

**Aspects of Signals and Signal Processing
in Echolocation by FM-Bats
for Target Range**

DISSERTATION

zur Erlangung des Grades eines Doktors
der Naturwissenschaften

der Fakultät für Biologie
der Eberhard Karls Universität Tübingen

vorgelegt von

Kristian Beedholm

aus Kopenhagen, Dänemark

2004

Tag der mündlichen Prüfung: 30.01.2004

Dekan

Prof. Dr. H.-U. Schnitzler

1. Berichterstatter

Prof. Dr. H.-U. Schnitzler

2. Berichterstatter

Prof. Bertel Møhl (Univ. Århus, Dänemark)

TABLE OF CONTENTS

GENERAL INTRODUCTION	5
CHAPTER I: Limits to jitter detection accuracy with the signals of echolocating FM bats (<i>Eptesicus fuscus</i>)	27
CHAPTER II: Evidence for an incoherent ranging mechanism in FM-bats	49
CHAPTER III: The “SCAT” receiver model of bat echolocation forms a coherent cross-correlation between input vectors	67
CHAPTER IV: A method for analyzing bat echolocation signals applied to the question of the time-frequency structure of <i>Eptesicus fuscus</i> calls in the laboratory	83
ZUSAMMENFASSUNG	107
ABSTRACT	110
ACKNOWLEDGEMENTS	113
LEBENS LAUF	115

GENERAL INTRODUCTION

Microbats, order Microchiroptera employ an active sensory modality (see Griffin, 1958), termed echolocation by the chief discoverer of the phenomenon, Donald R. Griffin (1944). Since its demonstration in the early 1940ies (e.g. Griffin and Galambos, 1941) several other species of mammals and birds have been shown to possess independently evolved sensory systems using the same principle.

Echolocation in Microchiropteran bats involves the vocal production of generally ultrasonic sound pulses, in the range from tens of kHz to 200 kHz in the extreme. Echolocation vocalizations by most species are relatively short, in the order of a submillisecond to tens of milliseconds, but a few specialized species produce sounds longer than 100 ms. Within each species there is generally a large variability in the range of both signal bandwidth and duration (e.g. Jensen and Miller 2000; Surlykke and Moss, 2000).

This thesis considers theoretical as well as some practical results of signal processing by echolocating bats, especially concerning these animals' determination of the distances to potential targets.

Much of the work presented in the following four chapters implicitly leans on or has consequences for a relatively long history of discussions concerning one hypothesis of why most – if not all – Microchiropteran bat species use a frequency sweep at least as a component of their echolocation calls. Initially some of the theory and data concerning the theory of a cross-correlation (CC) receiver in echolocation will be reviewed. Many of the points made below are found in the four following manuscripts, but often in a “hinted” form or a specific context. Whereas the chapters in the form of manuscripts address an audience that is somewhat familiar with this debate, I have tried to make the following outline accessible to a slightly broader audience.

Although this is a general introduction, also a novel analysis of some of the more historical data and assumptions is presented. This new analysis includes the notion that the receiver output is directly reflected in the performance curve in psychophysical experiments (Simmons, 1971; Simmons, 1973; Simmons, 1979; Simmons and Stein, 1980; Hackbarth, 1986; Menne & Hackbarth, 1986; Simmons and Grinnell, 1988; Simmons, 1990). This critical part of the following rather selective review may be viewed as polemic, and this is in parts intended. I have chosen to include it, since the historical interpretations of the data forms an important and often overlooked bridge to the discussion of more recent data.

I have chosen to exclude this type of analysis of the historical data from the following manuscripts, because it is a controversial type of problem and one, which is not (as will be argued below) actually concerned with relevant data, but rather with a pattern to the interpretation of these data.

As explained above the four following chapters are not concerned with the discussion of the connection between the performance curve and the receiver. I hope to end that part of the question with this introduction.

The notion of pulse compression in bat sonar research

The relationship between echolocation and sonar/radar technologies

I feel it is important to recognize that both sonar and radar principles were relatively well-developed, when echolocation in bats was discovered (Griffin and Galambos, 1941; Griffin, 1958). This means that initially the researchers knew about the principles that animals having this sensory ability were using. Rather, they were to some extent met with the task of trying to understand the degree to which radar and sonar were comparable with the echolocation of bats (e.g. Griffin, 1944).

For instance, when Hartridge (1945b) proposed that bats might use echo delay to judge the distance to a target, he could do so with the reassuring knowledge that this principle had been proven effective in the man-made equivalents.

To that date and beyond the situation is so, that although much funding has been invested in the field of echolocation research from militarily affiliated sources, the results in radar and sonar technologies by far exceed what is found or even proposed in echolocating animals. At least this has been the case with bats. On the other hand, it may very well be that the animals are very advanced in their processing, but that we have a limited ability to recognize “technology” that we do not grasp ourselves.

So most of the modes of echo processing that are suggested to apply to bats (and dolphins) are already implemented in man-made systems. This holds for Doppler compensation (Schnitzler, 1968; Skolnik, 1980), automatic gain control, AGC (Kick and Simmons, 1984; Hartley, 1992a, 1992b), time-delay spectrometry, TDS (Heyser, 1967; Biering and Pedersen, 1983; Pedersen and Miller, 1988), moving target indication, MTI (Altes, 1989, Moss and Simmons, 1993) energy detection in dolphins (Au et al., 1988, Urkowitz, 1967).

A possible exception to this pattern may be an early proposal of the overall nature of the navigational mode of bats – one that precedes Donald Griffin’s actual demonstrations of this phenomenon by at least 25 years. It was made by Hiram Maxim (as reviewed in Griffin, 1958 and Airepet’yants and Konstantinov, 1973). He suggested that bats might use a system like the one he envisioned might prevent accidents like the sinking of The Titanic. His proposed solution was an early form of underwater sonar.

I suggest here that this engineer may have searched for a solution to the engineering question and was able to see that his proposal could also represent an answer to the riddle of the sensory abilities of bats. This seems to me rather more likely, than if in searching for a practical navigational system, he came to think of one in nature that had not yet been demonstrated.

Perhaps Pye embodies the predominant direction of inspiration between technology and bat echolocation research, when he – concerning the formula for the minimum attainable limits to the accuracy in range determination (Burdic, 1968) in radar – writes: “These relations are based upon rigorous theory and must apply to bats. The difficulty arises in interpreting them for a range of bat signals.” (Pye, 1986, p. 167).

It is thus hard to determine to which – if any – extent ideas are flowing in the other direction so that man-made systems are directly improved by discoveries from the research in bat echolocation. Nevertheless, when observing bats catching prey in the air, it appears evident that there might well be some military or other applications associated with a deeper understanding of their sensorimotor operation. Also, radar and sonar engineers may well allow themselves to be inspired by the animal world in ways that are perhaps not revealed by their reference lists. At least the interest in animal sonar from radar engineers has traditionally been high (e.g. Cahlander, 1964; Kroszczyński, 1969; Altes and Titlebaum, 1970; Altes, 1981).

This is also noticeable when Strother (1961) in a note with widespread ramifications points out that the FM sweeps of most bat echolocation calls is also found in so-called chirp radar signals. It has been said that this hypothesis is a fruitful one in terms of papers concerned with it, and if fruitfulness is thus defined this is certainly true.

The chirp technique was developed to meet demands for a high signal to noise ratio (SNR) in peak power limited systems and it is described below.

The rationale and principle of chirp radar

The simplest way to make an effective sonar signal that is easy to localize is to hold it short and make it loud. Odontocete whales, which live in an incompressible medium where sound is easily conveyed over long distances, use a pulse that has a close to minimal duration for the given bandwidth of the signal (Wiersma, 1988; Au, 1990). The shorter a signal is, however, the less energy is contained and if the transmitter has limited peak power, this cannot be made up for by “turning up the volume”. If in a radar system a gated electromagnetic carrier is held very short the total energy will accordingly be low, leading to a limited SNR and consequently a short operational range. Increasing the output power may not always be possible, wherefore a longer signal is called for to increase the overall energy. Now, if the carrier frequency is kept constant this lengthening obviously results in a decreasing bandwidth. One answer to this problem then is to *not*

keep the carrier fixed, but to modulate it, i.e. to create a frequency-modulated (FM) signal, and for many applications a good solution is the above-mentioned chirp (Skolnik, 1980). This signal type covers a larger bandwidth and can at the same time be made to have a high energy with a limited instantaneous power.

For the receiver to fully utilize the bandwidth of the returning echoes it is

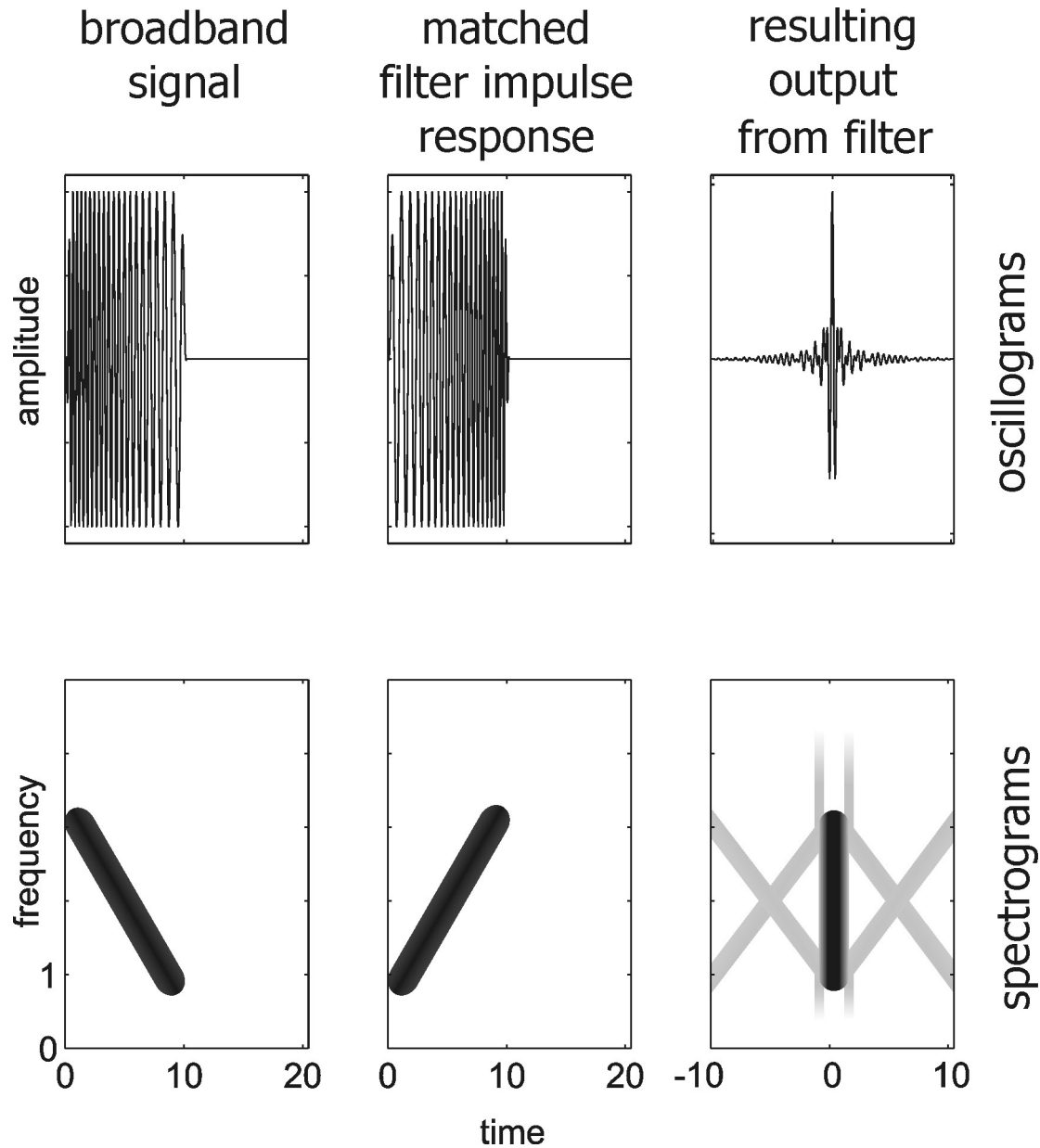


Fig. 0-1: Demonstration of the matched filtering process. Top row: Time series. Bottom row: Stylized spectrogram representation of the signals. Left column: A chirp signal with a bandwidth of more two octaves. Middle column: The impulse response of the filter, which is matched to the signal. Note that it is a time-reversed version of the chirp. Right column: The filter output, which in this case is also the auto-correlation of the input signal. Note also that although the ACF is much more localized in time, the time span in which the filter output is non-zero is the double of the equivalent measure in the input signal or the filter impulse response.

necessary to pass them through a filter that has an impulse response, which is the time-reversed waveform. This is a so-called matched filter (see Fig 0-1). The output from such a filter is the CC between the outgoing sound and the returning echo, and the system is thus also called a cross-correlation receiver (there are differences in the implementation that determine which name is more proper).

In essence, what the matched filter does to a chirp is to “dechirp” it; all the

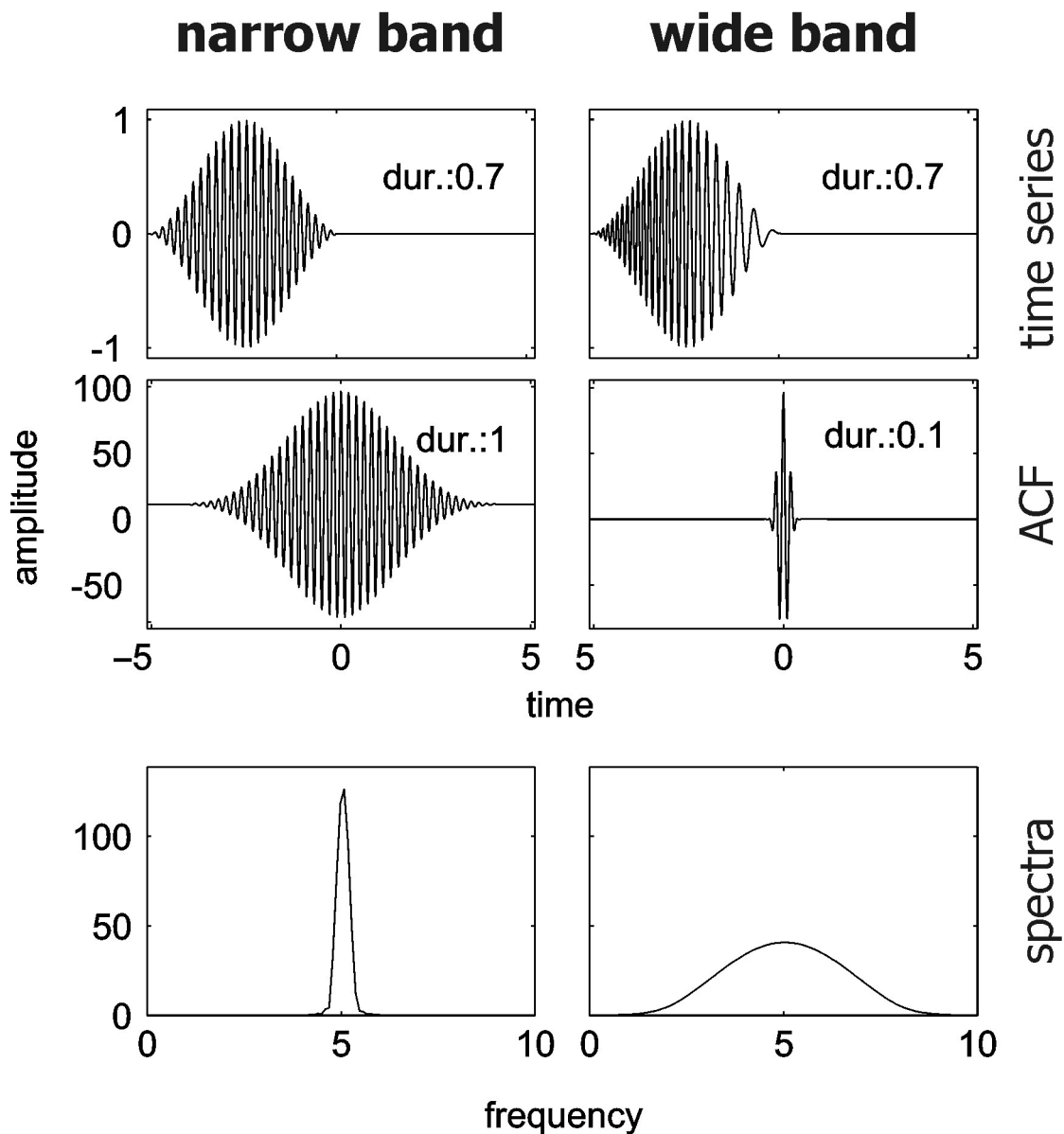


Fig. 0-2: The advantage of a large signal bandwidth in pulse compression for time determination. Top row shows the original signals. Middle row: Autocorrelation functions of the signals. Bottom row: Amplitude spectra with linear ordinate axis. Left column: narrow band signal, a cosine modulated by a raised cosine (Hann) window. Right column: wide band signal, linear frequency sweep multiplied by the same time window as the left column signal. The duration values are the RMS values. The compressed wideband signal is much easier to place accurately in time than the narrowband signal. But the peak amplitudes of the autocorrelation functions are the same, namely the total energy of the signals squared.

frequency components of the undistorted signal are made to occur in phase with each other as demonstrated in Fig. 0-1. For a wide-band signal this leads to pulse compression, but the opposite is the case with a narrow band signal, as seen in Fig. 0-2. The broadband signal becomes more localized in time (i.e. shorter) and can easier be timed, which is often (but not always) the point in having a large bandwidth. It can be shown that a matched filter receiver is the theoretically optimal processing solution for detection and ranging purposes, regardless of the signal used (Woodward, 1955).

In man-made radar and sonar the phase of the returning signal is often considered or expected to be a random factor prone to changes imposed by an unknown target surface structure and signal path, and it is therefore disregarded by forming the envelope of the output from the filter. This case is called a semicoherent cross-correlation receiver. To form the exact envelope, the phase of the returning signal, however, must be available to the receiver, a fact that is often overlooked (a mistake made in e.g. Simmons, 1971, 1973; Altes, 1980; 1981; Hartley, 1989). A matched filter receiver that does not form the envelope around the CC receiver output is a coherent CC receiver.

A short history of pulse compression in bats

As mentioned above the notion of the possibility of pulse compression in bat sonar first appears in print with a note by Strother in 1961, who simply noted that most bats use a frequency modulated sound, so-called FM-bats, which would be suited for an optimal processing mechanism like the ones found in man-made radar. It was followed up by other theoretical studies (e.g. Cahlander, 1964; Cahlander et al., 1964)

10 years later this note sparked a number of psychophysical investigations mainly by James A. Simmons, who in 1971 with a pioneering paper in *Science* showed that at least two species of FM-bats were able to discriminate target distances with an accuracy of ca. 1.2 – 1.3 cm equivalent to 70 – 75 μ s. He was able to successfully compare the bats' performance in his psychophysical experiment with that of the envelope of the autocorrelation function (ACF) of a cry produced by an individual animal of the species in question, when a correction for head movements (considered in detail below) had been applied to this envelope. In later work this finding was extended to several other species of bats, including bats with a constant frequency (CF) component in their calls, like *Rhinolopus ferrumequinum*. (Simmons, 1973, Simmons, 1978).

In the paper from 1973 Simmons also showed that the threshold for determining target delay differences is the same when a simulated target is used instead of a physical object. With the real object the bats must judge the *distance* to a target. Simulated targets are generated by picking up the animal's vocalizations and playing them back to it through a loudspeaker at a controlled *delay*, which is then to be evaluated by the animals. This result is important, since it shows definitely that delay really is the basis for determining distances to objects, thus proving the ranging theory of Hartridge

(1945a, 1945b) to be correct.

Simmons also reported that the absolute delay of the targets was without influence on the thresholds in the range from 30 to 240 cm.

The degree to which J. A. Simmons in these and later works found correspondence between his “theoretical” predictions and the actual psychometric functions is rarely encountered elsewhere in the life sciences. Some aspects concerning this will be discussed below.

In a new experimental paradigm Simons in 1979 showed with a paper to Science that the threshold for detecting *changes* in echo delay from call to call (jitter) of *Eptesicus* was in the order of ca. 2 μ s, although it says 0.5 μ s in the paper¹, and followed approximately not the envelope but the rectified ACF of a call produced by an individual of the species. The conclusion was that the bats retained the phase information and processed the echoes coherently.

This publication came at a time when criticism was rising over the conclusion from the previous work that the bats had and used a CC receiver for ranging. Schnitzler and Henson in a critique published in 1980 along with several other important points reflect on the fact that the animals’ thresholds to differences in target range were seemingly independent of the absolute distance to the target. This is inconsistent with the basic Woodward relation (Woodward, 1955) showing that the delay accuracy is dependant on the SNR, which should again be very predictably dependant on the absolute delay. This criticism has never been properly countered – perhaps in part due to the appearance of the “new data” (i.e. Simmons, 1979) – but it might be speculated that the notion of automatic gain control (AGC) in bats (Kick and Simmons, 1984; Hartley, 1992) was actually indirectly sparked as a result of this argument. If, namely, the perceived echo strength is kept constant, as is claimed by this theory, then it could explain the observation that the absolute distance to the target did not affect the ranging threshold. It should be noted though, that a prerequisite to this interpretation being valid, is that the bats are limited by internal noise.

In 1986 Menne and Hackbarth showed that the μ s jitter threshold of Simmons (1979) was not indicative of CC processing, and same year Hackbarth was able to demonstrate that the rough agreement between the psychometric function and the underlying proposed receiver output does not constitute a valid argument in favor of

¹ In the 1979 paper it is said that the threshold is about 0.5 μ s and in Simmons et al. (1990) the value is given as “less that 0.6 μ s”, but this seems to be a mistake. The first four data points in the performance curve of Fig. 1 of Simmons (1979) are spaced with 1 μ s between them. The third point (2 μ s) lies at 23-24 % errors, whereas the second (1 μ s) is around 33-34 % and the fourth (3 μ s) is about 18 %. Since the threshold is almost always defined as 75 % correct in this type of psychophysical paradigm, the actual threshold value should have read just below 2 μ s.

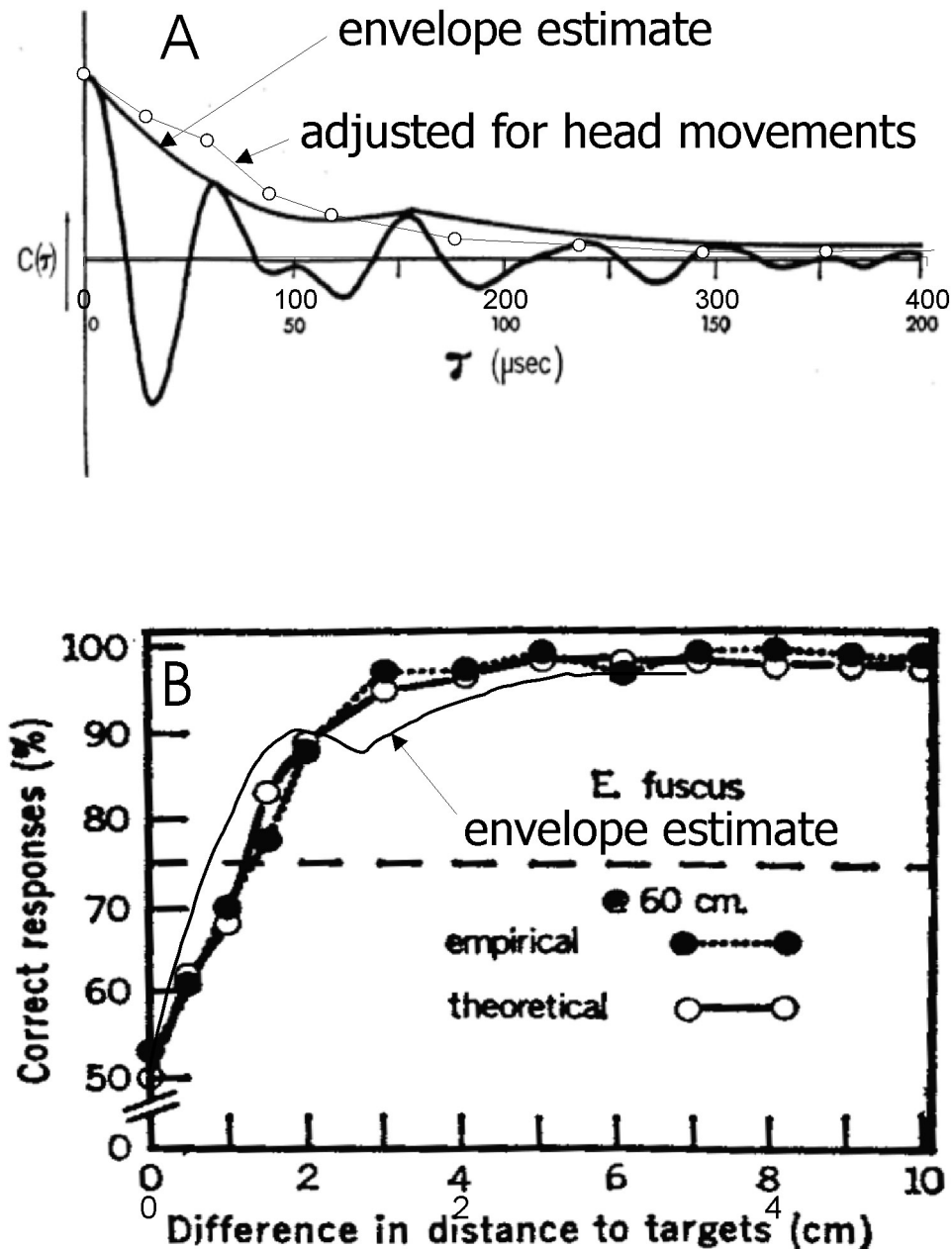


Fig. 0-3: Two graphs modified from Simmons (1971).

Upper graph is modified from Fig. 4 in Simmons (1971). Solid thick lines represent the autocorrelation function (ACF) of an *Eptesicus fuscus* call and its estimated envelope. The connected circles trace is added after grabbing the data from the "theoretical" trace in the lower graph. The upper of the set of X-axis values correspond to the trace added by me.

The lower graph is modified from Fig. 2 in Simmons (1971). It shows the ranging performance of *E. fuscus* ("empirical") and the performance predicted from the envelope of the ACF ("theoretical"). The thin continuous line added here represents the envelope trace from the upper graph translated as described in the text. The lower x-axis values belongs to trace added by me.

The double abscissa in both graphs is due to the translation of the ACF envelope mentioned in the text.

the coherent CC receiver, but might well indicate other receiver configurations. Also, in 1986 in two papers Møhl and coworkers (Troest and Møhl, 1986; Møhl, 1986) showed that *Eptesicus* and *Pipistrellus* did not use optimal reception in a detection task. The experiment by Møhl (1986) has been widely regarded as a falsification of the idea of an optimal receiver in bats.

Two interesting papers appeared in printing in 1989 with basis in this laboratory, namely Menne et al. (1989) and Moss and Schnitzler (1989). Both these works, using similar equipment, failed to find a threshold in a reproduction of the jitter experiment of Simmons (1979). The minimum detected jitter amplitudes, which were limited by the equipment, were 0.4 μ s.

Simmons and co-workers then in 1990 in a much-debated paper (e.g. Pollak, 1993; Simmons, 1993) reported that the actual jitter detection threshold was as low as 10 ns. This threshold, if valid, can only be indicative of coherent processing at the SNRs that can be estimated in the experiments (Altes, the Sandbjerg Meeting).

Although the by far most interesting and revolutionary data is this threshold, equivalent to a difference in distance of 1.7 μ m, most of that paper is devoted to comparisons between the psychometric function and the CC function.

Problems with the data supporting the cross-correlation hypothesis

As can be seen from the above chronological presentation of some of the more significant input to the discussion of the CC receiver, the arguments supporting the idea have changed over time, but none of these data have thus far been publicly denounced by the author(s) as not supporting the CC hypothesis.

Thus, in discussing the problem it is important to keep in mind what sort of data are being brought forth as evidence for the CC receiver. These generally fall into two categories: 1) the performance curve looks like the CC function or its envelope, and 2) the threshold found is so low that it can only imply perfect reception (\Leftrightarrow utilization of all information in the signals \Leftrightarrow matched filtering=CC receiver).

Taking the first point first: is the shape of the psychometric function, the performance curve, directly revealing the underlying receiver output upon which the animal's decision is based? Most of the data that have been proposed to support the theory of coherent or semicoherent reception in bats have been of the sort that demonstrated this kind of similarity between the receiver output and performance (Simmons, 1971; Simmons, 1973; Simmons and Stein, 1980; Simmons, 1979; Simmons and Grinnell, 1988; Altes, 1989; Simmons et al., 1990). However, the notion that the amount of errors that the animals make should closely follow the shape of the receiver output has been shown to be in error at realistic SNRs (Menne & Hackbarth, 1986). Only when SNR is extremely low, it is correct.

Below it will be shown that in fact the performance in Simmon's experiments

does not follow the shape of the CC functions in question (or their envelopes), which fact renders this part of the discussion irrelevant for the discussion of the CC receiver.

The second point – that the range determination threshold is so low that only a CC receiver can explain the performance – only applies to the data from Simmons et al. 1990, claiming the threshold to be 10 ns. The 200 times higher threshold of about 2 μ s obtained in 1979 has been demonstrated to be much less accurate, than what is predicted by both the coherent receiver and semicoherent receiver at ordinary laboratory SNR (Schnitzler and Henson, 1980; Menne and Hackbarth, 1986).

Together with Bertel Møhl I have previously shown that the 10 ns threshold was obtained with equipment that had some malfunction – we hypothesized impedance mismatch as a likely possibility – and which did therefore not warrant the trust that one

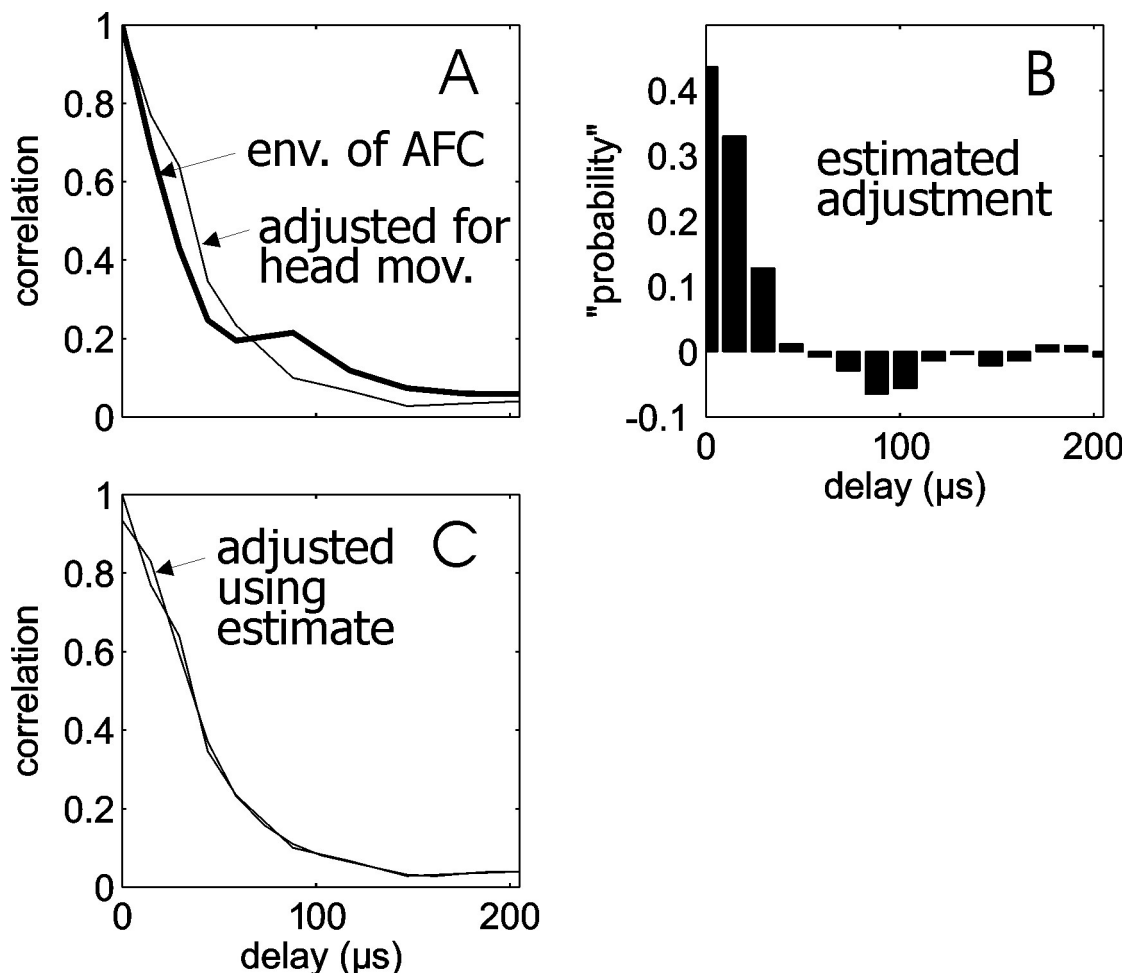


Fig. 0-4: Deconvolution operation to recover the head movement data from the corrected ACF envelope and the envelope itself (see Fig. 0-3). A. Thick line: Envelope of ACF of the *E. fuscus* call. Thin line: envelope corrected for head movements. B. Deconvolution result. It is plotted as a histogram in order to emphasize the interpretation of the result as a distribution of delays. C. The test of the success of the operation. The convolution between the deconvolution result in B and the untreated envelope (thick line in A) plotted together with the data corrected for head movements (thin line in A).

should be able to have in the equipment used to obtain the data supporting this rather claim (Beedholm and Møhl, 1998).

Performance data do not match the cross-correlation function of the cries

The ranging data – or at least their interpretation – antedating the 1979 jitter experiment have to some extent been overlooked in later discussions of the CC hypothesis. I feel that this is a mistake, especially in the light of the historical development or “arms race” between theory and data outlined above. The experiments that are being made today are largely shaped by the interpretations of the experiments made in the past, as should also be evident from the review.

Head movements and the predicted performance from the autocorrelation envelope

As mentioned above, in the Science paper by Simmons from 1971 the performance of the bats was compared with a “theoretical” prediction derived from the envelope of the ACF of the call. The procedure for correcting the envelopes is described in Simmons and Grinnell (1988), as follows: “The envelope only needed to be displaced in small steps and the results averaged, thus mimicking the head-movement artifact, (...)”. This is a description of a normalized convolution process. So the distribution describing the bat’s head positions relative to the targets is convolved with the envelope of the ACF. This is also consistent with an oral explanation given by Simmons at the Sandbjerg Meeting (1994).

In Fig. 4 of the 1971 Science paper is given the ACF as well as its envelope. This we can use to compare with the prediction derived from it to give us an idea of the amount of correction that has been applied due to the measured head movement data. The result of this comparison is shown in Fig. 0-3. Here is depicted the “theoretical” data scanned from Fig. 2 of the Science paper together with the CC function and its envelope scanned from Fig. 4. Also, in the lower panel I have placed the envelope data, suitably scaled and treated as described in Simmons (1973), inside the performance curve graph. The conversion between performance and envelope in the latter case consisted of converting time to distance and doubling the resulting distance scale; as explained in Simmons (1973), explanation of Fig. 15, the “degree of overlap” between two identical functions, as a function of the delay between them, is itself a copy of this function, but stretched to have the double abscissa values compared to the original. The peak of the ACF envelope is made to correspond to 50% correct and zero to 100% correct. This process was reversed in placing the data in the upper panel of Fig. 0-3.

