \phi_i = -\phi_{-i} \quad \text{and} \quad f_{-i} = -f_i.$$

Again as in Appendix A we define

$$\begin{aligned} u(t) &= \int_{-\infty}^{\infty} x(t-\tau) b(\tau) d\tau \\ &= \frac{1}{2} \sum_{i=1}^n \{A(f_i - f_0) D_i e^{j\phi_i} e^{2\pi j f_i t} \\ &\quad + A(-f_i + f_0) D_i e^{-j\phi_i} e^{-2\pi j f_i t}\}. \end{aligned}$$

Expressing the expansion in terms of cosines and phase angles and squaring we obtain

$$u^2(t) = \frac{1}{4} \left\{ \sum_{i=1}^n A(f_i - f_0) D_i \cos(2\pi f_i t + \phi_i) \right\}^2.$$

Similarly

$$\begin{aligned} v(t) &= \int_{-\infty}^{\infty} x(t-\tau) c(\tau) d\tau \\ &= \frac{1}{2} \sum_{i=1}^n \{A(f_i - f_0) D_i e^{j(\phi_i - \pi/2)} e^{2\pi j f_i t} \\ &\quad + A(-f_i + f_0) D_i e^{-j(\phi_i - \pi/2)} e^{-2\pi j f_i t}\} \end{aligned}$$

and

$$v^2(t) = \frac{1}{4} \left\{ \sum_{i=1}^n A(f_i - f_0) D_i \cos(2\pi f_i t + \phi_i - \pi/2) \right\}^2.$$

Adding $u^2(t)$ and $v^2(t)$ and performing some trigonometric manipulation we get

$$\begin{aligned} P(t; f_0) - E\{P(t; f_0)\} &= u^2(t) + v^2(t) - E\{u^2(t) + v^2(t)\} \\ &= \frac{1}{4} \sum_{\substack{i < j \\ i > 0}} A(f_i - f_0) D_i A(f_j - f_0) D_j \\ &\quad \cos[2\pi(f_i - f_j)t + \phi_i - \phi_j]. \quad (\text{D.1}) \end{aligned}$$

Thus again here as in the case of Gaussian noise the difference terms are the only ones present, the others have been cancelled out in the summation of $u^2(t)$ and $v^2(t)$.

Now if we have a periodic signal of fundamental frequency, ν , then (D.1) will yield a number of terms of frequency ν , a number of terms of frequency 2ν , etc.; over a range corresponding to the width of the interval on which $A(f)$ is not negligibly small. The total contribution to the power "spectrum" of $P(t; f_0)$ at the points $\nu, 2\nu, \dots$, will depend upon the relative phases of these terms. Hence as much as can be said in general is that the power "spectrum" of $u^2(t) + v^2(t)$ when $x(t)$ is periodic with fundamental frequency will have spikes at the points $\nu, 2\nu, 3\nu, \dots$ the value of the spikes depending on the phasing of the harmonics but being forced to zero as the frequency increases by the character of $A(f)$.

We will now treat the problem of sampling and averaging $P(t; f_0)$ to obtain a good estimate of $E\{P(t; f_0)\}$. If the fundamental frequency is such that there is only one harmonic in the pass band; that is, if the fundamental is larger than the interval over which $A(f)$ is non-negligible, then $P(t; f_0)$ will not vary and only one sample will be required. The value of $P(t; f_0)$ will not be the power of this harmonic but the product of this power and $\frac{|A(f-f_0)|^2}{2}$ at the position of the harmonic.

If more than one harmonic lies in the pass band then $P(t; f_0)$ will vary and this variance can be reduced by sampling and averaging. The spectrum of the sampled $P(t; f_0)$ has the same repeated form described for sampled noise spectra in Appendix C; that is, if $\text{Var}_{P(t; f_0)}(f)$ is the spectrum of $P(t; f_0)$ then

$$(1/\Delta T)^2 \sum_{j=-\infty}^{\infty} \text{Var}_{P(t; f_0)}(f - j/\Delta T)$$

is the spectrum of $P(t; f_0)$ sampled at a sampling period ΔT . Also, as in Appendix C, averaging over n samples

