THE HETERODYNE FILTER OF THE INSTITUTE OF PHONETICS

Jørgen Rischel and Svend-Erik Lystlund

In the 1966 report of the Institute a brief mention was made of a heterodyne bandpass filter under work (1). The filter has now by and large reached the stage of completion aimed at in the said report, and it has been taken in use in our research (see this volume, pp. 35 ff). On account of the rather special nature of the device it seems reasonable to give a more detailed description of it here.

A great variety of electronic wave filters are used in phonetic research, serving a number of different purposes: frequency selective analysis of the speech-wave (the sonagraph, etc.), pre-shaping of the signal applied to intensity meters or pitch extracting devices, synthesis of speech, amplitude distortion of natural speech (for perceptual tests), and so on. Obviously, the specifications to be met by a filter vary according to the purpose, and the concept of a "standard filter" has limited validity. For purposes involving rather unselective filtering (pre-emphasis) it is generally a relatively simple task to construct individual networks meeting the individual demands. However, when very sharp filtering is required as it may be the case in experiments with distorted speech or in frequency analysis of the speech-wave, construction of the devices becomes quite complicated, and it may be most expedient to have access to a versatile unit which can perform a sharp filtering (highpass or lowpass or bandpass) at any frequency desired. The heterodyne filter is intended as such a standard device for sharp filtering.

It is somewhat difficult to argue in a general way about the merits or drawbacks of a wave filter, since ideal filters do not exist. If a filter is entirely adequate from one point of view it is likely to be insatisfactory from another point of view (although the drawbacks may escape the attention of the user). In most cases one aspect of filter performance (e.g., the "sharpness" of cutoff) is so essential that the phonetician may choose to pay much less attention to other aspects (e.g. phase or amplitude distortion in the passband), but it goes without saying that the filter cannot be characterized in terms of one such property.

This is particularly true of a complicated system like the heterodyne filter.

In the following survey the first section explains the general principle of the heterodyne system, and the second section describes our actual setup. In the last two sections an attempt is made to characterize the performance of the filter from different aspects and to motivate various details of the design.

1. Filtering according to the heterodyne principle.

Fant has given a detailed description (2) of a filter employing the heterodyne principle. However, the strategy used in our setup differs somewhat from his, so it is necessary to go into some detail also in this report.

The motivation for choosing the heterodyne principle is that it makes it possible to combine a very sharp cutoff with a free choice of cutoff frequencies. The idea is that the filter is designed to work at one fixed frequency (which enormously facilitates the design of the filter itself): instead of moving the cutoff frequencies of the filter up and down one moves the signal up and down in frequency so that different parts of the spectrum are allowed to pass through the fixed filter. Bandpass filtering with mutually independent highpass and lowpass functions requires that the procedure of signal frequency shift be performed twice: first the signal is shifted to one frequency location and applied to a filter performing the highpass filtering, and afterwards it is shifted to another frequency location and applied to a filter performing the lowpass filtering. Finally the signal is shifted back to its original frequency location.

Each frequency shift is performed by heterodyning the audio signal with a sinewave from a variable oscillator, the cutoff frequencies of the filtering processes being controlled directly by the settings of the oscillators. The heterodyning of the speechwave is performed in each step by means of a double-balanced modulator. If the frequency limits of the audio signal are denoted by f_{min} (the lowest frequency of interest) and f_{max} (the highest frequency of interest), and the (variable) frequency of the oscillator is denoted by f_{osc}, the output of such a modulator is essentially composed of two frequency bands: the upper sideband reaching