The upper panel most clearly shows the nature of the differences between the envelope and the envelope adjusted for head movements. Of particular interest is that the “theoretical” (i.e. head-movements corrected envelope) values above ca. 75 μ s are lower than the uncorrected envelope data. To explore this phenomenon I have used a deconvolution technique to estimate how the distribution had appeared with which

the envelope had been convolved¹. The results are shown in Fig. 0-4. In the right-hand graph is shown the deconvolution result, graphed as a histogram to show that it should represent a distribution; it is the head movement data displayed with a delay axis. The lower left graph shows the result of convolution of the envelope data with our estimated head movements data compared with the “theoretical” curve, and the fit is very close (correlation coefficient 0.996), proving the deconvolution operation a success.

The interesting thing about the head movements estimate (Fig. 0-4) is that it has a substantial negative part between 50 μ s and 170 μ s. The interpretation that there was a negative probability of finding the bats’ head at the distance corresponding to this delay is not meaningful, and we can only conclude that *something is wrong*. There does exist the possibility that the corrections were made in another way, although the method used here is both relatively sound and described in the quote above. But whatever other method might have been used, it is not possible for me to account for the fact that some of the corrections to the envelope at some delays lead to an *improved* “theoretical” performance from the untreated CC function envelope.

The shape of auto- and crosscorrelation functions of *Eptesicus* calls in the laboratory

In 1979 and 1990 (see above) Simmons and Simmons et al., respectively, again compared the performance of *E. fuscus* in a range related experiment to the ACFs of their calls. In 1979 he used the half-wave rectified ACF to compare with the performance in the jitter experiment. One main argument in this and in later papers is the observation of a secondary dip in the performance of the bats of 10-15% occurring around 30 μ s² jitter. This is demonstrated to coincide approximately with the side-lobe peak of the ACF of an echolocation call.

But what is potentially available to the bats is not the ACF of their cries but the CC function between emitted and returned calls, which in a target simulator where the

¹ Deconvolution in practice is something of an art form. The problem is mainly that it involves division with a spectrum of a signal, which might contain zeroes (notches), leading to unbounded results. In the case at hand I have initially interpolated the data points to make the samples equally spaced in time. They were then mirrored to make the curve symmetric and the deconvolution was made in the lower frequency part only, where both signals had significant energy. The best way to evaluate the success of the operation is to make the corresponding convolution again and compare the result with the original dividend.

² In these later papers the delay axis of the CCF was not doubled when predicting the performance, as it was in 1971 and 1973. If this had been the case the prediction would have been a dip in performance around 60 μ s. Altes (1989) has also noted that this is expected, but he argues that under the assumption of a moving target indicator, the drop in performance again appears at 30 ms (provided that the RMS bandwidth is 33 kHz).

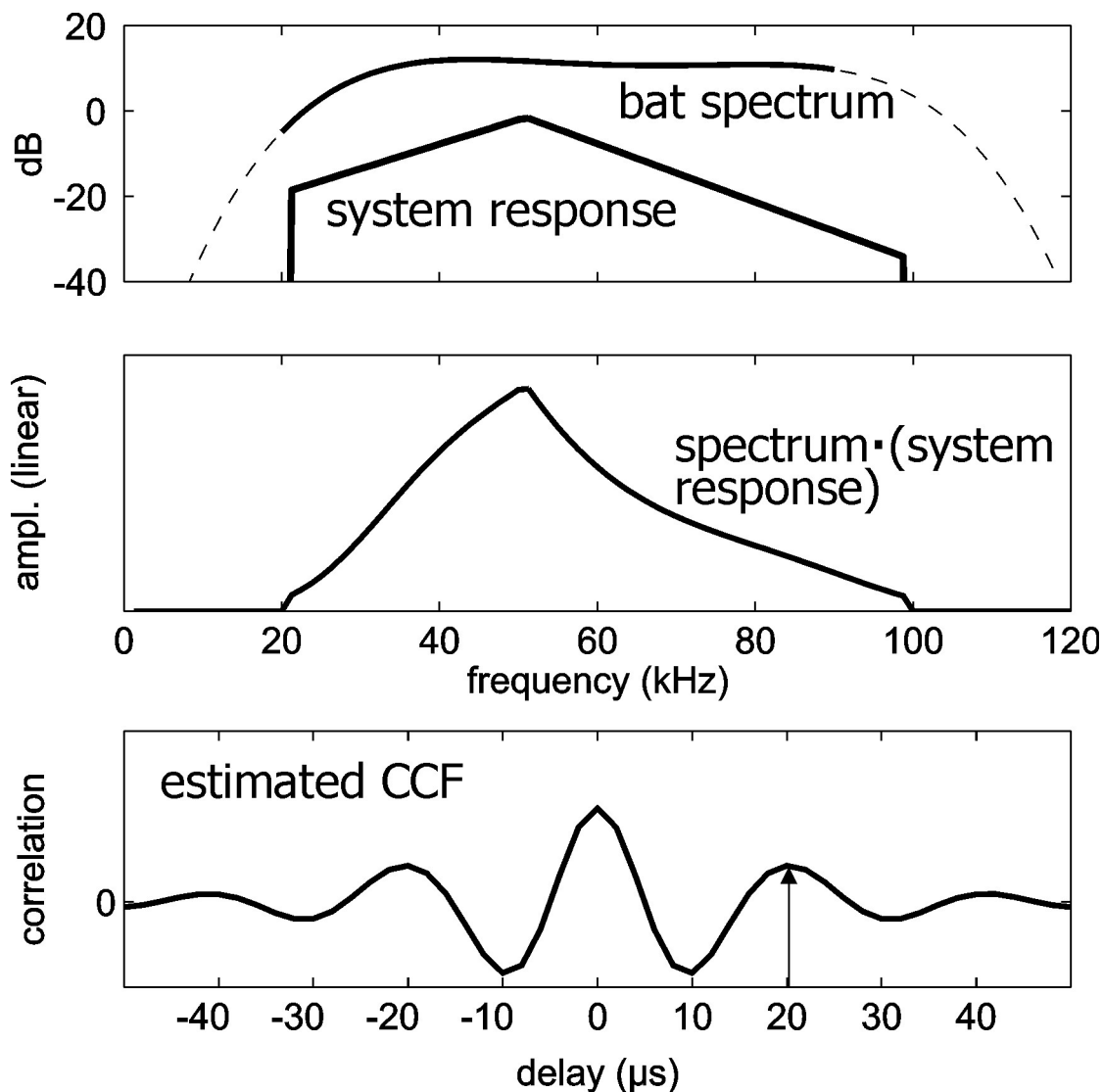


Fig. 0-5: The CCF between the average outgoing *E. fuscus* call in Simmons et al. (1990) and an estimated returning phantom echo. Top panel: Upper thick line is the average spectrum of 10 bat calls estimated from Simmons (1993), Fig. 7b. The dashed line is an extrapolation. Lower thick line is a rectilinear approximation (in dB) to the transfer functions of the loudspeakers in the target simulator and is taken from Fig. 5 of Simmons et al. (1990). Middle panel shows the product between the linear scale amplitude spectrum squared and the linear scale estimated transfer function. This is equal to the spectrum of the CCF of the signals in question. The bottom panel is the inverse Fourier transform of the data in the middle panel. This is the sought for estimate of the CCF. Note that it has its first side-lobe peak at 20 μ s.

transfer function is not flat, may be somewhat different from the ACF. We can estimate this by looking first at the transfer functions of the speakers given in Fig. 5 of Simmons et al. (1990). This estimate must next be multiplied with the energy spectrum (the amplitude spectrum squared) of the calls and the result inverse Fourier transformed to yield the CC function assuming the phase spectrum is overall zero. We obtain this

estimate of the average spectrum of the calls from Fig. 7b of Simmons (1993). The result of these estimations is shown in Fig. 0-5. The lower panel shows the finally CC function, which has its secondary peaks at 20 μs from the main peak. So, provided the argument – that the performance curve should directly reflect the CC function – holds, the dip in the performance curve is expected to occur at 20 μs , not 30 μs , as is the case in both Simmons (1979) and Simmons et al. (1990).

In the paper by Simmons et al. 1990 a special way of converting the behavioral data is used in the construction of a so-called compound curve (Kiang et al., 1965). This type of curve when used in electrophysiological contexts is constructed by displaying first the neural response to a given stimulus in a normal histogram of action

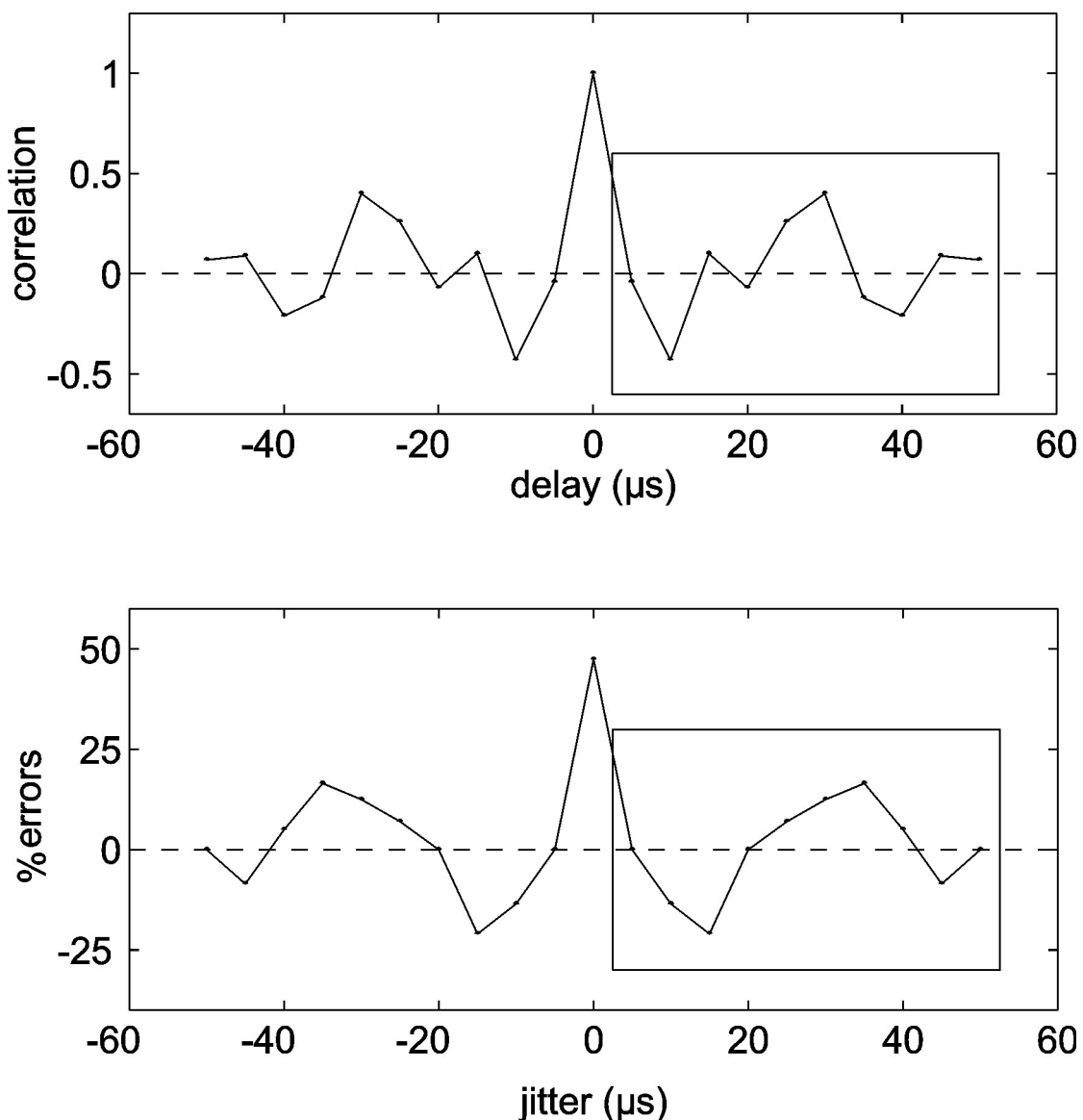


Fig. 0-6: The data from Fig. 21 of Simmons et al. (1990). The upper graph represents the ACF of a call by *E. fuscus* resampled at 20 kHz. The lower graph is the “compound jitter discrimination curve” for 3 bats. The boxes in each graph delimit the independent part of the data. See text for explanation.

potentials (peri-stimulus histogram, PSTH). Secondly, the response to a phase-reversed version of the stimulus is displayed as an inversed PSTH in the same graph. In difference to what is the case in Kiang et al. (1965), in Simmons et al. (1990) it is error percentages that are being displayed on the ordinate axis, not spike densities, and jitter values are used on the abscissa. Furthermore, the data are mirrored around zero jitter. The compound curve is finally compared to a mean ACF from the bats performing the task in Fig. 21 in Simmons et al. (1990). This comparison is reproduced here in Fig. 0-6.

As mentioned, the performance data in this graph is mirrored around zero (which is not necessary); the autocorrelation function is per definition symmetrical around zero, and we need not display the negative delay values. As a next step we recognize that at zero delay both the ACF and the “compound performance curve” are always one unless something is very wrong with the experiment. Since we investigate the similarity between the precise structure of the performance curve and the ACF (“sampled” at 20 kHz), we should exclude this trivial point from the analysis.

In the end, doing so leads to a comparison between the remaining data within the boxes drawn in Fig. 0-6, between the correlation values at delays 5 - 45 μ s and the compound curve in the same interval. This similarity can be numerically expressed by the correlation coefficient between the two data sets. In this case it is about 0.3, which indicates that these two data sets are rather remotely related.

I believe that I have now made the point that there is no demonstrated connection between the performance of the bats in range and jitter experiments and the ACF or CC function of their calls.

What are the jitter thresholds?

As we point out in chapter I, newer data have supplied interesting input to the discussion of whether the bats use ranging to perform the jitter task. Some relatively recent experiments by Surlykke (1992) on the ranging accuracy of *E. fuscus* showed that they had a rather high dependency on signal bandwidth. Limiting the returning echoes to basically the lowest harmonic (40 kHz) resulted in a reduction of at least 50 μ s in their ability to perform the task. This is particularly interesting when compared with what Moss and Schnitzler (1989) found in their Yes/No reproduction of the Simmons (1979) jitter experiment. These authors found the threshold to be less than 0.4 μ s when the full bandwidth of the target simulator was used. However, limiting the bandwidth of the calls returning to bats by low-pass filtering the signals to frequencies below 55 kHz resulted in a slight reduction in one animal’s ability to detect the jitter with a hinted threshold of slightly less than 0.4 μ s. When Surlykke’s low-pass filter was set to 55 kHz the threshold remained the same as with the full range utilized. It appears from the Surlykke (40 kHz low-pass) experiment that the bats are not limited by head-movements but possibly

limited by bandwidth. The 55 kHz low-pass data from both experiments hint at a different dependence on bandwidth in the two types of tasks, jitter detection vs. range estimation.

Hyperacuity

Some comments on the notion of hyperacuity in connection with the jitter results of Simmons (1979) and Simmons et al. (1990) are necessary. Altes (1989) published an analysis of the SNR relationship of different target configurations, and showed that the jittering target was an advantage in terms of range difference detection compared to a task of a simple range determination of point targets when a coherent CC process is involved. However, there is nothing magical about the jittering targets that are not included in the Woodward (1955) relation, and the simulated data published by Menne and Hackbarth (1986) do encompass this hyperacuity.

Alternatives to the cross-correlation hypothesis

The only viable alternative to the CC model around is the filter bank (FB) receiver, which has been treated by Hackbarth (1986), Menne (1988) and utilized by Grossetête and Moss (1998) and Masters and Raver (2000). This represents in many ways a formalized version of how many neurophysiologists think of the way in which a delay estimate may be formed in the nervous system of the bats.

This receiver forms a delay estimate from many frequency channels by averaging or integrating. The FB thus incorporates basic known facts about the mammalian hearing system, something that is not immediately true for the CC models.

Saillant et al. (1993) have proposed the so-called spectrogram correlation and transformation (SCAT) receiver, which we will show in Chapter III to be simply a phase sensitive implementation of the filter bank, which amounts to a coherent CC receiver. In this it does not represent a very novel model, but simply a CC receiver with a filter bank preprocessing and a peak-detector simulating phase locked spike generation.

The scope of this thesis

With this brief mention of alternative to the CC receivers we at last arrive at the main purpose of this introduction, which is of course to pave the way for the data and ideas presented in the following 4 chapters. As they represent laid out manuscripts all chapters should be individually readable, but sometimes we have allowed ourselves to reference other chapters within the thesis. The independence means that there are 4 introductions, but we have tried to vary them somewhat.

Chapter I shows that the physical limits to delay difference detection in a target simulator do allow for tens of nanoseconds jitter detection when real bats (*Eptesicus fuscus*) are used. We do however feel that the most important conclusion from that work is that the simulator has to be very close to perfect for this to be true. At least the simulator used in Simmons et al. (1990) had almost exactly the same phase characteristics that caused the problems faced when the loudspeaker response was not perfectly equalized. This result corroborates the findings by Beedholm and Møhl (1998), who demonstrated errors in the equipment used in Simmons et al. (1990), by showing that jitter discrimination of tens of nanoseconds was indeed not possible in that study.

Chapter II demonstrates that the above-mentioned FB receiver in form is in fact also a CC receiver, albeit stripped of all phase information. We discuss the immediate problems that follow from this simple fact in terms of deciding between receiver models in bats and we show that this can indeed be done and that such experiments have already been made.

The discussion also deals with a problem concerning the (at least nearly) Doppler tolerant waveform employed by many FM bats. That area of discussion has direct implications for the notion of the optimal (coherent CC) receiver, since only a very sensitive system would be affected notably by the errors introduced by Doppler changes at the flight speeds of most bats. We suggest a new explanation for the (almost) Doppler tolerant time-frequency structure based on the idea that Vespertilionid bats might have evolved a sweep rate filter to avoid jamming.

Lastly our simulation results suggest a new reason for shortening the signals as the bats are approaching sound reflecting objects.

Chapter III leans in many ways on the work from Chapter III in that it shows the relationship between the SCAT receiver (Saillant et al. 1993; Peremans and Hallam, 1998) and the FB receiver model mentioned above. We demonstrate for practical signals

that the SCAT receiver is in fact a form of coherent CC receiver implemented with filter bank pre-processing.

We discuss the physiological and practical relevance of the SCAT model of the sonar process.

Chapter IV is to some extent the presentation of an idea mainly related to signal processing of harmonically structured signals. We show how to separate signal components that overlap in time and frequency by using a cross-correlation technique implemented from both a time- and frequency domain perspective.

This technique is used to settle a discussion concerning the exact time-frequency structure used by *Eptesicus fuscus* in the laboratory.

We find that all the models for the time-frequency structure of the calls have problems with the initial part of the calls.

Finally, the findings, brought about by this novel technique, is discussed in terms of possible consequences for the FB receiver, which seems to warrant this specialization more than the (conventional) CC receivers.

Plans

All chapters are to be submitted as articles to The Journal of the Acoustical Society of America. Chapter III was thought as a companying paper to Chapter II. To the extent that he agrees to this, all articles are to be submitted coauthored by Prof. Dr. Hans-Ulrich Schnitzler, who was my supervisor during the project, has provided the equipment, and who has had significant influence on the work presented here.

References

- Altes, R. A. (1980). "Detection, estimation, and classification with spectrograms," J. Acoust. Soc. Am. **67**, 1232-1246.
- Altes, R. A. (1981). "Echo phase perception in bat sonar?" J. Acoust. Soc. Am. **69**, 505-508.
- Altes, R. A. and Titlebaum (1970). "Bat signals as optimal Doppler tolerant waveforms," J. Acoust. Soc. Am. **48**, 1014-1020.
- Au W. W. L., Moore, P. W. B. and Pawlowski, D. A. (1988). "Detection of complex echoes in noise by an echolocating dolphin," J. Acoust. Soc. Am. **83**, 662-668.
- Au, W. W. L. (1993). *The sonar of dolphins* (Springer-Verlag, New York).
- AAirepet'yants and Konstantinov (1973). *Echolocation in animals* (Israel Program for Scientific Translation, Jerusalem).
- Beedholm, K. and Møhl, B (1998). "Bat sonar: An alternative explanation for the 10 ns jitter result," J. Comp. Physiol. A **182**, 259-266.
- Biering, H. and Pedersen, O. Z. (1983). "System analysis and time delay spectrometry," Brüel and Kjær Technical Review **32**, 3-52.
- Burdic, W. S. (1968). *Radar signal analysis* (Prentice Halls, Engelwood Cliffs, NJ).
- Cahlander, D. A. (1964). "Echolocation with Wide-Band Waveforms: Bat Sonar Signals," MIT Lincoln Lab. Tech. Rep. **27**.
- Cahlander, D. A., McCue, J. J. G. and Webster, F. A. (1964). "The determination of distance by echolocating bats," Nature **201**, 544-546.
- Griffin, D. G. (1944). "Echolocation in blind men, bat and radar," Science **100**, 589-590.
- Griffin, D. G. (1958). *Listening in the Dark* (Yale University Press, New Haven).
- Griffin, D. G. and Galambos, R. (1941). "The sensory basis of obstacle avoidance by flying bats," J. Exp. Physiol. **86**, 481-506.
- Grossetête, A. and Moss, C.F. (1998). "Target flutter rate discrimination by bats using frequency modulated sonar sounds: Behavior and signal processing models," J. Acoust. Soc. Am. **193**, 2167-2176.
- Hackbarth, H. (1986). "Phase Evaluation in Hypothetical Receivers Simulating Ranging in Bats," Biol. Cybern. **54**, 281-287.

- Hartley, D. J. (1989). "The effect of atmospheric sound absorption on signal bandwidth and energy and some consequences for echolocation," *J. Acoust. Soc. Am.* **85**, 1338-1347.
- Hartley, D. J. (1992a). "Stabilization of perceived echo amplitudes in echolocating bats. I. Echo detection and automatic gain control in the big brown bat, *Eptesicus fuscus*, and the fishing bat, *Noctilio leporinus*," *J. Acoust. Soc. Am.* **91**, 1120-1132.
- Hartley, D. J. (1992b). "Stabilization of perceived echo amplitudes in echolocating bats. II. The acoustic behavior of the big brown bat, *Eptesicus fuscus*, when tracking moving prey," *J. Acoust. Soc. Am.* **91**, 1133-1149.
- Hartridge, H. (1920). "The Avoidance of Objects by Bats in their Flight," *J. Physiol* **54**, 54-57.
- Hartridge, H. (1945a). "Acoustic control in the flight of bats," *Nature* **156**, 490-494.
- Hartridge, H. (1945b). "Acoustic control in the flight of bats," *Nature* **156**, 692-693.
- Heyser, R. C. (1967). "Acoustical measurements by time delay spectrometry," *J. Audio Eng. Soc.* **15**, 370-382.
- Jensen, M. E. and Miller, L. A. (1999). "Echolocation signals of the bat *Eptesicus serotinus* recorded using a vertical microphone array: effect of flight altitude on searching signals," *Behav. Ecol. Sociobiol.* **47**, 60-69.
- Kick, S. A. and Simmons, J. A. (1984). "Automatic gain control in the echolocating bat, *Eptesicus fuscus*," *J. Comp. Physiol. A* **145**, 431-435.
- Kiang, N. Y. S., Watanabe, T., Thomas, C. and Clark, L.F. (1965). *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (MIT Press, Cambridge Mass.).
- Kroszczyński, J. J. (1969). "Pulse Compression by means of Linear Period Modulation," *Proc. IEEE* **57**, 1260-1266.
- Masters, W. M. and Raver, K. A. S. (2000). "Range discrimination by big brown bats (*Eptesicus fuscus*). using altered model echoes: Implications for signal processing," *J. Acoust. Soc. Am.* **107**, 625-637.
- Menne, D. and Hackbarth, H. (1986). "Accuracy of distance measurements in the bat *Eptesicus fuscus*: Theoretical aspects and computer simulations," *J. Acoust. Soc. Am.* **79**, 386-397.
- Menne, D., Kaipf, I., Wagner, I., Ostwald, J. and Schnitzler, H.-U. (1989). "Range estimation by echolocation in the bat *Eptesicus fuscus*: Trading of phase versus time cues," *J. Acoust. Soc. Am.* **85**, 2642-2650.

- Miller, L. A. and Pedersen, S. B. (1988). "Echoes from insects processed using time delay spectrometry (TDS)," in *Animal Sonar Processes and Performance*, edited by Nachtigall, P. E. and Moore, P.W. B. (Plenum Press, New York).
- Moss, C.F. and Schnitzler, H.U. (1989). "Accuracy of target ranging in echolocating bats: acoustic information processing," *J. Comp. Physiol. A* **165**, 383-393.
- Moss, C.F. and Simmons, J. A. (1993). "Acoustic images of a point target in the bat *Eptesicus fuscus*: Evidence for sensitivity to echo phase in bat sonar," *J. Acoust. Soc Am.* **93**, 1553-1662.
- Møhl, B. (1986). "Detection by a *Pipistrellus* bat of normal and reversed replica of its sonar pulses," *Acustica* **61**, 75-82.
- Peremans, H. and Hallam, J. (1998). "The SCAT receiver revisited," *J. Acoust. Soc. Am* **104**, 1101-1110.
- Pollak, G. D. (1993). "Some comments on the proposed perception of phase and nanosecond time disparities by echolocating bats," *J. Comp. Physiol.* **172**, 523-531.
- Pye, J.D. (1986). "Sonar Signals as Clues to System Performance," *Acustica* **61**, 166-175.
- Saillant, P.A., Simmons, J.A. and Dear, S.P. (1993). "A computational model of echo processing and acoustic imaging in frequency modulated echolocating bats: The spectrogram correlation and transformation receiver," *J. Acoust. Soc. Am.* **94**, 2691-2711.
- Schnitzler, H.-U, and Henson, O. W. Jr. (1980). "Performance of airborne animal sonar systems," in *Animal Sonar Systems*, edited by Busnel, R. G. and Fish, J. F. (Plenum Press, New York). pp 109-181.
- Simmons J. A. (1971). "Echolocation in Bats: Signal Processing of Echoes for Target Range," *Science* **171**, 925-928.
- Simmons J. A. (1973). "The resolution of target range by echolocating bats," *J. Acoust. Soc. Am.* **54**, 157-173.
- Simmons, J. A., Lavender, W. A., Lavender, B. A., Childs, J. E., Hulebak, K., Rigden, M. R., Sherman, J., Woolman, B. and O'Farrel, M. J. (1978). "Echolocation by free-tailed bats (*Tadarida brasiliensis*)," *J. Comp. Physiol.* **125**, 291-299.
- Simmons J. A. (1979). "Perception of Echo Phase Information in Bat Sonar," *Science* **204**, 1336-1338.
- Simmons, J. A. and Grinnell, A. D. (1988). "The performance of echolocation: Acoustic images perceived by echolocating bats," in *Animal Sonar Processes and*

Performance, edited by Nachtigall, P. E. and Moore, P.W. B. (Plenum Press, New York).

Simmons, J. A. and Stein, R. A. (1980). "Acoustic Imaging in Bat Sonar: Echolocation Signals and the Evolution of Echolocation," *J. Comp. Physiol.* **135**, 61-84.

Simmons, J. A., Ferragamo, M., Moss, C. F., Stevenson, S. B. and Altes, R. A. (1990). "Discrimination of jittered sonar echoes by the echolocating bat, *Eptesicus fuscus*: The shape of target images in echolocation," *J. Comp. Physiol. A* **167**, 589-616.

Simmons J. A. (1993). "Evidence for perception of fine echo delay and phase by the FM bat, *Eptesicus fuscus*," *J. Comp. Physiol. A* **171**, 533-547.

Skolnik (1980). *Introduction to Radar Systems* (McGraw-Hill, New York), 2. ed.

Strother, G. K. (1961). "Note on the Possible use of Ultrasonic Pulse Compression by Bats," *J. Acoust. Soc. Am.* **33**, 696-697.

Surlykke, A. (1992). "Target ranging and the role of time-frequency structure of synthetic echoes in big brown bat, *Eptesicus fuscus*," *J. Comp. Physiol. A* **170**, 83-92.

Surlykke, A. and Moss, C.F. (2000). "Echolocation behavior of big brown bats, *Eptesicus fuscus*, in the field and the laboratory," *J. Acoust. Soc. Am.* **108**, 2419-2429.

Troest, N. and Møhl, B. (1986). "The detection of phantom targets in noise by serotine bats; negative evidence for the coherent receiver" *J. Comp. Physiol.* **159**, 559-567.

Urkowitz, H. (1967). "Energy detection of unknown deterministic signals," *Proc. IEEE*, **55**, 523-531.

Woodward, P. M. (1955). *Probability and Information theory, with applications to radar* (Pergamon Press, New York).

CHAPTER I

Limits to jitter detection accuracy with the signals of echolocating FM bats (*Eptesicus fuscus*)

Introduction

Microchiropteran bats that use echolocation as a sensory system emit trains of ultrasonic pulses and evaluate the returning echoes for a multitude of parameters (Griffin, 1958; Pollak and Casseday, 1989).

With respect to the distance to sound reflecting objects it was long hypothesized that bats - in similarity with the man-made systems, radar and sonar - use echo delay to measure this parameter (Hartridge, 1945; Cahlander et al., 1964).

Simmons (1971) has elegantly showed in a series of psychophysical experiments that bats are indeed so good at determining the distance to a fixed target that this would seem to be something the animals were specialized in. He later finally settled the question by showing that the bats were equally good at determining the distance to a virtual target, i.e. the delay to an electronically produced replica of the outgoing pulse (Simmons, 1973), whereby is ruled out that the animals solved this task by other means such as differences in target bearing when they are moving from one position to the next or simply by using the associated slight amplitude differences associated with different distances for the sound to travel.

There has been a rather lengthy discussion as to whether bats use processing for target distance that can be modelled as a phase sensitive (coherent) cross-correlation receiver (Simmons, 1979; Menne and Hackbarth, 1986; Hackbarth, 1986; Altes, 1981; Saillant et al., 1993), which can be shown to constitute the optimal receiver for target range and detection (Woodward, 1955). The optimal receiver is one that uses all the available information in the signals to arrive at an estimate.

It was shown in a study by Menne and Hackbarth (1986) that even with very conservatively estimated signal to noise ratios (hereafter SNR) in the laboratory, only a range resolution in the area of tens of nanoseconds would constitute evidence of a coherent (phase preserving) receiver. Therefore this question has been more or less reduced

to revolving around the credibility of the data later published by Simmons et al. (1990) showing the limit of the bats' ability to determine a change in delay between successive calls (jitter) to be 10 ns - equivalent to 1.7 μm range difference (for discussion see Pollak, 1993; Simmons, 1993; Masters et al., 1997; Beedholm and Møhl, 1998).

Another receiver that has been discussed extensively in this context is the semicoherent receiver, which is identical to the coherent receiver, but the envelope of the output is formed before the delay measurement is made. Theoretically this has been the focus of some interest since it can be shown to be possible in principle to construct the output from this receiver from a spectrogram representation (Altes, 1980), a fact owing to a theorem by Ackroyd (1971).

In judging which of these two, coherent or semicoherent, might be inside an unknown sonar system, the important differences in performance are that the semicoherent receiver is insensitive to a constant phase shift across frequencies and that the ranging accuracy attainable is inversely proportional to the centralized RMS bandwidth (hereafter B_{CRMS}) rather than to the RMS bandwidth (B_{RMS}), which is the relevant quantity limiting the coherent receiver (see e.g. Menne and Hackbarth, 1986; Au, 1993). As this point is a difficult one it is discussed briefly below. As the exact expressions for the two measures of bandwidth are not particularly intuitive and are readily attainable from the mentioned sources if needed, we omit them here.

One important difference between these two measures of "bandwidth" is their very different degree of dependence on carrier frequency. The B_{CRMS} measure is independent of the location of the spectrum on the frequency axis and is always smaller than B_{RMS} . It is the second moment of the amplitude spectrum interpreted as a distribution. In contrast B_{RMS} in many instances is the carrier frequency. It is the first moment, or the mean, and is as such highly misleading if read as a measure of what is intuitively understood as bandwidth. An example of this taken from the bat research area: When CF-FM bats like rhinolophids encounter a novel object in their flight path, they increase the relative weight of the FM tail in their signals (Langenheinecke, 2000) and we rightly think of this as an increase in bandwidth. But the effect on the B_{RMS} is actually that it *decreases* as the center of gravity of the signal is moved towards the lower frequencies of the FM tail. Likewise, many FM-bats decrease the frequencies of the fundamental in their sweep during the final stages in the hunt (e.g. Siemers et al., 2001). This will increase the bandwidth in the

normal sense of the word (including the B_{crms} measure) but it decreases B_{rms} , which - if the animals use a coherent receiver - leads to lower ranging accuracy at this presumably critical stage of the pursuit.

The minimum standard deviation of range estimates obtainable with one of the two receivers mentioned can be calculated using (see Menne and Hackbarth, 1986)

$$\sigma = \frac{1}{2\pi\beta\sqrt{\frac{2E}{N_0}}}, \quad (\text{II-4})$$

where E is the energy of the echo, N_0 spectral noise level, and β is the “bandwidth” measure relevant to the receiver as discussed above.

It has been debated (e.g. Pollak, 1993; Simmons, 1993) how the bats could cope with the unavoidable head and body movements between pulse emission and echo return. It is another question altogether whether in a laboratory situation it is physically possible to determine range with nanosecond acuity using bat echolocation signals even if indeed the head movements can be taken into account.

In discussions on this subject with colleagues we have often met the view that this would not be possible in the face of microturbulence, temperature inhomogeneity and other factors inevitably present in the laboratory under normal circumstances and in particular in front of an active bat and a trainer.

In the present study we seek to answer the question of the minimum practically attainable accuracy as seen from the viewpoint of the bat without actually putting the question to the animals. Instead we sampled both the outgoing sounds and returning echoes in a psychophysical experiment and investigated how accurately we could measure the delay between these two sound instances using optimal processing of the echolocation signals. We investigate and explain the found dependence of the attained accuracies on the transfer function of the playback system.

Materials and methods

Animals and training

4 *Eptesicus fuscus* (2 males, 2 females, one of each sex caught in the wild, the other two born in captivity) were initially trained over a period of more than two years, typically 5 days a week.

