# DETECTION AND DISCRIMINATION OF MODULATION TYPE
## AT LOW MODULATION RATES

## A.P. SĘK

Institute of Acoustics,
Adam Mickiewicz University
(60-769 Poznań, ul. Matejki 48/49)

For a sinusoidal carrier at a frequency of 1 kHz, amplitude (AM) or frequency (FM) modulated by a sinusoidal modulator at a rate of $f_{\mathrm{mod}} = 2$, 5 or 10 Hz, psychometric functions for the detection and the discrimination of modulation type were measured as a function of an appropriate modulation index (i.e. $m$ or $\beta$). Stimuli were presented in quiet or with a band of noise chosen to mask the low- or high-frequency side of the excitation pattern produced by a modulated signal. In AM case $d'$ markedly depended on the presence of a noise that masked the high-frequency part of the excitation pattern. For FM signals, on the other hand, when modulation rate was equal to 2 Hz bands of noise did not influence $d'$ values. Probability of identification of modulation type (AM or FM) was the highest for the smallest modulation rate and it was nearly equal to the probability of modulation detection. Presence of any of these two bands of noise did not effect the modulation identification. The results suggest that there are two mechanisms underlying the detection and the discrimination of modulation type. One of them is based entirely on the changes in the excitation pattern level and operates for a whole range of carrier frequencies and modulation rates (place mechanism). However, for a low modulation rate there is another mechanism responsible for the detection of frequency changes only. This mechanism provides additional information about frequency changes and brings about markedly higher delectability $d'$ for the detection and discrimination of frequency changes. This mechanism is not based on the ability of the auditory system to compare a phase of the excitation pattern changes at different frequency areas. It seems that information about frequency changes at a low rate may be effectively coded in a time distribution of neural spikes.

## 1. Introduction

Investigations concerned with amplitude and frequency changes at very low modulation rates (i.e. less than 20 Hz) have provided a great variety of data about the auditory system and have been extensively discussed in the psychoacoustical literature, [1–3, 8, 11, 12, 16–19, 21, 26–28, 30, 32]. There are two basic hypotheses that have been advanced about the perception of amplitude and frequency changes in an acoustic signal. One of them, called the ZWICKER–MAIWALD model [11, 12, 30, 32] postulates that a single mechanism is responsible for the perception of changes in amplitude and frequency. The other one, presented by CONINX [2, 3] and FETH [5], assumes the existence of two independent mechanisms, one for amplitude changes and one for frequency changes.

The Zwicker–Maiwald model is essentially a place model based on the concept of the psychoacoustical excitation pattern. The excitation pattern evoked by a sound can be defined as the output of the auditory filters as a function of centre frequency, in response to that sound [13, 15]. A maximum of the excitation pattern evoked by a sinusoidal signal is observed at a characteristic frequency which corresponds to the frequency of exciting sinusoid. This excitation pattern is a lot steeper on the low-frequency side and decays gradually on the high-frequency side.

In a more recent version of this model [16, 18] a non-linear compression that takes place on the basilar membrane is included [23, 24], as well as some of the phenomena occurring in the auditory nerve [22]. An increase in amplitude of a signal gives rise to an increase in the maximum of the excitation pattern but also brings about a spread of the excitation pattern in frequency domain; more auditory filters are active.

When the auditory system is excited by AM signal then the excitation pattern changes in the same way as the amplitude of the stimulating signal; this is illustrated in the upper left panel of Fig. 1. To make the excitation pattern changes clearly visible two curves illustrate the excitation patterns for extreme values of AM signal and for AM index equal to 20% that corresponds to difference in level $\Delta L = 4.5\,\text{dB}$. It is worth noting that a bigger change in excitation level occurs on the high-frequency side of the excitation pattern. This happens because the excitation level on the high-frequency side of the pattern grows non-linearly with changes in level. The left lower panel of Fig. 1 presents a difference of the excitation patterns evoked by AM signal for extreme values of its amplitude for modulation depth of $m = 0.04$ ($\Delta L = 0.7\,\text{dB}$) which is close to the average threshold for modulation rate from the range of $f_{\text{mod}} = 2 - 50\,\text{Hz}$. Note that the difference on both sides of the maximum is in the phase and bigger on the high-frequency side.