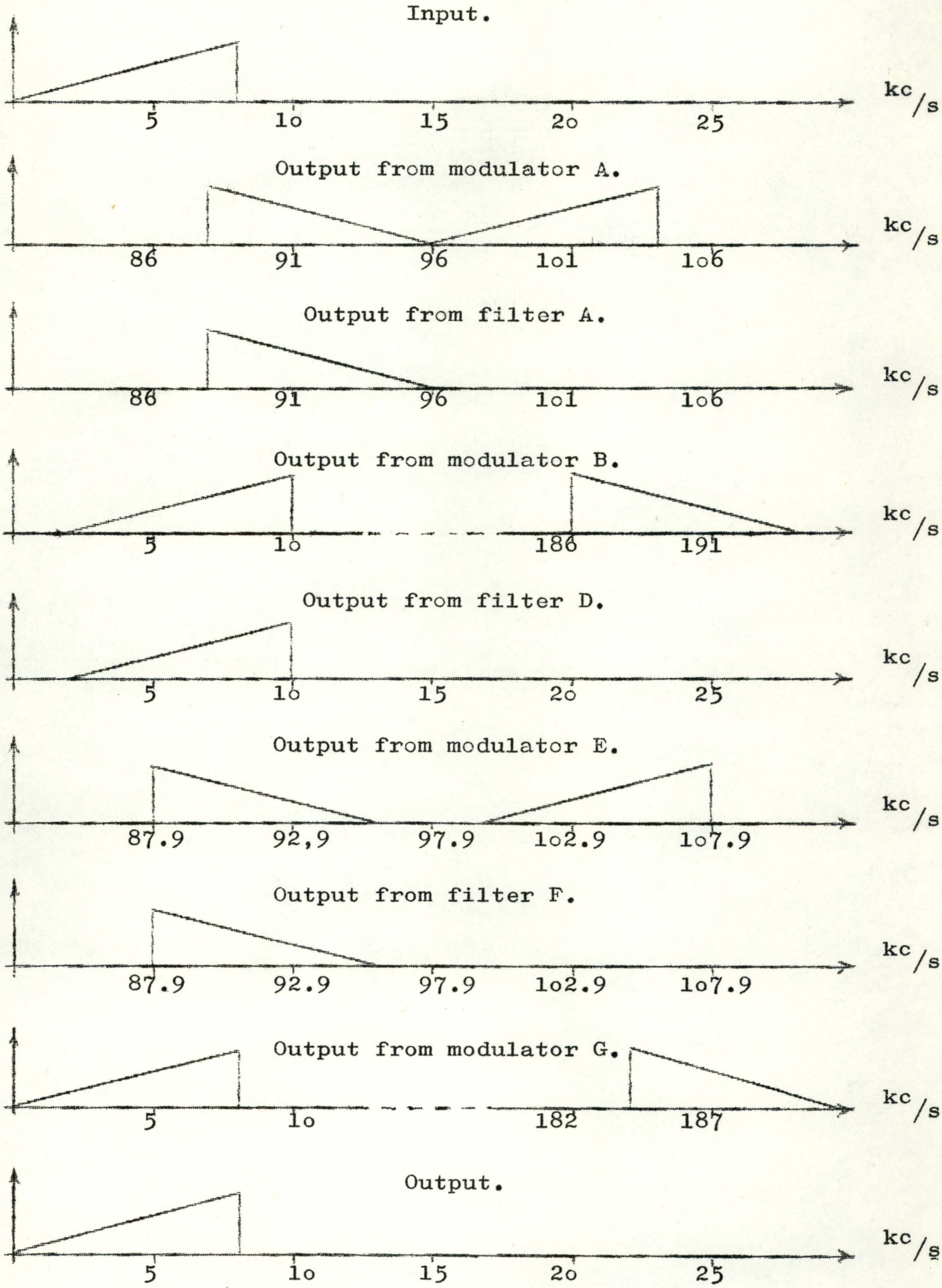


Fig. 1a. HETERODYNE SIGNAL PROCESSING
(cp. Block diagram Fig. 2)
Full frequency range