During collection of the data the bats were performing in an Y/N psychophysical detection experiment, in which they were rewarded for move towards the microphone/loudspeaker if a simulated target was present and turn away from it if no target was pre-

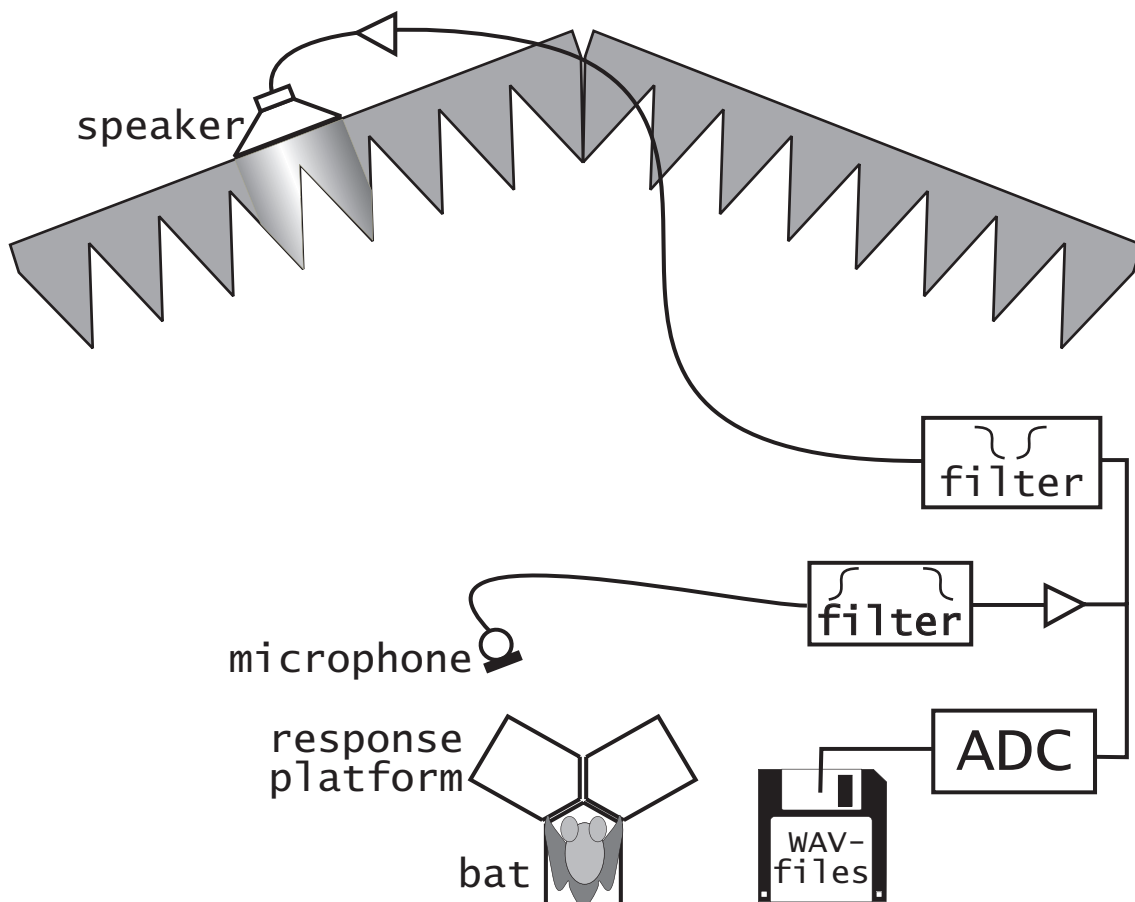


Fig. I-1: Schematic of the setup used for bat training, signal playback and sampling. See text for explanation.

sented. Correct responses were rewarded with half a mealworm. Only signals recorded during reinforced (target present) trials were included in the current investigation.

The normal training usually started with the playback gain set to -16 dB relative to the outgoing signal at which unnaturally high level the bats never made any mistakes once they were well trained. In order to get as high a SNR ratio as possible we chose this gain setting for the present analysis, and since bats were used to this level, they did not react in any extraordinary way during these trials.

Apparatus

We decided to let the signal part of the setup consist of analog equipment only. The electronic parts of the echo playback system consisted of a microphone, attenuators, amplifiers, analog filters and a loudspeaker and is shown in fig. I-1.

The electrostatic loudspeaker (Polaroid) had a transfer function with a peak around 60 kHz. This response was “flattened” by means of analog filters, one Krohn-Hite

(model 2650) set to band reject around 60 kHz and a simple 1st order passive high pass filter with -6 dB cutoff at 60 kHz. The resulting combined frequency response was flat within 6 dB over the range of 28 to 100 kHz. The microphone was a 1/8'' B&K (model 4138).

The loudspeaker that delivered the played back echoes was attached to the back of a chipboard wall covered with echo-attenuating foam and placed 1.5 m from the microphone. Echoes were delivered through a hole in this wall and simulated a target distance of 0.75 m and resulting in a propagation delay of 4.4 ms. Usually the microphone was attached to the training platform but for these recordings it was placed freely in front of the response platform to ensure mechanical independence of the two. See Fig. I-1.

Sampling

The signals picked up by the microphone positioned ca. 10 cm in front of the left leg of the platform served 2 purposes: They were played back through the loudspeaker after being attenuated to a controlled extent, and they were digitized and stored to disk for offline analysis. The experimenter stopped the continuous sampling by means of a foot switch when the bat seemed to have reached its decision and the samples made in the preceding 2 seconds were then saved (480 kHz, 16 bit). Trials usually lasted less than 2 seconds. For each bat 4 trials were run with these conditions.

Offline signal treatment

The sampled trial-sequences were split up so that all sequences consisted of consecutively emitted signals with a recorded level in excess of 90 dB peSPL and no overlap between pulse and echo. Signals below this level or signals so long that overlap resulted were not analyzed further. Within the resulting sequences of consecutively emitted sounds each pulse-echo pair was then split up in pulse and corresponding echo.

We were interested in the physical limits in a jitter paradigm, which means that the task is to investigate if the delay/position of each echo changes between successive calls, whereas the absolute delay and slow fluctuations are irrelevant. This means that the behavior of the animals is an integral part of the measurement: the higher the repetition rate, the less we must assume is changed in the sound conveying medium from call to call. Therefore we only used delay differences between successive pulse-echo pairs.

The cross-correlation (coherent receiver) as well as the envelope of the cross-correlation (semicoherent receiver) between the pulse and echo were formed. The delay between pulse and echo was evaluated using both these methods. The exact location of

the peak was found by expanding the sampling rate to 3.84 MHz and interpolating the three samples centered around the peak with a Lagrange polynomial which was differentiated and solved for the zero crossing (i.e the time of the peak)¹.

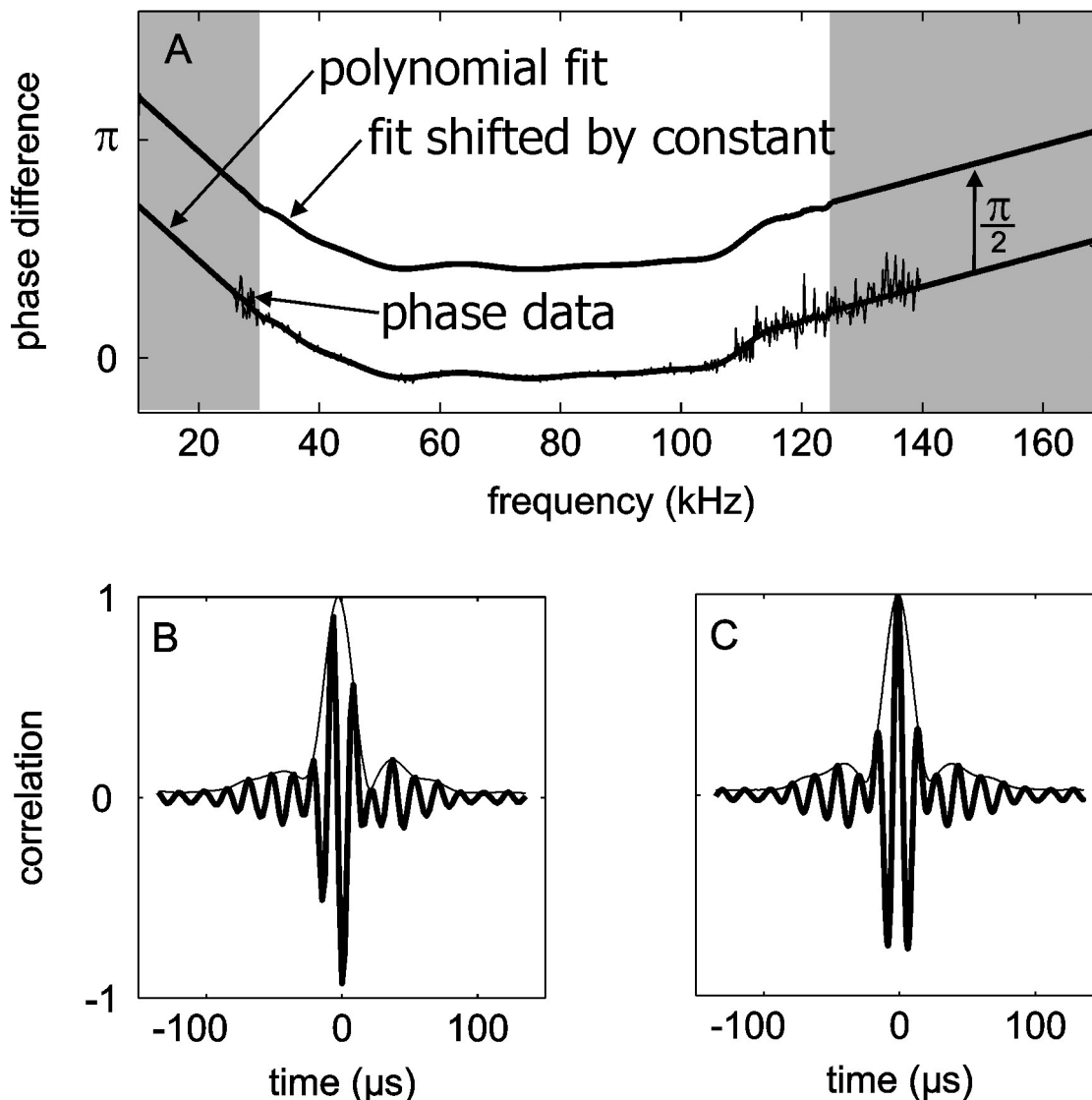


Fig. I-2: Phase functions and the effect of their application to the data. A: Phase spectrum corrected for linear trend (absolute delay), and fitting polynomials. The thin noisy line is the average phase difference for 8 signal-echo pairs. The lower thick line on top of the phase data is a polynomial fit to the transfer function. The grey areas indicate, where a linear fit is used to extrapolate the fit into the frequency range, where the spectrum is dominated by noise. The upper thick line is the fit with an additional constant phase shift of $\pi/2$. B: CCF of a cry made by EF4 and its envelope without the corrections to the transfer function. C: CCF and envelope with both group delay and phase shift accounted for.

¹ If $y(n)$ is an array of samples to be interpolated and n_0 is the peak value sample of y , the interpolated time of the extremum can be found from:

$$\frac{y(n_{-1})n_1^2 - y(n_{-1})n_0^2 - y(n_0)n_1^2 + y(n_0)n_{-1}^2 + y(n_1)n_0^2 - y(n_1)n_{-1}^2}{2(y(n_{-1})n_1 - y(n_{-1})n_0 - y(n_0)n_1 + y(n_0)n_{-1} - y(n_1)n_{-1} + y(n_1)n_0)}$$

which we give here in the hope that the reader may be spared an hour of calculus.

For all series of consecutively emitted cries the result was two arrays of range estimates (coherent and semicoherent). For each of these N-sized data sets the differences to the previous estimate were now calculated, resulting in N-1 estimates of the variation from call to call. These arrays of differences were finally pooled across cry series.

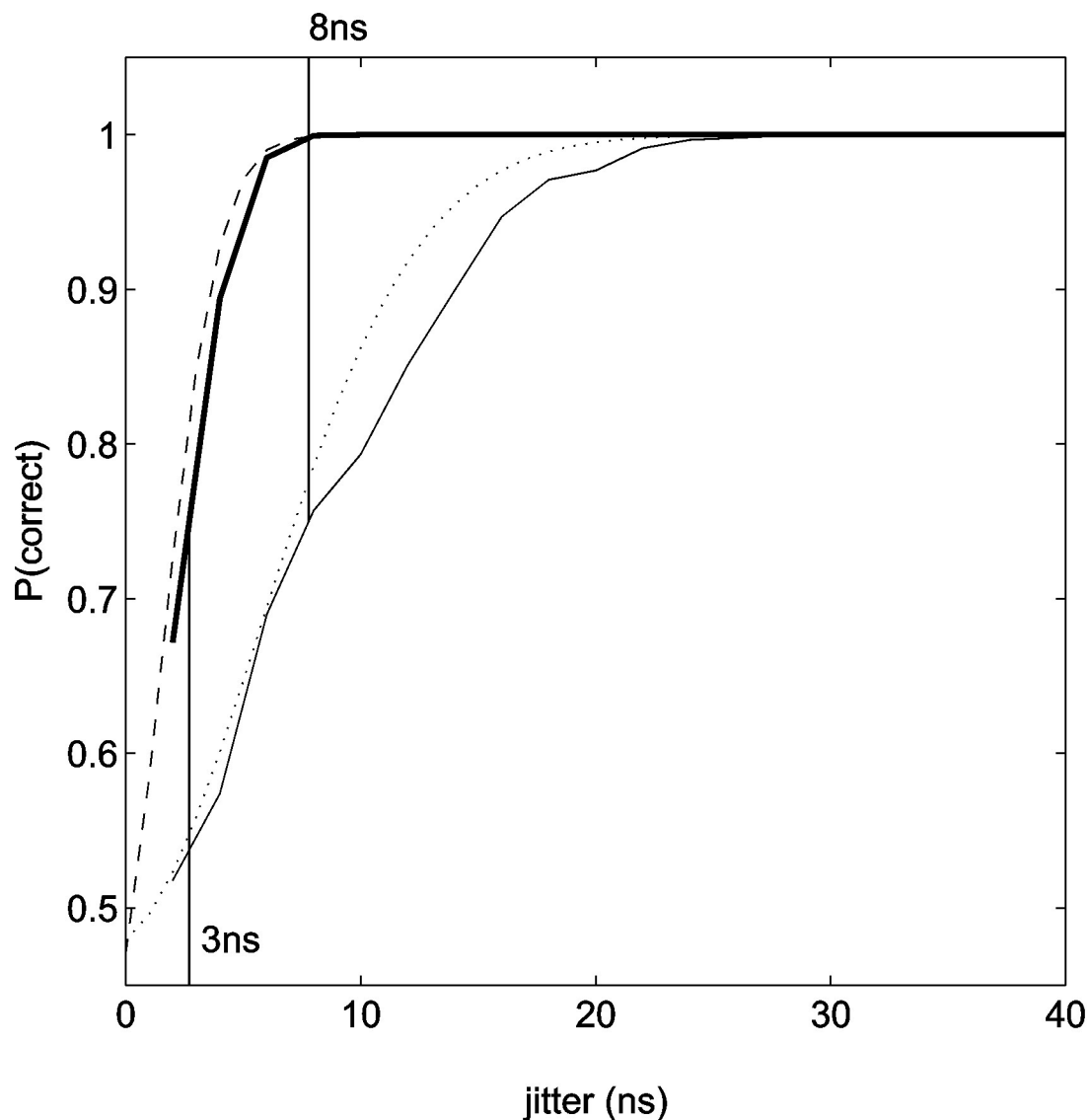


Fig. I-3: Performance curves resulting from a sequence of 20 simulated calls in white noise using the same procedures as with the real bat call sequences. Thick solid line: Performance curve of coherent estimates. Thin solid line: Semicoherent performance curve. Dashed line: Theoretical prediction of coherent performance using (I-1) and (I-3). Dotted line: Prediction of semicoherent performance. Compare with Fig. I-4d.

The transfer function of the system was determined from 8 pulse-echo pairs that were excluded from the analysis. This was used to determine high degree polynomials that approximated the average amplitude and unwrapped phase spectra closely in the frequency range where this was discernable. Beyond this range linear extrapolations were made (see Fig. I-2).

The group delay (the phenomenon that some frequency components are delayed more than others) was equalized as follows: 1) The CCF was calculated between the outgoing and returned signal. 2) The envelope was formed and the peak sample found. The CCF was delayed so that now $t=0$ marked the time of the peak of its envelope. 4) The phase of the FFT of the recentered CCF was calculated and a linear regression line to the phase data was subtracted to yield an estimate independent of absolute delay and constant phase shift. The average of these data for the 8 calls was formed and this phase shift was subtracted from the phase spectrum of all the echoes to be corrected.

A constant phase shift of $\pi/2$ was measured “by eye”. When calculated, the average was close to $\pi/2$ and we assumed some natural process to be responsible, like current rather than voltage driving the transducer or it being a velocity transducer (although neither of these two is very probable, see discussion below).

Six data sets were created using different modifications to the echoes. 1) A set without modifications, 2) one with variations in spectrum level equalized, 3) group delay made zero for all frequencies and 4) the constant phase shift across frequencies of $\pi/2$ eliminated. As 5) a data set with phase shift and group delay corrected and finally 6) one where all three corrections were applied.

For all echoes the theoretical lower limit to the standard deviation was calculated from the energy of the echo, E , spectral noise level, N_0 (measured only once), and from the bandwidth, using $(\beta=B_{\text{rms}})$ for the coherent receiver and $(\beta=B_{\text{crms}})$ for the semicoherent receiver estimate (see Menne and Hackbarth, 1986) by means of the relationship I-1.

We used the average values to make theoretical performance curves. These could then be compared with the experimental ones.

All methods were tested using sequences of synthetic calls for which the parameters were known (see Fig I-3).

Simulation procedure

The data sets were pooled for all bats recorded from. The resulting data sets were the basis for simulated 2 alternatives forced choice jitter discrimination experiments. For each of

the spectral modifications to the echoes mentioned above a coherent and a semicoherent performance curve was generated.

In order to generate the performance curves, the following procedure was used. The computer picked two random items, A and C from a data set of measured delay differences. The decision rule was that if

$$|A + jitter| > |C|, \quad (I-2)$$

a correct decision was counted. This was done 40.000 times for each jitter value and the probability of answering correctly was determined from the ratio of correct decisions to this number of trials. The jitter values were spaced 5 ns apart between 1 and 2000 ns. Curves are shown in Fig. I-4. With a little consideration one finds that if the distribution of the differences is normal, the expected curves are describable as

$$P_{correct}(jitter) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{jitter}{4\sigma}\right)^2, \quad (I-3)$$

which expression is the basis for the theoretical performance curves shown in Fig. I-3 and Fig. I-4d.

Results

Performance curves are shown in Fig. I-4 and the evaluated thresholds in Fig. I-5.

The performance curve for the simulated bat in Fig. I-3 shows the very low thresholds of 2.7 (coherent) and 7.8 ns (semicoherent) together with the theoretical curves calculated as for the real calls.

The lower limit to jitter detection was found to be 48 ns in our system (See Fig. I-4D), as defined by the coherent estimate when group delay, phase shift and amplitude were all equalized, which condition gave the lowest threshold. The theoretical limit as defined by the average SNR * B_{rms} product (see I-1) was as low as 4.3 ns.

When none of the corrections of the different imperfections of the transfer function were applied to the echoes, the performance limit is several hundred nanoseconds with both the coherent and semicoherent receiver (Fig. I-4A).

The semicoherent estimates are independent of correction of the constant phase shift (Fig I-4A,C).

When group delay is taken into account but not the constant phase shift, the semicoherent estimates are better than the coherent ones (Fig. I-4B).

Equalizing the amplitudes across frequencies has very little or no effect on performance. This is the case both for the estimates that are corrected for the other two parameters and the ones with no further corrections (Fig. 4-A and Fig. 5).

For the performance curves without the corrections to the transfer function it is evident that normality is not completely fulfilled. This is mainly the result of some outliers that occur in this condition, which cause convergence on 1 to result at a rather large jitter value. When all corrections are made the curves appear normally distributed.

Discussion

The results show the limits to the attainable accuracy in jitter measurements using the bats own signals under different degrees of transfer function optimization. These results are not comparable with similar experiment using man-made signals simulating the bat calls, since such signals typically lack the variation introduced to the calls by the bat. The measurements made here are independent of head movements since the stationary microphone, not the bat, is the point of reception of both the cross-correlated sound instances. Since the microphone was always at least 10 cm in front of the bat, most but maybe not all of the air movement associated with sound production and body movements by the bats is excluded from our measurements.

There are several possible factors that will lead to a larger variation than is expected from consideration of SNR and bandwidth alone. They fall into two classes.

One consists of the errors due to for instance temperature changes in the room between calls or an instable platform or microphone stand. This will affect coherent and semicoherent estimates equally. If this type of error is dominating we expect the absolute difference between the values for the coherent and semicoherent estimates to be the same as the difference between the two theoretical estimates: a variation of this type adds to both thresholds and thus does not increase the numerical differences between them.

The other type of error is rapid phase changes arising from factors that we can imagine typically to be turbulence, especially in the vicinity of the test animal and the experimenter. These errors are characterized by having the same effect as noise: Noise can be thought of as a phenomenon that adds a random phase component to the signal, independent of frequency. So, if this type of error is prevalent, we expect to find that the difference between the coherent and the semicoherent thresholds is different from the theoretical values, but instead the ratio between the two stays stable, which can be seen from (I-1): The ratio between the thresholds when $\beta=B_{\text{rms}}$ and $\beta=B_{\text{crms}}$ is independent of the noise level.

Comparing the differences between theoretical and predicted thresholds for the coherent and semicoherent receivers we can therefore get a rough estimate of the kind of

process that creates the differences. The theoretical ratio (coherent/semicoherent) is $(4.3/19.4) = 0.22$, and the difference (semicoherent-semicoherent) is 15.1 ns, whereas the ratio

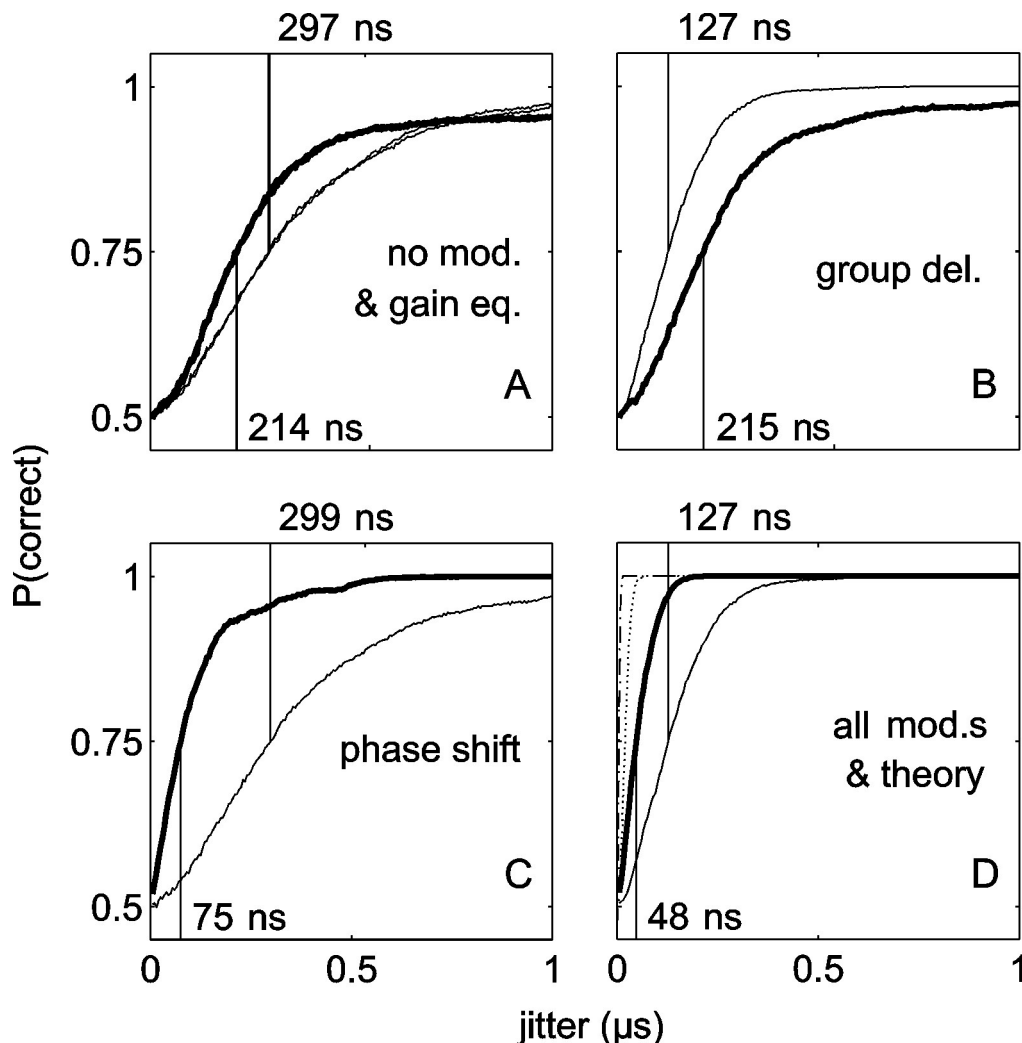


Fig. I-4: Performance curves for modifications to echo spectra as indicated inside graphs. A contains two sets of curves (practically identical) and in D also is shown the theoretical curve calculated from eq. (I-3) and the average SNRs and bandwidths according to eq (I-1). Thick solid lines: Coherent performance curve. Thin solid lines: Semicoherent performance curve. Dot-dash line: Theoretical coherent receiver. Dotted line: Theoretical semicoherent receiver. Vertical lines indicate the time of crossing the 75% correct threshold.

for the actually measured thresholds is 0.38, with the difference equal to 79.2 ns. So the errors behave rather more like we would expect if they added white noise to the channel, than if the slower processes, such as temperature changes, were dominating. In the jitter paradigm the “long term” errors are filtered out by the use of range estimate differences only, so this finding is as we expect it to be. A note should be made, however, on the bandwidths that have been calculated from the echoes. These do contain some harmonic

distortion products that add some bandwidth to the signals, but on the other hand are not helping the range determination. The influence of this phenomenon we have judged to be very small, less than 1 %.

The comparatively long propagation delay used here has the advantage that variations per unit distance become minimal, since most of the turbulence is expected to take place immediately in front of the bat and the experimenter. This ensures that the specific estimate of errors due to the imperfect channel properties is conservatively estimated (see below). Also, the slow, additive component is weighted less when this distance is long.

Artificial sequence and theoretical limit

The simulated sequence shown in Fig. I-3 illustrates several points. First of all it shows that the methods and programs used seem to be working as intended. The match between theoretical and simulated curves is very good. Since only 20 calls were in the analyzed sequence we conclude that in this study we have used sufficient sounds (184) to estimate the real variation. The thresholds are lower than the ones found to be the theoretical limit to the real bat calls; this is because both a higher SNR and bandwidth were used. The fact that we can measure a threshold of 3 ns in agreement with the theoretical prediction shows that the sampling rate boost and Lagrange interpolation procedures are reliable to at least this degree of precision.

Secondly it extends in a rather informal way the analysis by Menne and Hackbarth (1986) that showed the limit to jitter discrimination to be very low in both a coherent and a semicoherent system but using comparatively low bandwidth and SNR. In fact in our case the threshold for the semicoherent receiver is also in the vicinity of 10 ns, equal to the jitter threshold of Simmons et al. 1990, but as mentioned above for rather high SNR and B_{crms} .

The use of the rule (I-2) is based on the idea that the bat is not at all times (at least not at the beginning of a trial) aware of which of the two signals is supposed to be farther away, wherefore the absolute value of difference would appear to be the better measure.

Effects of equalizations

Perhaps the most interesting finding here is the strong dependence of the jitter threshold on the degree to which the transfer function of the system has been corrected. Below we seek to explain these large effects that the modifications to the transfer function have on the stability of the range estimates.

Phase shift correction

It surprised us slightly at first that the phase shift imposed on the echoes by the system had such a large negative effect on the stability of the estimates. Using simulated sequences this phenomenon did not occur (Fig. I-3). We found the answer to lie in the slight variations in bandwidth that will always be present in a biological system as opposed to the simulated situation.

It is a lucky coincidence that our loudspeaker had a phase shift of $\pi/2$. The width of the central peak of an ACF measured between zero crossings is very close to $1/(2B_{\text{rms}})$. This is relatively easily seen by considering that the ACF is the inverse Fourier transform

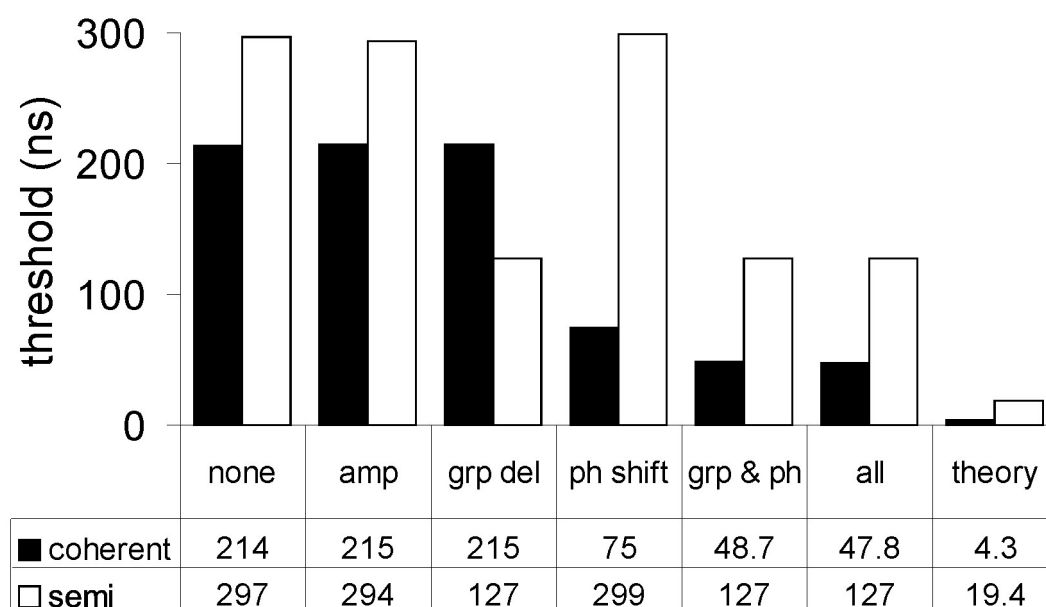


Fig. I-5: 75% thresholds for all data sets. amp; amplitude equalization; grp del: group delay correction; ph: phase shift elimination; ph shift: phase shift correction; grp & ph: both group delay and phase shift corrections.

of the power spectrum (Wiener-Khintchine theorem, see also Papoulis, 1962). If a CCF is formed between a bat signal and the same signal phase shifted by $\pi/2$, the highest peak occurs at the time of the first zero crossing measured from the peak in the ACF. This can be envisioned by considering that the ACF like any signal can be written as $A \cos(\psi)$, where A is the instantaneous amplitude and ψ is the phase function. When $\pi/2$ is subtracted from ψ , one can instead of $A \cos(\psi - \pi/2)$ write $A \sin(\psi)$ and so the peaks of the original signal become zero-crossings by this phase turn.

So changing the center frequency (B_{rms}) slightly from call to call leads to foreseeable changes in the position of the highest peak in the case where the echo is a phase-

shifted version of the transmitted signal (Fig. I-6). It is an expected finding that this modification does not influence the semicoherent estimates. These are identical to the uncorrected ones within rounding errors.

The physical explanation for the observed phase shift eludes us. The transducer is a condenser type and as such it should translate voltage into displacement. Whatever the explanation we are not the only ones to experience this phenomenon as is discussed below.

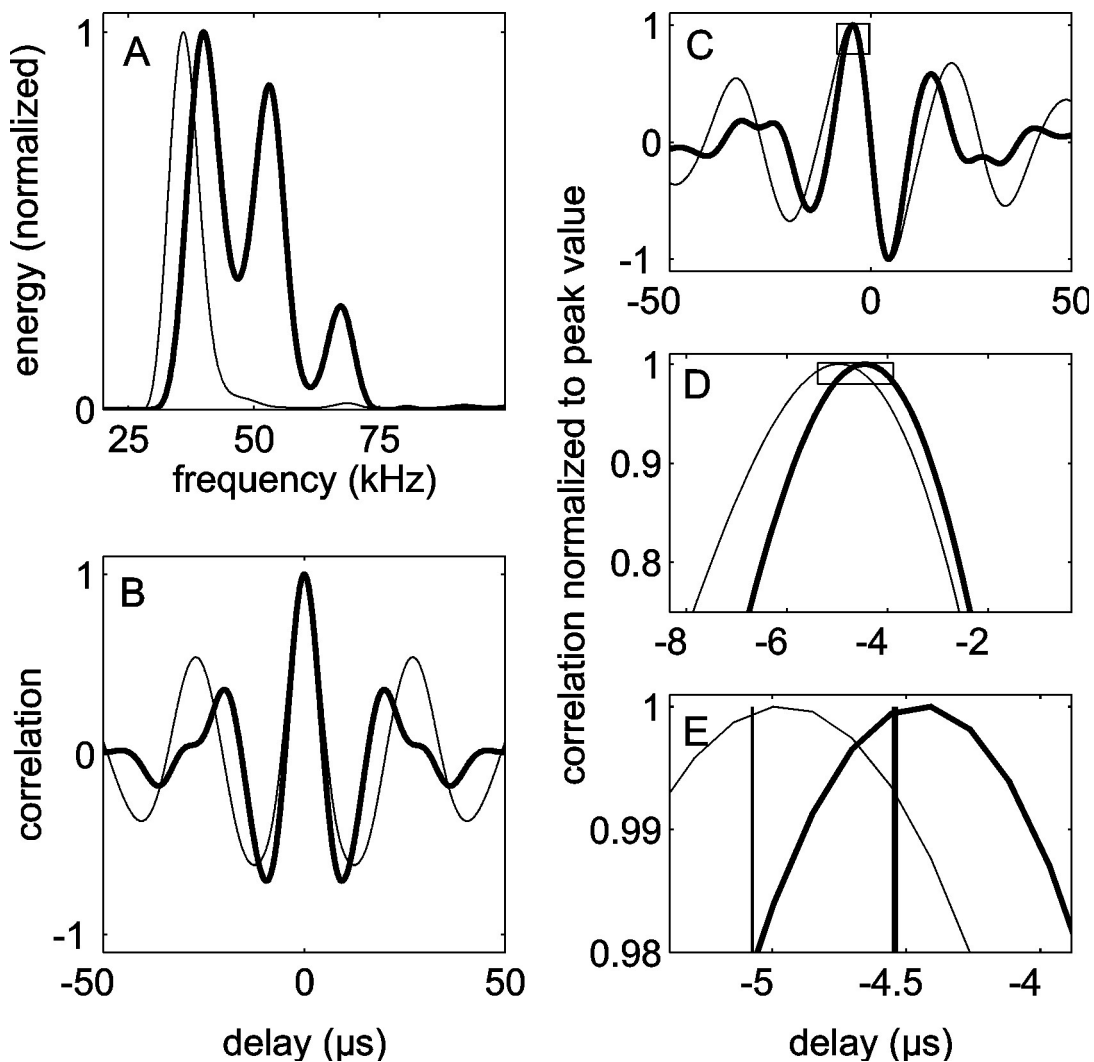


Fig. I-6: An example of the effect of a phase shifting transfer function on the peak location of the CCF between outgoing and returning signal. Thick and thin solid curves represent data for 2 signals emitted consecutively by EF3. A: Power spectrum B: ACFs. C-E: CCF between one copy of the signal phase shifted $\pi/2$ relative to an undistorted one. Squares in C and D indicate the range covered by D and E respectively. The vertical lines near the peak in E are the predicted values of the shifted peak location using $1/(4 B_{\text{rms}})$.