The excitation pattern produced by the FM signal has a constant maximum but it moves along a frequency axis as the signal frequency changes. The excitation patterns evoked by the FM signal at extreme values of its frequency are presented in the upper right panel of Fig. 1. To make the excitation pattern changes clearly visible deviation was equal to 100 Hz. For FM signals a bigger change in excitation level usually occurs on the low-frequency side of the pattern, where the slope is steepest. This is also shown in the lower right panel of Fig. 1. It shows the difference in the excitation patterns produced by FM signal at extreme frequencies for deviation of 4 Hz. This value is close to the average threshold for modulation rate from the range of $f_{\text{mod}} = 2 - 50\,\text{Hz}$.

In the simplest version Zwicker–Maiwald's model assumes that changes in either amplitude or frequency are detected by monitoring the single point on the excitation pattern that changes most. This is equivalent to monitoring a single auditory filter. The detection of changes in amplitude or frequency occurs when the change in the excitation pattern exceeds a criterion amount, which was empirically established to be approximately 1 dB. The model also assumes that amplitude and frequency changes cannot be distinguished at the detection threshold.

Some more recent models assume that information about amplitude and frequency changes can be combined over a certain region of the excitation pattern. This is equivalent to monitoring several auditory filters simultaneously [6, 14].
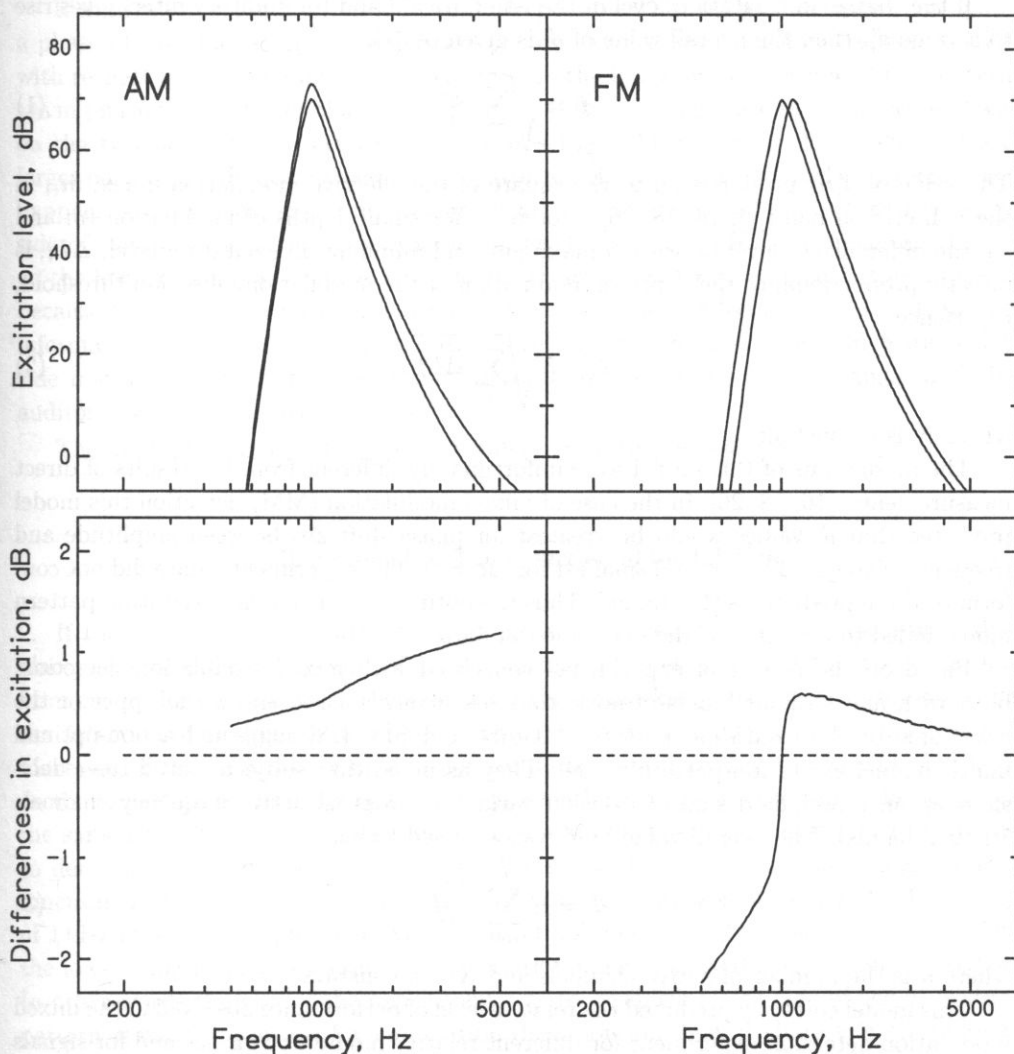
Fig. 1. Upper row shows the excitation patterns produced by amplitude (left panel) and frequency (right panel) modulated tones for extreme values of amplitude and frequency respectively, for suprathreshold values of appropriate modulation indices. Lower row shows the differences in excitation patterns produced by amplitude (left panel) and frequency (right panel) modulated tones for extreme values of amplitude and frequency respectively, for nearthreshold values of appropriate modulation indices.