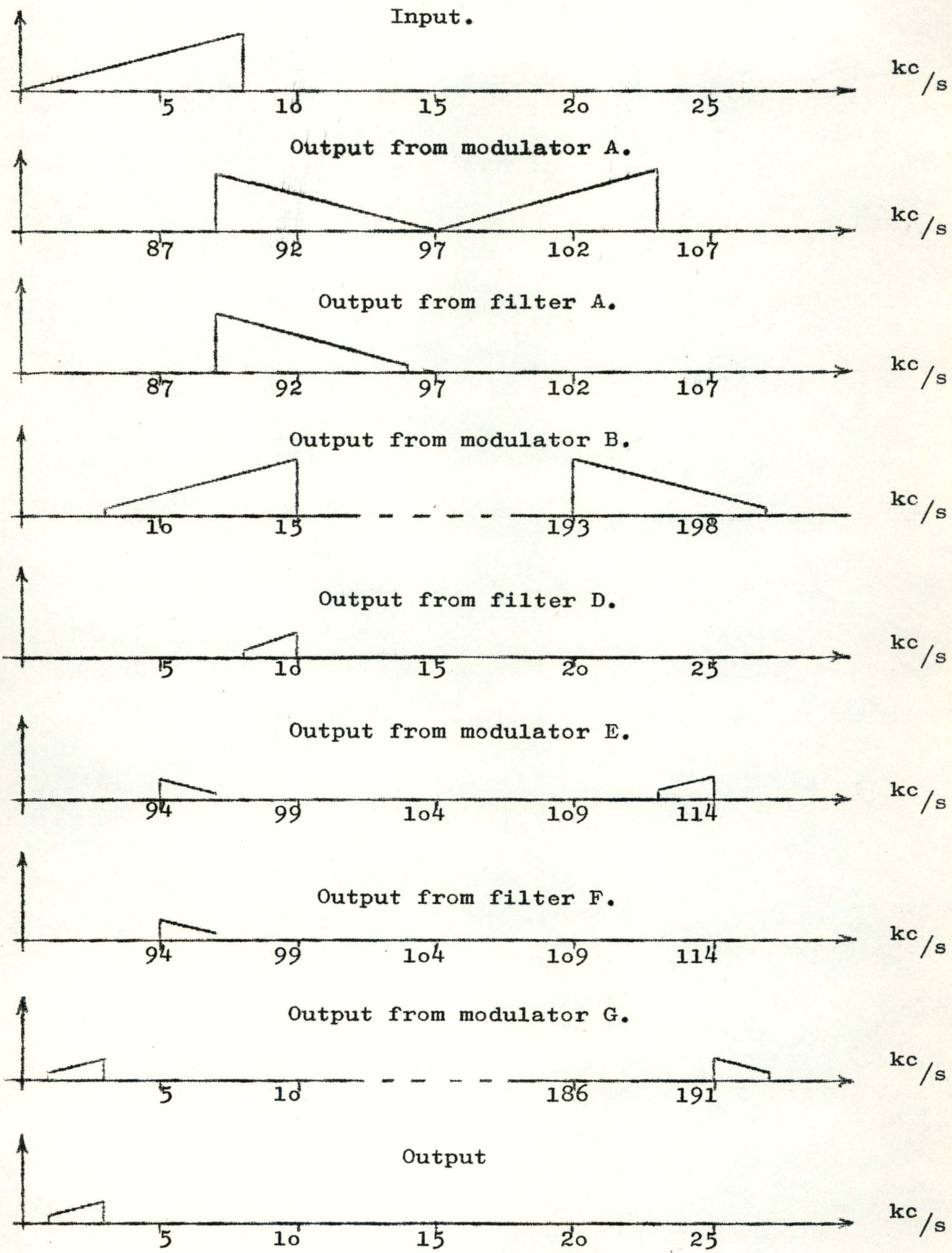


Fig. 1b. HETERODYNE SIGNAL PROCESSING
(cp. Block diagram Fig. 2)
Bandpass 1 kcps - 3 kcps

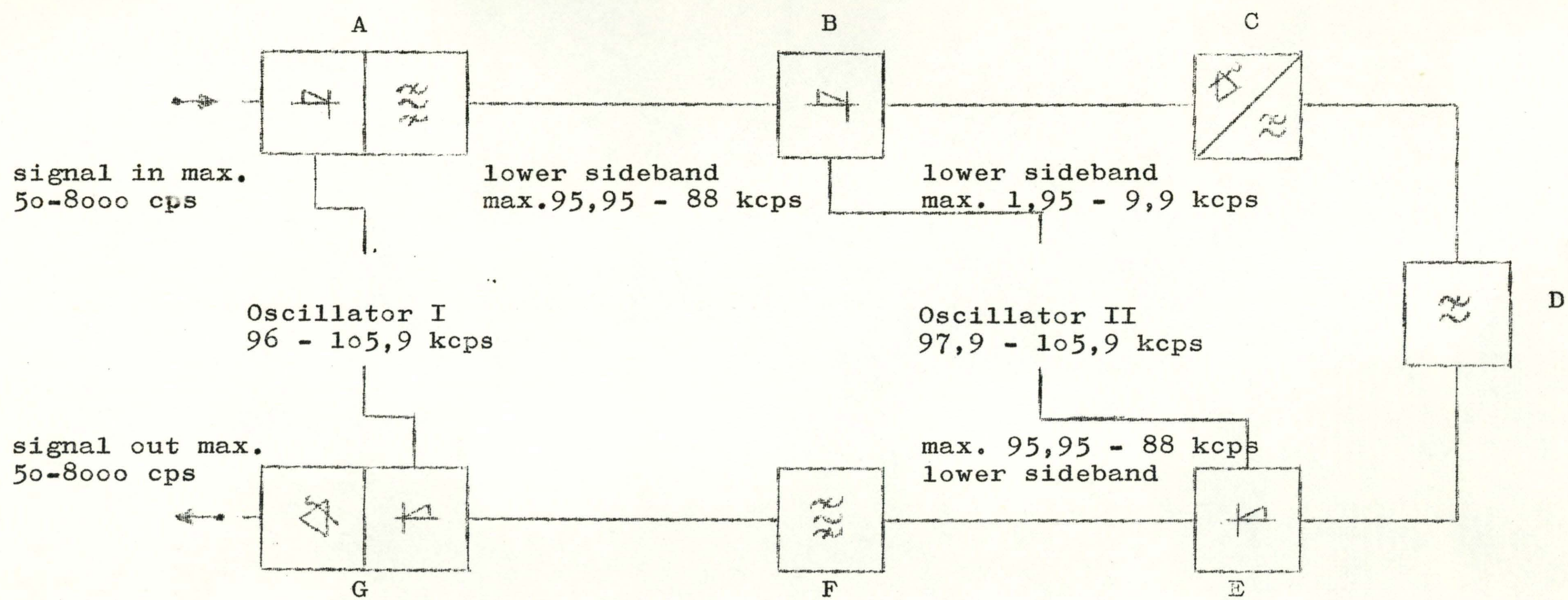
from $f_{osc}+f_{min}$ to $f_{osc}+f_{max}$, and the lower sideband reaching downwards from $f_{osc}-f_{min}$ to $f_{osc}-f_{max}$. Since the entire signal information is contained in each sideband, it is possible to filter out either the upper or the lower sideband and to transmit that one alone through the system, the original signal being restored at the output. In our setup the lower sideband (oscillator frequency minus signal frequency) is used throughout. This entails that at some points in the system the frequency spectrum is inverted, i.e. turned upside down ($f_{osc}-f_{min}$ being a higher frequency than $f_{osc}-f_{max}$). One consequence of this is that the highpass filtering of the signal is in fact performed by means of a lowpass filter.

