Bioacoustics
The International Journal of Animal Sound and its Recording, 1988, Vol. 1, pp. 187-208
0952-4622/88 $10
© 1988 AB Academic Publishers—Printed in Great Britain

SPECTROGRAPHIC ANALYSIS OF ANIMAL VOCALIZATIONS: IMPLICATIONS OF THE "UNCERTAINTY PRINCIPLE"

MICHAEL D. BEECHER

Animal Behavior Program, Department of Psychology, University of Washington, Seattle, Washington 98195, USA

ABSTRACT

Many animal vocalizations are non-periodic, frequency-modulated signals. Because this type of signal varies simultaneously in two dimensions, time and frequency, spectrographic measurement is constrained by the "uncertainty principle": to increase accuracy of measurement in one dimension we must sacrifice accuracy of measurement in the other dimension. Although this trade-off is unavoidable, inherent in the measurement of frequency, for any particular frequency-modulated, non-periodic signal, there is an intermediate, optimal setting of spectrographic bandwidth, equal to the square root of the average rate of change of the measured signal. This optimal bandwidth minimizes the time-frequency smear, and thus permits the most accurate measurement of the instantaneous frequency. Investigators analyzing the microstructure of animal vocal signals therefore should choose their analyzer bandwidths to match the signals under study.

INTRODUCTION

The "uncertainty principle" is a fundamental rule of communication theory with broad implications for the analysis of the vocal signals of animals. The principle states that our ability to measure the frequency of a continuous, periodic event (such as a tone) is constrained by the time interval over which we measure it. That is, precision in the measurement of frequency (Δf) is gained only by measuring over a longer time interval (Δt) and thus is gained at the expense of precision in the measurement of time. This uncertainty relationship is roughly $\Delta f = 1/\Delta t$ (Gabor 1946, Joos 1948, Brillouin 1962, Cherry 1965, Greenewalt 1968).

An intuitive explanation of the uncertainty relationship is given in Figure 1 in which we measure the frequency of a tone by counting positive peaks occurring within the interval Δt. Then if $\Delta t = 10$ msec, we cannot distinguish, for example, 1000 Hz from 1050 Hz, for both give 10 positive peaks in 10 msec. We can just distinguish 1000 Hz and 1100 Hz (10 vs 11 peaks) and in general, $\Delta f = 100$ Hz. If we increase Δt to 100 msec, we can now just distinguish 1000 Hz (100 peaks) from 1010 Hz (101 peaks), and in

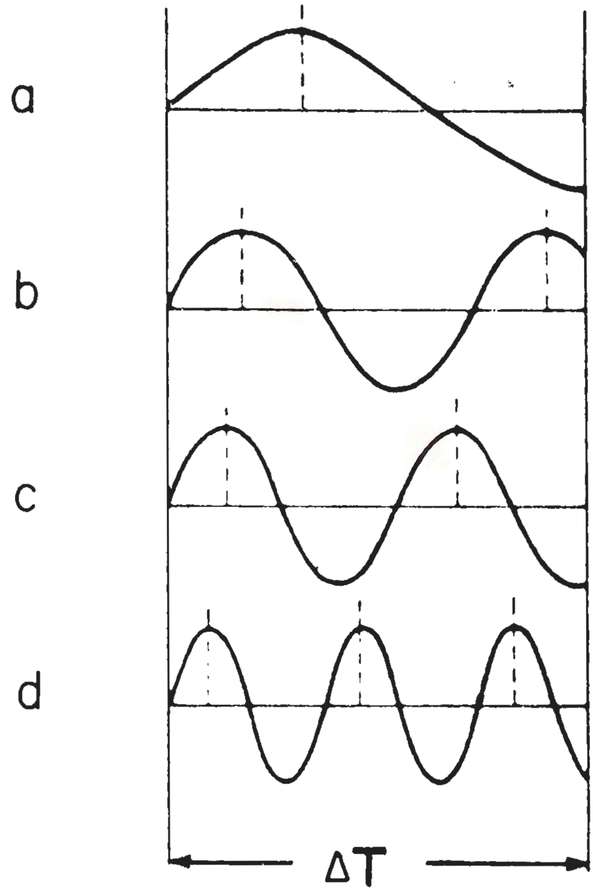

Figure 1. Different frequency sine waves showing the number of positive peaks within a time window Δt (assume Δt begins on a zero crossing). In case **a**, we count 1 peak, hence the frequency is $1/\Delta t$. In cases **b** and **c**, we count 2 peaks, hence the frequency is $2/\Delta t$. In case **d**, we count 3 peaks, and the frequency is $3/\Delta t$. Clearly, we cannot measure frequency to greater accuracy than $1/\Delta t = \Delta f$. Note that we cannot improve our accuracy and distinguish between cases **b** and **c** by extrapolation without assuming that the signal continues unchanged beyond Δt.

general $\Delta f = 10$ Hz. But this ten-fold decrease in frequency uncertainty has been gained at the cost of a ten-fold increase in temporal uncertainty. That is, $\Delta f = 1/\Delta t$.

The sound spectrograph has been the instrument of choice for the analysis of animal vocal signals. The spectrograph consists of a bandpass filter with a variable centre frequency. The signal is analyzed by sweeping the filter through the desired frequency range as a continuous recorded loop of the signal is replayed. The spectrograph can be characterized in terms of the bandwidth (B) of frequencies it passes and its time constant or time window (T). These two characteristics are related by approximately $B = 1/T$. Thus the spectrograph's ability to resolve frequency or time differences is fully subject to the constraint of the uncertainty principle. Most spectrographs have at least a "broad-band" and "narrow-band" setting. For a given bandwidth setting, bandwidth is constant across frequency.

It is widely understood that a narrow band setting is appropriate when frequency information is to be extracted from spectrograms of the signals, and a broad-band setting when temporal information is to be extracted. This rule of thumb, however, is actually correct only in the two extreme cases, where the information in the communication signal is (a) entirely in the time domain or (b) entirely in the frequency domain. Many animal vocalizations, however, are non-periodic, frequency-modulated (FM) signals, in which frequency changes rapidly over time, and the pattern of frequency modulation is not periodic and thus is not predictable. This is pre-eminently true of bird sounds, both calls and song elements. Broad-band clicks and hisses, constant frequency sounds and pure (periodic) FM sounds are less common bird sounds. The purpose of this paper is to outline the special problems encountered in the analysis of non-periodic FM signals by an instrument such as the sound spectrograph which is constrained by the uncertainty principle. While several previous papers have examined the special problems posed by spectrographic analysis (Watkins 1967, Greenewalt 1968, Marler 1969, Staddon et al. 1978, Hall-Craggs 1979), none has dealt with the critical implications of the "uncertainty principle" for spectrographic analysis of non-periodic FM signals.

Figure 2 portrays simple examples of three different classes of communication signal. In case *a* information is encoded entirely in the frequency domain. The signals shown are similar to the pure vowel sounds of speech, which differ primarily in terms of the frequencies of the first three "formants". In this case we distinguish between different signals entirely in terms of frequency information. Thus precision in a spectrogram is obtained by setting B small (narrow). As Figure 2a shows, the uncertainty of our frequency measurement Δf is represented in a spectrogram by the degree of smear in the vertical (frequency) dimension, which is approximately B. In Figure 2b, information is encoded entirely in the time domain. Signals consist of broad-band sound pulses, and one signal differs from another only in the time between pulses. The conventional Morse Code is a familiar system in which information is encoded entirely in the time domain. Here precision in a spectrogram is obtained by setting T small. For signals of this type, the uncertainty of our time measurements Δt is represented in a spectrogram by the degree of smear in the horizontal (time) dimension, and is approximately $T = 1/B$.