Group delay correction

Understanding the influence of a non-zero group delay on the estimate is also relatively simple when one considers the variability of the transmitted signal. If the group delay is not constant across frequencies, then some frequencies in the signal being reflected arrive with more delay than others. If the relative weight of the frequencies within calls

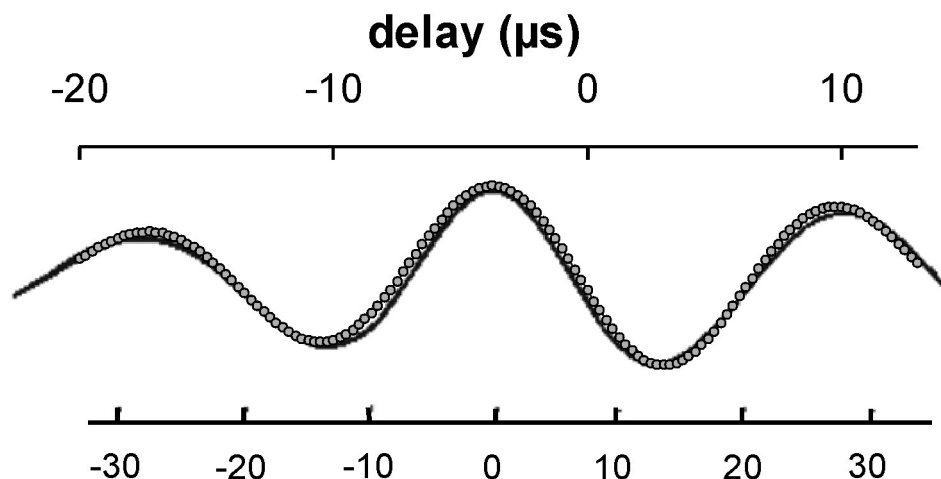


Fig. I-7: Comparison between the system responses of the apparatus used here and the one in Simmons et al. (1990). The solid line is obtained by scanning Fig. 8 in Simmons (1993) showing the central part of the CCF between an artificial bat call transmitted through the target simulator used in Simmons et al. (1990) and the returning echo. The “string of pearls” trace is the central part of a CCF between a bat call (EF3) and the returning echo from the present study without the phase shift correction. The upper abscissa is the time centered on the peak of the *envelope* of the “string of pearls” trace only. The lower axis came with the scanned data and we presume that it has been centered on the peak of the data without taking the envelope first.

shifts between emissions, then the time of arrival of the peak energy of the CCF changes accordingly, but less predictably than was the case with the phase shift above. In this error type more frequencies have an influence on the estimate than is the case in the above treated case of a constant phase shift on the coherent estimate.

Amplitude equalization

The finding that making the amplitude part of the transfer function flat did not improve the accuracy, is explained by the fact that this did not lead to any increase in SNR. The microphone noise is dominating our estimates, and equalizing the echoes *after sampling* changes the noise level equally. Correcting this factor prior to transmission through the speakers (or in general: prior to the limiting noise source) would presumably have reduced the variation in range estimate differences slightly.

Comparison with real jitter experiments

The jitter experiments made by Simmons (1979), Moss and Schnitzler (1989) and Menne et al. (1989) as well as that carried out by Masters et al, (1997) all had their lowest detected values above $0.4 \mu\text{s}$ and thus operated well or just outside the range where the

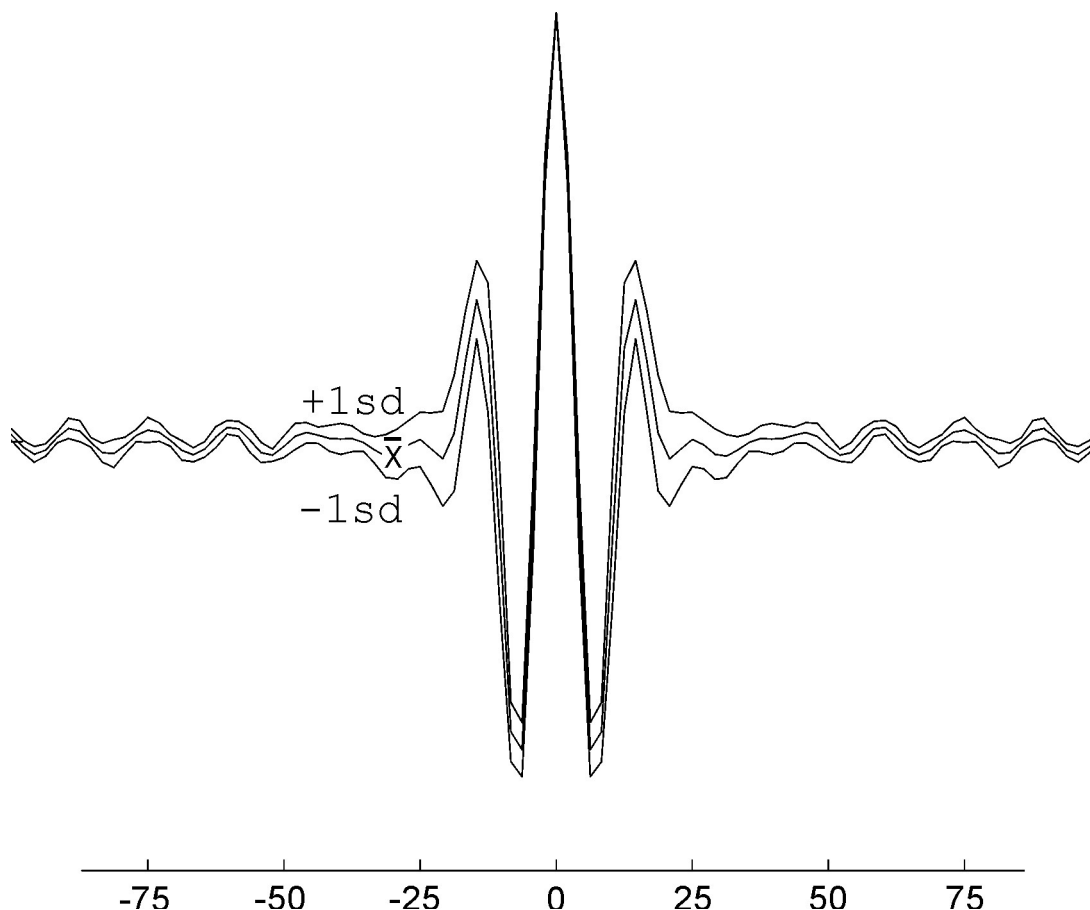


Fig. I-8: Mean and standard deviation of ACFs of 50 calls emitted by EF2. When the phase shift between call and echo is $\pi/2$, the zero crossing to the left of the central peak becomes the highest point. The variation that would be critical in a ranging or jitter task is therefore at this point. The calls are not all consecutive. This graph was made for comparison with with Fig. 20b in Simmons et al. (1990).

effects reported on here become important. The jitter experiments in 1990 by Simmons et al. yielded a threshold about 10 ns. Although we believe the results in that experiment to have been influenced by other factors (Beedholm and Møhl, 1998), we find here that this accuracy is physically attainable under conditions like the ones in our laboratory. In Simmons et al. (1990) the air path was 1.6 ms, whereas in ours it was 4.4 ms. Since we found that the errors are attributable to the channel, we divide the 48 ns jitter threshold found here by the ratio (4.4/1.6) yielding 17 ns as a prediction to what would be the threshold

with that shorter air path. The 48 ns (and the corrected 17 ns) were found with the minimum number of echoes, so using a larger number of estimates before each decision could easily bring the threshold below the 10 ns.

However, 10 ns is not a realizable threshold if the system has a phase shift (coherent receiver) or a considerable group delay (semicoherent receiver) as described and discussed above. We do not know if group delays were important in the setup of Simmons et al. (1990), although the transfer functions (Fig. 5 of Simmons et al. 1990) strongly suggest so, knowing that the phase of a minimum phase system can be found from the Hilbert transform of the logarithm of the amplitude spectrum (see e.g. Oppenheim and Schaefer, 1989): The system response is not flat and thus it can be expected to have a non-zero group delay. A minimum phase assumption is a simple one. If the system is not minimum phase the relationship will be another, but unless specific, special measures are taken to eliminate differences in group delays they will be different from zero. As mentioned above, the minimum errors of less than 48 ns found here are probably reduced when a shorter air path is used, but this does not apply to the errors introduced by the imperfect transfer function in combination with the variation in the echolocation calls. These errors are additive as is evident from Fig. I-5.

The more important error source for the coherent receiver as found in this study is the phase shift and this is definitely present in the Simmons et al. (1990), which is revealed by Fig. 8 in Simmons 1993, showing the central part of a CCF between outgoing (simulated) cries and returning echoes for the system used in the 1990 setup. We plotted an arbitrarily selected uncorrected CCF from this study over a scan of that graph and the result is shown in Fig. I-7. The amazing similarity is in part coincidental, since some signals in this study have a more complex ACF, but the common phase shift does reveal that the analysis made here is highly relevant to the Simmons et al. (1990) study.

The bats could possibly be imagined to have taken measures to eliminate the influence on the delay estimation of the phase shift resulting from the variation in the average frequency of their calls. This would, however, mean that they were aware that such an operation should be performed before evaluating the receiver output. In nature a perfect phase shift would rarely if ever occur and we must assume that this means that the animals do not have evolved the ability to cope with it.

It could justifiably be argued that in the present study the task was one of detection and an extraordinarily easy one at that. Therefore the bats might not keep the pulse-to-pulse variation in bandwidth as low as they could if they were performing in a sub-micro-

second jitter task. To survey this factor we computed the autocorrelation function for 50 non-consecutive calls emitted during the task. Comparing these with the similar data of Fig. 20b in Simmons et al. (1990) we felt convinced that the variation was not much larger, if larger at all, in the present investigation. (Fig. I-8).

Relationship of jitter with ranging experiments

As an aside we next discuss the relationship between the ranging and jitter detection experiments. The jitter experiments were originally thought up as a way to exclude the head movements that seemed to be the limiting factor in ranging experiments where the animals must simply indicate the closer of two targets (e.g. Simmons, 1971; Masters and Jacobs, 1989) or that the target is farther away than a standard position (e.g. Masters and Raver, 2000). A very important experiment by Surlykke (1992), however, shows that in these experiments the bats are limited by bandwidth (in the usual sense of the word), rather than by head movements. In that study the returning echoes were filtered (low- or high- pass), and as a result the thresholds increased very markedly in both instances, more so with the high-pass filtered signals. In contrast, the study by Moss and Schnitzler (1989) indicated that this was not the case for the jitter experiments, or at least to a dramatically different extent. These investigators did not find a threshold even when the jitter was at $0.4 \mu\text{s}$ with normal simulated echoes, which was the limit of the system, but for one bat operating with low-pass filtered echoes the threshold seemed to lie around this value. This means that the reduction in bandwidth lowered the jitter resolution by $0.4 \mu\text{s}$ at the most. These two investigations taken together then seem to demonstrate that the two types of experiments, jitter and ranging, are dealing with different processes and that head movements are not the key to this difference.

Conclusions

In conclusion we have shown that jitter detection in the range of tens of nanoseconds is physically possible with bat echolocation signals under conditions found in the laboratory.

We found that for a coherent ranging mechanism in bats to be feasible the target must not have phase shifting characteristics, since the variation in the calls of the bats is so that this leads to errors in the peak location. For the semicoherent receiver a very similar problem with group delays exists.

From the viewpoint of echolocation modeling it is of some interest that the vari-

ation in the emitted calls play a comparatively large role in the quite common situation where the target is not a perfect point target. The emitter variation is rarely considered, simply because it can be taken into account (and this is implicitly done so by the CCF), but the present findings suggest that also knowledge of the reflector is necessary, which is what is sought by the active sensory process.

In general it does not pose any technical problem to reproduce the same outgoing signal over and over again in radar and sonar systems, but what is said here about emitter variation is of course also true for variations occurring in the channel between emissions, which could hint at a possible practical significance of this result. This is, however, the experience leading to the selection of a semicoherent receiver over a coherent one.

We have shown that the claim of Simmons et al. (1990) that bats are able to discern a simulated jitter of only 10 ns is not reliable since there is evidence that in this earlier investigation there was a phase shift associated with the transfer function, similar to what was found to be the case in our investigation.

References

- Ackroyd, H. M. (1971). "Short-Time Spectra and Time-Frequency Distributions," J. Acoust. Soc. Am. **50**, 1229-1231.
- Altes, R. A. (1980). "Detection, estimation, and classification with spectrograms" J. Acoust. Soc. Am. **67**, 1232-1246.
- Altes, R. A. (1981). "Echo phase perception in bat sonar?" J. Acoust. Soc. Am. **69**, 505-508.
- Au, W. W. L. (1993). *The Sonar of Dolphins*, Springer-Verlag, New York.
- Beedholm, K. and Møhl, B (1998). "Bat sonar: An alternative explanation for the 10 ns jitter result," J. Comp. Physiol. A **182**, 259-266.
- Cahlander, D. A., McCue, J. J. G. and Webster, F. A. (1964). "The determination of distance by echolocating bats," Nature **201**, 544-546.
- Griffin, D. G. (1958). *Listening in the Dark*, Yale University Press, New Haven.
- Hackbarth, H. (1986). "Phase Evaluation in Hypothetical Receivers Simulating Ranging in Bats," Biol. Cybern. **54**, 281-287.
- Langeheinecke, E. J. (2000). *Guidance Behaviour and Orientation Reaction in Commuting Greater Horseshoe Bats (Rhinolophus ferrumequinum)*, Dissertation, University of Tübingen.
- Masters, W. M., Jacobs, S. C. (1989). "Target detection and range resolution by the big brown bat (*Eptesicus fuscus*). using normal and time-reversed model echoes," J. Comp. Physiol. A **166**, 65-73.
- Masters, W. M., Raver, K. A. S., Kornacker, K. and Burnett, S. C. (1997). "Detection of jitter in intertarget spacing by the big brown bat *Eptesicus fuscus*," J. Comp. Physiol. A **181**, 279-290.
- Masters, W. M., Raver, K. A. S. (2000). "Range discrimination by big brown bats (*Eptesicus fuscus*). using altered model echoes: Implications for signal processing," J. Acoust. Soc. Am. **107**, 625-637.
- Menne, D. and Hackbarth, H (1986). "Accuracy of distance measurements in the bat *Eptesicus fuscus*: Theoretical aspects and computer simulations," J. Acoust. Soc. Am. **79**, 386-397.
- Menne, D., Kaipf, I., Waner, I., Ostwald, J. and Schnitzler, H.-U. (1989). "Range estimation by echolocation in the bat *Eptesicus fuscus*: Trading of phase versus time cues," J. Acoust. Soc. Am. **85**, 2642-2650.
- Moss, C.F. and Schnitzler, H.U. (1989). "Accuracy of target ranging in echolocating bats: acoustic information processing," J. Comp. Physiol. A **165**, 383-393.

- Oppenheim, A. V. and Schaffer, R. W. (1989). *Discrete time Signal Processing*, (Prentice-Hall, Englewood Cliffs, New Jersey).
- Pollak, G. D. and Casseday, J. H. (1989). *The Neural Basis of Echolocation in Bats*, (Springer-Verlag, Berlin).
- Pollak, G. D. (1993). "Some comments on the proposed perception of phase and nano-second time disparities by echolocating bats," *J. Comp. Physiol.* **172**, 523-531.
- Papoulis, A. (1962). *The Fourier Integral and its Applications* (McGraw-Hill, New York) pp. 240-242.
- Saillant, P.A., Simmons, J.A. and Dear, S.P. (1993). "A computational model of echo processing and acoustic imaging in frequency modulated echolocating bats: The spectrogram correlation and transformation receiver," *J. Acoust. Soc. Am.* **94**, 2691-2711.
- Siemers, B. M., Kalko, K. V. and Schnitzler, H.-U. (2001). "Echolocation behavior and signal plasticity in the Neotropical bat *Myotis Nigrans*, (Schinz, 1821). (Vespertilionidae): a convergent case with European species of *Pipistrellus*?" *Behav. Ecol. Sociobiol.* **50**, 317-328
- Simmons J. A. (1971). "Echolocation in Bats: Signal Processing of Echoes for Target ,Range," *Science* **171**, 925-928.
- Simmons J. A. (1973). "The resolution of target range by echolocating bats," *J. Acoust. Soc. Am.* **54**, 157-173.
- Simmons J. A. (1979). "Perception of Echo Phase Information in Bat Sonar," *Science* **204**, 1336-1338.
- Simmons, J. A. (1993). "Evidence for perception of fine echo delay and phase by the FM bat, *Eptesicus fuscus*," *J. Comp. Physiol. A* **171**, 533-547.
- Simmons, J. A., Ferragamo, M., Moss, C. F., Stevenson, S. B., Altes, R. A. (1990). "Discrimination of jittered sonar echoes by the echolocating bat, *Eptesicus fuscus*: The shape of target images in echolocation," *J. Comp. Physiol. A* **182**, 259-266.
- Surlykke, A. (1992). "Target ranging and the role of time-frequency structure of synthetic echoes in big brown bat, *Eptesicus fuscus*," *J. Comp. Physiol. A* **170**, 83-92.
- Woodward, P. M. (1955). *Probability and Information Theory with Applications to Radar* (McGraw-Hill, New York)

CHAPTER II

Evidence for an incoherent ranging mechanism in FM-bats

Introduction

In bat sonar research a lot of attention has been paid to cross-correlation models taken from radar and man-made sonar to describe the performance of these animals in range discrimination tasks (Strother, 1961; Simmons, 1973; Simmons, 1979; Simmons, 1990). However, most neurobiologists have implicitly endorsed the type of ranging mechanism, the “filter bank” receiver, treated computationally by Hackbarth (1986) and Menne (1988a).

In a recent psychophysical experiment, Masters and Raver (2000) elegantly address the problem of which of the hitherto suggested receiver categories best models the bats’ ranging performance.

In these experiments the bats (big brown bats, *Eptesicus fuscus*) were rewarded for indicating if the delay of triggered echoes was shorter than a standard delay. The echo signals were a standardized version of the call they naturally emitted in the same situation. Several of the parameters in the model echoes were then changed to determine the effect on the ranging capabilities of the bats. Of particular interest here is the case where the function describing the instantaneous frequency in the fundamental of the sweeps was modified. These changes consisted of either an equally large shift of all frequencies (up or down), a Doppler factor of 1.1 or a modified “curvature” (both directions) The instantaneous frequency was modeled as a decaying exponential and the curvature was defined as the time constant.

Only a modification of the signal curvature had any significant effect on the delay resolution thresholds found. The authors conclude that no existing model of ranging in bats can account for this result. This may not be so. Below we will present an analysis which emphasizes the degree to which the most popular receiver models can be told apart, and it is pointed out that sensitivity to a change in the instantaneous sweep rate affects a so-called “filter bank model” but not a coherent cross correlation receiver (or the SCAT-receiver (Saillant et al., 1993; Peremans and Hallam, 1998) for that matter, since this is

also a coherent cross-correlation receiver (see Chapter III)).

We will also discuss an alternative possibility, namely that more specialized sensitivity to sweep rates has evolved as a mechanism for jamming avoidance, and that the Doppler tolerant waveforms (near-hyperbolic FM-sweeps, Kroszczyński, 1969) might be an adaptation to this mechanism.

Analysis

We will show that a *coherent* cross-correlation receiver may indeed very well operate on a filter bank output and that the filters making up the bank are only very mildly restricted. We will also point out that the “filter bank” receiver model (Menne & Hackbarth, 1986) is one type of cross-correlation receiver, where on the other hand the filter design is crucial for performance (Menne, 1988). In doing so we shall bring the measurements of time of the peaks in the channels on a form that is in itself a function of time that we can then cross-correlate.

The cross-correlation receiver (Papoulis, 1962),

$$r_{pe}(\tau) = \int_{-\infty}^{\infty} p(t) \cdot e(t+\tau) dt \quad (\text{II-1a})$$

is a coherent receiver without information loss if the functions of time p (pulse) and e (echo) describe pressure values (in acoustics that is). We may see intuitively that if e and p are only differing by a delay, α , and an attenuation factor, then the function has its highest value at $\tau=\alpha$ (see e.g. Altes, 1981). It will be an advantage later if we now introduce the digital variant,

$$r_{pe}(d) = \sum_{n=-\infty}^{\infty} p(n) \cdot e(n+d) \quad (\text{II-1b})$$

using n to indicate sample number instead of time, t , and d instead of τ as the delay variable. Here, the functions e and p may in this expression be read as the correctly sampled identically named functions of time in (II-1a).

Suppose that we have a filter bank consisting of M filters, so that for all frequencies, f of significant energy in e or p , the property

$$\sum_{i=1}^M |H_i(f)|^2 = k \quad (\text{II-2})$$

holds. We could state this as “the summed energy spectrum of the impulse response of the individual filters, $h_1 \dots h_M$ is flat”. We let both pulse and the echo pass through a filter, h_i in the bank and form the cross-correlation between the two outputs (responses to the pulse

and response to the echo) and then define

$$o_{i,p}(n) = h_i(n) * p(n) \quad (\text{II-3a})$$

and

$$o_{i,e}(n) = h_i(n) * e(n) \quad (\text{II-3b})$$

where * means convolution. In the frequency domain we can express the cross-correlation between these two as

$$R_{o_{i,p} o_{i,e}}(f) = [H_i(f) \cdot P(f)] \cdot [\overline{H_i \cdot E(f)}], \quad (\text{II-4})$$

where uppercase symbols indicate that we are dealing with the Fourier transform of the signals, and $\overline{\quad}$ indicates a complex conjugation. Note specifically that R is the Fourier transform of the cross-correlation result, r . Doing this for all channels and summing over all filters we get

$$\begin{aligned} \sum_i R_{o_{i,p} o_{i,e}}(f) &= \sum_i |H_i(f)|^2 \cdot P(f) \cdot \overline{E(f)} \\ &= k \cdot P(f) \cdot \overline{E(f)} \\ &= k \cdot R_{pe}(f) \end{aligned}$$

Since Fourier transformation is a linear process, we can finally write the equivalent operations in the time (delay) domain:

$$\sum_i r_{o_{i,p} o_{i,e}}(d) = k \cdot \sum_i \sum_n o_{i,p}(n) \cdot o_{i,e}(n+d) \quad (\text{II-5})$$

$$= k \cdot r_{pe}(d) \quad (\text{II-6})$$

So if the condition (II-2) is fulfilled, one must simply sum the cross-correlated outputs of the filters in the bank to get the overall cross-correlation of the input functions p and e , and this is of course true for both frequency and time domain. This may all seem somewhat trivial and the only important thing to notice is that the condition (II-2) sets no constraints on the number of filters, M , or on the degree to which the filters overlap. For instance in the limit $M=1$ there is no real filter bank, as this is an all-pass filter (since it must still obey (II-2)) in the frequency range of interest, and this leads then directly to (II-6).

We now turn to what may be dubbed the “classical” filter bank receiver model as described by Hackbarth (1986), which was termed “cross-channel average” by Masters and Raver (2000). The term “average” is somewhat unfortunate, since there are several other ways of evaluating the collection of estimates from the filters. In light of the considerations above the term “filter bank receiver” seems ambiguous and fairly non-descriptive, hence the quotation marks whenever the term is used above, but it will be used here for lack of anything better (abbreviated FB) and – with a definition coming up – without the quotation marks.

This receiver uses – not just copes with – a FB and forms the envelope, $|\cdot|$, of the individual outputs. In each filter the delay between the peak output for the pulse and echo is measured and the estimates are displayed in a histogram.

One way to do this is first for each channel to place a “1” at the time of peak power for echo and for the pulse. This is conveniently done by shifting now to a digital version of the signals (by sampling them correctly) and applying the Kronecker delta function, $\delta(\cdot)$, so that the sample of peak intensity becomes 1. One could for instance apply the following way of converting the now digital output to the desired time marker¹:

$$\hat{o}_{i,p}(n) \equiv \lim_{l \rightarrow \infty} \left\{ \frac{|o_{i,p}(n)|}{\max[|o_{i,p}(n)|]} \right\}^l \quad (\text{II-7})$$

with n as sample number, which can also be written as

$$\hat{o}_{i,p}(n) = \delta(n - \text{peaksample}_{|o_{i,p}|}) \quad (\text{II-8a})$$

for the responses to the pulse, and

$$\hat{o}_{i,e}(n) = \delta(n - \text{peaksample}_{|o_{i,e}|}) \quad (\text{II-8b})$$

for the responses to the echo. This is perhaps a rather clumsy way to write a simple thing. It reads “the output of the i th channel is a “1” at the sample of maximum envelope amplitude”. Now we cross-correlate these functions and sum over all filters:

$$\sum_i r_{\hat{o}_{i,p} \hat{o}_{i,e}}(d) = \sum_i \sum_n \hat{o}_{i,p}(n) \cdot \hat{o}_{i,e}(n+d) \quad (\text{II-9})$$

Comparing this expression to (II-5) it should be evident that this indeed is a digital cross-correlation process. It is also a histogram over delay estimates in the channels,

$$\begin{aligned} &= \sum_i \sum_n \delta(n - \text{peaksample}_{|o_{i,p}|}) \cdot \delta(n - \text{peaksample}_{|o_{i,e}|} + d) \\ &= \sum_i \delta(n - \text{peaksample}_{|o_{i,p}|} + \text{peaksample}_{|o_{i,e}|}) \\ &\approx \text{histogram}(d) \end{aligned} \quad (\text{II-10})$$

since (eq. II-10) is a delta function delayed by the number of samples separating the peak of the echo from the peak of the pulse in this filter channel. So summing over all filter channels produces something like the promised histogram with a bin-width equal to the sampling interval in this case, which is yet another bonus of the digital variant.

¹ This definition would not have been possible if we had stuck to the analog signals.

If the index variable i can be read as an ordered frequency parameter, (eq. II-9) can be seen in its form to be a digital spectrogram correlation carried out along the time axes only (for a practical example see Mellinger & Clark (2000)). To point out this fact is the purpose of introducing the rather clumsy definitions (eq. II-8), allowing us to see the time markers as the transformation (eq. II-7) of the filter outputs, which again allows us to express the measurement of delays as the correlation process (eq. II-9).

So, the FB receiver can be seen as a cross-correlation receiver where a considerable amount of information has been removed from the input vector.

If the envelopes of the filter outputs were left untouched (not reduced to a peak-time marker), we would have (eq. II-9) as a spectrogram correlation receiver (Mellinger and Clark, 2000; Masters and Raver, 2000), which clearly leaves more information intact. We shall not discuss this receiver further, but note without further ado that it has a dependence on filter bank design, which property we now shall proceed to investigate for the FB receiver. The incoherent spectrogram correlation receiver was never discussed to any great extent in the bat literature, but it has been shown that the autocorrelation function can in principle be retrieved from a spectrogram representation. (Altes, 1980, 1981).

In this context one important qualitative difference exists between the FB receiver (eq. II-9) and the coherent cross-correlation receiver (eq. II-1), namely the independence of FB design noted as a property of the coherent version above, which is gone from the FB receiver. This can be illustrated by once again considering the limit case of $M=1$, where the FB receiver reduces to a peak signal envelope receiver (touched on by Hackbarth, 1986), which is not particularly sophisticated.

Menne (1988a; 1988b) has treated the subject of filter bandwidth. He showed in his 1988a analysis that with a FB receiver, a given bandwidth and a given sweep rate there exists an optimal filter bandwidth to use in the filter bank for target ranging purposes.

That result builds among other things on the fact that the narrower the band of the filters, the more independent estimates fit into a given bandwidth. We here extend the theoretical findings of Menne (1988a) by showing in a simulated example that given one filter (or any fixed number) there exists a region of SNRs, where deviations from the optimum sweep rate is relatively costly in terms of an increase in the standard deviation of the timing estimates (Fig. II-1).

The differences to the analysis by Menne (1988a) arises because we have included the “Barankin region” where SNRs are so low that the established relationships between SNR and the accuracy limit of time determinations are no longer valid (Boman, 2000). It

is very difficult to treat this region analytically, hence the simulation method.

Discussion

Deciding between the receivers

It should be clear from the above analysis that it is by no means a simple matter to decide between the proposed receiver types in bat sonar research. Almost any call parameter change will affect the different models in the same direction, since they are all some form of cross-correlators. It thus seems difficult to design an experiment that qualitatively decides between the receivers, since they differ only by the amount of information excluded from the filter bank output functions.

One nice example is the experiment by Møhl (1986) concerning detection, showing the threshold to be independent of a change in sweep direction, which is in direct contradiction to the predictions of any of the above-mentioned receivers (apart from the peak signal envelope detector). Masters and Jacobs (1989) later confirmed the observation but extended the experiment by testing the degree to which ranging performance was also unaffected by this dramatic phase transformation. The performance degraded considerably and based on this they conclude that a degree of matching is present.

But this is really a badly formed question. If these two tasks are compared with a fictitious human psychophysics experiment, detection could be the equivalent to this situation: The experimental subject says “now” and “won” is returned. The person detects the letters and answers correctly that a word was returned. Ranging, in this analogy is comparable to the following situation: The person says “now” and is supposed to detect exactly when that letter sequence returns. But again what is returned is a “won”. The task is clearly not well defined, since the “correct” sequence is never returned.

A coherent cross-correlation receiver offers a special, highly efficient tool for answering the question of the presence of the transmitted “letters”, which is clearly not relevant to bats. In a ranging task any modification to the signal structure will degrade performance, and this is the case for almost any receiver, as this example is meant to show.

Another problem is the quantitative discrimination between the receivers. Here the problem is that one could claim that a given less than optimal detection or ranging performance was limited by internal noise and that the underlying receiver could in fact be coherent as well as semicoherent. But as an analogy to the phenomenon of sampling

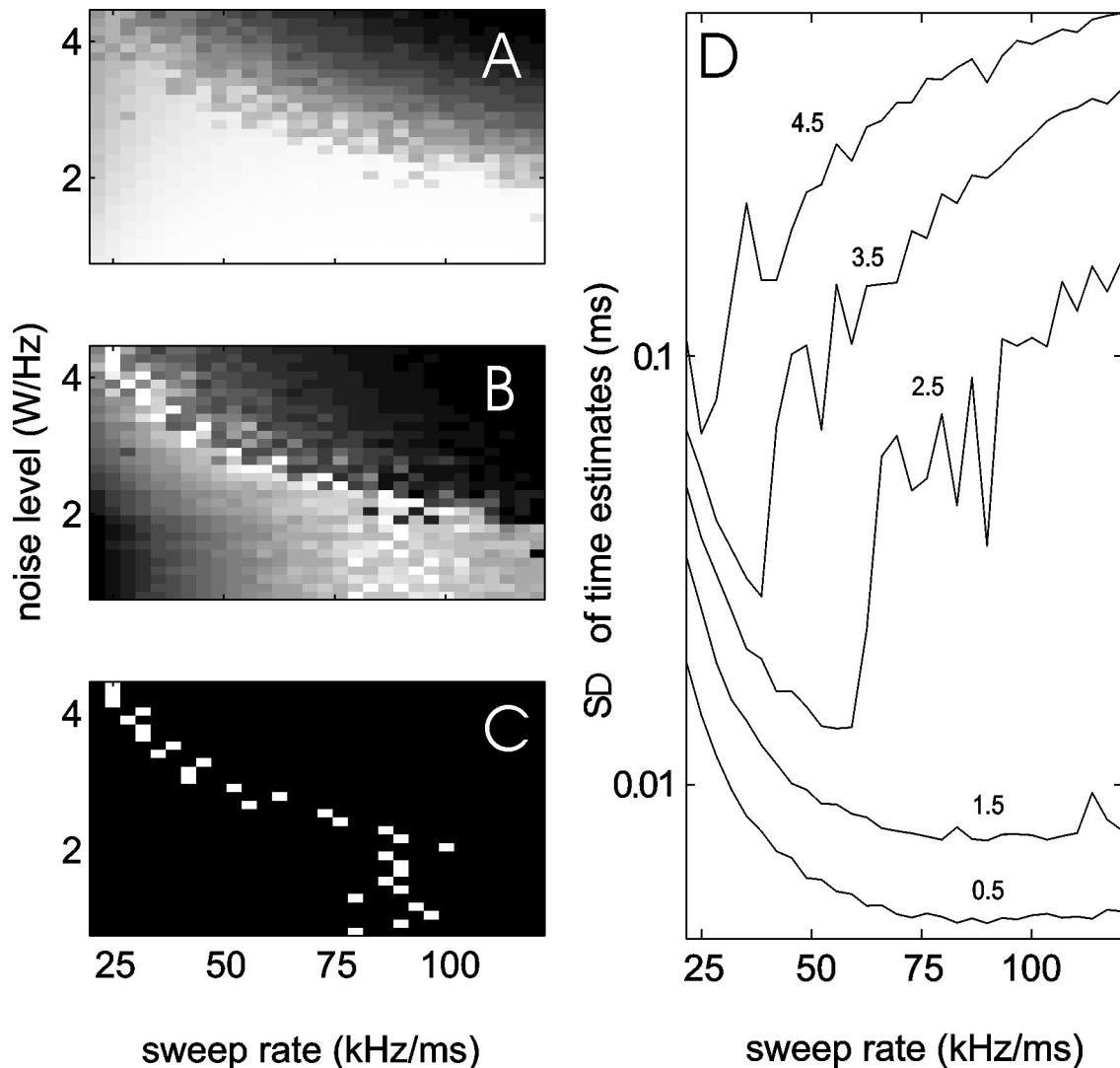


Fig. II-1: Simulation study: Standard deviation of peak time determination of the envelope of the output from a bandpass filter as a function of the noise level and of the sweep rate of the input signal. The filter was centered on 33.5 kHz and had a $Q_{10\text{dB}}$ of 4.8. The signal was a linear sweep from a variable start frequency down to 10 kHz over 2 ms. “Sampling rate” was 500 kHz. The sweep had fixed power, which is the more relevant factor here. Keeping the total signal energy constant would not have reflected the limitations that bats face, since the sounds would become unnaturally intense, as they became very short.

A: Untreated data SD data. Darker color indicates higher standard deviation values. B: Same as A but normalized so that for each noise level the lowest SD value is set to 0 and the highest is set to 1. C: Same as B but all minima set to zero and all other values set to 1. Note that below 2 W/Hz the minima occur at about the same sweep rate. Above ca. 2 W/Hz the “optimal” sweep rate decreases with increasing noise power. D: SD dependance of signal sweep rate for five noise powers (indicated above traces). Note logarithmic ordinate axis.

Other filter parameters would have resulted in another position of the landscape in sweep rate - noise power space, but the general shape does not change.