2. The actual setup.

Fig. 2 gives a block diagram of our heterodyne filter in its present form.

The idea governing this project was to use standard components as building blocks to the greatest possible extent. As reported previously, the Danish Post and Telegraph Department has generously put a good deal of equipment for carrier telephony and radio telephony at our disposal. This equipment included a number of modulators and fixed bandpass filters. Various arrangements of these have been tried out, but since the filtering process implies that the carrier frequency oscillators are tuned off the frequencies for which the units are designed, it has proved difficult to obtain a satisfactory performance at all cutoff frequencies. Whistle interference among the many frequencies generated in the modulators is a very disturbing phenomenon likely to occur unless the utmost care is taken in the choice of frequency bands.

The solution we prefer at present is to perform the highpass filtering (which is physically performed as a lowpass filtering) at a high frequency, and to perform the lowpass filtering at a low frequency. For the former purpose a very sharp filter with a cutoff frequency of 95,95 kcps (block F in Fig. 2) is at our disposal. For the latter purpose we employ a filter (block C-D in Fig. 2) with a sharp cutoff at 9,8 kcps (which of course dictates the upper limit of the entire frequency range of the signal). This filter was designed and built by us at the Institute, whereas all



A Siemens modulator "1" with filter.
Rel 14R 168 al

B Siemens modulator "2"
Rel 15C 71o b

C Equalizing network for filter

D 9,9 kcps lowpass filter

E Siemens modulator "3"
Rel 15C 7o6b

F Siemens filter
Rel str 14R 181 a dt/en

G Siemens modulator "4" with 1o kcps
lowpass filter
Rel 14R 174 al

Fig. 2. HETERODYNE FILTER

other sections (including the 95,95 kcps filter) are commercial equipment.

The maximum **usable** bandwidth of the system is at present 8 kcps. (We intend to replace modulators 2 and 3 by a different type of modulator, which will extend the **usable** frequency range, cp. p. 32.) It is possible, for example, to transmit the frequency range 5o cps - 8 kcps and to suppress frequencies outside this range, cp. Fig. 1a. The carrier frequency of the first modulator is then chosen to be 96 kcps. When audio frequencies are applied to the modulator and converted to sidebands around the carrier the lower of these sidebands (which is inverted) will pass through the 95,95 kcps lowpass filter except for the 5o cps range closest to the carrier frequency, i.e. the sideband will be limited in a way corresponding to a 5o cps highpass filtering of the original audio signal (5o cps being changed to 96-o,o5 = 95,95 kcps, 8 kcps being changed to 96-8 = 88 kcps, and so on). The upper sideband is eliminated. If now the carrier frequency of the second modulator is chosen to be 97,8 kcps, the sideband extending from 95,95 to 88 kcps is moved back to a low frequency location, and turned upside down again. The 9,8 kcps lowpass filter situated after the second modulator limits the sideband in a way corresponding to an 8 kcps lowpass filtering of the original audio signal (95,95 kcps being changed to 97,8o-95,95 = 1,85 kcps, 88 kcps being changed to 97,8o - 88 = 9,8o kcps, and so on). Modulators 3 and 4 reverse this procedure, so that we end up with the original audio signal, only limited to the range 5o cps - 8 kcps. A lowpass filter after modulator 3 (block F in the diagram, Fig. 2) cuts at 95,95 kcps (like the one after modulator 1) and removes the unwanted sideband and distortion products at this stage, and a lowpass filter in the output amplifier ensures that only audio frequency components reach the output. It is seen that at least four filters are needed in the heterodyne system.

If the carrier frequency applied to modulators 1 and 4 ($f_{osc\ 1}$) is raised, obviously a greater portion of the audio spectrum is cut off, since a greater portion of the lower sideband generated in the modulator will be raised above the cutoff frequency of the 95,95 kcps lowpass filter. This corresponds to highpass filtering with a higher cutoff frequency. E.g., a carrier **frequency** of 96,95 kcps conditions an effective highpass filtering at 96,95-95,95 = 1 kcps, and a carrier frequency of 1oo,95 kcps conditions a highpass filtering at 1oo,95-95,95 = 5 kcps.