For example, the model proposed by FLORENTINE and BUUS [6] assumes that information from the entire excitation pattern, i.e. from each excited critical band or active auditory filter, is combined in an *optimal* way (so-called optimal multi-channel model, [16]), from the point of view of the signal detection theory (so-called integration model) [7].

If the change in excitation level in the $i$-th critical band (or auditory filter) gives rise to a value $d_i'$, then the overall value of $d'$ is given by [7]:

$$d' = \sqrt{\sum_i d_i'^2}\,. \tag{1}$$

The value of $d_i'$ is proportional to the square of the effective modulation index, $m_i$, in the $i$-th critical band [9, 16, 18, 26]. Moreover, for small depths of modulation ($m$ and $\beta$), the difference in level between a maximum and minimum in excitation level, $\Delta L_i$, is directly proportional to the depth of modulation at the modulation detection threshold, $m_i$. Hence,

$$d' = K\sqrt{\sum_i \Delta L_i^4}\,, \tag{2}$$

where $K$ is a constant.

The predictions of this model were unfortunately different from the results of direct measurements, [16, 18, 26]. In the case of mixed modulation (MM) detection this model predicted that $d'$ values would be greatest for phase shift $\Delta\phi$ between amplitude and frequency changes $\Delta\phi = \pi$, and smallest for $\Delta\phi = 0$. The experimental data did not conform to these predictions [16, 18, 26]. Thus the optimal multi-channel excitation-pattern model failed to account for the experimental data.

Based on the results of experiments concerned with mixed modulation detection, both with and without bands of noise that selectively masked either the upper or the lower side of the excitation pattern, MOORE and SĘK [18] suggested a non-optimal multi-channel excitation-pattern model. They assumed that subjects based their decisions on an unweighted sum of decision variables across all active frequency channels (critical bands). The overall value of $d'$ was assumed to be:

$$d' = K\sum_{i=1}^{n} \frac{\Delta L_i^2}{\sqrt{n}}\,, \tag{3}$$

where $n$ is the number of active channels and $K$ is a constant.