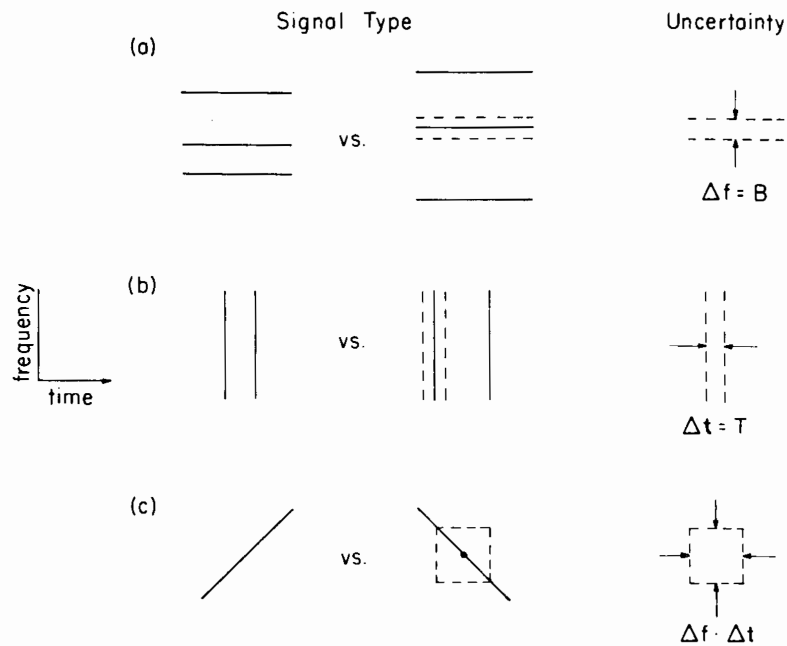

Figure 2. Three types of signals and their associated uncertainties. In each case, two representative signals of the particular type are contrasted on the left, and the nature of the uncertainty is indicated on the right. Note that the space between the dotted lines indicating the uncertainty would be filled by smear on an actual spectrogram. (a) Signal differences entirely in the frequency domain. (b) Signal differences entirely in the time domain. (c) Signal differences in the pattern of instantaneous frequencies. Note that as uncertainty is reduced toward zero in the frequency dimension (case a) or in the temporal dimension (case b), uncertainty approaches infinity in the orthogonal dimension.

In these two extreme cases, then, where we are concerned strictly with one-dimensional quantities (time or frequency), the uncertainty of Δf or Δt (characteristics of our measurements) are equivalent to B or T, respectively (characteristics of our measuring instrument).

Non-periodic FM signals are quite a different case. One example, a simple frequency sweep, is shown in Figure 2c. Here and throughout, the discussion will be confined to *linear* rates of change of frequency over time; this will simplify the presentation with no significant loss of generality. Note that the two FM signals we wish to distinguish are identical with respect to frequency spectrum and duration. They differ only in terms of the instant in the signal when a particular frequency is reached. What we wish to measure accurately, therefore, is *instantaneous* frequency. This is a two-dimensional quantity which can be specified by a time-frequency coordinate, e.g. (t_1,f_1). The spectrograph estimates instantaneous frequencies by doing a running computation of the frequency spectrum of the signal to give us the familiar time-frequency plot (spectrogram or sonogram). While the resulting spectrogram gives us a picture of the time-varying frequency pattern, each time-frequency point (t_1,f_1) is inevitably embedded in a two-dimensional smear, of area $\Delta t \Delta f$, which represents the precision (or uncertainty) of our estimate of that coordinate (Figure 2c). The dimensions of the smear are our uncertainty Δt as to the instant f_1 was present and our uncertanty Δf as to the frequency that was present at the instant t_1.

Clearly, when measuring a non-periodic FM signal, we want to minimize the uncertainty $\Delta t \Delta f$. How does this uncertainty relate to our B/T setting on the spectrograph? Intuitively, it is clear that a time-varying frequency requires some B-T compromise. Consider first the frequency uncertainty Δf. Since frequency is changing from one instant to the next, we must make T short if we are to "capture" the frequency before it changes, but if we set T short we make B large and thereby increase our uncertainty about frequency over the interval. That is, the *effective* bandwidth, Δf, has two components in the case of an FM signal: B, as defined above, and the frequency change of the signal over the interval T. That is, assuming that frequency changes at a constant rate w (Hz/sec), over the interval T,

$$\Delta f = B + wT$$

or, since $T = 1/B$,

$$\Delta f = B + w/B \tag{1}$$

The effective time uncertainty follows directly from the fact that $\Delta f / \Delta t = w$, so

$$\Delta t = \Delta f / w$$

Hence the joint uncertainty is

$$\Delta t \Delta f = (B + w/B)^2 / w \tag{2}$$

A graphical illustration of this argument is given in Figure 3. The function given by Equation 2 is plotted for several different filter bandwidths in Figure 4. Consideration of Equation 2 and Figure 4 leads to the major conclusion of our analysis: a change in B has opposing effects on the two terms in the equation, and there is a minimum in the function at some intermediate value of B. Taking the derivative of the function and setting it equal to zero, we find that $\Delta f \, \Delta t$ is minimal when

$$B^2 = B/T = w$$

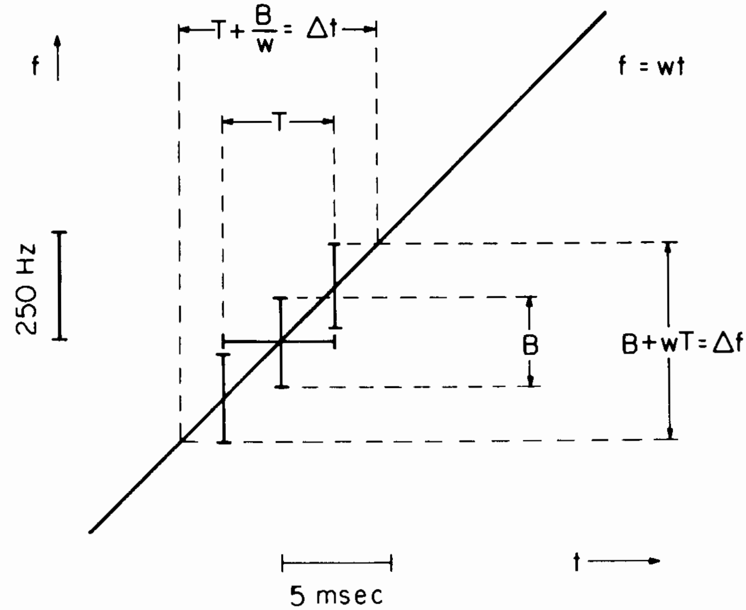

Figure 3. Heuristic derivation of Δf and Δt for a linear frequency sweep. The example diagrammed here is B = 200 Hz, T = 5 msec, and w = 50 Hz/msec. This B/T is near the optimum for this w of 224 Hz/4.47 msec (Equation 2). Scale markers are 250 Hz and 5 msec. The diagram shows that Δf is the sum of B and the frequency range covered in the interval T, i.e. B + wT. Similarly, Δt is T plus the time taken to cover the frequency range B, i.e. T + B/w.

Thus the optimum value of B is

$$B = \sqrt{w} \tag{3}$$

That is, the bandwidth setting of the spectrograph filter should be matched to typical rates of change in the FM signal to be measured. This will be referred to as the "rate-matching prediction". To put the conclusion slightly differently, any given (fixed) filter performs best—in the sense of having relatively low $\Delta f \Delta t$—over a particular range of w, and filters will differ with respect to their best ranges.