If the carrier frequency applied to modulators 2 and 3 ($f_{osc\ 2}$) is raised, it is also clear that an increasing portion of the audio spectrum is cut off, since a greater portion of the lower sideband generated in this modulator will be raised above the cutoff frequency of the 9,8 kcps lowpass filter. This corresponds to lowpass filtering with a lower cutoff frequency. E.g., provided that $f_{osc\ 1}$ is kept at 96 kcps (lower sideband 95,95 to 88 kcps), and $f_{osc\ 2}$ is raised to 1oo,8 kcps, we get an effective lowpass filtering at 5 kcps, since an audio signal of 5 kcps is first moved to 96-5 = 91 kcps, and afterwards to 1oo,8-91 = 9,8 kcps so that it just passes through the 9,8 kcps lowpass filter, whereas higher frequencies would be suppressed.

It is, of course, not the absolute frequency of $f_{osc\ 2}$, but the increment of this frequency above $f_{osc\ 1}$ that determinesthe lowpass cutoff. Thus with the carrier of the first modulator set at 1oo,95 kcps (highpass 5 kcps) the carrier of the second modulator must be raised above 1oo,95+5 = 1o5,95 kcps in order to get any signal through the system, 1o5,95 kcps being in this case the proper setting for lowpass cutoff at 5 kcps.

It is seen that the setting of the oscillators demands a bit of calculation work, although an exact adjustment of the carrier frequencies can be carried out more easily by measuring the response of the system while turning the knobs of the oscillators. It is a definite drawback of the system that it is not possible to calibrate the oscillator dials in true cutoff frequency, this making the heterodyne filter less expedient for purposes requiring a fast change of cutoff freguencies. In return, the heterodyne filter offers the possibility of continuously varying the cutoff frequencies and to place the frequency cuts at exactly the frequencies desired, if only the user takes the time to adjust it properly.

3. Performance of the heterodyne filter: frequency response.

Filters are generally characterized as devices that pass certain frequencies but suppress others. It is, however, to some extent a matter of definition whether a particular frequency is "passed" or "suppressed" by the filter. Nevertheless, in order to obtain a condensed characterization of the filter response it is often necessary to distinguish categorically between frequencies that are passed and frequencies that are not passed.

With filters characterized by monotonically increasing attenuation beyond a certain frequency it is customary to define the passband as the frequency range in which the attenuation exhibits values between 0 and 3 dB, i.e., the cutoff frequency equals the "-3 dB point". The performance in the "attenuation band" may then be specified in either two ways: (a) in terms of rate of increasing attenuation (expressed in decibels per octave) or (b) in terms of an absolute (minimum) value of attenuation required to suppress unwanted frequencies: "A_{min}". With the latter specification the attenuation band is split up into two: a transition region (immediately adjacent to the passband) with attenuation values between 3 dB and "A_{min}", and a stopband with attenuation values exceeding "A_{min}". When very sharp filtering is required the latter specification is definitely more meaningful. It might be desired, for example, to remove F1 of a vowel while keeping F2 and higher formants intact in order to study the effect of this removal on perception. It might then be estimated that all frequencies within the effective range of F1 should be attenuated at least 40 dB (or possibly more), whereas the envelope of F2 should be minimally distorted, i.e. the entire transition region between cutoff and A_{min} (= 40 dB attenuation) should fall in the valley between the formants.

Sharp filtering obviously requires that the transition region between cutoff and the A_{min} point be as narrow as possible. In filtering of speech it is often reasonable to employ a filter whose sharpness of cutoff is proportional to the cutoff frequency (A_{min} being reached, for example, at one third of an octave from the - 3 dB point), and this will normally be true of stepwise variable filters (including the LC highpass filter of the Institute of Phonetics). The heterodyne filter , however, operates with fixed filters, i.e. the sharpness of cutoff as expressed in cps is constant irrespective of the effective cutoff frequency (determined by the carrier oscillator). This means that the fixed filters must be extremely frequency selective in order that the cutoff be sufficiently sharp at low frequency settings. For our purpose it is sometimes interesting to "cut" between two harmonics of a vowel sound (particularly in the low frequency end of the spectrum), i.e. the transition region between passband and stopband should not - at least for the highpass section - be more than some 1oo cps wide.

Another important consideration is the required minimum attenuation in the stop band: A_{min}. For purposes involving perceptual tests a difference in level between the effective passband and the effective stopband of some 4o-5o dB is often necessary. Obviously, if a very high value of A_{min} is required, it is more difficult to obtain a narrow region between the passband and the stopband.

In the heterodyne system the situation is particularly complicated, since the filters perform a double function: (a) to highpass and/or lowpass filter the audio signal, (b) to remove unwanted frequencies generated in the modulators, i.e. to eliminate nonlinear distortion. The latter frequencies are partly situated rather far from the useful frequency bands, but on the other hand they must be very effectively suppressed. The filters employed are well suited for this double purpose, since in addition to a sharp cutoff they exhibit a rising attenuation throughout part of the stopband, the final attenuation exceeding 60 dB. For the user, however, the behaviour near cutoff is most interesting. Therefore, in the specifications given below, we have arbitrarily defined A_{min} as 40 dB attenuation, and described the sharpness of cutoff with reference to this value of stopband attenuation.