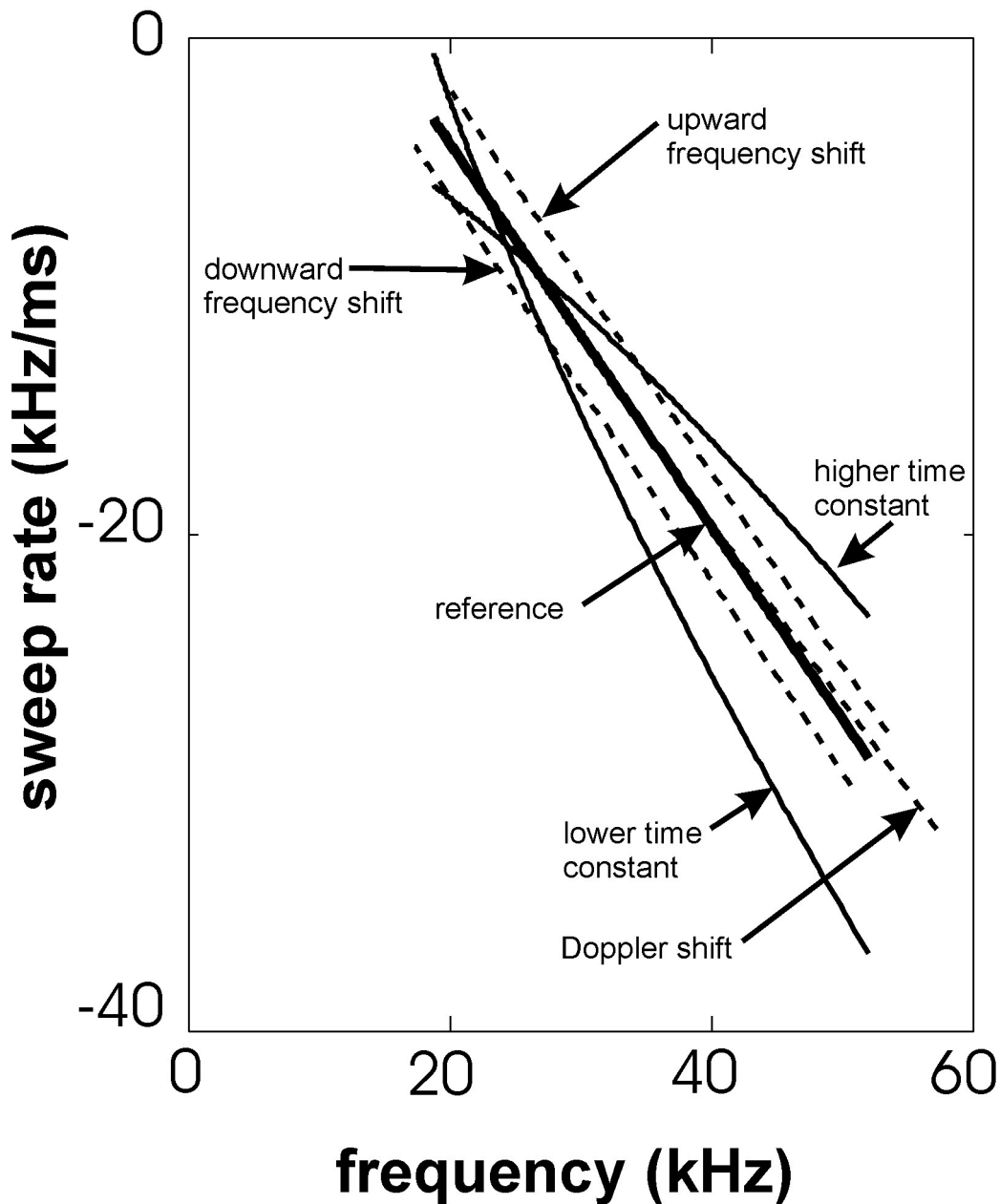


Fig. II-2: Sweep rates as a function of frequency for the different manipulated signals used in Masters and Raver (2000) as indicated with arrows. The “time constant” is the decay constant for the exponential function that described the time frequency course of the signals in Masters and Raver (2000). The graphs were calculated by manipulations of the formulae given by Masters and Raver (2000) – and their derivatives – with the different experimental parameters given for one of the bats (“Giggles”). The time constant changes led to an increased amount of errors by the bats in Masters and Raver (2000). It is seen in this graph that the change of time constant also led to the most severe changes in local sweep rates of the signals.

noise, the reduction in information content that results from a process like envelope formation or peak detection can be viewed as equivalent to adding internal noise to the signal. However, this noise cannot be assumed to be white and Gaussian.

Performance dependence on filter design

The above treatment and the simulation results shown in Fig. II-1 indicate that sensitivity to sweep rate changes or lack of such a phenomenon occurring would indicate that incoherent or coherent reception was involved, respectively.

In other words: if the sweep rate as seen by the individual filters is modified whilst holding other signal parameters constant a test might be possible that could decide between coherent and non-coherent reception under the almost trivial assumption that pre-processing, in which the signal is split up into frequency components in a filter bank, has taken place.

As stated in the introduction, one purpose of this letter is to show that Masters and Raver (2000) have already elegantly made such an investigation. The three most interesting echo manipulations used by these authors involved modifications of the time-frequency structure of the echoes. These were made in a way so that the width of the noise-free cross-correlation function were approximately the same, or – equivalently – the variation in the delay estimates across individual filters in a FB receiver was similar for these (still noise-free) signals.

In absolute accordance with the fact that all the receivers treated here are in some way cross-correlation receivers, Masters and Raver (2000) found no quantitative difference in the way these would be affected by the manipulations. The triggering variation of playback of the signals was found to be rather larger than the thresholds of ca 35 μ s found, so it seems reasonable to assume that this factor limited performance.

In fact, since no noise was added to the signals they all should in principle produce a limit approaching infinitely accurate delay measurements: The manipulations made do not change the reproducibility of a given estimate unless noise is added (Menne & Hackbarth 1986; Hackbarth 1986; Menne 1988). So what we are looking for to explain the results is not a mechanism which enable the bats to cope with a given type of modification, but something that causes it to produce larger errors than the trigger problem.

In Fig. II-2 is shown the instantaneous sweep rates for the first harmonic of the four signals – the manipulated ones and the reference – as a function of frequency. The data for the two signals in which the time constant (“signal curvature”) of the exponential

time-frequency course has been modified clearly stand out from the others in the region of the higher frequencies. So one interpretation of these experiments could indeed be that instantaneous sweep rate does seem to be a decisive parameter in range discrimination by Vespertillionid bats.

This interpretation, however, is probably only true in a rather qualitative manner. The actual size of the influence on the accuracy of a given (realistic) filter bandwidth is rather limited over a large range of sweep rates and SNRs, and it will take “a good deal” of internal noise to increase the errors beyond the 35 μs that was found to be the threshold for the unmodified signals. This fact casts some shadows, not only on the present interpretation of the result of Masters and Raver (2000), but also on what may be dubbed “the matched filter bank theory” concerning why many Vespertillionid bats’ calls are hyperbolic or near-hyperbolic (Menne, 1988a, 1988b).

Thus it would seem that the effect on delay estimation accuracy observed by Masters and Raver (2000) might be too large to be just the result of filter-bandwidth vs. sweep rate mismatch. One could instead speculate that the bats used the instantaneous sweep rate as a way of filtering out irrelevant signals.

To pave the way for an explanation of this idea, we next discuss the Doppler tolerant waveform.

Doppler-tolerance

A lot of attention has been given to the notion of Doppler tolerance in bats that use hyperbolic or linear period modulated (LPM) calls (Kroszczyński, 1969; Altes and Titlebaum, 1971). This has a direct connection to the discussion of pulse-compression in bats, since the way this parameter has been measured has been with the help of ambiguity diagrams (Cahlander, 1964; Masters et al., 1989), which for a Doppler factor of one is equal to the autocorrelation function and therefore the noise-free cross-correlation function.

Kroszczyński (1969) who was the first to notice that the LPM signal is in fact Doppler tolerant – a notion inspired by the call structure of bats – writes that at 22 m/s (~50 mph) some bats would experience up to 10 dB reduction in SNR. If we assume a SNR at the receiver of 40 dB, the bat with the signal parameters in the example used by Kroszczyński would for a coherent receiver have a standard deviation in its range estimate of 0.4 μm . Reducing the SNR by 10 dB worsens the accuracy to 4 μm . This is hardly a key factor in the evolution of the time frequency structure of bat calls. It is also difficult to imagine a situation where a Vespertillionid bat experiences range rates of this

magnitude.

However, the property that gives the hyperbolic signal its Doppler tolerance is the fact that the sweep rate at any given frequency does not change as a function of the amount of Doppler shift introduced.

This property can be demonstrated as follows: Let $f(t)$ describe the frequency at time t and let sf be the start frequency, ef the end frequency and D the Doppler factor. We can now express the echo of an LPM signal as

$$f(t) = \frac{D}{D\left(\frac{1}{ef} - \frac{1}{sf}\right)t + \frac{1}{sf}} \quad (\text{II-11})$$

The inverse function is then

$$t(f) = \frac{\frac{D}{f} - \frac{1}{sf}}{\left(\frac{1}{ef} - \frac{1}{sf}\right)D} \quad (\text{II-12})$$

and the first derivative is

$$f'(t) = \frac{-D^2}{\left[D\left(\frac{1}{ef} - \frac{1}{sf}\right)t + \frac{1}{sf}\right]^2} \left(\frac{1}{ef} - \frac{1}{sf}\right) \quad (\text{II-13})$$

We seek an answer to the question: what is the first derivative at any given frequency, i.e. at any given moment as a function of frequency. We therefore express (II-13) as a function of (II-12) which gives at first

$$f'(t(f)) = \frac{-D^2}{\left[D\left(\frac{1}{ef} - \frac{1}{sf}\right)\left(\frac{\frac{D}{f} - \frac{1}{sf}}{\left(\frac{1}{ef} - \frac{1}{sf}\right)D} + \frac{1}{sf}\right)\right]^2} \left(\frac{1}{ef} - \frac{1}{sf}\right) \quad (\text{II-14})$$

which immediately reduces to

$$= -f^2 \left(\frac{1}{ef} - \frac{1}{sf}\right). \quad (\text{II-15})$$

In the surprisingly simple expression (II-15) the Doppler factor, D , no longer appears, and so Doppler does not influence the local sweep rate of a returning echo, as long as the transmitted signal is LPM.

The above special property of the LPM signal, which is in a way also the “secret” behind its Doppler tolerance, was what lead us to ask the question: apart from the biologically irrelevant need for extremely high accuracy in the delay measurements, does this property offer any alternative reasons for the use of this signal type?

One such possibility is the direct involvement of sweep rates in discriminating the bats own emissions from that of other bats, conspecific or not. For the development of this idea a comparison with the CF-FM bat *Pteronotus parnellii* is useful. These bats

emit a powerful second harmonic and a comparatively weak fundamental. However, the ranging mechanism as reported by Suga and coworkers (Suga et al., 1978; Suga and O'Neill, 1979; O'Neill and Suga, 1982) seems to indicate that the “ranging neurons” are determining the delay between the first and the higher harmonic, so that the first harmonic starts the stopwatch and the second harmonic (or higher) stops it. It was suggested for these bats, that a jamming avoidance mechanism is present in the fact that the bat itself is the only one likely to hear the emitted weak fundamental, so that only the bat itself can trigger the neural measuring apparatus.

To a Vespertilionid bat, this “trick” is not available, since these bats rely on powerful first harmonic signals. The frequencies contained in the higher harmonics occur several times during the call (see Chapter IV), so these are not unambiguous time markers.

We therefore suggest that one way to obtain an advantage similar to what *Pteronotus* enjoys, is using the local sweep rate as a form of “stamp” indicating to which echo and harmonic a given event in each channel belongs.

Since FM bats change their call structure continuously during most of their hunting sequence, it is rather unlikely that another bat, close enough to be potentially a jamming risk, will be at exactly the same stage of the hunt and thus use the same sweep rates at the same frequencies.

Another ecological driving factor that might promote the use of such a system is the clicks emitted by some moths (Blest et al., 1963; Dunning, 1965). These are by some authors (Fullard et al., 1979; 1994) believed to pose a risk of jamming the bats' ranging mechanism (Miller, 1991). But if the sweep rate is critical, these sounds are filtered out, since they have essentially an infinite sweep rate.

With these scenarios there is in principle no limit to the usefulness of sharpening of the putative “sweep-rate-filters”. Additional potential jamming situations can be coped with, the sharper the sweep-rate filter is. However, an evolutionary hurdle arises when the filters are sharp enough to exclude slightly Doppler shifted versions of the emitted calls. This is when the near-hyperbolic or LPM signal type becomes attractive, since the mechanism suggested here allows for this waveform to be Doppler shifted without affecting the filter criterion.

The errors that the bat must experience (or perhaps cope with) due to the fact that they are travelling between emitting the call and receiving the echo, are much larger than the bias introduced by the signals being shifted slightly (Boonman et al., 2003). So what we propose here seems to us a rather more plausible explanation for the Doppler tolerant

waveform.

If this theory were correct, one would predict that there exists a general sweep rate map in the cortex of the bat. For the present question the interpretation would be that the bats in the study by Masters and Raver (2000) fail to perform as well as with the other modified echo types, because part of the echoes returned tend to fall outside the range of local sweep rates that are accepted by the neural system as belonging to echoes of the bats' own cries.

As an observation that does corroborate this idea, we note the fact that most bats that have a CF component in their calls use linear sweeps in the FM part of their echo-location signals. It would be in line with our idea, if these animals were able to identify their own calls by means of a "personal" CF frequency and thus not need the sweep rate filters.

Sweep rates and hunting behavior

As an aside we note that the result shown in Fig. II-1 points to a role in Vespertilionids for the continuously shortening signal durations during approach to a target observed by most if not all bats (e.g. Schnitzler et al. 1987; Hartley, 1989).

As noted and discussed above, Menne (1988a; 1988b) argued that there exists an optimum sweep rate for a given auditory filter. We find that the optimal sweep rate is actually dependent on the noise level, when a single filter is considered. But since the FM-bats have only a single component in their calls, a downward frequency modulation, they increase the local sweep rate, as the target gets nearer (certainly if start and end frequencies are kept constant at least). It is evident from Fig. II-1 that at bad reception conditions (high noise levels) it is "optimal" to use relatively low sweep rates. With decreasing noise levels the optimal sweep rate increases until the minimum described by Menne (1988a) is reached and then this then stays constant.

For the bats, with increasing proximity of the prey, signal to noise ratio is bound to increase, and this could then be seen as one reason for the shortening of the calls as the target is approached. The calls emitted during the buzz, when the target is very close, do not seem to shorten.

Conclusion

Evidence of an effect that might be attributable to mismatched sweep rates through the auditory filters of the bats has been shown to exist in the literature. We have in a simulation experiment shown that a region of noise levels exists where a well-defined optimal (for time determination) sweep rate can be discerned. Together these findings point to the existence of an incoherent FB mechanism in Vespertillionid bats.

As an alternative interpretation we show that hypersensitivity to mismatched local sweep rates might be the result of an anti-jamming strategy promoting the development of Doppler sensitive echolocation signals.

Finally, it was proposed that the increase in sweep rates (decreasing signal duration) seen in hunting Vespertillionid bats as they approach objects or prey might be an adaptation to high (presumably internal) noise levels.

References

- Altes, R. A. (1980). "Detection, estimation, and classification with spectrograms" J. Acoust. Soc. Am. **67**, 1232-1246.
- Altes, R. A. and Titlebaum, (1970). "Bat signals as optimal Doppler tolerant waveforms," J. Acoust. Soc. Am. **48**, 1014-1020.
- Altes, R. A. (1981). "Echo phase perception in bat sonar?" J. Acoust. Soc. Am. **69**, 505-508.
- Blest, A. D., Collett, T. S. and Pye, J. D (1963). "The generation of ultrasonic signals by a New World arctiid moth," Proc. R. Soc. Lond. B **158**, 196-207.
- Boman, K. (2000). "Low-Angle Estimation: Models, Methods and Bounds," Licentiate thesis, Uppsala University.
- Boonman, A. M., Parsons, S. and Jones, G. (2003). "The influence of flight speed on the ranging performance of bats using frequency modulated echolocation," J. Acoust. Soc. Am **113**, 117-128.
- Dunning, D. C. and Roeder, K. D., (1965). "Moth sounds and the insect catching behaviour of bats" Science **147**, 173-174
- Fullard, J. H., Fenton, M. B. and Simmons, J. A. (1979). "Jamming bat echolocation: the clicks of arctiid moth," Can. J. Zool. **57**, 647-649
- Fullard, J. H., Simmons, J. A. and Saillant, P. A. (1994). "Jamming bat echolocation: the dogbane tiger moth *Cycnia tenera* times its clicks to the terminal attack calls of the big brown bat *Eptesicus fuscus*," J. Exp. Biol. **194**, 285-298.
- Griffin, D. G. (1958). *Listening in the Dark* (Yale University Press, New Haven).
- Hackbarth, H. (1986). "Phase Evaluation in Hypothetical Receivers Simulating Ranging in Bats," Biol. Cybern. **54**, 281-287.
- Hartley, D. J. (1992). "Stabilization of perceived echo amplitudes in echolocating bats. II. The acoustic behavior of the big brown bat, *Eptesicus fuscus*, when tracking moving prey," J. Acoust. Soc. Am. **91**, 1133-1149.
- Kroszczyński, J. J. (1969). "Pulse Compression by means of Linear Period Modulation," Proc. IEEE **57**, 1260-1266.
- Masters, W. M., Jacobs, S. C. (1991). "The structure of echolocation sounds used by the big brown bat *Eptesicus fuscus*: Some consequences for echo processing," J. Acoust. Soc. Am. **89**, 1402-1413.

- Masters, W. M., Raver, K. A. S. (2000). "Range discrimination by big brown bats (*Eptesicus fuscus*). using altered model echoes: Implications for signal processing," J. Acoust. Soc. Am. **107**, 625-637.
- Mellinger, D.K. & Clark, C.W. (2000). "Recognizing transient low-frequency whale sounds by spectrogram correlation," J. Acoust. Soc. Am. **107**:3518-3529.
- Menne, D (1988a). "A matched filter bank for time delay estimation in bats," in *Animal Sonar Processes and Performance* (Eds Nachtigall, P. E. and Moore, P.W. B.). Plenum Press, New York.
- Menne, D (1988b). "Is the structure of bat echolocation calls an adaptation to the mammalian hearing system?" J. Acoust. Soc. Am. **83**, 2447-2449.
- Menne, D. and Hacbarth, H (1986). "Accuracy of distance measurements in the bat *Eptesicus fuscus*: Theoretical aspects and computer simulations," J. Acoust. Soc. Am. **79**, 386-397
- Miller, L. A. (1991). Arctiid moth clicks can degrade the accuracy of range difference-discrimination in big brown bats (*Eptesicus fuscus*). J. Comp. Physiol. A **168**, 571-579.
- Møhl, B. (1986). "Detection by a *Pipistrellus* bat of normal and reversed replica of its sonar pulses," Acoustica **61**, 75-82.
- O'Neill, W. E., Suga, N. (1982). "Encoding of target range and its representation in the auditory cortex of the mustached bat." J. Neurosci. **2**, 17-31.
- Papoulis A (1962). *The Fourier Integral and its Applications* (McGraw-Hill, New York) pp. 240-242.
- Peremans, H. and Hallam, J. (1998). "The SCAT receiver revisited," J. Acoust. Soc. Am **104**, 1101-1110.
- Saillant, P.A., Simmons, J.A. and Dear, S.P. (1993). "A computational model of echo processing and acoustic imaging in frequency modulated echolocating bats: The spectrogram correlation and transformation receiver," J. Acoust. Soc. Am. **94**, 2691-2711.
- Schnitzler H.-U., Kalko, E. K. V., Miller, L. A. and Surlykke, A. (1987). "The echolocation and hunting behavior of the bat *Pipistrellus kuhli*," J. Comp. Physiol. A **161**, 267-274.
- Simmons J. A. (1973). "The resolution of target range by echolocating bats," J. Acoust. Soc. Am. **54**, 157-173.
- Simmons J. A. (1979). "Perception of Echo Phase Information in Bat Sonar," Science **204**, 1336-1338.

- Simmons, J. A., Ferragamo, M., Moss, C. F., Stevenson, S. B., Altes, R. A. (1990). "Discrimination of jittered sonar echoes by the echolocating bat, *Eptesicus fuscus*: The shape of target images in echolocation," J. Comp. Physiol. A **167**, 589-616.
- Strother, G. K. (1961). "Note on the Possible use of Ultrasonic Pulse Compression by Bats," J. Acoust. Soc. Am. **33**, 696-697.
- Suga, N and O'Neill, W. E. (1979). "Neural axis representing target range in the auditory cortex of the mustache bat," Science **206**, 351-353.
- Suga, N., O'Neill, W. E. and Manabe, T. (1978). "Cortical neurons sensitive to combinations of information-bearing elements of biosonar signals in the mustache bat," Science, **200**:778-781.

CHAPTER III

The “SCAT” receiver model of bat echolocation forms a coherent cross-correlation between input vectors

Introduction

A standing discussion in the field of bat echolocation is concerned with whether the animals might employ a cross-correlation receiver for use in ranging tasks (e.g. Strother, 1961; Altes, 1981; Simmons, 1973; Pye, 1986) and perhaps other perceptual tasks as well (Simmons et al., 1990b).

The alternative hypothesis to the mathematically simple but physiologically exceedingly extravagant coherent cross-correlation receiver has mainly been the relatively loosely formulated assumption that bats determine distance to a target by measuring the timing difference of the responses to an outgoing cry and returning echo independently in several auditory filters. This idea is perhaps most explicitly formulated by Hackbarth (1986), and is usually referred to as the “filter bank model” (hereafter FB). The two models, however, are far from being mutually exclusive, and the real difference between them lies solely in the amount of information discarded immediately after the “peripheral” processing, i.e. the band-pass filtering (see Chapter II).

Thus, the filter bank receiver is also a cross-correlation receiver, albeit with phase information irretrievably discarded, and the cross-correlation function (CCF) can readily be formed from the output of a filter bank living up to certain rather trivial criteria (see chapter II again).

Saillant et al. (1993) have proposed the so-called spectrogram correlation and transformation (SCAT) receiver to explain and model findings by J. A. Simmons et al. (1990a), interpreted to mean that echolocating bats use a coherent cross-correlation receiver to determine the target range and transform any spectrally derived information to the a range axis as well.

The SCAT receiver consists of two parallel blocks after a common initial filter

bank (called the “cochlear block”). One operates on the envelopes of the output from the filters in the bank (spectrogram transformation or “spectral block”) and one operates on “spikes” generated as response to the ripples on the envelopes (temporal block).

The spectral block of this receiver shall not be treated to any large extent here. For the sake of completeness, however, we note that the input to this part of the receiver is the one-way rectified and subsequently low-pass filtered output from the “cochlear block”. This part of the receiver model operates on the envelopes, without a spike generating mechanism.

As is noted by the authors the temporal block of the receiver is closely associated with the Licklider model (1951) of low-frequency pitch perception and as such with the FB receiver as well. The difference lies in the operation on the basis of very high-frequency phase locked responses. This difference is mathematically unimportant but in terms of physiology it is of course of paramount significance (see below).

Peremanns and Hallam (1998) studied the proposed SCAT algorithm further in terms of performance, but to our knowledge no-one has placed it in a context where its relationship with other previously discussed receiver models for bat echolocation has been analyzed. This task is what we wish to accomplish with the present short communication.

Boonman et al. (2003) has discussed the SCAT model as a filter bank receiver akin to the model of Hackbarth (1986) in a comparison with a coherent cross-correlation receiver. This also appears to be the way Masters and Raver (2000) perceived the essence of this proposed model. Since this is not correct (see below), there seems to be some confusion as to the role of the SCAT receiver in the landscape of receiver models in bat echolocation research, which fact has prompted the present investigation.

It is demonstrated below that the summed cross-correlations between two outputs from the filter bank before peak detection corresponds to the CCF in the usual definition, a result that does not really need a simulation method as it is straightforward to prove (as we do in Chapter II). It is, however, useful for comparison purposes as it illustrates the importance of processes taking place after the initial band-pass filtering.

Analysis

The correlation part of the SCAT model involves phase-locking timing elements that follow a set ($n=81$) of narrow ($5 < Q_{10dB} < 25$) band-pass filters (constant $-10dB$ bandwidth = 4 kHz, $CF=20-100$) simulating auditory filters. The “neurons” are able to fire a

spike whenever a positive peak in the output has a larger amplitude than the preceding one (Peremans and Hallam, 1998; Saillant et al., 1993) and the net effect is that these timing elements peak-detect the output from the auditory filters on the rising slopes of the envelope (see Peremans and Hallam, 1998). This response is then “coincidence detected” (=cross-correlated) and the results of these processes are summed across channels.

We here use a somewhat simplified version of the SCAT receiver in order to show its relationship with an ordinary coherent cross-correlation receiver. The simplifications made, compared with the original, are a reduced number of filters making up the bank,

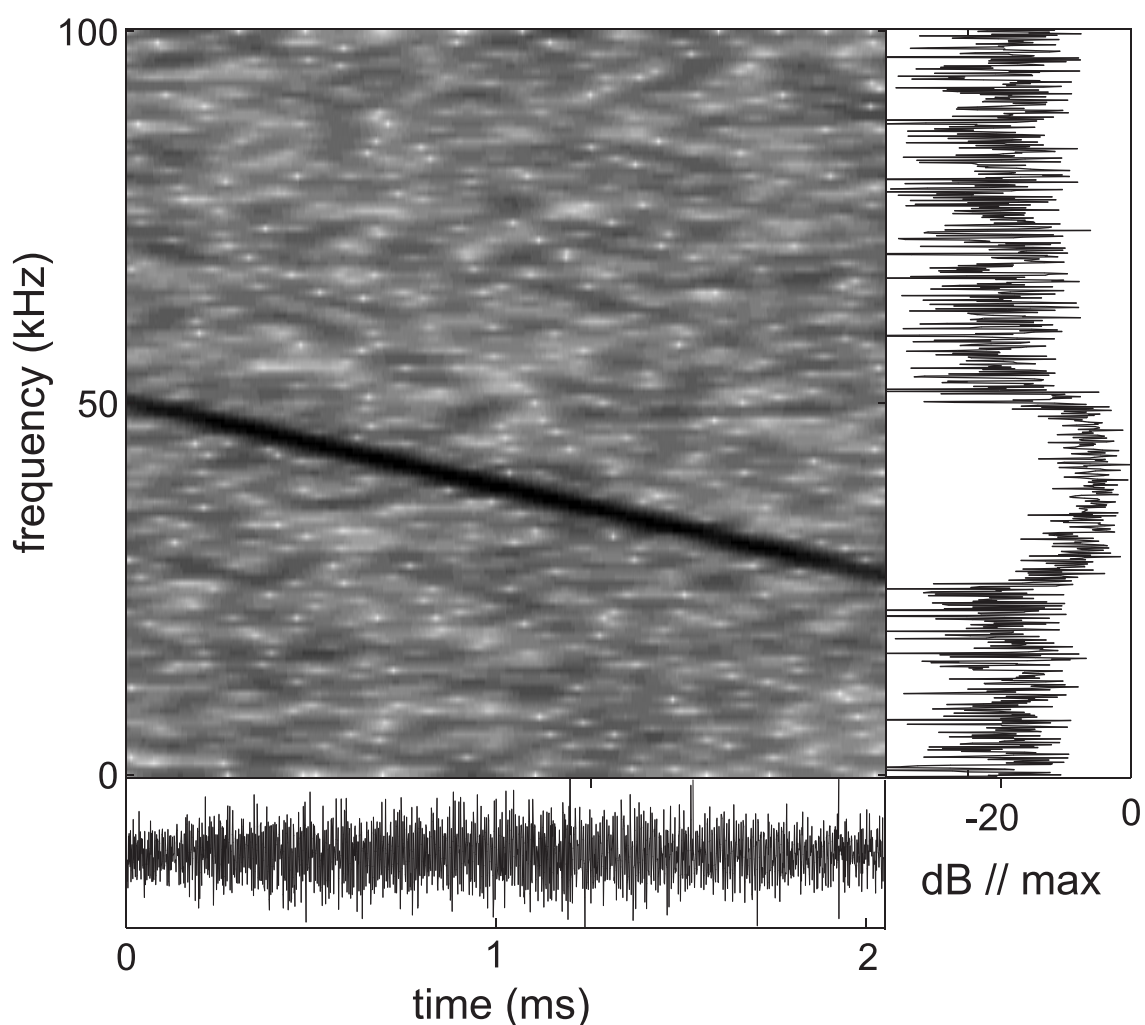


Fig. III-1: The signal used as input to all the receivers investigated here. Bottom panel: time series. Central panel: spectrogram representation. Right vertical frame: amplitude spectrum.

For the purpose of graphical display the signal has been down-sampled by a factor 5 - the actual waveform used as input has 1 Mhz sampling rate.

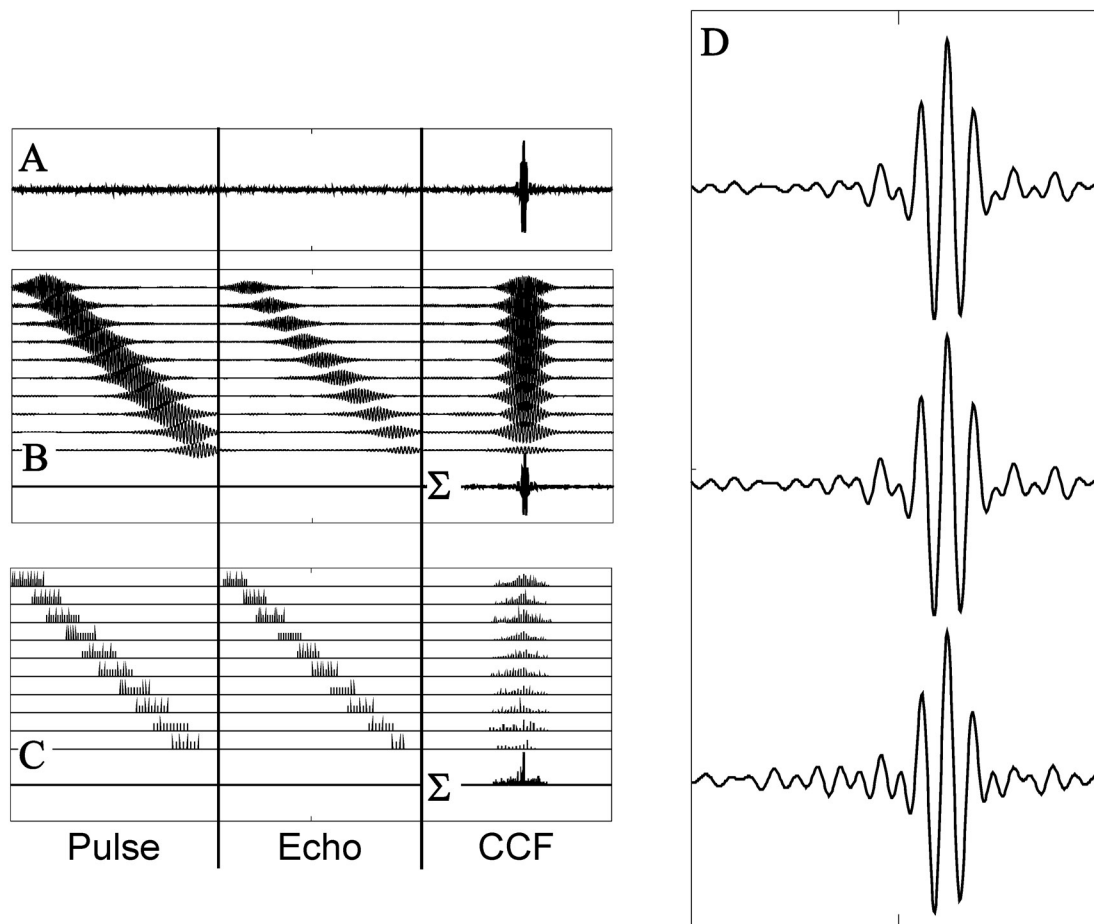


Fig. III-2: Similarities between the outputs from the three versions of the cross-correlation receiver discussed in the text. A: “Normal” cross-correlation process between the pulse and echo parts of the trace. The input signals are not visible in the noise. They are similar to the one seen in Fig. III-1. B: The cross-correlation performed on the output from individual filters in a filter bank of 10 filters that together cover the frequency range of the signals, when stimulated with the signals used in A. C: Same as B but with the filter outputs treated according to the SCAT receiver rules described in the text.

The trace in D at the level of the final zero lines is a blow up of the central part of the resulting functions after varying modifications: Top trace is band-pass filtered to cover the same range as the filter bank in B. Bottom trace is band-pass filtered with the same filter and has been weighted with a factor of $1/f^2$. See text for more explanation.

linear spacing of center frequencies and a slightly simpler “spiking algorithm”, which ignores some aspects of the “refractory period”, since this has mainly symbolic value for the signals considered here. For very slowly sweeping FM signals as input it may be important (see Boonman et al., 2003), but we do not use such input signals here. Also, the threshold for firing a “spike” in the SCAT formulation of Saillant et al. (1993) is dependent on the input noise level, but it is not completely clear to us how it is implemented. We have used a fixed threshold for simplicity.

The filter bank in this investigation consists of 10 filters with fixed bandwidth (5 kHz). Inputs were a “pulse” and an “echo” (=attenuated) version of the signal shown in Figs. III-1 and III-2a with noise added. This signal is a linear downward sweep, spanning ca. 1 octave (50-25 kHz).

As an intermediate step between the “normal” coherent cross-correlation receiver and the SCAT receiver we have in Fig. III-2b included the cross-correlated outputs from the filter bank as they appear before spike generation. The filter outputs as response to the input signals were cross-correlated and the resulting CCFs were summed over all frequency channels (filters) (Fig. III-2b).

For the SCAT receiver, peak detection followed band-pass filtering *on the rising slopes* of the envelope of the individual outputs when a fixed threshold was exceeded. We did not rectify and lowpass filter the output, as it was done in the original SCAT receiver, since this is without consequence as revealed in Peremans and Hallam (1998). The peak detected FB outputs were cross-correlated and the results were summed over channels (Fig. III-2c).

Differences between outputs

Normally, when converting a sampled signal to an analog waveform an interpolation or reconstruction filter is used, which is usually a lowpass filter. For the signals here a band-pass filter is used.

To eliminate any differences between the CCF and the sum of the cross-correlated outputs from the filter bank, the former signal may be bandpass filtered to have the same bandwidth as the frequency range covered by the bank using the reconstruction filter. After this the summed result (middle trace in Fig III-2d) is indistinguishable from the now bandlimited “normal” coherent CCF (upper trace in Fig. III-2d).

In the case of the differences between the “normal” CCF (Fig. III-2a) and the output from the SCAT receiver (Fig. III-2c), they are mostly due to the fact that the SCAT receiver “undersamples” by only marking positive peaks. We reconstruct something very closely resembling the CCF by again band-pass filtering around the frequencies known to be present before “sampling” (peak detecting) and we apply a frequency dependent weight.

The most drastic modification of the SCAT output here to make it similar to the “normal” CCF is this weighting of the spectrum, namely with a factor of $1/f^2$. This has to do with our choice of a linear spacing of filters with a fixed bandwidth. When “sam-

pling” a structured signal by peak detection, then clearly more samples are collected at high signal frequencies than at low frequencies (given a sufficiently high time-bandwidth product) and the weight is shifted towards high frequencies in this process. In Appendix III-A this factor is discussed more thoroughly. For the *cross-correlation* of the spike trains, the factor $1/f^2$ is the relevant size, because the signal has been treated by the weight, f , twice.

It is perhaps not surprising that the SCAT receiver is thus able to deliver the basis for the CCF between pulse and echo, when the “intermediate” result is taken into account that the sum of the individually cross-correlated outputs from a filter bank (with flat summed response) is identical to the CCF between the input signals (Chapter II, Fig. III-a).

What remains is to justify that the peak detecting and thresholding of signals retains most of the information in the outputs.