To summarize the argument, the conventional directive to use a narrow-band filter to measure frequency and a wide-band filter to measure time is not helpful when it comes to analyzing non-periodic FM signals. For this class of signal we are interested in frequency at a particular point in time, and we want to minimize our error in measuring this "instantaneous frequency". As just shown, this will entail a particular narrow-band/wide-band compromise. The notion that wide-band analysis is always the most appropriate for bird songs or calls is

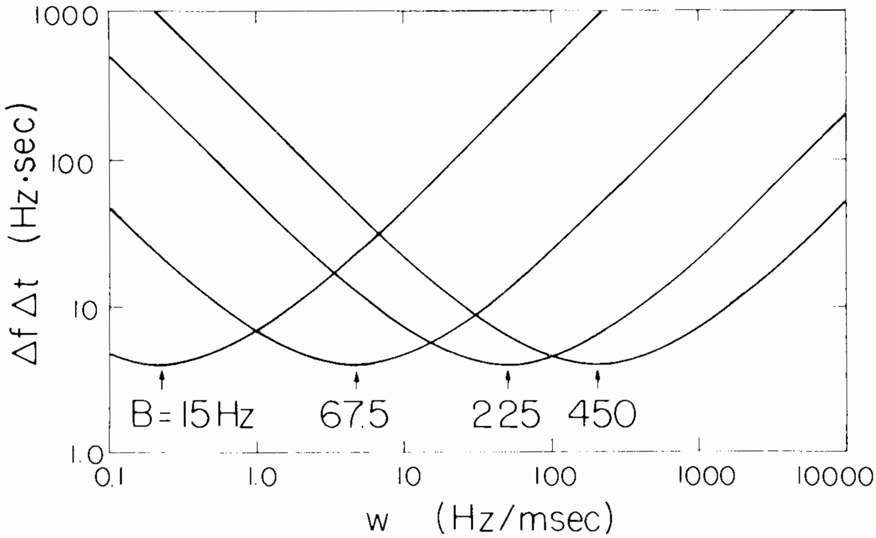

Figure 4. Joint uncertainty ($\Delta f\,\Delta t$) vs rate of frequency change (w), derived from Equation 2, for 4 particular filter bandwidths (B) (log log scale). The filters illustrated correspond to the 4 Kay plug-in filters (labels are "sonagram bandwidths") used for the subsequent measurements in Figures 5-11.

erroneous. The best filter bandwidth will be the one that most closely matches the typical rates of change of these signals; whether this best filter bandwidth is relatively wide-band or narrow-band will be determined by whether the typical rates in question are relatively fast or slow.

METHODS AND RESULTS

Response of spectrograph to linear frequency sweeps

Do the theoretical $\Delta f\,\Delta t$ values given by Equation 2 and plotted in Figure 4 accurately characterize the response of the spectrograph to FM signals? These values were checked empirically using a Kay Sonagraph and 4 plug-in filters, with nominal bandwidths 10, 45, 90 and 300 Hz. The 45 and 300 Hz filters are the familiar "Narrow Band" and "Wide Band" filters of most Kay Sonagraphs. Bear in mind that these bandwidth values apply to the "standard" drum speed of 2.4 sec (8000 Hz frequency range). In the past, some investigators have failed to note that these nominal bandwidth values change in proportion to changes of the drum speed (frequency range) or tape speed (e.g., doubling drum speed or tape speed doubles the bandwidth). The nominal bandwidths of the Kay Sonagraph are measured at the 3 db down, or half-power, points on a spectrum ("section") of a pure tone. On my machine, these values checked out precisely. For actual measurements of "smear", however, we must distinguish between this nominal bandwidth and what will be called the "sonagram bandwidth", determined at the output of the machine, i.e. directly off sonagrams. The sonagram bandwidth is the height of the smear around a pure tone, while the sonagram time window is the width of the smear around a wide-band, punctate stimulus such as a click. On the machine used in the present analysis, the bandwidths so measured were approximately 1.5 times larger than the nominal values, while the measured time windows were approximately 1.5 times smaller than the figure given by the reciprocal of the nominal bandwidth. While the height and width of a spectrogram line for a given signal could be jointly increased or decreased by suitable manipulations of the Gain, Darkness or Automatic Gain controls, the "sonagram" and "nominal" bandwidth values kept the same 1.5 to 1 ratio. Therefore Equation 3 (the rate-matching prediction), when expressed in terms of *nominal* bandwidths, is

$$B = \sqrt{w/1.5} \tag{4}$$

Hereafter reference will be made, as appropriate, to either the sonagram bandwidth or nominal bandwidth designations for the four filters, as well as to the conventional "Wide Band" (WB) or "Narrow Band" (NB) designations. Thus the sonagram bandwidth of the WB-1 filter = 450 Hz (nominal bandwidth = 300 Hz), WB-2 = 225 Hz (150 Hz), NB-1 = 67.5 Hz (45 Hz), and NB-2 = 15 Hz (10 Hz). The numbers 1 or 2 arbitrarily designate the two different Kay plug-in filter combinations (as mentioned above, most Kay Sonagraphs are equipped with the WB-1/NB-1 filter combination). You will note that the four bandwidth values selected as parameters in Figure 4 correspond to the sonagram bandwidths of our four filters.

Two types of constant-rate FM signals were used to estimate $\Delta f \Delta t$ from actual sonagrams. The first was a triangular wave modulated from 3000 to 5000 Hz (range or R = 2000 Hz) at modulation frequencies (f_m) of 0.25, 1, 4 and 16 Hz: this gives rates of 1000, 4000, 16000 and 64000 Hz/sec, respectively ($w = 2Rf_m$). The second signal was a linear frequency sweep from 20 to 20000 Hz at a rate of 70000 Hz/sec (repetition period 285 msec). This signal was tape-recorded and played back at various speeds to give rates of 1094, 4375, 17500, 70000 and 280000 Hz/sec. Strictly speaking, both of these signals are periodic, but in both cases the period is long compared to the time window of the filter, and so the periodicity is immaterial. That is, for all practical purposes the signals are equivalent to non-repeating, linear frequency sweeps. The only exception is the highest rate as seen by the narrowest-band filter: the fastest triangular wave (16 Hz) has a period of 62.5 msec, the fastest sweep a period of 71.2 msec. Since in each case the signal period is shorter than the

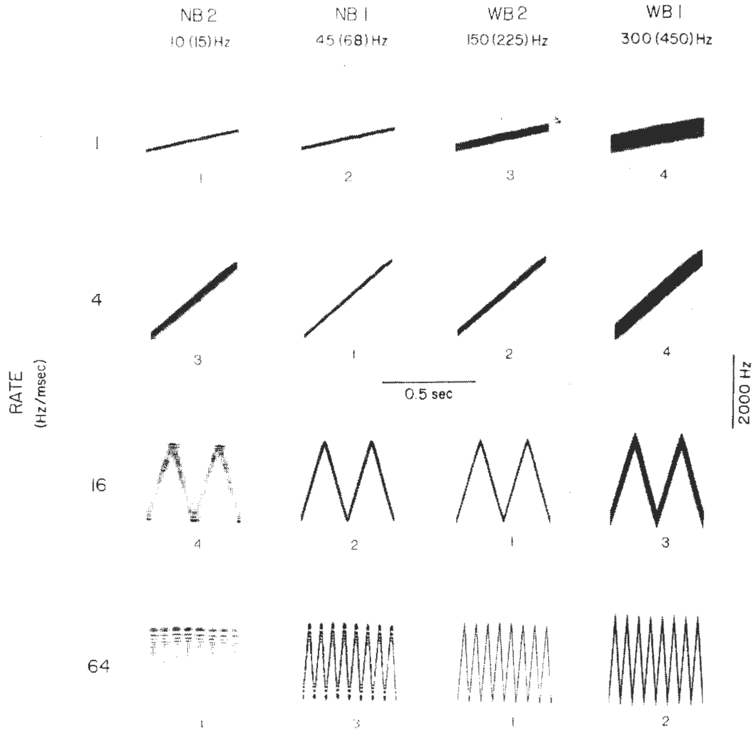

Figure 5. Response of 4 filters (columns) to triangular FM signals of 4 rates (rows). Nominal bandwidths (and sonagram bandwidths) indicated. Predicted $\Delta f\,\Delta t$ ranks (from Table 1) indicated under each sonagram; comparisons are across rows.