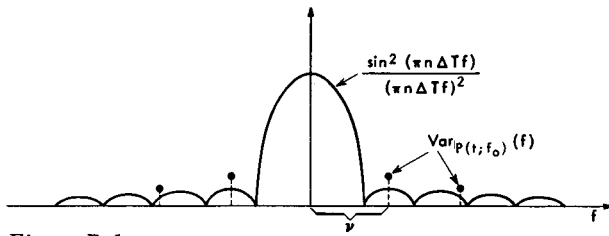


Figure D.1

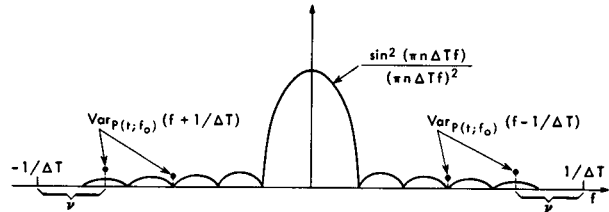


Figure D.2

corresponds to the passing of the sampled $P(t; f_0)$ through a filter whose pass band has the shape

$$(\Delta T)^2 \frac{\sin^2(\pi n \Delta T f)}{(\pi n \Delta T f)^2}.$$

Now in the design of a sampling procedure one must select the variables ΔT and n . In our use of the program under conditions of periodic signal input these were selected as follows. We knew some bounds for the fundamental frequency. The quantity $n\Delta T$ was set large enough so that the response of the averaging filter

$$(\Delta T)^2 \frac{\sin^2(\pi n \Delta T f)}{(\pi n \Delta T f)^2}$$

was down to the desired amount in the region where the fundamental, v , was known to lie. For example, in its first side lobe, $\sin^2 x/x^2$ is down, at a minimum, approximately 13 db, in its second side lobe approximately 17 db. In our experience with periodic speech signals the placing of the fundamental in the first lobe resulted in variations in our estimates of $E\{P(t; f_0)\}$ of about 1 db. This selection of $n\Delta T$ is illustrated in Fig. D.1 where the fundamental has been placed in the first side lobe of the averaging filter. Now with $n\Delta T$ selected, we further select ΔT small enough so that the terms $\text{Var}_{P(t; f_0)}(f + 1/\Delta T)$ and $\text{Var}_{P(t; f_0)}(f - 1/\Delta T)$ in the spectrum of the sampled $P(t; f_0)$ do not have any contributions in the main lobe or lobes of the averaging filter. This means choosing ΔT sufficiently small so that $1/\Delta T$ is larger than the total non-negligible width of $A(f)$. This selection of ΔT is illustrated in Fig. D.2. Thus choosing ΔT , n is determined through the previous choice of $n\Delta T$ and the procedure is defined.

For periodic inputs a normalization with peak spectral window heights of unity and adjacent windows crossed at their half power points was found desirable. With a smoothly varying frequency scale addition of such nor-

malized spectral windows gives an almost flat response of unit height making possible easy estimates of total power in any selected region.

Appendix E: Some additional remarks on quantization of weighting functions and the sample time function

This problem is discussed in Section 5 from the standpoint of spectral dynamic range determination. In this Appendix we will discuss further the particular effects of the two types of quantization, weighting function and sample time function.

The effect of the quantization of the weighting function is to produce instead of the nonrandom spectral window

$$Q(f; f_0) = |Q_{1,u}(f; f_0)|^2 + |Q_{1,v}(f; f_0)|^2$$

the random spectral window

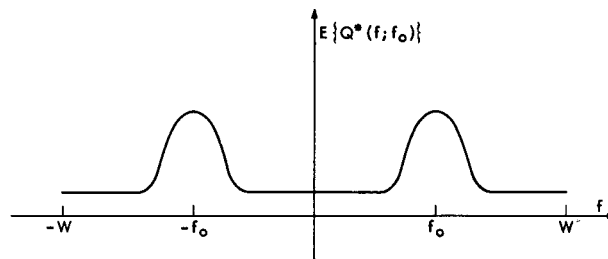
$$\begin{aligned} Q^*(f; f_0) = & \left| Q_{1,u}(f; f_0) + 1/2W \left\{ \sum_{m=-M}^M e_{1m} \cos 2\pi \frac{m}{2W} f \right. \right. \\ & \left. \left. + e_{1m} \sin 2\pi \frac{m}{2W} f \right\} \right|^2 \\ & + \left| Q_{1,v}(f; f_0) + 1/2W \left\{ \sum_{m=-M}^M e_{2m} \cos 2\pi \frac{m}{2W} f \right. \right. \\ & \left. \left. + e_{2m} \sin 2\pi \frac{m}{2W} f \right\} \right|^2 \quad \text{for } |f| < W \\ = & 0 \quad \text{otherwise.} \end{aligned}$$