Finally, the behaviour of the filters in their passbands may deserve special attention. Above, the passband was defined as the frequency band with attenuation between 0 and 3 dB. However, the filter system exhibits a certain amplitude ripple in the passband, and with this type of behaviour the above measure does not characterize the passband response well. It goes without saying that an amplitude ripple of the magnitude of 3 dB throughout the passband is not desirable, although it would not violate the definition of passband. It is therefore essential to specify also the permissible ripple or amplitude fluctuation in the passband, and accordingly to specify a "linear portion" of the passband, in which the response should keep within predescribed limits. There is no general standard as to the permissible ripple. Considering the irregularities found with other pieces of equipment used in acoustic-phonetic studies (in particular, microphones and loudspeakers), a ripple of 2 dB or even more might perhaps be tolerated. However, we have strived to reduce the ripple to one dB, i.e. to keep the amplitude response in the passband within $\pm$ 0,5 dB of the average value. Accordingly, if the maximum value is denoted "0 dB", we define the cutoff frequency of the linear portion of the passband as the point at which the attenuation exceeds 1 dB (cp. reference (3)). By designing the filter in such a way that the -1 dB point is quite close in frequency to the -3dB point, a nearly ideal response is obtained: with a strictly linear passband and a stopband, separated by a very narrow region of transition.

By introducing the -1 dB point as an additional measure of cutoff we get a filter response that is advantageous when filtered speech is used for tests on perception, since for such a purpose a faithful reproduction of frequencies near cutoff is just as interesting as the reproduction of frequencies further off in the passband. Filters with a slowly increasing attenuation from 0 to 3 dB in the passband introduce an error which may not be entirely negligible.

Fig. 3 shows the frequency response of the whole heterodyne system with some different choices of cutoff frequencies. It is seen that the passband as well as the cutoff slopes have a very regular form, except that in cases of narrowband filtering the two cutoff slopes are not quite equal, the highpass section having a considerably sharper cutoff than the lowpass section. On the highpass side the -3 dB point is reached some 15 cps from the -1 dB point, and the -40 dB point is reached some 40 cps from the -3 dB point. On the lowpass side the -3 dB point is reached some 20 cps from the -1 dB point, and the -40 dB point is reached only some 240 cps from the -3 dB point. This difference is in many cases of no consequence, since the highpass filtering is most likely to be used at rather low frequencies, and the lowpass filtering at somewhat higher frequencies, so that the relative sharpness of filtering will often be approximately the same.

The reason for the difference between the two filter sections is that we have built the lowpass section with ordinary coils and condensers and altogether without the use of very advanced (and expensive) techniques, whereas - as stated above - the highpass section includes a commercial high quality crystal filter designed for single sideband transmission. In order to be able to compete with this filter we designed the lowpass filter section for the lowest cutoff frequency compatible with the desire to obtain a sufficient maximum bandwidth, and this is the major reason for modulating down to a low frequency range in modulator "2" (see Fig. 1). Since, everything else being equal, the "sharpness" of a filter will be proportional to the cutoff frequency, the absolute distance in cps between passband and stopband can be reduced by designing the filter for a low cutoff frequency. However, even in the 10 kcps range it is quite a problem to design a sufficiently sharp filter. A distance of 260 cps between the -1 and-40 dB points entails that A_{min} be obtained at a frequency that is only approximately 1.026 times the cutoff frequency (defined as the -1 dB point). We have obtained this only in a roundabout way.