100 msec time window of the NB-2 filter, this filter displays the signal partially in the frequency domain (see discussion of the response of the spectrograph to periodic FM signals in the next section).

A total of 36 sonagrams was made for the different filter-rate combinations (4 rates × 4 filters for triangular signals, and 5 rates × 4 filters for sweep signals). Measurements of $\Delta f\,\Delta t$ were made directly from the sonagrams: Δf is the height and Δt the width of the sonagram smear at a given time-frequency point. As mentioned earlier, these dimensions are affected by the particular Gain, Darkness and AGC settings; moreover, the measurements are somewhat subjective. For this reason, the comparison of interest is not between predicted and observed $\Delta f\,\Delta t$ values *per se* but between predicted and observed *ranks* of the $\Delta f\,\Delta t$ values across filters and rates. That is, it would not matter if all observed values were systematically larger than the predicted values, but it would matter if, for example, the lowest $\Delta f\,\Delta t$ did not occur as predicted at about B/T.

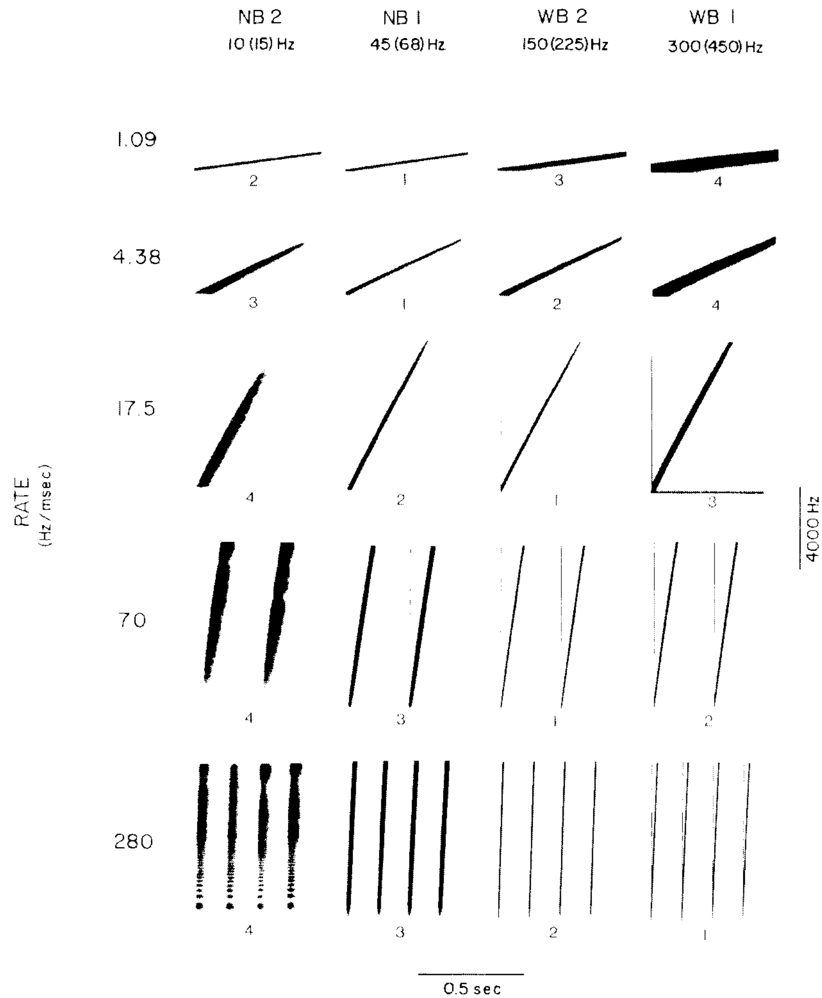

Figure 6. Response of 4 filters (columns) to frequency sweep signals of 5 rates (rows). Nominal (and sonagram) bandwidths indicated. Predicted ranks indicated under each sonagram. Note: Because of a null in the signal at about 1.5, 3, 6, 12 and 24 kHz for the 5 rates respectively (visible in the 3 lowest rates as a tapering off of sonagram width), measurements were taken at a constant point with respect to the null, namely at 1, 2, 4, 8 and 16 kHz, respectively.

The sonagrams are given in Figure 5 (triangular signals) and Figure 6 (sweep signals). Empirical measurements of $\Delta f \Delta t$ from these sonagrams are compared with the predicted values (Equation 2) in Table 1. Consider first our key prediction that for any given rate of change there will be an optimum filter (the precise predictions for these four filters are shown in Figure 4). The *predicted* ranks of the four filters for each given

TABLE 1

Predicted $\Delta f \Delta t$ (Equation 2, sonagram bandwidths) and observed $\Delta f \Delta t$ (Figures 5 & 6) and B (Equation 3, sonagram bandwidths) for filters of bandwidth 15, 62.5, 225 or 450 Hz (sonagram bandwidths) as a function of rate of frequency sweep (w, in Hz/msec).

		Predicted $\Delta f \Delta t$				*Observed* $\Delta f \Delta t$			
w	B	15	62.5	225	450	15	62.5	225	450
1	32	6.7	6.8	52.6	204	16.9	19.2	108	624
4	63	19.8	4.0	14.7	52.6	51.8	8.9	26.9	195
16	126	73.0	5.8	5.5	14.7	138	12.8	8.9	33.2
64	253	286	16.1	4.1	5.5	*	25.0	4.3	9.6
1.09	33	7.1	6.4	48.3	187	8.3	8.3	44.2	224
4.38	66	21.5	4.0	13.6	48.4	27.4	4.3	11.5	67.2
17.5	132	79.8	6.1	5.2	13.7	109	9.6	5.1	17.9
70	264	313	17.4	4.1	5.2	448	36.4	4.5	5.8
280	529	1246	63.4	7.7	4.1	*	88.0	12.0	4.8

*Sonagram trace too diffuse for accurate measurement (see Figures 5 & 6)

rate (these ranks are taken directly from the predicted values in Table 1) are indicated directly under the sonagrams in Figures 5 and 6. Note that the particular rates used cover a range in which each of the four filters is predicted to be optimal (see Figure 4): WB-1 should be best for the 280 Hz/msec sweep, WB-2 slightly better than WB-1 at 70 and 64 Hz/msec, WB-2 slightly better than NB-1 at 17.5 and 16 Hz/msec, NB-1 best at 4.38 and 4 Hz/msec, NB-1 slightly better than NB-2 at 1.09 Hz/msec, and NB-2 slightly better than NB-1 at 1 Hz/msec. It is readily apparent from Figures 5 and 6 and Table 1 that the matching prediction is borne out. For example, at 4 or 4.38 Hz/msec, the NB-1 filter is indeed the best, the WB-1 filter the worst, and so on. For the triangular wave series, all 16 rank predictions are met. For the sweep series, all 20 rank predictions are met, except for the tie between NB-1 and NB-2 at 1.09 Hz/msec. It can be seen from Table 1 that the measured values were systematically higher than the predicted values, with only a few exceptions. Observed and predicted values, however, were highly correlated: product-moment correlations were $r = 0.986$ for the triangular wave series and $r = 0.995$ for the sweep series.