Here $Q_{1/2,u}(f; f_0)$ and $Q_{1/2,v}(f; f_0)$ are the transforms of the functions $b(t)$ and $c(t)$ of Appendix A while the e_{im} are random variables described in Section 5. If we take the expectation of $Q^*(f; f_0)$ and operate on the result as we did the second term of (5.3) we get

$$E\{Q^*(f; f_0)\} = Q(f; f_0) + (1/2W) k_0 \{w(t)\} / 3N_1^2,$$

where again k_0 and N_1 are constants described in Section 5. Thus with the quantization of the weighting functions the spectral window does not reject more and more strongly as one moves away from the main lobe, but rather there is an average limit to the total rejection attained anywhere in the frequency range. This is illustrated in Fig. E.1. The error introduced by this quantization (i.e., the variance in the estimator $P(t; f_0)$) is not

Figure E.1



reduced by the averaging procedure which produces $P(f_0)$.

The quantization of the sample time function has the same effect as sampling the function

$$x^*(t) = x(t) + e_3(t),$$

where $e_3(t)$ is independent of $x(t)$, has the uniform distribution described in Section 1.5 and a uniform power spectrum from $-W$ to W ($\Delta t = \text{sampling period} = 1/2W$). The power spectrum of $x^*(t)$ is

$$P^*(f) = P(f) + \text{Var}(e_3)/2W.$$

Thus we are effectively looking at a power spectrum which is the one we are interested in plus a uniform noise spectrum. Since $x(t)$ has a much greater range than $e_3(t)$, if $x(t)$ is normally distributed then $x^*(t)$ will be approximately normally distributed. Thus the error in estimating $P^*(f)$ will be given by the discussion of Sections 3 and 4. This sample time function quantization will contribute to the dynamic range limitation described in Section 5 since the confidence limits on the estimation of $P^*(f)$ (and hence on the estimation of $P(f)$ after removal of the bias) have as lower limits, as $P(f)$ approaches zero, those for the estimation of $\text{Var}(e_3)/2W$. Thus the (mean)²-to-variance ratios obtained in Sections 3 and 4 do not apply in regions where the magnitude of $P(f)$ approaches $P^*(f)$. This quantization error, unlike that for the weighting functions, does get smaller with averaging.

References and footnotes

1. R. B. Blackman and J. W. Tukey, "The Measurement of Power Spectra from the Point of View of Communications Engineering—Parts I and II," *Bell System Technical Journal*, January 1958, March 1958; and Dover Publication, 1959.
2. This question of side lobe rejection is discussed further in Appendix E from the standpoint of the quantization of the functions $w_i(t)\cos 2\pi f_0 t$ and $w_i(t)\sin 2\pi f_0 t$.
3. By the *resolution band* of the estimator we mean the region over which $Q(f; f_0)$ is non-negligible.
4. Some discussion of the effects of the two types of quantization error is contained in Appendix E.
5. G. L. Shultz, IBM Research Report RC-37, "The Use of the IBM 704 in the Simulation of Speech Recognition Systems," December, 1957.
6. Here $\text{sinc } x = (\sin \pi x)/\pi x$, as in P. M. Woodward, *Probability and Information Theory, with Applications to Radar*, McGraw-Hill, 1953.
7. Woodward, *op. cit.*, Ch.2.
8. The symbol j is used in this paper both for the imaginary unit and as an indexing variable. In each case the proper interpretation will be evident.
9. W. B. Davenport, Jr. and W. L. Root, *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill, 1958, p. 31.
10. This will happen for any $u(t)$ and $v(t)$ such that $b(t)$ and $c(t)$ are Hilbert transforms of one another.
11. The function $\delta(x)$ is the Dirac delta function.

Revised manuscript received June 22, 1960