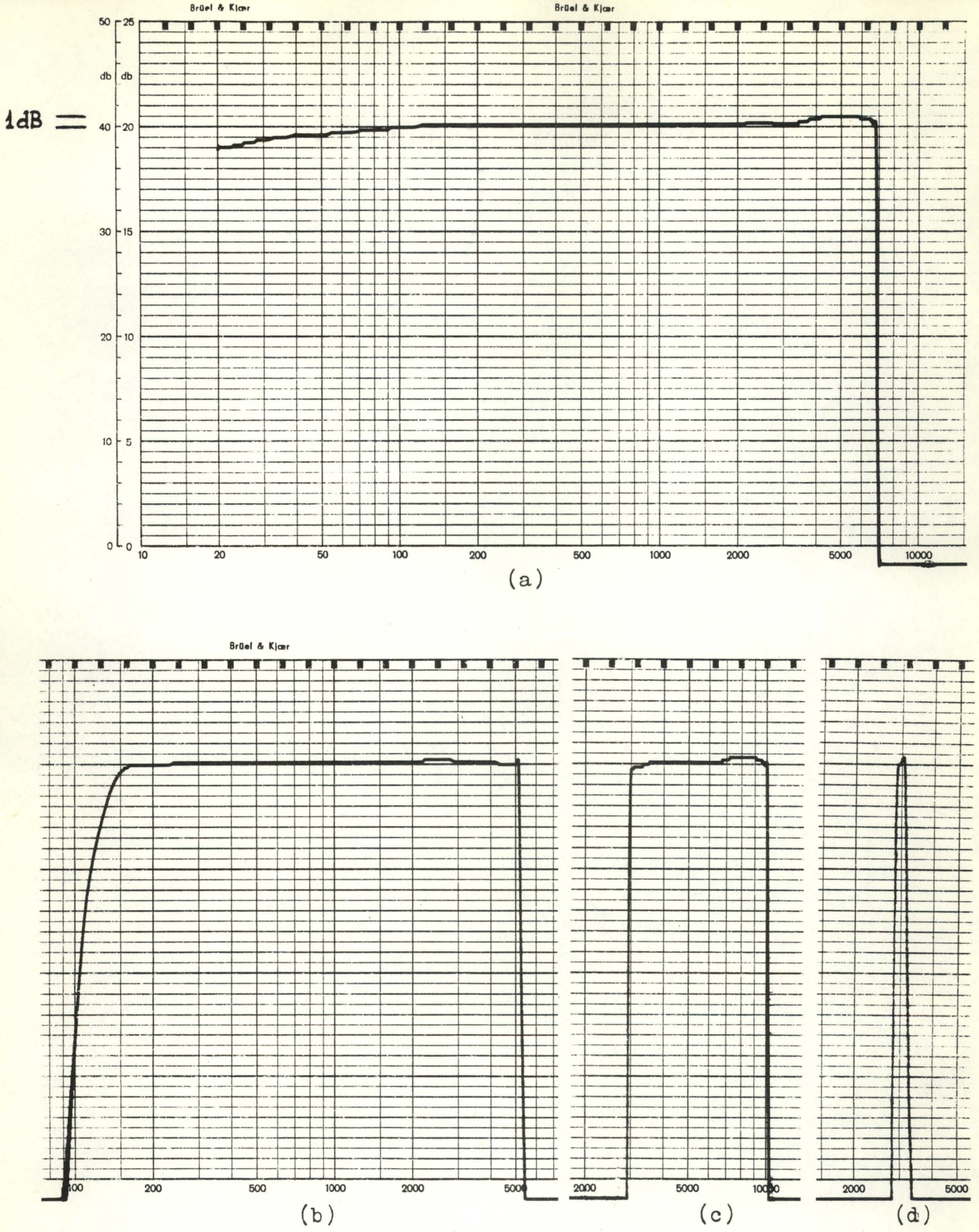


Fig. 3.

Frequency response of heterodyne filter at various settings. Passband amplitude ripple adjusted to be less than ±0,5 dB; paper range 50 dB.

(a) wide band response (-1 dB points at 90 cps and approx. 7000 cps; notice that the amplitude is only 3 dB down at 20 cps).

(b) & (c) various bandpass settings (-1 dB points in (b) at approx. 150 cps and 5000 cps, in (c) at 3000 and 10000 cps).

(d) narrow band response (-1 dB points at 2900 and 3100 cps; this corresponds to a bandwidth of some 235 cps between the -3 dB points. Notice that the bandwidth at the 40 dB attenuation level is only some 500 cps).

The lowpass filter consists of two sections. The first of these (Fig. 2 : C) is a pre-shaping unit, whereas the second (Fig. 2 : D) is a passive LC filter of the conventional constant-k type (image parameter filter) with seven sections, which gives a very high attenuation in the stop band. This filter was designed for a nominal cutoff frequency of 1o kcps. In view of the great number of sections a certain attenuation at cutoff might be expected (due to lossy components). However, the actual attenuation at cutoff is as high as 12 dB and thus exceeds what might be predicted, possibly because there are alignment difficulties with a filter of this complexity. In order to improve the behaviour of the filter near cutoff we have added the amplitude correction equalizer (pre-shaping network) before the filter. This network, which was constructed by a cut-and-try procedure, includes an LC pole circuit resonating at a frequency just below the nominal cutoff, an LC zero circuit resonating at a frequency just above the nominal cutoff, and some RC circuits improving the response at frequencies well below cutoff. When carefully adjusted the equalizer provides a satisfactory overall response for the combined system. (It is possible here to trade some of the linearity in the passband for better sharpness of cutoff, the maximum sharpness obtainable being -4o dB at a distance of 195 cps from the -3 dB point; this gives a passband ripple of slightly more than 2 dB.)

4. Performance of the heterodyne filter: noise, nonlinear distortion, and phase distortion.

The signal-to-noise ratio of the heterodyne filter is very good. With sinewave testing it is possible to obtain a maximum output (before clipping) of o,32 volts RMS with o,44 volts RMS applied to the input terminals. The background noise at the output terminals is more than 65 dB below maximum signal output. However, with maximum input voltage the distortion is relatively high, and in order to minimize distortion it is preferable to reduce the input voltage to max. 1oo millivolts. Even so, the signal-to-noise ratio exceeds 5o dB, which competes well with most recording equipment.