Response of spectrograph to periodic frequency modulated signals

Before turning to spectrographic analysis of bird vocalizations, we must make an important distinction between periodic and non-periodic FM signals. Non-periodic FM must be described in terms of the moment-to-moment pattern of instantaneous frequencies, an accurate representa-

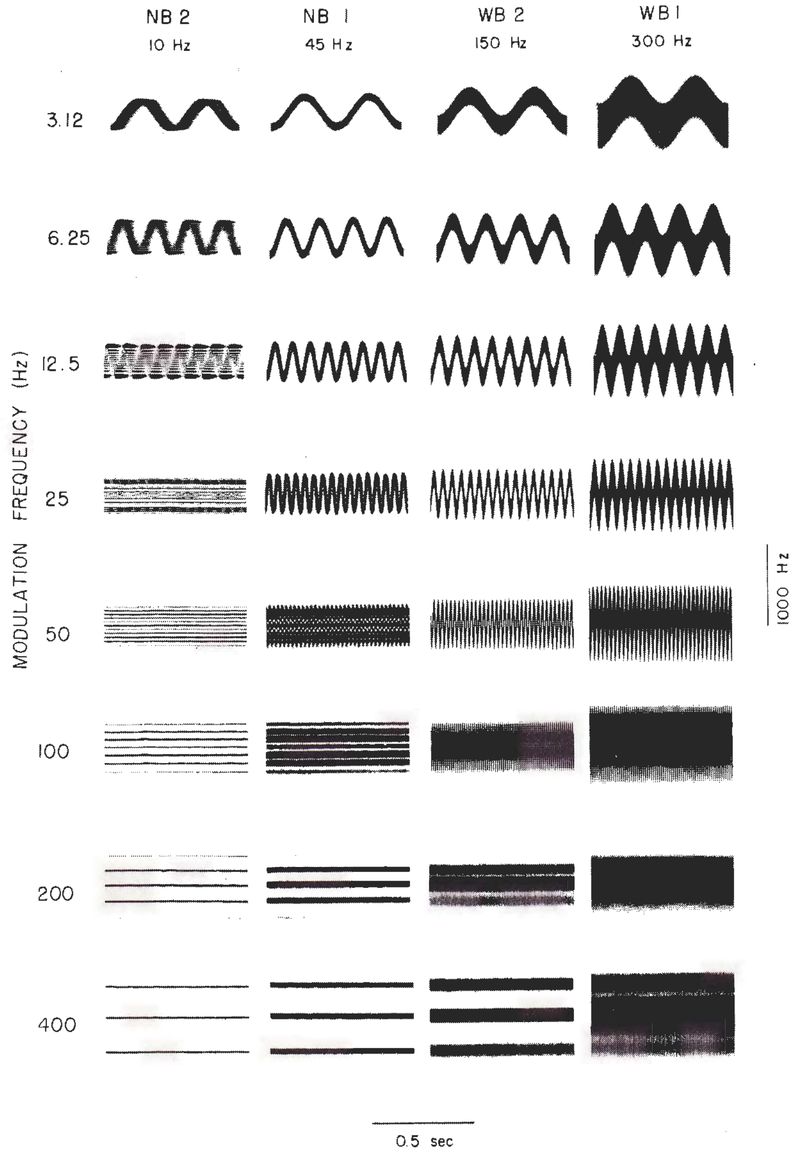

Figure 7. Sonagrams of periodic FM signals for 4 filters (columns) and 8 modulation frequencies (rows). Range = 400 Hz. Instantaneous sweep rates can be calculated from w = 2Rf. Nominal filter bandwidths are indicated, as these predict the point at which the filter switches from time-domain to frequency-domain analysis.

tion of which is critically dependent on an appropriate bandwidth setting, as shown above. Periodic FM, on the other hand, can be characterized completely in terms of the *overall* features: modulation frequency (f_m) and frequency range (maximum minus minimum frequency, R). The optimal filter argument is much less critical for a periodic FM signal, for such a signal can be regarded either as a highly redundant, time-varying frequency, *or* as an unchanging spectrum of frequencies with particular relationships (sideband structure) determined by the modulation frequency (Hund 1942). Thus with periodic FM signals we can set the filter wide, so that T is short compared to the period of the FM signal, and measure the period (and hence f_m) in the time domain, *or* we can set the filter narrow and measure f_m in the frequency domain, i.e., as the separation of side bands. Figure 7 shows a sinusoidal FM signal of varying modulation frequency (and constant range) measured on the different filters. Looking down the columns it can be seen that any particular filter shifts between time-domain analysis and frequency-domain analysis at approximately $f_m = B$ (the nominal B, not the "sonagram" B). Because these periodic FM signals are redundant, f_m can be measured fairly accurately from *any* of the time-domain or frequency-domain displays (although it is clear from Figure 7 that displays near the shift point between time-domain and frequency-domain analysis are difficult to read). Nevertheless, for most precise measurements, there are two optimum settings. If f_m is measured in the *time domain*, the optimum is given by the rate-matching principle (provided B is not too near f_m). (The bandwidth will actually only be critical if the time scale is stretched and measurements must be made over a few cycles, i.e., not the conditions of Figure 7.) The $\Delta f \Delta t$ calculations are not given here as the top part of Figure 7 is essentially redundant with Figure 5: looking across the rows in Figures 5 and 7 give essentially the same picture (the triangular modulation of Figure 5 was chosen for the quantitative analysis because it has a constant rate of frequency change). If f_m is measured in the *frequency-domain*, on the other hand, then the more narrow-band the filter the better (bottom part of Figure 7). These same arguments also apply to R.

In summary, Figure 7 demonstrates that while the filter setting can dramatically affect the appearance of a sonagram of a periodic FM signal, in general it does not strongly affect our ability to extract information about f_m and R, and so this sort of signal stands outside the argument of this paper. A good discussion of spectrographic analysis of periodic FM elements in bird vocalizations can be found in Marler (1969).