Reconstruction from peak detected signals

The problem of reconstruction from a peak detected signal clearly has a strong connection to the theorem of Logan (1977) who treats the retrieval of a signal from its zero crossings. (or other level-crossing). That result states that it is theoretically possible to reconstruct a “severely clipped” signal, as long as the bandwidth of the signal does not exceed one octave (and there are no zeros shared between the original signal and its Hilbert transform). It was, however, also found by Logan that the retrieval is very difficult in praxis. (In practically encountered clipped signals the reconstruction may sometimes be straightforward as we show in Appendix III-B).

In the case of the CCF here, the problem is made a lot easier by the fact that a filter bank is involved. This arrangement guaranties that the higher a signal component, the higher the overall number of spikes are elicited across the filters: A powerful frequency component will elicit spikes in the band-pass filters having neighboring center frequencies, whereas a weak component only triggers its specific channel. In this way a degree of sampling resolution above the “one bit” trigger is achieved.

The reconstruction of a peak-detected signal is illustrated in Fig. III-3. We see here that there is also a connection to the sampling theorem of Shannon (1949) stating that the sampling frequency must exceed the reciprocal of the bandwidth (i.e. not the Nyquist frequency) for the sampled representation to be unambiguous.

When sampling the output from a narrow-band filter in this way, the output is thus

weighted almost exactly by a factor equal to the center frequency, hence the correction factor of $1/f^2$ for the CCF discussed above. In the SCAT algorithm this fact is remedied by spacing the constant bandwidth filters hyperbolically (linearly in reciprocal center frequency). This makes the summed response of the filter bank a counterweighing low frequency response to this kind of signals (see Appendix III-A). As we have spaced the filters of our miniature SCAT-receiver linearly (for illustrative purposes) we needed to apply the weight afterwards, if our output should match the “standard” cross-correlation result. This is not necessary in the original SCAT model.

An alternative approach to the hyperbolic spacing in center frequency would have been the use of a constant Q filter bank with equal overlap between neighboring filters¹. Why this – perhaps more physiologically justifiable – approach was not taken by Saillant et al. (1993) is not completely clear to us, but it is presumably due to the demands of the spectral block.

The similarities between the coherent CCF and the SCAT receiver output are shown here for input signals that are limited to one octave. It seems evident that for higher bandwidths the similarities will be less obvious. On the other hand all the important features are still present and the same amount of information, including all phase information, is still used. It will, however, no longer be as easy (if possible) to reconstruct the CCF from the summed response.

Discussion

We believe that we have shown that the SCAT receiver is evidently a cross-correlation receiver in its form; that it has full phase preservation and that at least for the class of band-limited signals used here, there is an easy way to reconstruct the CCF from the SCAT output, demonstrating the close correspondence and preservation of signal information in both receivers.

These findings shed light on the recent investigation by Boonman et al. (2003), who showed, using the SCAT implementation of Peremans and Hallam (1998), that unless the signal is a very shallow sweep, i.e. almost a constant frequency signal, the performance of the SCAT receiver is very close to that of the cross-correlation receiver.

In discussing the relevance of the model we consider two issues, namely the

¹ Hackbarth (1986) uses a filter bank with constant Q and linear spacing of center frequencies, which also gives rise to a filter bank with an overall high-pass characteristic.

algorithm as a model of the physiology of echolocating bats especially with respect to the phase locking properties, and, very briefly, its potential practical implications in for instance robotic sonar systems, since this was the interest of Peremans and Hallam (1998).

Physiology of FM bats and phase-locking

It is generally believed that the peripheral auditory system of non-CF bat species is not dramatically different from other mammals except from the higher frequencies involved. The lack of phase locked responses at high frequencies in auditory neurons can be successfully modeled as a number (~ 3) of independent first order lowpass filters in series

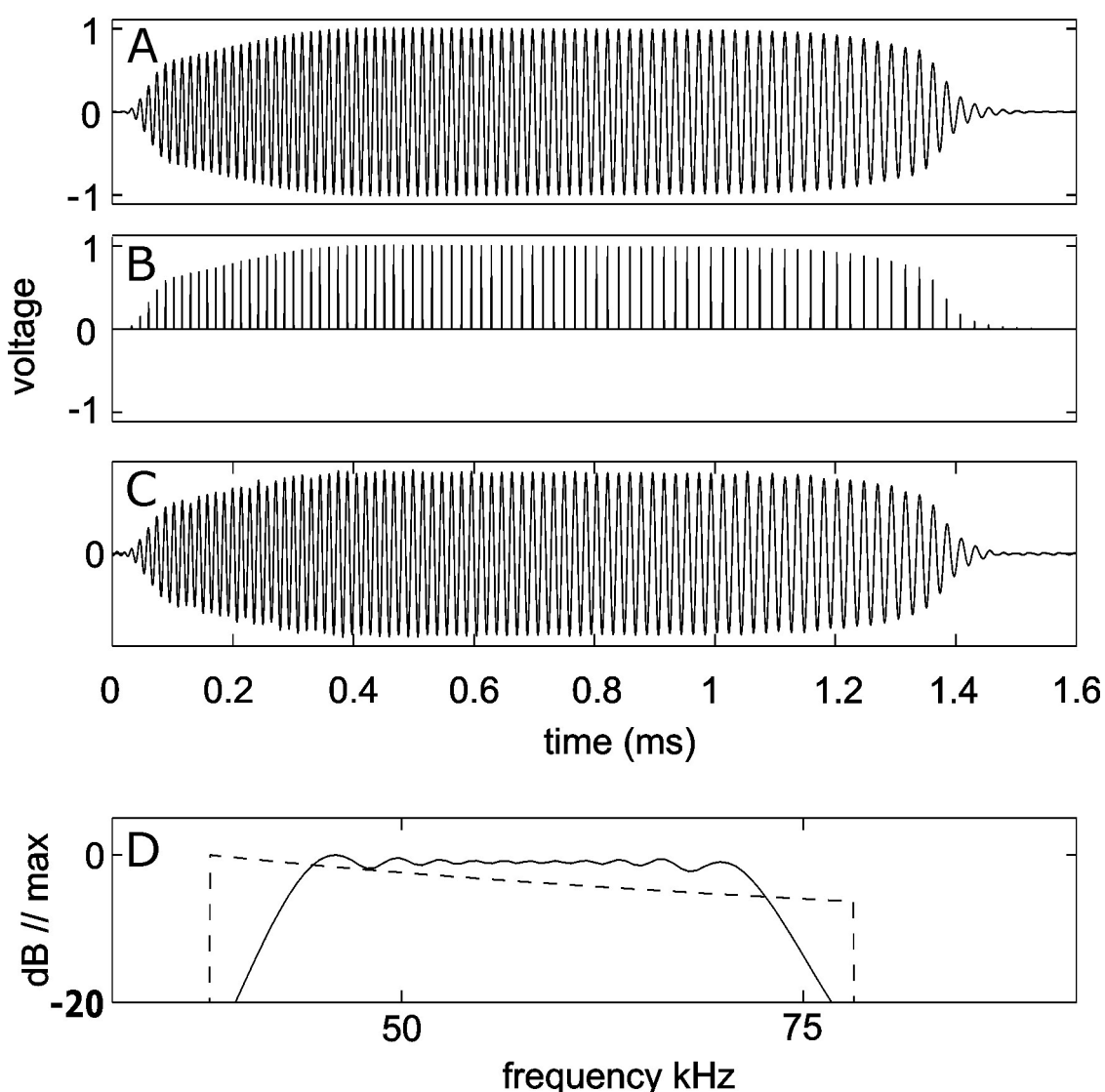


Fig. III-3: Isolating a signal represented by its positive peaks. A: Original signal (linear sweep ca. 1 octave). B: Positive peaks of signal in A. C: Reconstructed signal from band-limiting and weighting. D: Spectrum of signal in A (solid line) and filter transfer function used to isolate the signal in C (dashed line).

(Kidd & Weiss 1990). These filters reflect basic neural properties such as membrane capacitance and calcium buffering. It is therefore not very likely that the bats' 8th nerve neurons can produce phase locked responses at high frequencies (see e.g. Palmer and Russell, 1986; Palmer et al. 1986; Charlet de Sauvage et al., 2000); nevertheless we shall proceed for a while as if this were in fact the case.

Of course the phase locking limitation is not to be understood as the obvious fact that a single neuron will not fire spikes at a rate higher than approximately 1 kHz. The SCAT receiver is an algorithmic implementation of a hypothetical auditory system and naturally simplifications have been made (or it would indeed deal with an enormous data load). One of these simplifications is the pooling of neural responses from many filters into a single "afferent" channel. In this channel the spike rate is allowed to dramatically exceed 1 kHz in recognition of the consequences of the volley principle.

For computational reasons this is understandable. The problem is the underlying assumption that the firing probabilities in the thereby pooled neural responses are not independent. In a "wet-ware" implementation of the SCAT receiver one would imagine that several neural fibers carry the sequence of phase locked responses to (an) integrating neuron(s) for each filter in the bank.

Still, no individual fiber must necessarily fire at any particularly high rate as long as the number of fibers is sufficiently high. But it is crucial to the working of the SCAT receiver that what we might term the "phase event" marking spikes are distributed among the neurons so that at every single one of these phase events, there is at least one neuron representing it. But in a real auditory system all neurons having a synapse with the inner hair cells making up the filter channel will have a high probability of firing at the onset of the stimulus, i.e. at the first, second or third phase event above the threshold. After this all neurons in that frequency neighborhood will be in an absolute refractory state for at the very least 0.6 ms (Charlet de Sauvage, 2000). When this amount of time has elapsed the rising slope of an auditory filter (with a Q_{10dB} of 10 and characteristic frequency of e.g. 25 kHz filtering a typical 3-5 ms FM bat signal) is over and no more spikes (or possibly some in the end) will be fired as response to this signal.

One might attempt to make up for this by making the synaptic strength between the inner hair cells in the filter and the neurons carrying the individual phase events very variable. But this strategy can easily be envisioned to have the disadvantage that it will work only for one particular and rather narrow range of echo (or vocalization) amplitudes. Since vocalizations are presumably almost always a stronger stimulus than the returning

echoes it is hard to imagine a spread in connection strengths large enough to encompass both of the signals to be correlated in a correlated manner.

Connection to the FB receiver model

The physiological data used to justify the phase locking property of the simulated auditory filters are taken from a review by Russel and Palmer (1986). These authors mention that in the cat, auditory nerve fibers have a synchronization index of 0.5 at 2.9 kHz. This is interpreted to mean that the “ripple” of the envelope of the output from the filters falls off as if it were filtered by a 1st order lowpass filter with -6 dB point=3 kHz. In the SCAT receiver this assumption results in a synchronization index of 1 at 100 kHz. The data reviewed by Russell and Palmer (1986) also show that the synchronization index falls to 0 at 6 kHz in the cat.

Since this evaluation of the data on phase preservation in the mammalian cochlea may seem to be pressing the limits, it could be interesting to see what would happen, if the usually assumed physiological limits to phase locking are introduced into the neural circuit making up the SCAT receiver.

The responses in this case would have to be reduced to a single spike per channel, but we could allow each response to be at the peak of the now almost smooth envelope. The cross-correlation between the two spikes in each channel – one for the vocalization and one for the echo – is a single spike at the delay between them. Summing, we get a histogram of delay estimates in the bank. So, in this case the temporal block is actually reduced to the FB receiver in the formulation of Hackbarth (1986) (for a more thorough argument, see Chapter II).

Practical implications

Peremans and Hallam (1998) considered the receiver algorithm from the viewpoint of practical applications in robotic navigation by means of sonar. Given the above analysis the question is whether or not a rather more computationally efficient design would utilize simply the zero-crossings and/or the peaks in the signal for the temporal block without the filter bank preceding the correlation process. By using many overlapping filters the SCAT receiver as suggested by Saillant et al. (1993) does increase the resulting effective sampling depth of the input to the correlation process, but we would suggest that the increase in accuracy resulting using the 81 filters over just one (or two) is very modest.

Appendix III-A

When we investigated the spectra of bandpass signals represented by their values at the time instances of positive peaks only, we found that there exists an interesting relationship between the phase spectrum and the amplitude spectrum for such signals that does represent an ambiguity in the representation.

The original to such a “peak transformed” signal can be more or less reconstructed simply by limiting the spectrum to the relevant octave and adjusting the amplitude with a frequency dependant weight.

We find that for an original signal with a smooth phase function the spectrum of the transformed signal is weighted by a factor of f , which is easily understood, since the number of peaks, and therefore the weight of that component, is related in this way. However, this is only true when the signal in question has a phase spectrum that is sufficiently smooth. For an FM sweep at high signal to noise ratios the factor is relevant. For a noise signal of comparable bandwidth this factor must not be applied to revert to the original signal. If the type of signal responsible for the original is not known, an ambiguity exists in the reconstructed signal.

In the case of the individual filters in the SCAT receiver this is not an issue, since they are so narrow that within each filter very little frequency dependant variation can take place. However, it is still highly relevant to consider the weight relative to the rest of the filters in the bank. This weight is then proportional to the center frequency.

Appendix III-B

As an illustration of the information contained in the peaks or the zero-crossings of a bandpass signal we describe here the following reconstruction possibility for clipped signals. It does clearly relate to the problem of the information contained in the SCAT receiver, albeit indirectly. As a means of signal reconstruction it is very simple. The conditions that must be fulfilled, however are seldom met, when most practical signals are digital, sampled at a not too high rate, and at a not too high resolution. But consider a signal stored in an analog form with a frequency band not exceeding one octave. If that signal is passed through a circuit with a limiting characteristic before storage, it is in praxis relatively easy to reconstruct the original signal to a degree of accuracy that is essentially limited by the sampling equipment at hand.

If no DC offset is present in the original signal, the clipping, no matter how severe, does not affect the rate of change at the times of the zero-crossing, since the clip-

ping would then have to be infinite. So the idea is to use only these points to restore the original. In order to do this it is necessary to sample the given analog signal at a very high (see below) rate and determine the numerical difference between the two samples surrounding each zero-crossing. This difference is a measure of the intensity of the envelope of the signal at that specific time instance. These data can be used to reconstruct the original signal simply by band-limiting the resulting digital signal, which is the collection of zero-crossings and their amplitudes. Since the measured data represent the differentiated signal (rates of change) a factor of $|1/f|$ must be applied to the spectrum before converting to the time domain.

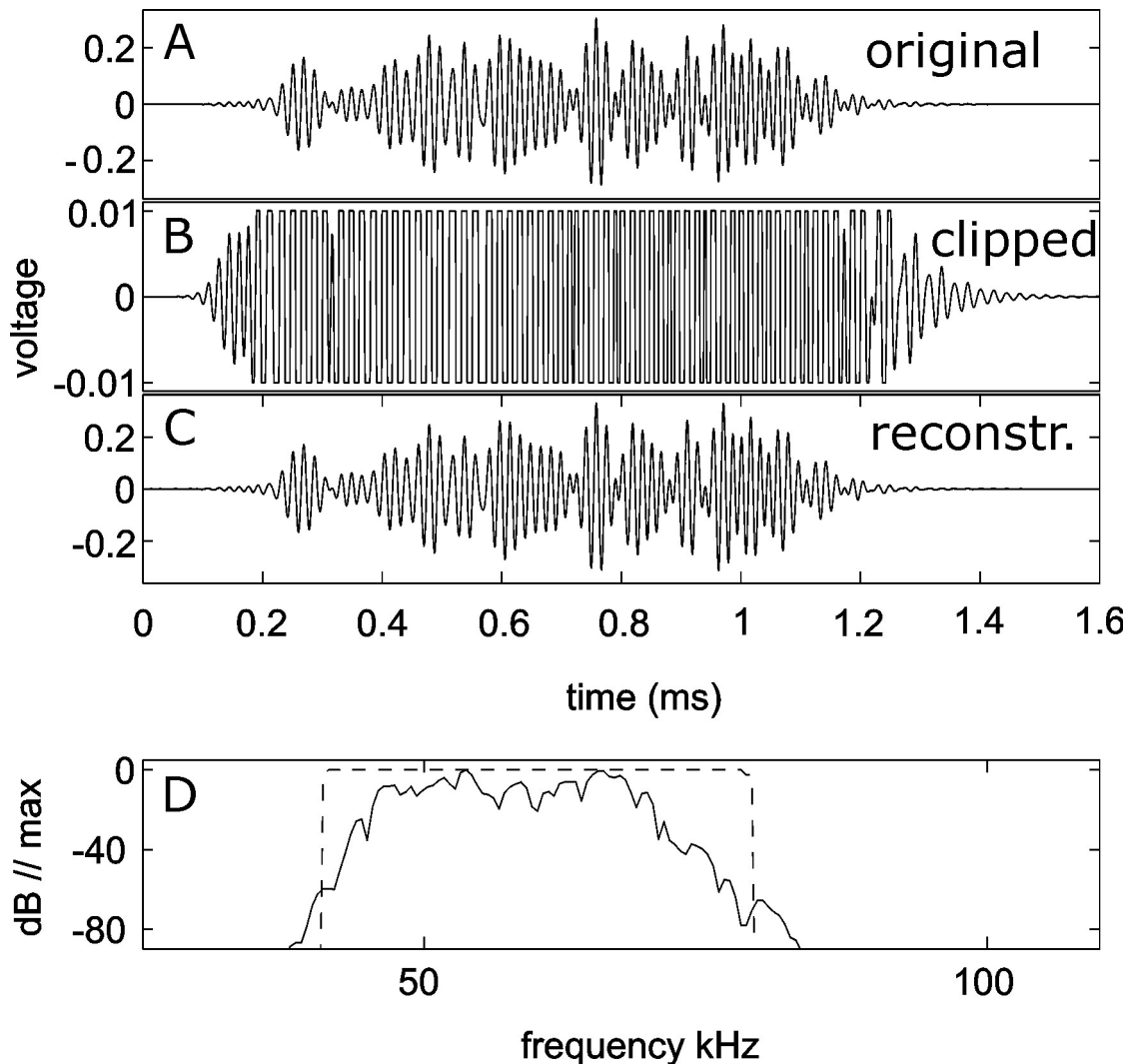


Fig. III-B1: Reconstructing a clipped signal from the zero-crossings. A: Original signal, bandpass-filtered white noise with a Hann window. B: The signal in A limited to ± 0.01 . C: Reconstruction based on the method described in the text. D: Amplitude spectrum of the original signal together with the window used to isolate the reconstruction result.

The problem of determining if the minimum sampling rate has been used is easily solved: If the amplitude of any one of the sample pairs surrounding a zero-crossing equals the level to which the signal is clipped the sampling rate must be increased.

The weight factor considered in Appendix III-A is also relevant for this type of reconstruction.

In practice, although the very nice reconstruction shown in Fig. III-B1 is only possible with signals really limited to 1 octave bandwidth, it is often possible to retrieve signals with a bandwidth of 2 or more octaves. Whether this is feasible depends on the signal in question.

References

- Boonman, A. M., Parsons, S. and Jones, G. (2003). "The influence of flight speed on the ranging performance of bats using frequency modulated echolocation," J. Acoust. Soc. Am **113**, 117-128.
- Charlet de Sauvage, R., Erre, J.-P. and Aran, J.-M. (2000). "Discharge rate of the auditory nerve during noise revealed by electrocochlear stimulation," Hear. Res. **142**, 141-158.
- Hackbarth, H. (1986). "Phase Evaluation in Hypothetical Receivers Simulating Ranging in Bats," Biol. Cybern. **54**, 281-287.
- Kidd, R. C. and Weiss, T. F. (1990). "Mechanisms that degrade timing information in the cochlea," Hear. Res. **49**, 181-208.
- Licklider, J. C. R. (1951). "A duplex Theory of Pitch Perception," Experientia **7**, 128-133.
- Logan, B. F. Jr. (1977). "Information in the zero-crossings of bandpass signals," Bell System **56**, 487-510.
- Masters, W. M., Raver, K. A. S. (2000). "Range discrimination by big brown bats (*Eptesicus fuscus*) using altered model echoes: Implications for signal processing," J. Acoust. Soc. Am. **107**, 625-637.
- Menne, D. and Hackbarth, H (1986). "Accuracy of distance measurements in the bat *Eptesicus fuscus*: Theoretical aspects and computer simulations," J. Acoust. Soc. Am. **79**, 386-397.
- Palmer, A. R. Winter, I. M. and Darwin, C. J. (1986). "The representation of steady-state vowel sounds in the temporal discharge patterns of the guinea pig cochlear nerve and primarylike cochlear nucleus neurons," J. Acoust.Soc.Am. **79**, 100-113.
- Palmer, A. R., (1986). "Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells," Hear. Res. **24**, 1-15.
- Peremans, H. and Hallam, J. (1998). "The SCAT receiver revisited," J. Acoust.Soc.Am. **104**, 1101-1110.
- Pye, J.D. (1986). "Sonar Signals as Clues to System Performance," Acustica **61**, 166-175.
- Russel, I. and Palmer, A. R. (1986). "Filtering due to inner ear hair-cell membrane properties and its relation to the phase-locking limit in cochlear nerve fibres" in *Auditory Frequency Selectivity* (Eds: Moore, B. C. J & Patterson, R. D.) Plenum Press, New York , pp 199-206.

- Saillant, P.A., Simmons, J.A. and Dear, S.P. (1993). "A computational model of echo processing and acoustic imaging in frequency modulated echolocating bats: The spectrogram correlation and transformation receiver," *J. Acoust. Soc. Am.* **94**, 2691-2711
- Shannon, C. E., and Weaver, W. (1949). *The Mathematical Theory of Communication* (Univ. of Illinois Press, Urbana).
- Simmons J. A. (1973) "The resolution of target range by echolocating bats," *J. Acoust. Soc. Am.* **54**, 157-173.
- Simmons, J. A., Ferragamo, M., Moss, C. F., Stevenson, S. B. and Altes, R. A. (1990a). "Discrimination of jittered sonar echoes by the echolocating bat, *Eptesicus fuscus*: The shape of target images in echolocation," *J. Comp. Physiol. A* **167**, 589-616.
- Simmons, J. A., Moss, C. F. and Ferragamo, M. (1990b). "Convergence of spectral and temporal information into acoustic images of complex sonar targets perceived by the echolocating bat, *Eptesicus fuscus*," *J. Comp. Physiol. A* **166**, 449-470.
- Strother, G. K. (1961). "Note on the Possible use of Ultrasonic Pulse Compression by Bats," *J. Acoust. Soc. Am.* **33**, 696-697.

CHAPTER IV

A method for analyzing bat echolocation signals applied to the question of the time-frequency structure of *Eptesicus fuscus* calls in the laboratory

Introduction

Microchiropteran bats use echolocation to find their way and catch prey (Griffin & Galambos, 1941). Bat echolocation calls are generally ultrasonic, frequency modulated and they very often have a harmonic structure (Simmons & Stein, 1981). Harmonics is a natural consequence of the way the vocal apparatus functions in many terrestrial mammals. They emit band-pass filtered click trains and the inverse of the rate of production makes up the fundamental frequency (Zbinden, 1988). For the Vespertilionid bats, which all have frequency modulated calls as their most ubiquitous signal type (Simmons and Stein, 1981), there is often a considerable overlap in the frequency content of the different harmonics, particularly in the laboratory, where signals are generally steeply frequency modulated. This fact makes the exact analysis of the individual harmonics difficult.

Several investigators have attempted to find a function that satisfactorily (or just best) describes the frequency-time course of the calls of the most common experimental animal, *Eptesicus fuscus* or tried to otherwise ascribe significance to parameters of the signal structure that requires frequencies to be known at a specific point in time (Menne, 1989). These investigations all divided the signals into shorter chunks, which were then analyzed for peak frequency. Masters et al. (1989) used band-pass filtered (centered on a single harmonic) chunks that were analyzed for average zero-crossing intervals to achieve an estimate of “instantaneous frequency” at the time-center. Waters (2000) uses the FFT of the chunk to determine the peak power frequency, whereas Menne (1989) used a maximum entropy method. The latter is probably less suited for the type of signal in question, since each chunk contains an unknown number of harmonics (>1) and this method returns a frequency centroid – not the peak(s), at least as long as only one “pole” is searched for.

There are some problems with chopping up the signals and determining the average fre-

quency content in each chunk as it was done in all these investigations, whatever method of frequency determination was used. As a rule, if the frequency is not a linear function of time or the envelope is not flat, one is in error if one assumes that the frequency estimate found when analyzing the chunk can be ascribed to the time center. To find the correct time one must know the time course, but this is what is sought by the analysis! In some cases this may be overcome by taking the roughly estimated time course into account and repeating the process until no further improvement is seen, but this has not been done in any of the above mentioned studies.

Another problem is that fast frequency variations are averaged out over the interval analyzed, which could lead to some systematic steep modulation features being overlooked. One may often assume these errors to be negligible, but it is hard to know this for certain.

In the study by Masters et al. (1991) several different mathematical models were fitted to the estimated time frequency course, and relatively small differences between the goodness of the fits with these different models are found. Small errors of the sort mentioned above might make a crucial difference as to which function gives the lowest residual error.

In principle the ideal way to overcome these problems is to calculate the analytical signals and then take the derivative of the phase function. This method is unfortunately very sensitive to noise and does not work at all in the presence of strong harmonics, which is problematic for bat signals as touched on above. Below we introduce what we have coined the “cross-correlation extraction method” and a slightly alternative “morphing method”, which overcome these obstacles for reasonably good signal to noise ratios (SNR). One primary aim here is thus to introduce these methods of analysis.

We address the problem of the precise call structure of *E. fuscus* in the laboratory using these methods, which is the other major purpose of the present communication.

Significance of call structure

The notion that some Vespertilionid bats like *Myotis lucifugus* seem to be producing Doppler tolerant calls stems from Cahlander (1964) and Altes & Tiltlebaum (1971). The Doppler tolerant waveform is either a Dirac impulse or has a hyperbolic time-frequency structure, i.e. linear period modulation (Kruszczynsky, 1968). The use of such a waveform might imply that the animals must have need for very high precision in their timing estimation. For this property of the hyperbolic pulse to be useful, the bats must reference

their delay estimates to a time somewhat prior to emission of the pulse (Pye, 1986; Lin, 1988).

If the bats use an exponentially decaying time-frequency structure as was reported in Masters and Raver (2000) it could point to a basis for the sweep in the sound production mechanism. If the muscles stretching the vocal folds are tensed and then made to relax during the call, it would be expected that the time course resulting would indeed be exponential. If this were the case it would point to the near hyperbolic waveform as possibly a by-product of the mode of sound production.

Masters et al. (1991) found the best fit to be a “logarithmic time” function, originally proposed by Simmons (1987). The interpretation of this finding in both these sources is that this frequency-time cause makes the spectra of the returning echoes equally affected by the attenuation by the middle ear muscles, which is released following a linear function in a double logarithmic plot (Kick & Simmons, 1984). It is indeed in this context compelling that the duration of the calls is the same when plotted against logarithmic time since emission (Masters et al. 1991, Fig. 8). As long as this rule is obeyed it ensures that the attenuation within the echo of any given frequency over any other frequency in the same call is constant for all calls regardless of the sweep used. The logarithmic sweep is not needed to obtain this advantage, so this interpretation is not valid.

In the following we have tested the degree to which these three functions fit the calls in two ways.

First, to analyze the results of the previous investigations by Masters et al. (1991) we used a Monte Carlo simulation scheme in which we sought an answer to the question: How well do the different function types that are tested fit a signal known to actually have a time-frequency course corresponding to that function type. For instance, how well can a signal generated with a hyperbolic function be fitted to an exponential, logarithmic or hyperbolic function, when they are analyzed approximately like in the previous experiments? This question is particularly relevant when considering that the hyperbolic function has 2 parameters, whereas the EF and LT functions have three, rendering these rather more flexible. One interesting possibility is that the signals are actually hyperbolic but the inevitable noise in the recordings makes a better fit with e.g. the LT function possible.

Second, we use one of the techniques described below to analyze the question using bat signals recorded during a detection experiment with *E. fuscus* in the laboratory by means of the instantaneous frequency of the two lowest harmonics.

Methods

A. Cross-correlation extraction and morphing

In the following, unless otherwise stated, signals are to be understood as digital (sampled). The spectrograms are made as follows: Every 8 samples a 128 point Hann window is used to pick out a chunk of the signal, which is zero-padded up to 512 samples. The energy spectrum of this is calculated. The positive-frequency part of the amplitude spectrum is displayed as a color map with frequency on the Y-axis and the time corresponding to the center sample of the chunk on the X-axis. A logarithmic gray-scale Z-axis is used. Sampling rate is always 480 kHz.

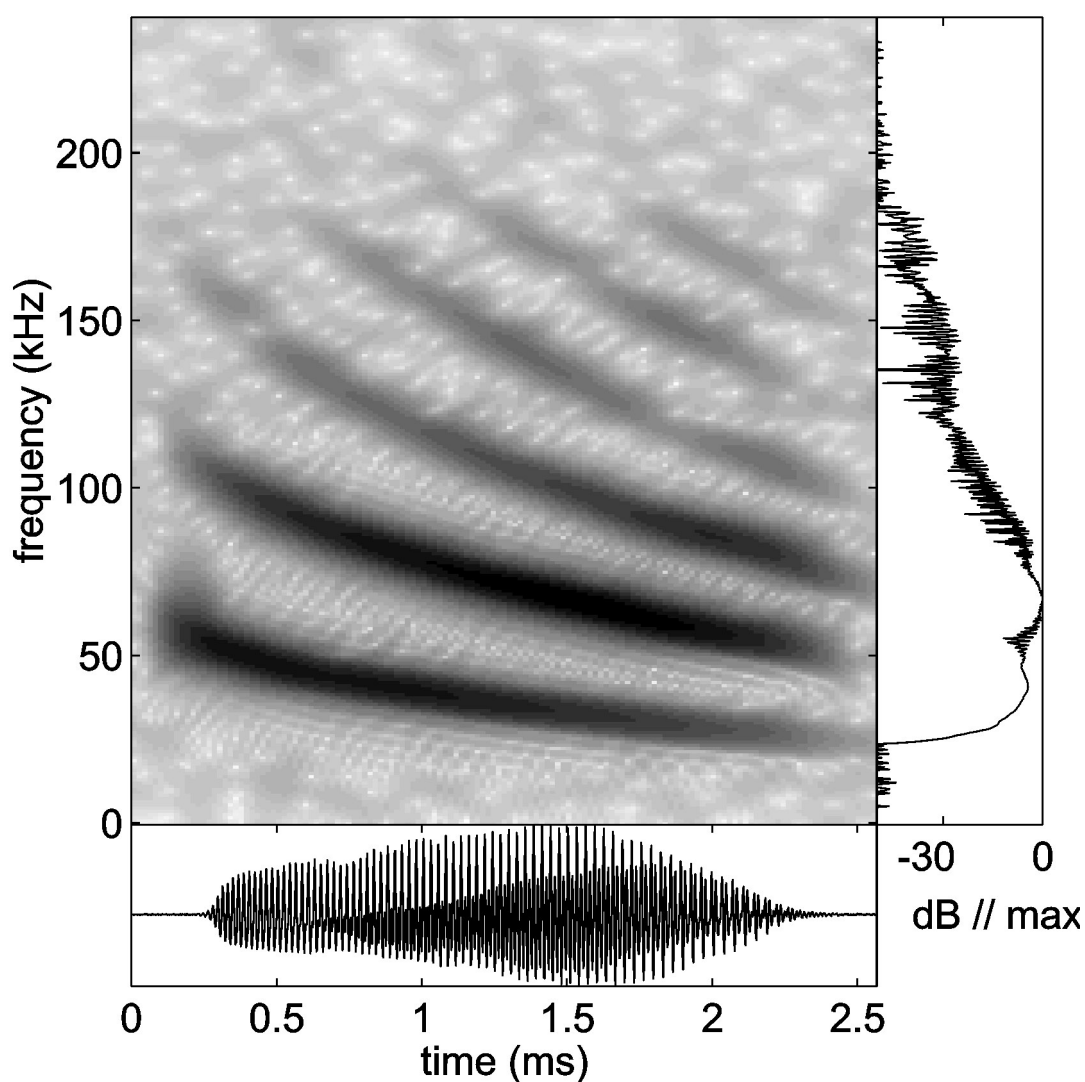


Fig IV-1: Spectrogram, spectrum and time-series of a call emitted by *Eptesicus fuscus* (EF1) under the conditions described here. The high number of harmonics that are visible in the spectrogram is due to the high signal to noise ratio of the recording, compared to a recording in the wild.

When looking at a spectrogram of a bat echolocation call (as in Fig. IV-1) it is immediately obvious that it is impossible to separate the harmonic components by simply band-pass filtering the whole signal. On the other hand it is also relatively self-evident that it should not be impossible to separate them either, since at all time instances the frequency components of the individual harmonics lie well apart per definition. At least one other study has used a separated harmonic (Capus and Brown 2003; Altes, pers. comm.).

Cross-correlation extraction

Our first solution was to make use of a cross-correlation technique to compress the individual harmonics in a way that uses a rough estimate of the frequency-time structure as well as the envelope of the calls. The result is a signal where the harmonic components are separated in time, since highest maximal correlation between a harmonically structured signal and a sweep will occur at different delays for the different harmonics. One can then isolate the individual harmonics with a time window and restore it to the original time course by deconvolution.

In the first place this approach requires that the frequency-time structure of the analyzed call be known to a certain extent – the more accurate the better. We made this estimate simply by inspecting the spectrogram visually and estimating a linear (or e.g. hyperbolic) FM-sweep that roughly followed the harmonic. It was then made into an “artificial” signal, the time-reversed version of which is then used as a filter on the call to be analyzed. As expected, in the resulting filtered signal the harmonics are much more time local and well separated (if the sweep was estimated reasonably well) and can be windowed out as is shown in Fig. IV-2.