Whistle interference among some of the several frequencies generated in the modulators, is apt to be a disturbing factor. In our first setup we used a lower carrier frequency for modulators 2 and 3, the lowpass filtering being performed (on the inverted sideband, i.e., physically as a highpass filtering) by a Siemens bandpass filter with cutoff at 6okcps. However, it appeared that audible interference tones were generated at certain settings of the carrier oscillators (the interference being between higher harmonics of the two carriers). With the present choice of frequencies we have more or less eliminated this problem, although the transposition of the signal to a low frequency range in modulator 2 (see Fig. 1) creates problems, too. It has proved impossible to utilize the entire passband of the 9,8 kcps lowpass filter, because the occurrence of very low frequencies at this place in the system creates strange distortion products. We do not know at present to what extent this is due to the design of the modulators (which are intended for operation at higher frequencies), or to other factors as well. For better exploitation of the passband modulators 2 and 3 must probably be replaced by other units.

As stated above the suppression of signals outside the passband is very effective. E.g., a o,44 volts signal at 11 kcps does not give any detectable contribution to the output voltage, i.e. the attenuation of the lowpass section at this frequency is (considerably) better than 6o dB. The performance of the highpass section is similarly high, however, at first sight the suppression of the signals below the highpass cutoff frequency does not seem to exceed some 48 dB. This is due to the presence of distortion products. When a sine wave is applied to the heterodyne filter, the inevitable nonlinearity of the modulators conditions the presence of a (weak) third harmonic of the signal, which eventually passes through even if the signal itself is suppressed.

When complex signals are filtered, both harmonic distortion and intermodulation among the components of the signal will occur. The presence of these types of distortion is an inherent limitation of the heterodyne system, although their effect can be reduced by careful design and operation of the filter. In our present setup the distortion due to nonlinearity is typically of the order of 1 to 2%. This is not a shocking figure compared to the performance of medium quality tape recorders and amplifiers, but it should be kept in mind that the generation of distortion products in the stopbands may seriously blur the picture of the filtering function. This consideration also motivates the modest requirement on minimum attenuation in the stopbands which was formulated in the preceding section.

In this context it may be added that according to our experience filtered signals should be recorded and manipulated with the utmost care. It goes without saying that a lowgrade tape recorder may more or less spoil the result. For example, if the fundamental of a harmonic spectrum (e.g., a vowel) is cut off by highpass filtering, intermodulation among the higher harmonics may regenerate the fundamental, the result being an apparently imperfect filtering. This kind of distortion may pass unnoticed under normal circumstances, i.e. with unfiltered speech, since the frequencies generated by intermodulation are already there, so that the contribution of the distortion products to the intensity level is negligible.

Finally, the response of the filter in the time domain should be considered. For analyses of transient sounds (consonants) it is important that the filter does not "smear out" the signal, i.e. the ringing effect must be within narrow limits. With quasiperiodic sounds including vowels the effect of ringing may be less harmful, since the ear seems to be rather insensitive to phase. However, visual inspection of the waveform reveals the distortion. It is well-known that a sharp cutoff brings with it an extremely nonconstant delay of different frequencies within the passband, i.e. a sharp filter inevitably distorts the waveshape of the signal, especially because the phase shift of the system changes rapidly near cutoff.

The phase characteristics of the heterodyne filter are on the whole poor. The rate of phase shift of the lowpass section is rather uniform, and square wave testing shows a rather good reproduction of the waveform, but the highpass section exhibits an enormous shift from one end of the passband to the other. Thus the complete heterodyne system badly distorts the waveshape of signals, and it must be concluded that it is not suitable for analyses involving a closer inspection of phenomena in the <u>time</u> domain, although it is well suited for slowly changing phenomena like vowel stimuli used in perception tests. In spite of the very outstanding frequency discriminating properties of the highpass filter section, a modification of the heterodyne system might profitably start with a redesign of this section, possible with a shift of the operating range to a much lower frequency, where a simpler filter might do the same job with less phase shift. However, it remains to be decided to what extent high fidelity in the time domain can be important for our experiments with filtered speech.

9. Acknowledgements.

Our thanks are due to the Post and Telegraph Department, which has put most of the equipment used in the filter at our disposal, to the State Institute of Speech Defects, which has put a wave analyzer used for control of the system at our disposal, and to the Danish Council for Scientific and Industrial Research (Danmarks teknisk-videnskabelige Forskningsråd), whose financial support has made the project possible.

References:

(1) Jørgen Rischel, "Heterodyne Filter", ARIPUC 1/1966, pp.13-14.

(2) Gunnar M. Fant, The Heterodyne Filter (Stockholm, 1952).

(3) A.I. Zverev, Handbook of Filter Synthesis (New York, London, Sydney, 1967), pp. 63 ff.

SPEECH SYNTHESIZER

Jørgen Rischel and Svend Erik Lystlund

In the past year we have been experimenting with a method to generate control voltages (varying as a function of time) for the formant circuits of a terminal analog synthesizer. Also a modulator system for the synthesizer is being tried out. The work is still in a too early phase for presentation. An account of it will probably appear in the next report of our Institute.

Acknowledgements.

The work is supported by a grant from the Danish Council for Scientific and Industrial Research.