Response of spectrograph to bird calls and songs

Non-periodic frequency modulation is common in animal vocalizations

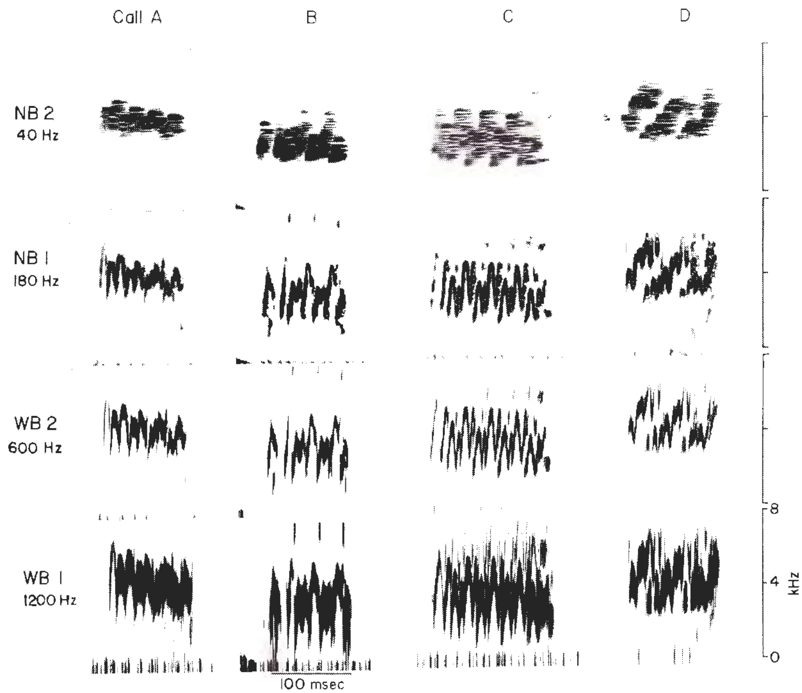

Figure 8. Calls of 4 different Bank Swallows (columns) analyzed by 4 different filters (rows). Nominal bandwidths indicated. Tape at half speed, drum on 16 kHz range.

and is the pre-eminent feature of bird vocalizations. For example, individual Indigo Bunting songs consist of 5 or so figures, taken from a repertoire of 90-odd figure types which differ primarily in their non-periodic time-frequency patterns (Thompson 1970). Thus the ability to distinguish between particular songs ultimately hinges on the ability to resolve these different time-frequency patterns. Similarly, young Bank Swallows are recognized by their parents via their individually distinctive ("signature") calls (Beecher et al. 1981, Beecher 1982). The calls of different individuals differ primarily in terms of their microstructure, specifically the time-frequency patterns of the basic elements of their calls (overall features such as call spectrum and duration are not reliable cues to identity). The smaller elements of the signature calls are typically 10–20 msec in duration, or about an order of magnitude smaller than the figures of the typical song, and the rates of frequency change are an order of magnitude greater than those of typical song figures.

These two examples will be used to illustrate the general problem the frequency modulated signals of birds present for a spectrographic analysis. Consider first the Bank Swallow signature calls. The calls of

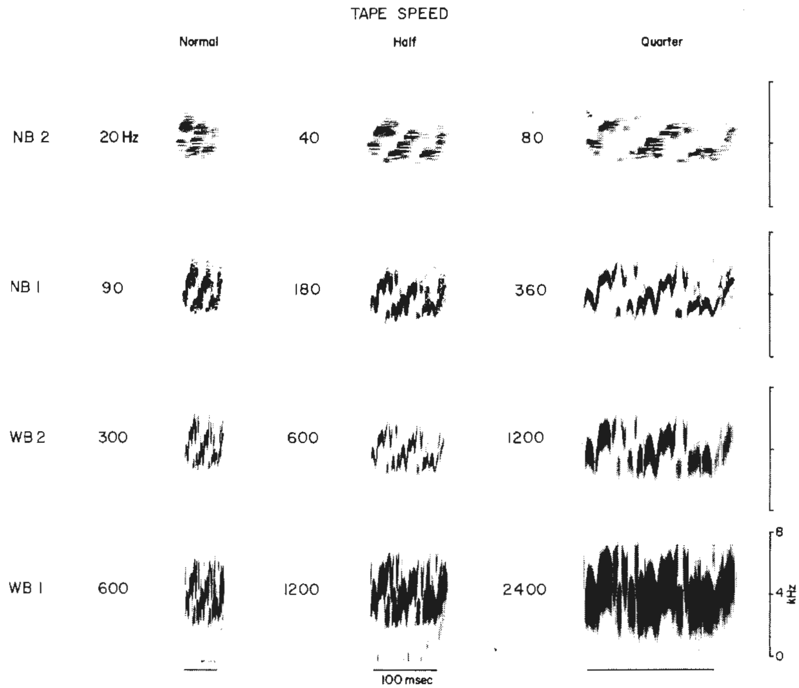

Figure 9. One of the calls from Figure 8 analyzed at 12 different filter values (4 filters × 3 tape speeds). Nominal bandwidths indicated. 16 kHz range.

four individuals are shown in Figure 8. These calls have been sonagraphed on the 16 kHz (1.2 sec) range at half tape speed with the scale magnifier set at 25% (gives 8 kHz as the true upper limit). Thus the true nominal bandwidth of, for example, the NB-1 filter is 4 × 45 = 180 Hz (see earlier discussion of bandwidth determination). I originally chose these settings simply so that the call would fill the typical sonagram sheet (3 × 12 inches). I first tried the standard Kay wide-band filter (WB-1); under these conditions this gives a true nominal bandwidth of 1200 Hz. The result is shown in the bottom row of Figure 8. While the four calls are clearly different, the microstructure is blurred; if many calls sonagraphed at this setting are compared, it is difficult to identify individuals. The most narrow-band filter (NB-2, 40 Hz, top row, Figure 8) is even worse. At the two intermediate bandwidth settings (middle rows, Figure 8), on the other hand, the call microstructure is quite clear, and individual birds can easily be distinguished. Thus the optimum bandwidth appears to be in the 180–600 Hz range. In Figure 9, one of these calls has been analyzed over a wider range of bandwidths, by combining three tape speeds with the four filters (again, reducing tape speed by some factor reduces T, and increases B, by that same factor). It is even more apparent from Figure 9

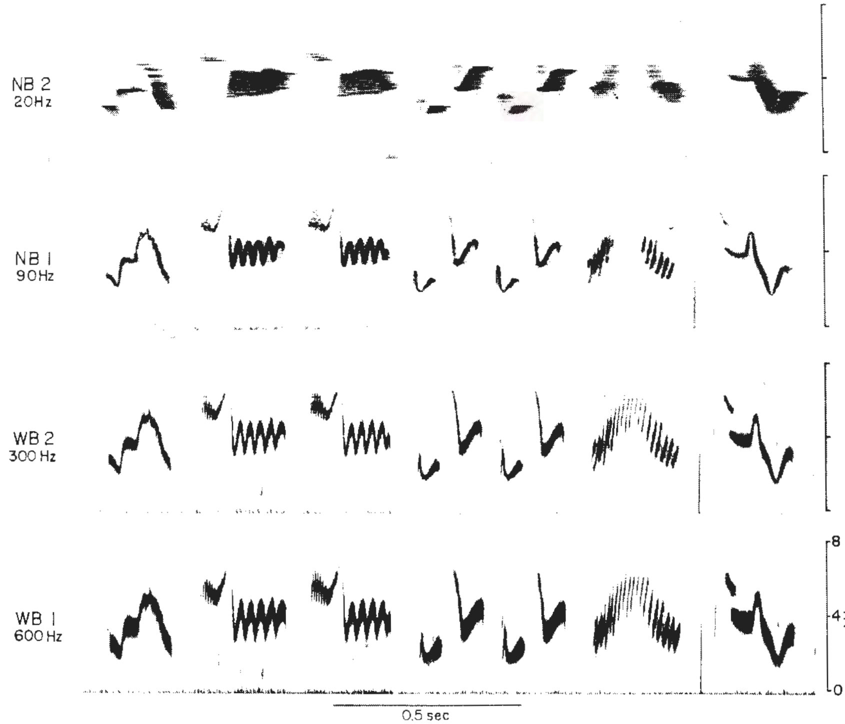

Figure 10. Indigo Bunting song analyzed at half speed on 8 kHz range by the 4 filters. Nominal bandwidths indicated. Five figures are shown (second and third are each repeated). Note: The song is not complete.

out (with a 36 Hz modulation). As noted in the previous section, periodic FM can be analyzed in either the time domain (wide band) or, via sideband separation, in the frequency domain (narrow band), and the filter setting is not critical. Considering the purely non-periodic figures (first, third and fifth), the average rate of change is 47 Hz/msec. By Equations 3 and 4, then, the optimal nominal bandwidth is 144 Hz (optimal sonagram bandwidth 216 Hz). While this is the optimum match, calculations from Equation 2 reveal that the 90 and 300 Hz filters should give good and roughly equivalent resolution, the 600 Hz filter slightly poorer resolution and the 20 Hz filter much the poorest ($\Delta f \Delta t$ values 4.97, 6.54, 19.3 and 54.2 respectively). Figure 10 confirms this prediction in a general way: the 90 and 300 Hz filters are clearly the best, with the 600 Hz filter close behind and the 20 Hz filter clearly unacceptable. Note also that since rates of change vary in these song figures, the wider band filters tend to look better for the faster rates, and the narrower band filters better for the slower rates.