In order to undo the time-compression one must now deconvolve the transformed signal with the same filter that was used to make the compression. This operation is not as unproblematic as it might seem. Dividing by the FFT of the “filter signal” will, unless this signal covers the whole frequency range, involve dividing by very small quantities. If the filtered signal is left unchanged this of course does no harm, as is readily seen from the formula, but in our case we have multiplied by a windowing function to null signal components that belonged to unwanted harmonics and noise. The windowing function introduces new non-zero elements in the Fourier spectrum of the filtered signal, and these components are likely to be amplified enormously by the deconvolution operation. To avoid this, it is necessary to introduce a threshold: at frequencies where the magnitude of the Fourier transform of the filter signal (the divisor) drops below this threshold the

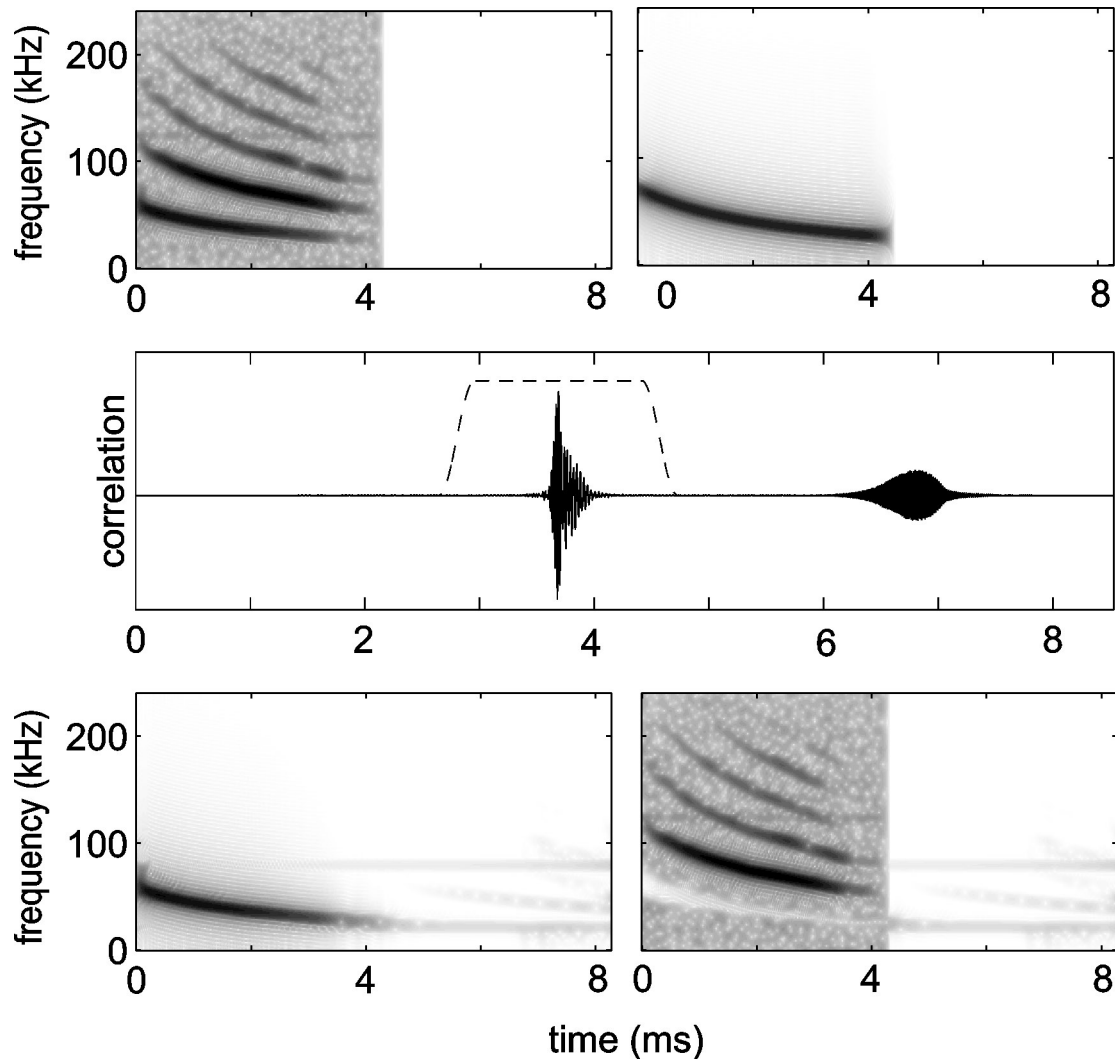


Fig IV-2. Cross-correlation extraction method for isolating harmonics of a bat call (EF1). Top left frame: Spectrogram of the zero-padded bat signal to be taken apart. Top right frame: Artificial call with hyperbolic time course estimated roughly from the fundamental of the signal. Center panel: Time series of the result of a cross-correlation between the two signals above. Also shown is the window used to isolate the first harmonic in time (dashed). The part of the signal belonging to the second harmonic is clearly visible between 6-8 ms delay. Bottom left: Deconvolution result for the isolated 1st harmonic. Bottom right: Bat call minus the isolated harmonic. From here another round can be made to get the second harmonic and so forth. See text for further explanation.

magnitude is set to 1.

Morphing method

In order to avoid the non-linear threshold element we devised another technique, starting out by thinking of the cross-correlation extraction method as a form of signal “morphing”. If one starts out in the Fourier domain the inconvenience of having to set a threshold value can be avoided and this approach constitutes an in some ways preferable method. In this case one once again defines a suitable linear (or other) function that matches the harmonic

in question. The slope of the sweep is then translated into a frequency dependent phase shift which is applied to the Fourier transformed signal. The resulting signal is thus controllably “morphed” or phased, and if the slope was chosen well it is now possible to use a single conventional band-pass filter and time window to isolate the harmonic. The resulting isolated harmonic is then Fourier transformed, the phases are moved back into place and once again the signal is transformed into the time domain, where it can be subtracted from the original signal to give more “room” (time-frequency gaps) for windowing the

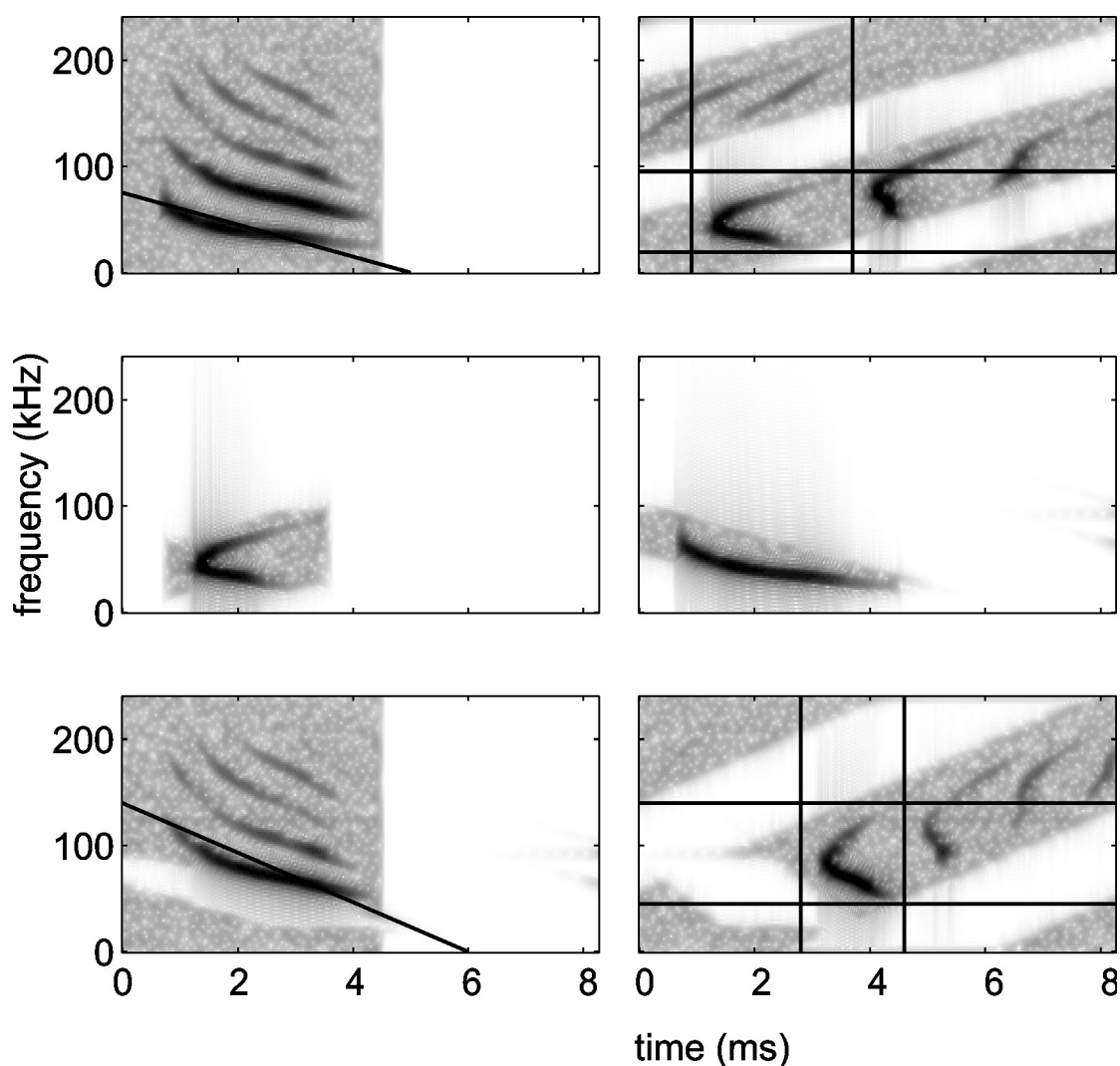


Fig. IV-3. Morphing method for isolating overlapping frequency components of a signal. Top left: Signal to be taken apart (produced by EF4) together with the time-frequency profile used to distort the signal (the solid line). Top right: The result of the morphing and the time and frequency limits used to isolate the fundamental. Center left: The isolated morphed harmonic. Center right: Isolated fundamental after reversing the distortion. Bottom panel: Same as top panel, but for the second harmonic after subtraction of the fundamental from the signal.

next harmonic.

The main advantage of the morphing technique over the cross-correlation method is that the frequency-smear resulting from the time window and the time smear stemming from the filtering is well defined and predictably depending on the apodization functions. On the other hand, in the cross-correlation case where the filter is a finite duration signal, generated in the time domain, one has easy control over the envelope of the impulse response of the filter, which in difficult cases (like very steep sweeps with abrupt onsets) may be an advantage, since better time compression can be achieved through a closer matching signal. It is relatively simple to automate this way of separation. The disadvantage is the mentioned nonlinear operation necessary to contain the deconvolution result.

B. Simulations

We generated three digital signals, one according to each of three function types (linear in logarithmic time (LT), exponential frequency (ET) and linear period (LP)) that were used to fit the time frequency structure of the bat calls in Masters et al (1991). We used the average signal parameters in Table I, (“Average” column, Detection Task) of that paper. Signals were windowed with a Welch apodization function.

In each run white Gaussian noise was added to the signals and they were then chopped up into 64 samples long chunks (no apodization function). These were zero-padded up to 1024 samples and the peak frequency of the magnitude of the FFT was determined. The resulting array of peak intensity frequencies was then fitted with the three function types; EF and LT using MATLABs nonlinear fitting routine and LP using linear regression on the reciprocal values. The hyperbolic function should also have been fitted with a nonlinear method, but this approach tended to be unstable, converging on solutions far from the global minimum. The differences in the parameter values, which we found with the nonlinear fitting procedure (when it occasionally worked well) and the ones found with the linear fit to the reciprocal data, were minimal. In the experiments that we are simulating here, namely the ones reported on by Masters et al. (1991), also used this method.

1000 runs were made at each of the signal to noise ratios, 36 and 48 dB. The mean squared errors (MSEs) of the fits were finally used in forming a mean MSE that is the basis for our comparison with the corresponding real bat values in Table II of Masters et al. (1991).

C. Analysis of bat calls

The calls analyzed here were sampled during a psychophysical detection experiment involving four *Eptesicus fuscus*, 2 males 2 females, one of each sex caught in the wild, the others born in captivity. The signals were picked up by a B&K 4138 microphone, amplified and sampled with 480 kHz (16 bit) using a custom build sampling device and software from Menne BioMed. The experimenter stopped the continuous sampling by means of a foot switch when the bat seemed to have reached a decision and the samples made in the preceding 2 seconds were then saved to a hard disk. Trials usually lasted less than 2 seconds. For each bat 5 signals were picked for analysis based on a signal to noise ratio criterion. The lowest two harmonics were extracted with the morphing method as described above and the analytical signals were formed for each. For each of the two harmonics the numerical derivative of the unwrapped phase function (the instantaneous frequency) was calculated and the average (with the second harmonic values divided by 2) was formed with a weight for each sample equal to the instantaneous envelope squared (=intensity) of the harmonic in question. This weighted average was fitted with the three functions, LP, EF and LT as described for the simulation part above. We also fitted the functions to the same calls where the first 300 μ s of the data were excluded. For the morphing process we used a linear time frequency profile for the extraction since in this way we could avoid creating a signal with one of the models tested.

No statistic tests of the quality of the fits were made. We looked for systematic errors in the fit, which might indicate the existence of an alternative “true” function or an alternative strategy. If indeed the correct function were used for the fit, the true instantaneous frequency data minus the fit should resemble white noise.

Results

A. Simulations

The results of the simulation are shown in table IV-I. The results of Masters et al (1991) show mean MSEs in the order of 0.2 (0.137 - 0.291) for the best fits, which in that study is invariably the LT function type. At 48 dB SNR our simulations yield the lowest values between 0.038 and 0.053, which indicates a considerably better fit. When the lower SNR was used the function type that gave the closest fit, was the function used to create the noise-free signal. So here a LP function type fits the LP sweep best of the three functions tested, and so on.

This is not the case when the SNR is reduced to 36 dB. In this case we arrive at

Table IV-1. The mean square errors (MSE) of the fits to the simulated signals using the indicated functions with parameters from Masters et al. (1991) under two different signal to noise ratios. The values in Masters et al. (1991) range from 0.137 to 0.291 for the LT fit (lowest MSEs) in the detection task.

Fitting function	SNR	
	48 dB	36 dB
	logarithmic time (LT) sweep	
LT	0.049	0.199
EF	0.091	0.247
LP	0.149	0.317
	exponential frequency (EFT) sweep	
LT	0.095	0.278
EF	0.037	0.225
LP	0.162	0.366
	linear period (LP) sweep	
LT	0.054	0.210
EF	0.057	0.218
LP	0.053	0.226

best fits between 0.199 and 0.210, i.e. in the same range as the actual bat data from Masters et al. (1991) with which we compare. The hyperbolic signal with noise is now fitted marginally closer with the LT function than with the LP function. The other two signals are fitted best with their own respective function types. In all these cases the LP function gives the least close fit (0.23-0.32).

In Fig IV-4 the noise has been band-pass filtered to 25 - 100 kHz in order to give a realistic idea of how a signal with these SNRs looks like in the time domain. Visual inspection of the signal with noise in the 36 dB SNR case, when our results lie in the same range as the real data, seems to exclude that the previous investigators have operated at a SNR ratio anywhere near this noise level. At 48 dB SNR the noise is still considerably above what is often found in laboratory recordings.

B. Analysis of real bat calls

The calls analyzed could for the most part be fitted closely by all three functions and the deviations were generally contained within a few kHz see Fig. IV-5. The MSEs found are in the same range as the ones found by Masters et al. (1991). 16 of 20 calls are fitted closer

with the LT function compared to the EF function. In one instance the fit was better with the LP function than the LT function. The LP fit was never best of the three.

When the first 300 μ s are excluded from the call all the fits improve. In average

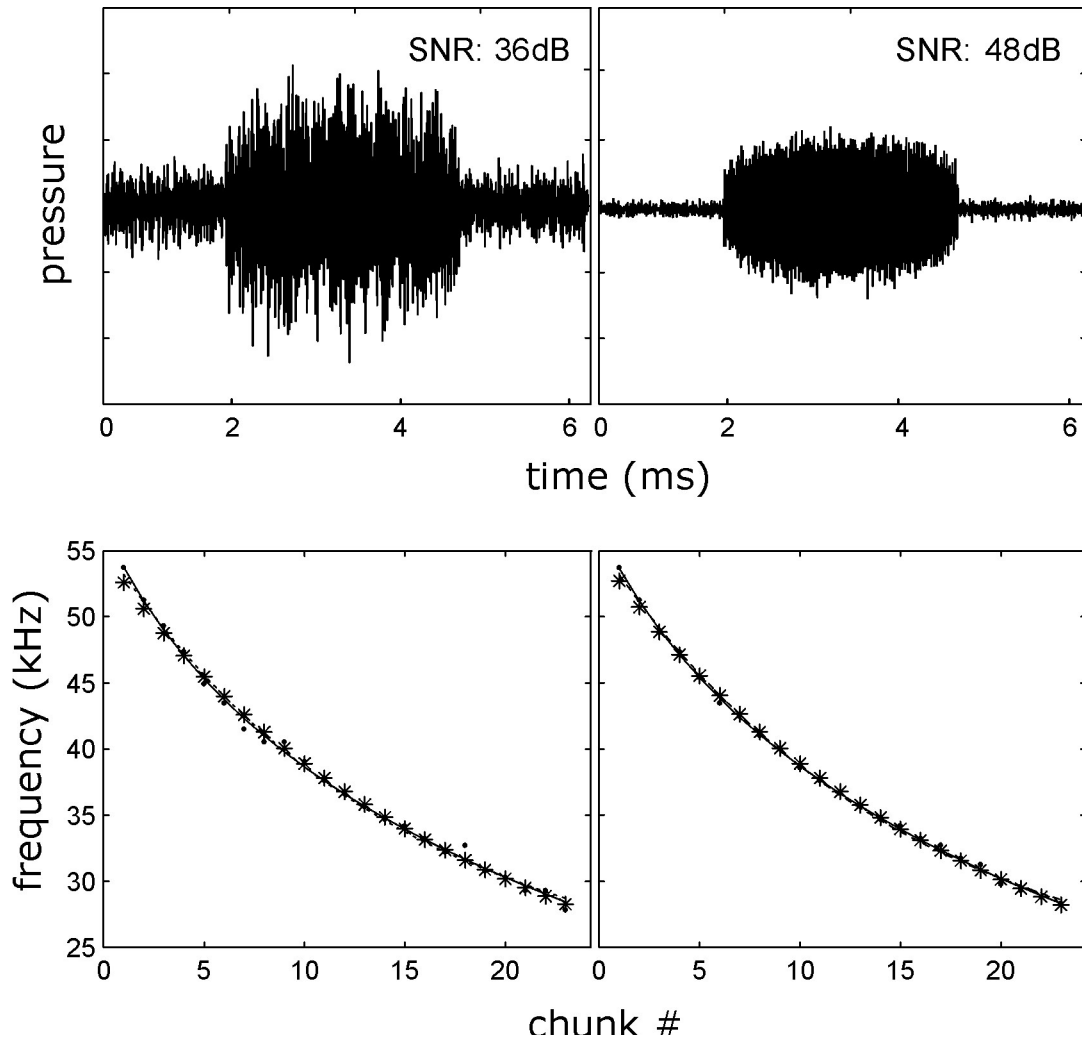


Fig. IV-4. Logarithmic time (LT) signals at two different signal to noise ratios (top panel) and fits to the center frequency values (bottom panel) as described in the text.

Bullets: The frequency data. Stars: The hyperbolic fit. Dashed line: The exponential frequency (EF) fit. Solid line: The logarithmic time (LT) fit.

the ratio of full-length signal fit MSE to truncated fit MSE is 0.4 ± 0.23 , calculated over all three fitting functions. In one instance this ratio is as low as 0.03. In 11 cases the EF function fitted better than the LT function after truncation.

The deviations at the beginning of the full length calls are in all but five out of 60 fits an overshoot from the model followed by an undershoot, indicating that the frequencies here are sweeping faster than can be accounted for by any of the models. In the five

remaining cases (LP fits of EF2 and EF3) there was an initial undershoot. These five fits also had the highest MSEs of all fits. When the calls were truncated both over and under-shoot generally disappeared.

Discussion

A. Simulations

In terms of a discussion of a “sweep-function” underlying the time-frequency structure of *E. fuscus* signals in the laboratory the conclusion from these simulations is quite clearly that the bats do not use any of the functions with which Masters et al. (1991) fitted their time-frequency data. If they had done so the lowest MSEs in their experiments would

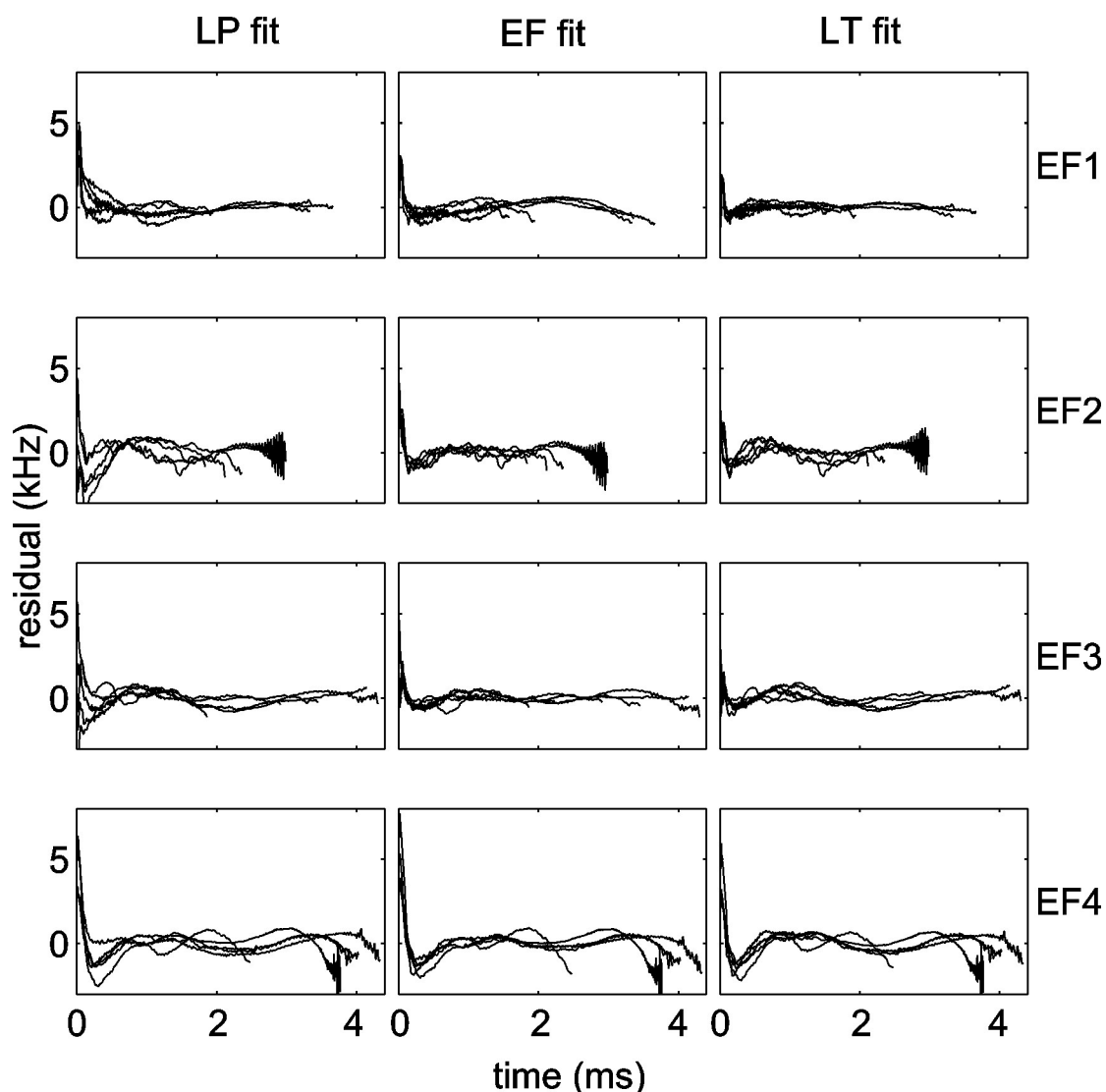


Fig. IV-5. Residues from the fits. The fitting function type is indicated above the top panel. The bat that emitted the calls that is fitted is indicated to the right of the graphs. See text for further explanation.

have been much lower.

Our initial idea was that the bats were actually using hyperbolic signals, but that the fit with the LT function nevertheless was better because of its higher flexibility. The latter function has 3 parameters instead of 2, which allows for noise to be taken better into account. This idea is evidently not true since only at the unrealistically low SNR of 36 dB do the LT (and EF) functions give a closer fit to the LP sweep than the LP fitting function itself, and only marginally so.

B. Analysis of bat calls

Pertaining to a yes/no question of whether these bats use hyperbolic signals or not, we seem forced to answer that they do not, however reluctantly. Their calls are clearly fitted

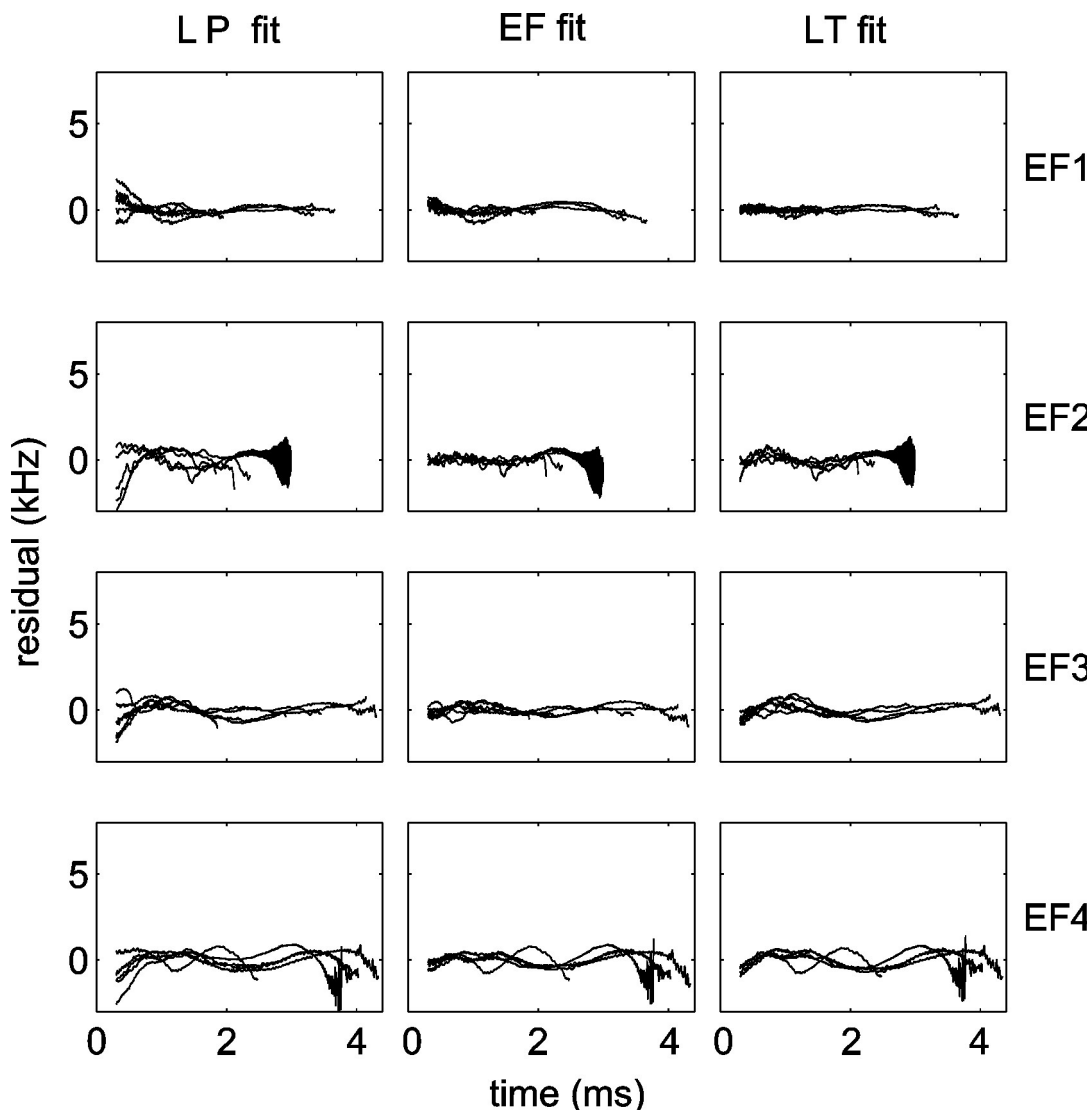


Fig. IV-6. Same as Fig. IV-5, but with the first 300 μ s excluded from the fit.

better with more flexible functions, also if the first segment of the sound is excluded from the analysis. In agreement with Masters et al. (1991) we also find that most calls can be fitted “rather well” with the hyperbolic function, only not as closely as with the other two functions.

Doppler tolerance implies that at any one frequency the first derivative of the frequency vs. time function of a returning echo stays relatively unchanged regardless of the Doppler shift imposed on it (see Chapter II). Several possible interpretations exist pertaining to this fact (Kroszczyński, 1969; Altes and Titlebaum, 1970; Menne, 1988), which we discuss in another context (Chapter II).

With respect to the functional significance of the findings it seems that the LT function fits the calls better than the EF function due to this function’s ability to model a disproportionately steep initial sweep rate compared with the EF model. When the initial portion of the calls is removed from the analysis these two functions perform about equally well in accounting for the sweep structure of the calls with an exponential fit doing very marginally better.

This perhaps hints at a role for the sound production mechanism in shaping the exact sweep function. We speculate that the initial steep part was also present in the study by Masters et al. (1991) since their overall fitting quality is very comparable to ours before the initial 300 μ s were excluded from analysis (Fig. IV-6). Interestingly in the study by Masters and Raver (2000) the calls of both the two bats studied there were better modeled by an exponential function using reportedly similar analysis methods. We have no suggestions as to the reasons for this.

Parssons and Jones (2000) also report an exponential function to give a better fit, which – as they used outdoor recordings – might be in accordance with the notion that the energy of the initial steep part of the calls is low and is likely to be absent from recordings of the calls made at a distance from the animal, and they might therefore be comparable to the truncated calls analyzed here.

We also see that the calls of some bats (e.g. EF 4) do not conform particularly well to any of the functions tested here. EF 4 thus seems to be using a “steep-shallow-steep sweep” pattern. This observation is also in better thread with a role for the sound production system than with a carefully “planned” sweep course to meet the requirements of some form of advanced signal processing mechanism. If this were the case all the bats

tested would be expected to use it.

C. Significance of the initial steep portion of the calls

We cannot find any significance of a logarithmic time-frequency structure as mentioned in the introduction, which of course does not mean that there is not any. We do not believe that the bats really do use such an exact describable frequency course, and we summarize our proposal here:

The calls are shaped from the relaxing of the muscles tensing the vocal folds. The resulting pattern is variable from bat to bat but may be closely described by an exponen-

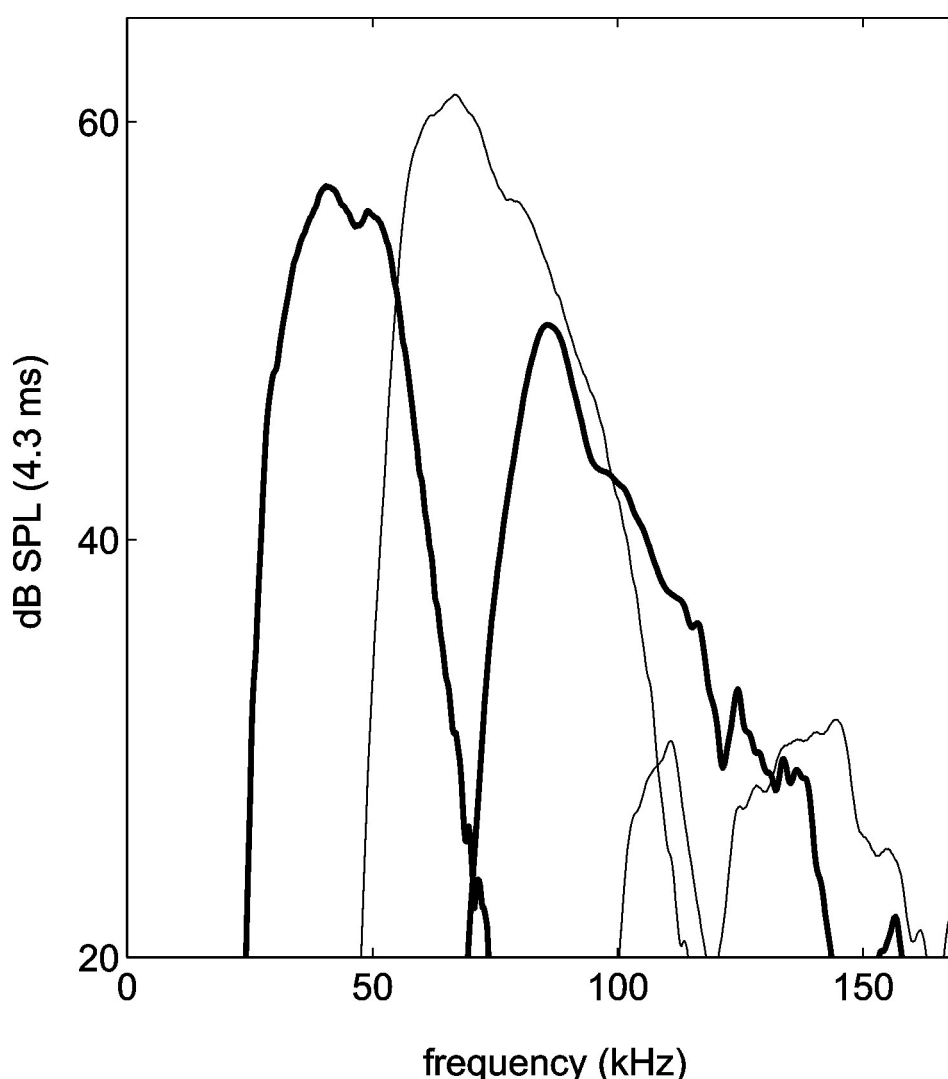


Fig. IV-7. Spectra of individual harmonics of the call shown in Fig IV-1. The scale is dB SPL with the integration window equal to 4.3 ms. The peSPL value (Troest and Møhl, 1986) of this call is 93 dB.

Thick lines indicate odd harmonics; thin lines show the even harmonics. Note that the first and third harmonics meet around 35 dB down from the peak of the first harmonic.

tial function over a large portion of the call. At the beginning and ends of the calls deviations from this pattern are more likely to occur as the folds open or close. The fact that at the very beginning of the call the animals recorded from here seem to be using sweep rates that stand out from the general frequency course of the call in that it is rather steeper, which explains why a logarithmic function fits the calls well, since with this function very steep initial sweep rates can be modeled.

We next present a theory of a possible functional role for this initial part of the signal.

The calls of *Eptesicus* and many other species of bats have several harmonics.

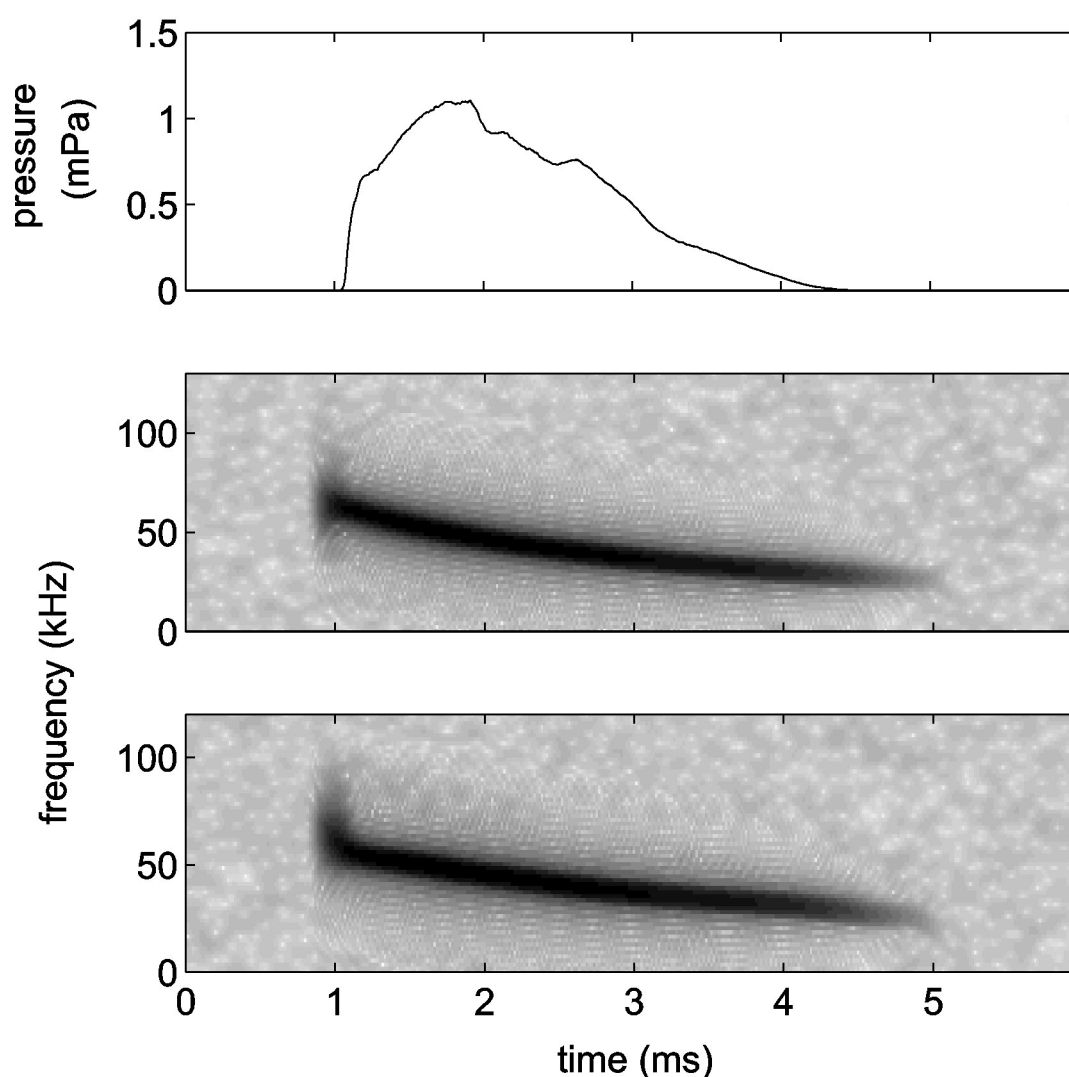


Fig. IV-8. The effect of signal onset on the uniqueness of frequency components of the first harmonic. Data are from a signal, which was not part of the analysis, emitted by EF4 and was chosen, because it had a particularly clear initial steep portion. Top panel shows envelope of fundamental. Middle panel shows the spectrogram of the envelope modulated by the best hyperbolic (LP) fit to the instantaneous frequency. Bottom panel shows the spectrogram of the harmonic without further modification..