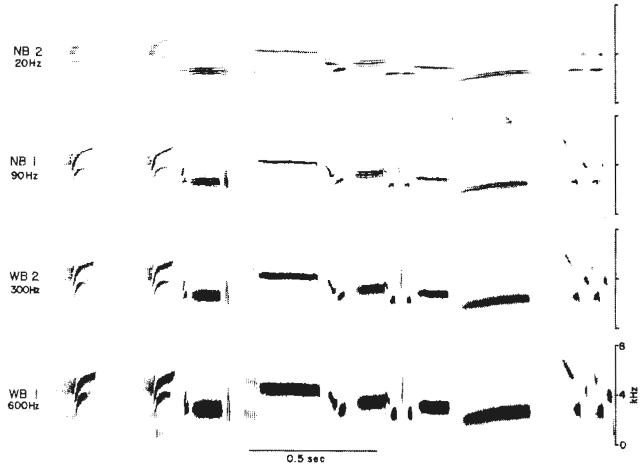

Figure 11. Bewick's Wren song analyzed at half speed on 8 kHz range by the 4 filters. Nominal bandwidths indicated. Most of the figures are constant frequency or periodic FM tones. An exception is the first figure, repeated, which also provides an example of the independent use of the two voices. Note: The song is not quite complete.

Finally, the Bewick's Wren song in Figure 11 is included as an example of a less typical passerine song, containing mostly constant frequency and periodic FM figures. This song is actually atypical even of Bewick's Wrens (see Kroodsma 1974). As explained earlier, the constant frequency figures argue for a narrow-band setting while the FM in the periodic FM figures can be reasonably well represented by either a wide-band (time domain) or narrow-band (frequency domain) record. Since non-periodic figures are rare in this song, it would make sense, for the purposes of measurement, to follow the common practice of taking both a narrow-band record (to measure the centre frequencies of the tonal figures) and a wide-band record (to measure pure temporal features such as the silent intervals between notes). If a single record is required for the sake of representation (e.g., in a publication) other considerations entirely may determine the "optimal" setting here: with the 600 Hz filter, B is close to the modulation range of the periodic FM figures and given the compressed time scale the constant frequency and periodic FM figures are not readily distinguishable, especially in the publication-size reproduction. With the 90 Hz filter, however, the sideband patterns clearly distinguish modulated from constant frequency tones, yet the bandwidth is not so narrow as to blur the few non-periodic FM figures or the segmental time structure of the song, as the 20 Hz filter (T = 50 msec) does. The major point of the Bewick's Wren example, however, is simply to illustrate that the optimality argument of this paper is critical only for non-periodic FM signals. When the vocalizations being analyzed are not of this sort, then other factors will determine the best filter setting.

DISCUSSION

The study of animal vocalizations has flourished despite no great concern over the problem of analyzing filter bandwidths; indeed, until recently, many papers did not even give the filter bandwidth. The study of bird song in particular has not been seriously impeded by this lack of concern, as is clearly illustrated by Kroodsma and Miller's recent two-volume summary of research on bird vocalizations (1982), profusely illustrated with spectrograms. This state of affairs is also reflected in the paucity of discussions of analyzer problems (notable exceptions are Greenewalt 1968, Marler 1969, Staddon et al. 1978, Hall-Craggs 1979). Bird song researchers have been untroubled by the optimal filter question, I suggest, thanks to a happy historical coincidence: the "standard" filter settings of the spectrograph, developed for analysis of the human voice, just happened to match, more or less, the typical rates of change found in the non-periodic FM notes of bird song. Perusal of Greenewalt's (1968) song catalogue shows that these rates of change are usually within the range 10–100 Hz/msec which is best matched by the nominal bandwidth range of 66–210 Hz (Equation 4, which assumes that my Kay Sonagraph is more or less typical of those used in these studies). Most spectrographs have a "narrow-band" setting of roughly 50 Hz and a "wide-band" setting of roughly 300 Hz on the 8 kHz (2.4 sec) scale, and, typically, the 300-Hz setting is used. While in general either of these "standard" settings will be suitable, the analysis does suggest that an intermediate filter bandwidth often would be somewhat better. In fact, in our lab we prefer the 150 Hz filter for song birds such as Indigo Buntings and Song Sparrows (cf. Stoddard et al. 1988). Moreover, the analysis argues against the natural tendency to use a "standard" setting across different species. It suggests, rather, that when we switch between species with relatively slow and relatively fast FM figures, we may want to switch between narrower- and wider-band filters. Rather than "standardizing" one particular filter, we should opt for whichever filter offers the best precision for the particular signal being analyzed. Finally, perusal of Greenewalt's catalogue suggests that the songs of the minority of birds which have very fast FM song figures are poorly displayed with the conventional settings (see also the discussion of bird calls below). While these points are relatively minor when viewed in the context of traditional bird song research, they promise to become more significant with the movement toward more sophisticated, finer-grained analyses. In particular, with the new focus on the micro-structure of vocal signals (e.g. Marler and Pickert 1984, Gouzoules et al. 1984, Clark, Marler and Beeman 1987), increasingly finer measurements are going to require an increasing precision of measurement. Thus the fairly subtle differences between the 90, 300 and 600 Hz filters seen in Figure 10 may become significant when we are trying to place a song figure in a catalogue, to quantify the goodness of fit between an imitation and a model, or to distinguish subtle variants within a vocalization class.

In contrast to the relative success of bird song research, research on bird calls is at a much more primitive level. I suspect that this research has been impeded in part by the failure of the "standard bird" Sonagraph settings to reveal the microstructure of these short vocalizations and thus to reveal the diversity and information capacity of calls. For example, one can begin to correlate call type with behavioural/situational context only if the different call types have been accurately distinguished to begin with.

The measured "bandwidth" of the trace is often used in sonagraphic analyses of bird vocalizations (e.g. Marler and Pickert 1984). It should be apparent from this discussion that it is not generally a valid parameter, in that it is a characteristic of the measuring instrument, not the voice, and is a joint function of the rate of change (a voice variable) and the filter bandwidth (an instrument parameter) as described by Equation 1.

There is an alternative way to analyze FM signals, fundamentally different from the spectrograph, that is free from the constraints of the uncertainty principle. The zero-crossing analyzer (Greenewalt 1968, Staddon et al. 1978) measures the time between successive zero crossings of the signal waveform. This is the instantaneous period of the signal, and its reciprocal is the instantaneous frequency. While this instantaneous frequency has a rather different meaning from that of conventional methods of Fourier analysis, when it is plotted over time we get a "sonagram" very similar to (often better than) that obtained from a spectrograph. The only real limitation of zero-crossing analysis is that it cannot distinguish two different frequencies occurring at the same time without careful pre-filtering. It is relatively common in bird vocalizations for two independent voices to produce rapid FM signals with overlapping frequency ranges (for example see the first figure in the Bewick's Wren song of Figure 11). In such cases there is no way the zero-crossing analyzer, with a fixed filter, can produce a single representative record of the time-frequency pattern, and this record must be reconstructed from several records using different filter cutoffs (see the many examples in Greenewalt 1968, Staddon et al. 1978). Nevertheless, the zero-crossing analyzer largely circumvents the spectrograph problems discussed in this paper and deserves to be more widely used than it is at present.

In recent years we have seen increasing usage of real-time spectrographs based on the FFT spectrum analyzer (first described in Hopkins et al. 1974). This type of spectrograph is fully subject to the constraint of the uncertainty principle. Although it appears that the next generation of real-time (and hybrid) machines will provide flexible control of bandwidth, most of the real-time spectrographs in use at this time are somewhat more limited than conventional Kay-type spectrographs based on the FFT spectrum analyzer (first described in (bins) over the frequency range. On a 8000 Hz range, therefore, the bandwidth is 31.2 Hz (256 lines) or 15.6 Hz (512 lines). Clearly this is too narrow-band for bird vocalizations. Going to the 16 kHz range, or halving the tape speed, doubles the bandwidth. For the 256-line machine, this is 62.5 Hz and is comparable to the narrow-band setting (8 kHz range) on a conventional spectrograph. To get satisfactory precision on a 256-line machine for vocalizations such as those in Figures 8 and 9, however, the analysis must be done at quarter-speed *and* on the 16 kHz scale (giving a 250 Hz bandwidth). Achieving the optimal bandwidth by doubling the frequency range and decreasing the tape speed four-fold has the undesirable effect of compressing the signal into the bottom eighth of the display, and, these machines presently have no scale magnifier to correct for this problem. The message, then, is that new users of these real-time analyzers will discover rather disappointing resolution if they simply regard the machine as analogous to a standard spectrograph (such as a Kay) and use the 8 kHz scale at normal speed. If they attempt to increase bandwidth and precision by manipulating frequency range and tape speed they will run into the problem of display compression. Again we come to an historical explanation, for the FFT spectrum analyzers on which real-time spectrographs are based were designed for narrow-band applications, not for the analysis of rapidly changing bird vocalizations (or human speech).

The solution to the problem, however, may be at hand. The newest generation of real-time analyzers, just becoming available at the time of writing, allow manipulation of bandwidth independently of frequency of time scale. With such a machine, the bioacoustician will be able to select a bandwidth approximating to the optimal bandwidth for the signal being analyzed. Such direct manipulation of bandwidth is clearly preferable to indirect manipulation via tape speed or frequency range, as the expansion or compression of the time or frequency scale is generally not desired by the bioacoustician.

I have focused on bird vocalizations since they are the context in which I came to understand the problems discussed in this paper. Non-periodic FM signals, however, occur in many animal vocal communication systems (e.g., for two recently-studied examples see Leger et al. 1984, and Robinson 1984). I would argue, therefore, that we should not automatically use the traditional bandwidth settings that are the heritage of human speech research (Kay-type machines) or of spectrum analysis (FFT real-time machines), for implicit in this action (or inaction) is the assumption that there are one or two ideal settings that will fit all animal signals, and that these are the one or two available on the machine at our disposal. Rather, we should attempt to "tune" our instruments to the rates of change (and other characteristics) of the animal signals we are studying. We will need to do this if we hope to detect the fine structure of these vocal signals that is likely perceived by the animals themselves.

ACKNOWLEDGEMENTS

My thanks to Philip Stoddard and Patricia Loesche for discussing these ideas with me, poring over sonagrams, and reviewing various drafts of this paper. Steve Crump of Kay Elemetrics provided technical information on the Kay Sonagraph and discussed some of these points with me. This work has been supported by grants from the National Science Foundation.

REFERENCES

Beecher, M. D., Beecher, I. M. & Hahn, S. (1981). Parent-offspring recognition in Bank Swallows (*Riparia riparia*): II: Development and acoustic basis. *Anim. Behav.*, **29**, 95-101.

Beecher, M. D. (1982). Signature systems and kin recognition. *Amer. Zool.*, **22**, 447-490.

Brillouin, L. (1962). *Science and Information Theory.* Acad. Press; New York.

Cherry, C. (1965). *On Human Communication.* M.I.T. Press; Cambridge, Mass.

Clark, C. W., Marler, P. & Beeman, K. (1987). Quantitative analysis of animal vocal phonology: An application to Swamp Sparrow song. *Ethology,* **76**, 101-115.

Gabor, D. (1946). Theory of communication. *J. Inst. Elect. Eng. (London),* **93**, 429-435.

Gouzoules, S., Gouzoules, H. & Marler, P. (1984). Rhesus Monkey (*Macaca mulatta*) screams: representational signalling in the recruitment of agonistic aid. *Anim. Behav.,* **32**, 182-193.

Greenewalt, C. H. (1968). *Bird Song: Acoustics and Physiology.* Smithsonian; Washington, DC.

Hall-Craggs, J. (1979). Sound spectrographic analysis: Suggestions for facilitating auditory imagery. *Condor,* **81**, 185-192.

Hopkins, C. D., Rosetto, M. & Lutjen, A. (1974). A continuous sound spectrum analyzer for animal sounds. *Z. Tierpsychol.,* **34**, 313-320.

Hund, A. (1942). *Frequency modulation.* McGraw-Hill; New York.

Joos, M. (1948). Acoustic phonetics. *Language Suppl.,* **24**, 1-136.

Kroodsma, D. E. (1974). Song learning, dialects, and dispersal in the Bewick's Wren. *Z. Tierpsychol.,* **35**, 352-380.

Kroosdma, D. E. & Miller, E. H. (1982). *Acoustic Communication in Birds.* Academic Press; New York.

Leger, D. W., Berney-Key, S. D. & Sherman, P. W. (1984). Vocalizations of Belding's Ground Squirrels (*Spermophilus beldingi*). *Anim. Behav.,* **32**, 753-764.

Marler, P. (1969). Tonal quality of bird sounds. In *Bird Vocalizations. Their Relations to Current Problems in Biology and Psychology* (Hinde, R. A., ed.). Cambridge University Press; London and New York, pp. 5-18.

Marler, P. & Pickert, R. (1984). Species-universal micro-structure in the learned song of the Swamp Sparrow, *Melospiza georgiana. Anim. Behav.,* **32**, 673-689.

Robinson, J. G. (1984). Syntactic structures in the vocalizations of Wedge-capped Capuchin Monkeys, *Cebus olivaceus. Behav.,* **90**, 46-79.

Staddon, J. E. R., McGeorge, L. W., Bruce, R. A. & Klein, F. F. (1978). A simple method for the rapid analysis of animal sounds. *Z. Tierpsychol.,* **48**, 306-330.

Stoddard, P. K., Beecher, M. D. & Willis, M. S. (1988). Response of territorial male Song Sparrows to song types and variations. *Behav. Ecol. Sociobiol.,* **22**, 125-130.

Thompson, W. L. (1970). Song variation in a population of Indigo Buntings. *Auk,* **87**, 58-71.

Watkins, W. A. (1967). The harmonic interval; fact or artifact in spectral analysis of pulse trains. In *Marine Bioacoustics*, Vol. 2 (Tavolga, W. N., ed.) Pergamon Press; New York, pp. 15-43.