This can perhaps be seen as a way of creating a high bandwidth “for free”, since a voiced vocalization will naturally tend to results in a harmonic structure. The advantage of the harmonics is however reduced if there are “blind spots” in the spectra where the harmonics do not close the gab. We assume this to be the reason for the bats to make sweeps that generally are just over one octave outside of the search phase where sweeps are shallower, presumably for detection purposes (e.g. Jensen and Miller, 2001).

The presence of harmonics means that for the higher frequencies in the calls there

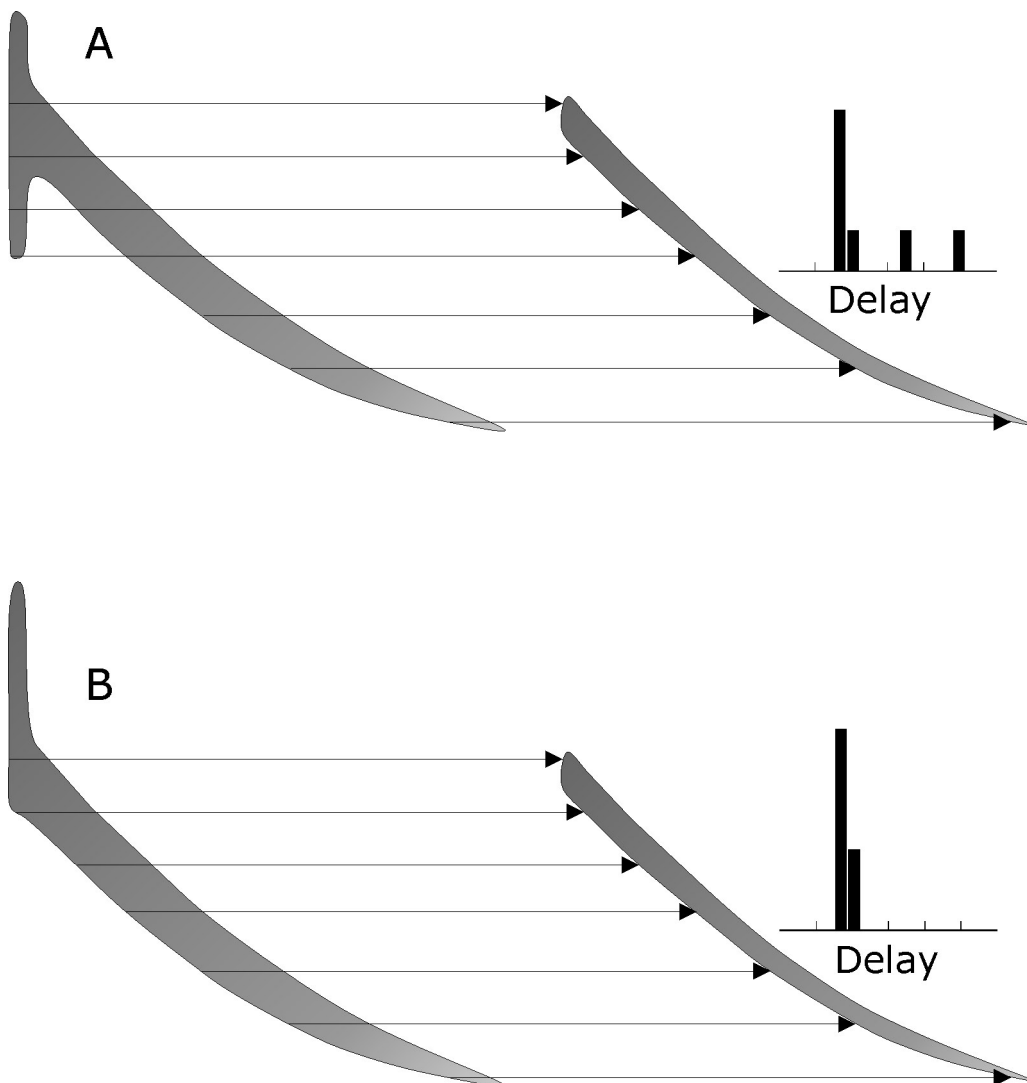


Fig. IV-9. Schematized illustration of the potential negative effects of an abrupt signal onset and its remedy by using a very high modulation rate at the very beginning of the call. The small histograms show the distribution of the lengths of the arrows “measuring” the delays between threshold crossing for the calls and the returning echoes. A: A first harmonic with abrupt onset and no steep sweep in the initial portion. B: A similar signal but with an initial steep sweep rate.

are often several time instances when the same frequency component occurs. For example, the low frequency part of the 3rd harmonic will occupy a frequency range where the 2nd harmonic also has some energy. The first harmonic is free from such overlap up to an octave above the lowest frequency in the sweep. Thus, it is perhaps not surprising that Surlykke (1993) in a psychophysical experiment found the fundamental harmonic to be more critical to the bat in a ranging task than were the higher harmonics together. For ranging purposes it is essential that the frequencies of the transmitted signal occur only once, otherwise ambiguities arise.

The analysis made here also has the potential of shedding some new light on this point in that we can see from the spectra of the individual harmonics (Fig. IV-1) that at a level -35 dB down from the peak of the first harmonic, the energy of the low-frequency part of the third harmonic overlaps with the high-frequency part of the fundamental. So there is no frequency for which the second harmonic is the only one to potentially activate the hearing system and thus be an unambiguous time marker.

The envelopes of the initial part of the first harmonic of the calls analyzed here are rising very steeply, typically in 120 to 150 μ s. We believe this may be a side product of the sound production mechanism: Making a “soft” onset to a sound that lasts 2-3 ms and has a peak intensity of ca. 100 dB SPL seems a very tough task indeed. The consequence is that the envelope of the signal actually has a considerable bandwidth at the onset. This can be seen in a spectrogram if one modulates the envelope suitably as it is done in Fig. IV-8. If this frequency “splatter” is allowed to enter the frequency range occupied by the lower frequencies of the fundamental, the advantage of this frequency range – that it is free from ambiguities – is partly lost.

Moreover, this broadband initial component will most likely only be perceived in the vocalization and not as a part of the echo, since it contains relatively little energy compared to the frequency modulated parts of the signal. It therefore might well create the situation that for some frequencies the signal components that trigger the “stop watch” in the signal are different from the components that stop it. This may result in erroneous delay estimates. We have illustrated this phenomenon in Fig. IV-9.

By starting at a very high instantaneous frequency the frequency splatter can be “delivered” above the general course of the first harmonic and the problem described above can thus be avoided.

Conclusion

The cross-correlation extraction and the introduced morphing technique for bat signals in this study has been shown to be a superior analyzing tool for the problem of the time-frequency course and exact envelope of bats signals for signals recorded with good signal-to-noise ratios.

As of yet no proposed function can account perfectly for the time-frequency course of the calls investigated here and most of the deviations from the fitting functions seem to be explainable by a very steep initial sweep rate. We propose that this initial part is an adaptation to the presence of harmonics that may introduce ambiguities in the bats range estimate procedure.

Appendix III-A

```

function [harmonics]=harmsepa(varargin);
%
% "Manual" bat call signal harmonics separation.
% harmsep('wavfilename',sr) allows the separation of the harmonic
% components in the single bat call stored in the wavfile. The optional
% parameter 'sr' specifies a sampling rate for display of time and
% frequency parameters. Default is 1.
%
% Instead of 'wavfilename' an array containing the samples of the call
% may be used as the first input parameter. Several further parameters
% must be entered during the run, when spectrogram displays of the
% signals are displayed as guides.
%
% If succesful separation is to be achieved a linear time-frequency
% profile must be chosen wisely.

if nargin==0
    error('You must specify a wav-file or a signal array');
else
    try
        sig=wavread(varargin{1});
    catch
        %okay if wav-file
        if ~ischar(varargin{1}) %it might be an array
            try
                sig=varargin{1};
                %OK if array as input
                sig(100);
                %error if length of array < 101
            catch
                error('doesn''t appear to be a proper bat signal');
            end;
        else
            error(['file '' varargin{1} '.wav'' not found']);
        end;
    end;
end;
%no more error handling in this script

```

```

si=size(sig);          %force array into a row
if si(2)==1
    sig=sig';
end; clear si;

try
    sr=varargin{2};    %second input parameter optionally sampling rate
catch
    sr=1;
    disp('using default sampling rate of 1 ...')
end;

clf;                   %new graph window

specgram(sig,512,sr,128,120);    %first display the signal
zdyn=get(gca,'CLim');           %save dynamic range
input('press enter to zero pad');
                                %a small pause to behold the signal

OL=length(sig);              %original length
sig(2^(ceil(log(OL)/log(2))))=0;
                                %zero pad signal up to next power
                                %of 2 for real FFT use
L=length(sig)                 %new temporary length
result=zeros(1,L);           %initially the result is meagre
HarmonicNumber=0;           %and there are no harmonics yet

while 1                      %repeat until user break
    HarmonicNumber=HarmonicNumber+1;
    sig=sig-result;
        %subtract the isolated harmonics from signal
    SIG=fft(sig).*[ones(L/2,1)' zeros(L/2,1)']; %onesided spectrum

    specgram(sig,512,sr,128,120);    %display signal
    set(gca,'CLim',zdyn);           %use old dynamic range

    if HarmonicNumber==1
        disp('A profile must be specified for the lowest harmonic.')
    end;

    f0=input('Enter the frequency at time zero (0 to exit): ');
    if ~f0
        break                        %a graceful way out
    end;

    vt0=input('time of 0 kHz: ');

    PhaseShift=cumsum(((t0/2*sr+(t0/f0)*[1:L/2]*sr/L)*2*pi)/L*sr);
                                %This is the only tricky bit
    PhaseShift(L)=0;           %ignore negative frequencies

    morph=2*real(ifft(abs(SIG).*(cos(phase(SIG)-PhaseShift)...
        +i*sin(phase(SIG)-PhaseShift))));
                                %the profile has now been erected

    subplot(2,2,1)
    specgram(sig,512,sr,128,120)
    set(gca,'CLim',zdyn);
    line([0 t0],[f0 0]);        %Display the profile

```

```

subplot(2,2,2)
specgram(morph,512,sr,128,120)           %show the morphed signal
set(gca,'CLim',zdyn);

if input('recenter? (1 for yes 0 for no): ');
    morph=fftshift(morph);
        %Sometimes necessary
    shiftbool=1;                          %Flag if it was done
else
    shiftbool=0;                          %... or not
end;

clf;
specgram(morph,512,sr,128,120)   %big picture
set(gca,'CLim',zdyn);

if HarmonicNumber==1
    disp('An area with the morphed harmonic must be specified.')
end;
TimeWindow=zeros(1,L);
        %Now time limit morphed signal
TimeWindow(round(sr*input('time window begin: '))...
    :round(sr*input('time window end: ')))=1;
        %square window for simplicity
timelimmorph=morph.*TimeWindow; %time limited morphed signal
TIMELIMMORPH=fft(timelimmorph); %spectrum of morphed signal (again)
FrequencyWindow=zeros(1,L); %mostly zero
FrequencyWindow(round(input('Lowest frequency in passband: ')*...
    L/sr):round(input('Highest frequency in passband: ')*L/sr))=1;
        %square window for simplicity
TIMELIMBANDMORPH=FrequencyWindow.*TIMELIMMORPH;
        %bandlimit signal
result=2*real(ifft(abs(TIMELIMBANDMORPH).*...
    (cos(phase(TIMELIMBANDMORPH)+PhaseShift)+...
    i*sin(phase(TIMELIMBANDMORPH)+PhaseShift))));
        %undo the morphing for the pruned signal
if shiftbool
    result=fftshift(result); %if it was shifted - shift it back
end;
subplot(2,2,1)
specgram(sig-result,512,sr,128,120)
        %something to look at in the end

subplot(2,2,2)
specgram(morph,512,sr,128,120)
subplot(2,2,3)
specgram(2*real(ifft(TIMELIMBANDMORPH)),512,sr,128,120)
subplot(2,2,4)
specgram(result,512,sr,128,120)
for plots=1:4
    subplot(2,2,5-plots);
    set(gca,'CLim',zdyn);
end;
harmonics(HarmonicNumber,[1:OL])=result(1:OL);
        %return a matrix with the values
end;

```

References

- Altes, R. A. and Titlebaum, E. L. (1970). "Bat signals as optimally Doppler tolerant waveforms," *J. Acoust. Soc. Am.* **67**, 1232-1246.
- Cahlander, D. A. (1964). "Echolocation with Wide-Band Waveforms: Bat Sonar Signals," MIT Lincoln Lab. Tech. Rep. **27**, 1964.
- Cahlander, D. A., McCue, J. J. G. and Webster, F. A. (1964). "The determination of distance by echolocating bats," *Nature* **201**, 544-546.
- Capus, C. and Brown, K. (2003). "Short-time fractional Fourier methods for the time-frequency representation of chirp signals," *J. Acoust. Soc. Am.* **113**, 3253-3263.
- Grossetête, A. and Moss, C. F. (1998). "Target flutter rate discrimination by bats using frequency-modulated sonar sounds: Behavior and signal processing models," *J. Acoust. Soc. Am.* **103**, 2167-2176.
- Masters, M. W., Jacobs, S. C. and Simmons, J. A. (1991). "The structure of echolocating sounds used by the big brown bat *Eptesicus fuscus*: Some consequences for echo processing," *J. Acoust. Soc. Am.* **89**, 1402-1413.
- Masters, W. M., Raver, K. A. S. (2000). "Range discrimination by big brown bats (*Eptesicus fuscus*). using altered model echoes: Implications fro signal processing," *J. Acoust. Soc. Am.* **107**, 625-637.
- Parssons, S. and Jones, G. (2000). "Acoustic identification of echolocating bat by discriminant function analysis and artificial neural networks," *J. Exp. Biol.* **203**, 2641-2656.
- Kalko, E. K. V. and Schnitzler, H.-U. (1993). "Plasticity in echolocation signals of European pipistrelle bats in search flight: implications for habitat use and prey detection," *Behav. Ecol. Sociobiol.* **33**, 415-428.
- Kick, S. A. and Simmons, J. A. (1984). "Automatic gain control in the bat's sonar receiver and the neuroethology of echolocation," *J. Neurosci.* **4**, 2725-2737.
- Kroszczyński, J. J. (1969). "Pulse Compression by means of Linear Period Modulation," *Proc. IEEE* **57**, 1260-1266.
- Lin, Z.-B. (1988). "Wideband ambiguity function of broadband signals," *J. Acoust. Soc. Am.* **83**, 2109-2116.
- Menne, D (1988). "Is the structure of bat echolocation calls an adaptation to the mammalian hearing system?" *J. Acoust. Soc. Am.* **83**, 2447-2449.
- Simmons, J. A. (1987). "Behavior as a component of an auditory signal-processing system: Stabilization of echo spectra in bat sonar". *Assoc. Res. Otolaryniol. Abstr.* **10**, 82.

-
- Simmons, J. A. and Stein, R. A. (1980). "Acoustic imaging in bat sonar: echolocation signals and the evolution of echolocation," *J. Comp. Physiol. A* **135**, 61-84.
- Surlykke, A. (1992). "Target ranging and the role of time-frequency structure of synthetic echoes in big brown bat, *Eptesicus fuscus*," *J. Comp. Physiol. A* **170**, 83-92.
- Troest, N. and Møhl, B. (1986). "The detection of phantom targets in noise by serotine bats; negative evidence for the coherent receiver" *J. Comp. Physiol.* **159**, 559-567.

ZUSAMMENFASSUNG

Das Ziel dieser Arbeit ist die Klärung der Frage, wie gut Fledermäuse aus der Echolaufzeit die Zielentfernung schätzen können. Ein großer Teil der Arbeit befasst sich mit theoretischen Überlegungen zur Entfernungsmessung, ein weiterer Teil präsentiert neue Methoden für die Analyse der Ortungslaute von Fledermäusen.

Viele Veröffentlichungen befassen sich mit der Hypothese, dass Fledermäuse für die Entfernungsschätzung alle Echoparameter inklusiv der Phase analysieren. Diese Hypothese ist umstritten, da einerseits bei den Ortungslauten der Fledermäuse fast alle Energie im Ultraschallbereich über 20 kHz liegt, andererseits aber bei keiner Tierart eine Phasenempfindlichkeit über 9 kHz festgestellt wurde.

Die Ergebnisse der Arbeit werden unter Berücksichtigung dieser und anderer Hypothesen diskutiert.

In der **Einleitung** werden die bisher veröffentlichten Untersuchungen zur der Theorie zusammengefasst, dass Fledermäuse für die Entfernungsmessung einen voll kohärenten Kreuzkorrelations-Empfänger einsetzen. Der wichtigste Befürworter dieser Theorie ist James A. Simmons. In den 70er und 80er Jahren veröffentlichte er die Ergebnisse einer Serie von psychophysischen Experimenten, die seiner Meinung nach belegen, dass der Verlauf der Autokorrelationskurve eines Ortungslautes mit der Zahl der Fehler bei der Abstandsmessung (psychometrische Leistungskurve) direkt korreliert. Es wird gezeigt, dass diese Beziehung nicht existiert und dass die Autokorrelationsdaten nur schwach mit den veröffentlichten Leistungskurven übereinstimmen.

Kapitel 1 befasst sich kritisch mit den Ergebnissen einer Veröffentlichung von Simmons (1990), in der eine Auflösungsschwelle von 10 ns bei der Laufzeitunterscheidung beschrieben wurde. In einem range jitter Experiment (2AFC) hatten die Fledermäuse (*Eptesicus fuscus*) die Aufgabe, ein stationäres Ziel von einem zwischen zwei Positionen hin und her bewegten Ziel (Jitterziel) zu unterscheiden. Um zu überprüfen, ob eine Auflösungsgrenze von 10 ns überhaupt physikalisch möglich ist, wurde experimentell überprüft, wie genau bei fast perfekten Phantomzielen der Laufzeitunterschied mit Hilfe von Kreuzkorrelation gemessen werden kann.

Bei einer einfachen Echoortungsaufgabe wurden Ortungslaute und künstliche Echos mit einem Mikrophon aufgenommen, digital gespeichert und analysiert. Mit den aufgenommenen Signalen war in einer simulierten jitter Unterscheidung eine Auflösung im Bereich von 10ns nur möglich, wenn eine systeminterne Phasenverschiebung in der Übertragungsfunktion des elektronischen Systems mit $\pi/2$ korrigiert wurde. Ohne

diese Korrektur konnte nur eine jitter Auflösung von etwa 100 ns erreicht werden. Es gibt Hinweise darauf, dass so eine systeminterne Phasenverschiebung auch bei den Experimenten von Simmons zu finden war. Das legt nahe, dass seine Fledermäuse andere cues als ausschließlich die Laufzeitdifferenz benutzt haben, um die Schwellenleistung von 10 ns zu erreichen.

Kapitel 2 setzt sich mit einer Veröffentlichung von Masters und Raver (2000) auseinander, in der die Autoren beschreiben, dass die Laufzeitunterscheidung bei *Eptesicus fuscus* durch Manipulation der Krümmung der Laute schlechter wurde, nicht jedoch bei Veränderung anderer Lauteigenschaften.

Es wird gezeigt, dass der Krümmungsgrad eines Lautes einen Einfluss auf die Leistung eines Filterbankempfängers hat, nicht jedoch auf die eines phasenempfindlichen Kreuzkorrelationsempfängers. Hieraus kann geschlossen werden, dass für Fledermäuse das Filterbankmodell das wahrscheinlichere Empfängermodell ist. Es wird zudem gezeigt, dass die hyperbolische Modulation von Ortungslauten als Anpassung an bisher noch hypothetische Frequenzmodulationsfilter im auditorischen System von Fledermäusen erklärt werden könnte.

Kapitel 3 setzt sich mit einem Modell auseinander, das von Saillant et al 1993 vorgestellt wurde. Dieses Modell beinhaltet die Simulation eines fledermaustypischen auditorischen Systems, das eine hohe Laufzeitauflösung durch Kreuzkorrelation mit darauf folgender Summierung von simulierten sehr kurzen Nervenimpulsen erreicht. Das Modell sollte u.a. die von J.A. Simmons gemessene Unterscheidungsschwelle von 10 ns erklären.

Anhand einer vereinfachten Version des Saillant Modells wird gezeigt, dass das Modell bei zu geringer Abtastrate eine phasentreue Kreuzkorrelation durchführt. Durch Bandpassfilterung dieser Kreuzkorrelation verschwindet jedoch der Unterschied zwischen dieser und einer gewöhnlichen Kreuzkorrelation. Es wird dargestellt, dass die in diesem Modell angenommene phasentreue Antwort der Hörfasern keine Unterstützung in der Literatur findet. Abschließend wird die physiologische Relevanz dieses Modells diskutiert. Eine interessante Methode der Signalabtastung durch „peak detektion“ wird im Anhang beschrieben.

Kapitel 4 befasst sich mit Methoden der Lautanalyse und der Lautbeschreibung und erläutert diese am Beispiel der Struktur der Ortungslaute von *Eptesicus fuscus*.

Im ersten Teil wird mit einer Monte Carlo Simulation untersucht, inwiefern die Daten einer früheren Studie von Masters et al (1991) verifiziert werden können. In dieser Studie, wurde die Zeit-Frequenzstruktur folgendermaßen bestimmt: Das Signal wurde in kleinere Stücke aufgeteilt und jedes Stück um die relevante Harmonische herum mit einem Bandpass gefiltert. Für jedes dieser Stücke des Signals wurde die Bestfrequenz

vermessen und daraus der Frequenz-Zeitverlauf ermittelt. Der beste Fit auf diese Daten war eine Funktion auf logarithmischer Zeitbasis. Die Ergebnisse meiner Monte Carlo Simulation zeigen, dass das Signal-Rausch Verhältnis hoch sein muss, um mit dieser Methode zuverlässige Ergebnisse zu erhalten.

Im zweiten Teil dieses Kapitels werden 2 einander ähnliche Methoden zur Trennung der harmonischen Komponenten von Fledermauslauten beschrieben. Diese Methoden beruhen beide auf einer kontrollierten Verschiebung der Gruppenlaufzeit. Das Signal wird dabei so transformiert, dass ein rechteckiges Zeit-Frequenz Fenster benutzt werden kann, um die Harmonischen voneinander zu isolieren. Das Resultat wird zurück transformiert. Diese neuen Methoden ermöglichen es, die instantanen Frequenzen jeder einzelnen Harmonischen zu bestimmen. Eine dieser Techniken wird angewandt um im Detail die Zeit-Frequenzstruktur der Laute von vier *E. fuscus* zu analysieren. Die Ergebnisse zeigen starke interindividuelle Variationen in der Lautstruktur. Ein weiteres Ergebnis ist, dass der Anfangsteil der Rufe steiler als bisher angenommen moduliert ist. Dieser Befund würde erklären, warum die von Masters et al benutzte logarithmische Zeitfunktion den besten Fit ergibt, da sie sehr steilen Anfangsmodulationen von Lauten gut folgen kann. Es wird die Theorie aufgestellt, dass dieser steile Anfangsteil dazu dient, die spektrale Verschmierung, die durch einen abrupten Amplitudenanstieg des Signals entsteht, auf die höheren Frequenzen zu verschieben. Dies würde dazu führen, dass innerhalb jeder Harmonischen jedes Frequenzereignis nur einmal vorkommt, was für die Abstandsmessung vorteilhaft ist.

SUMMARY

The main thrust of this thesis is concerned with how bats analyze the returning echoes for range information and the hypotheses proposed by researchers to explain bat echolocation behavior. The analysis of bat echolocation signals by both bats and researchers is covered. Most parts deal with theories of how the bats analyze echoes for target range.

Many elements of the thesis directly touch on or have consequence for an ongoing debate concerning the theory that bats are able to analyze echoes and utilize all information contained in these signals, including phase. Since the echolocation calls of many bat species contain significant energy at frequencies above 100 kHz, and since no known animal has a demonstrated auditory phase preservation mechanism above 9 kHz, the theory is a highly controversial.

In the **introduction** I summarize some of the historical data that have been presented for and against the theory of a cross-correlation (CC) receiver in bats. The main proponent of the theory is James A. Simmons. In the 1970ies and 1980ies Dr. Simmons published a series of psychophysical experiments claiming that the shape of the autocorrelation function of a bat call is related directly in the number of errors the animals make when judging the range to a target (performance curve). I show that there is no such relation, and that the autocorrelation data do in fact correspond rather poorly to the performance curve in these experiments.

Chapter I is concerned with a result that was published in 1990, also by J. A. Simmons. In that paper a threshold of only 10 ns was found in a so-called range-jitter experiment where the bats (*Eptesicus fuscus*) must judge which of two simulated targets is changing its delay between successive calls. The present investigation concerns the lower limits to what can be achieved in a target simulation system if perfect CC processing is used. Bat sounds and artificial echoes are collected using a single microphone during a simple echolocation task and the calls and echoes are digitized directly to disk and analyzed afterwards. It is found that a jitter resolution on the order of tens of nanoseconds was possible, but only after a phase shift of $\pi/2$ in the system was corrected. Without this correction a jitter resolution of about 100 ns could be achieved. There is very strong evidence for the existence of such a phase shift in the system used by Simmons et al. (1990) and therefore my results agree well with the notion that other cues might have been available to the bats apart from the pure delay difference.

Chapter II represents an interpretation of data published in 2000 by Masters and Raver. These authors reported that the delay discrimination by the bat *E. fuscus* deteriorated when the curvature of the calls was manipulated, but not when other aspects of the calls were changed. It is shown that this curvature parameter has influence on performance in a so-called filter bank receiver but not on a phase preserving CC receiver, leading to the interpretation that the filter bank is the more relevant model for bats. Also it is shown that the existence of an »instantaneous sweep rate filter« in the auditory system of bats would explain the near hyperbolic call structure of many bats.

Chapter III is concerned with a model proposed by Saillant et al. in 1993. This model of the auditory system of bats achieves a high delay resolution through the CC and summation of impulses in a simulated auditory system. In a simplified version of that model it is shown here that in essence it performs an under-sampled phase preserving CC between call and echo waveforms. A very close match with the usual CC function is achieved through band-pass filtering the output. The physiological relevance of the model is discussed. It is noted that the phase-locked responses of the auditory »fibers« that are an important part of the model has no base in the literature. Some interesting aspects of sampling a signal by peak-detection are mentioned in the appendices.

Chapter IV returns to the subject of the sweep structure of bat echolocation sounds, specifically for *E. fuscus*. The chapter falls in two main parts. First, using a Monte Carlo simulation method, I investigated to what extent the results from a previous study by Masters et al. (1991) were reliable. In that study the estimates of the time-frequency structure were obtained by partitioning the signal, band-pass filtering around the first harmonic, and finally estimating the mean frequency in each part. The resulting time-frequency array was then fitted to different types of functions. The best fit in the Masters et al. study was found with a »logarithmic time« function. I demonstrate that the signal to noise ratios must be rather high for the results to be reliable using the above method. In the second part of the chapter, two closely related methods for separating bat calls into its harmonic components are devised. The methods utilize a controlled change in group delay which »morphs« the signal so that a rectilinear time-frequency window can be used to isolate each harmonic in turn. The result is then morphed back. These new methods allow one to calculate the instantaneous frequency for each harmonic. One of these techniques is then used to investigate in detail the time-frequency structure of the bat calls from 4 *E. fuscus*. The main results show considerable individual variation in the call structure, and that the initial portion of the calls is very steeply frequency modulated. This last finding would explain why the logarithmic time function gives the best fit, as it can take on a very steep initial segment. I speculate that the initial steep portion serves the purpose of »lifting«

the frequency splatter, which results form a very abrupt onset of the call, out of the lower frequency range covered by the first harmonic. This helps to ensure that each frequency event occurs only once in the range covered by that harmonic, which is a great advantage if the signals are used in ranging.

ACKNOWLEDGMENTS

I would like first of all to direct my gratitude toward Prof. Dr. Hans-Ulrich Schnitzler, who has most generously let me use his know-how, funding, man-hours, laboratory space, workshops and other facilities during my stay in his laboratory. I am amazed by the workload this man allows himself to be thrown upon him – and he still gets the job done! I was never an easy student to supervise – thanks for putting up with me ...

I am particularly grateful to Richard A. Altes and Bertel Møhl who have reviewed the thesis in part. I have received invaluable competent comments.

The role of the electronic workshop (Adolf Cwein, Anton Thorwart and Thomas Steinbrenner) during my time in Tübingen can hardly be underrated. Always helpful and competent (sounds a bit like a boy scout thing).

A special thank toward Jörg Kindermann, whose incredible expertise in matters of electronics never ceased to impress me. But most amazingly, he never backs out of a chance to help you - a rather dangerous trait, that one.

The work of Herr Schneider defies description. I have experienced other accomplished technicians, but never anyone who made the end product so thoroughly beautiful.

Dr. Arjan M. Boonman deserves all the credits he can carry on his athletic(ly looking) shoulders for being that extremely positive, energetic, helpful, and generally likeable person that he is. He has been a great help with – among lots of other stuff – his large knowledge of statistics and the calls of European bat species. He also was very helpful in translating the summary into german

Peter Stilz also deserves lots of credit for his willingness and competence in discussing matters that were probably mostly of interest to me. I hope to be able to make it up to you one day. No mention of shoulders necessary there.

Dr. Annette Denzinger, Dr. Peter Pilz and Dr. Dieter Menne helped out in matters of target simulation and in other practicalities.

Dr. Björn Siemers was always a great help, regardless of the problem you presented him with. Even out of his depths, that man will (and can) lend you a helping hand ...

Ingrid Kaipf provided me constantly with her always-insightful opinions on animal care,

training and all sorts of other matters regarding bats. She also helped in other practical ways, besides being a very likeable human being, which is always a helpful character trait.

Andrea Schaub played a crucial role in translating the summary into German. And she helped in operating a tape recorder, but besides, sharing an office with her is great fun in a sort of subtle, cynical way that defies precise description.

Dr. Joachim Ostwald was always a keen listener and an insightful adviser whenever asked. He often was.

Bertel Møhl, Dr. Peter T. Madsen and Dr. Magnus Walberg of the University of Århus are all due special thanks for lending me an ear (and several books, now I come to think of it) when I needed one in the final stages of this work.

The thesis of Dr. Esther J. Langeheinecke unwittingly provided me with very direct guidance in the layout of this thesis.

In many other ways I am in particularly indebted to Susan Sümer, Dietrich von Staden, Erika Sampaio, Eric Wosname, Eva Lohmann, Jacob Fahr, Lars Nissen and Rolf Müller, as well as all contemporary members of the Graduierten Kolleg Neurobiologie Tübingen and possibly many others, who I do hope will never notice, if I left them out. I did not mean to! Honestly!!

Thanks to my parents, my brother and my parents in law for all kinds of valuable and highly valued support.

One thanks that I could never hope to express with a fitting weight goes to my wife who willingly went abroad with me in spite of all foreseeable costs to her own career, and who has never even started to question the general sense of this whole project. She has supported me morally, economically and in all sorts of other ways during the last 19 months. (And she read proofs expertly, but that sort of fades in comparison ...)

My daughter, Miriam, made everything a great adventure. If I had not done this for myself, I would gladly have done it for you.

This work was supported by: The Danish National Research Council; DFG (SFB 550); Graduiertenkolleg Neurobiologie Tübingen and Centre for Sound Communication, University of Southern Denmark.

LEBENS LAUF

Name: Kristian Beedholm
 Geboren am: 18. Juni 1968 in Kobenhagen, Dänemark
 Eltern: Jette Gynther Beedholm, geb. Larsen und
 Henrik Bidstrup Beedholm

Schulbildung

1975 – 1982 Kollerup skole, Jelling, Dänemark
 1982 – 1984 Bredagerskolen, Jelling, Dänemark
 8/1984 – 6/1987 Rødkilde Gymnasium, Vejle, Dänemark
 18/6 1987 Abitur

Militärdienst

1988 „Zivilverteidigung“ – Dienst abgeschlossen mit Rang eines
 Feldwebels

Studium

1988 – 1989 Allgemeine informationstechnologische Kurse
 1989 – 1990 Kurse in Chemie, Physik, Mathematik
 1990 – 1992 Grundstudium Biologie an der Universität Århus
 1992 – 1997 Hauptstudium Biologie an der Universität Århus
 09/1996 – 03/1997 Wissenschaftliche Hilfskraft, Århus Universität
 03/1997 Diplom in Biologie an der Universität Århus Institut für
 Tierphysiologie
 (Bertel Møhl) “Bat sonar: an alternative explanation for the
 10 ns jitter result”
 07/1999 – 10/1999 Wissenschaftlicher Assistent am Institut für
 Tierphysiologie, Århus Universität
 12/1999 – 02/2003 Arbeit an der Dissertation am Lehrstuhl Tierphysiologie
 Tübingen bei Prof. Dr. H.-U. Schnitzler
 12/1998 – 01/1999 Promotionsstipendium des Graduiertenkollegs
 Neurobiologie Tübingen
 12/1998 – 12/2000 Mitglied des Graduiertenkollegs Neurobiologie Tübingen
 02/1999 – 01/2002 Promotionsstipendium des Nationalen Dänischen
 Naturwissenschaftlichen Forschungsrats (SNF)
 09/2003 – 12/2003 Wissenschaftliche Hilfskraft, Århus Universität
 12/2003 – 01/2004 Software Entwickler, Tierphysiologie, Århus Universität

Im Selbstverlag herausgegeben von:

Beedholm, Kristian

Stenaldervej 43

8220 Brabrand

DÄNEMARK