This model correctly predicted the relative level of performance observed in the mixed modulation detection experiment for different relative modulator phases and for signals presented in quiet or with noise bands. Correlation of the predictions and measured $d'$ values reached about 90% [18, 27]. Based on this model it is possible to interpret a monotonic increase in difference limens for AM and FM signals [20, 29].

The non-optimal multi-channel excitation-pattern model also predicts that the discrimination of modulation type (or modulation identification) is impossible at the detection threshold of modulation. The ability of subjects to identify the modulation type, i.e. AM or FM was studied by DEMANY and SEMAL [4] for very low modulation rates. They showed that the identification performance was almost equal to the detection performance when the modulation rate was less than 5 Hz which cannot be explained by the non-optimal multi-channel. Thus is there any additional mechanism that enhances sensitivity of the auditory system to very low changes in physical parameters of an acoustic signal?

Identification of modulation type would be possible if the subject was able to compare a phase of the changes in the excitation pattern on the low- and high-frequency sides with respect to the maximum. If the changes on the lower and upper side of the pattern are in phase this indicates that AM is present. If these changes are in an opposite phase on the two sides of the pattern this indicates that FM is present. If the changes are larger on the high-frequency side of the pattern, this indicates that AM is present (this happens because of expansive non-linear growth of excitation level on the high-frequency side of the patterns with increase in signal level; see Fig. 1). If the changes are larger on the low-frequency side of the pattern this indicates that FM is present (this happens because the excitation pattern is usually steeper on the low-frequency side). Thus the information about the phase of excitation level changes on the low- and high-frequency side may be a source of an additional information about the signal changes and the auditory system may effectively use it.

The main purpose of this paper is verification of this hypothesis.

## 2. Experiment 1. The detection of modulation

If the detection (or identification) of modulation depended on comparison of displacements of different regions of basilar membrane then the limiting of information from one of the active areas of the membrane (using masking band of noise) should make the detection of modulation much more difficult and consequently increase in threshold. Moreover, if the detection of modulation for all modulation rates was based exclusively on the excitation pattern changes, then using masking band of noise should bring about the same effect for all modulation rates. The main purpose of the first experiment was to determine the detectability $d'$ for amplitude and frequency modulated signals as a function of $m$ or $\beta$ respectively. It was carried out for a sinusoidal carrier at a frequency of 1 kHz presented in quiet and with bands of noise designed to mask selectively either the low- or high-frequency side of the excitation pattern evoked by a modulated signal. In other words the aim of the experiment was to establish which side of the excitation pattern is more important for modulation detection and how it changes with modulation rate for AM and FM. The results gathered in this experiment were also a starting point for the next experiment concerned with the modulation type discrimination (see Sec. 3).

### 2.1. Method

Psychometric functions for the detection of amplitude and frequency modulated signals were determined in two separate experiments using a two-alternative forced-choice (2AFC) method. On each trial two successive signals were presented. One of them was modulated (AM or FM) and the other one was a pure tone. The order of the signals was random and subject's task was to indicate which of two signal in a pair was modulated. Each signal had an overall duration of 1000 ms including raised-cosine rise/fall times of

50 ms. The long duration was chosen so that several cycles would occur in each stimulus. The time interval between two successive signals was 500 ms.

A run consisted of 55 trials. Five different modulation depths, i.e. $m$ or $\beta$ for AM and FM respectively, were used in each run, and they were used in random order. The highest value of an appropriate modulation index was chosen to be easily detectable, giving typically 90–95% correct responses, and the smallest one was chosen to be difficult to detect, typically giving 55–60% correct responses. These values were established in several pilot runs individually for each subject. Twenty blocks of trials were run for each modulation rate and modulation type (AM and FM), so any point on each psychometric function is based on at least 200 judgements.

In separate experimental sessions, using the above described method, psychometric functions for detection AM and FM signals presented with bands of noise chosen to mask either the lower or the upper part of the excitation pattern were measured.

The carrier was a sinusoid with a frequency of $f_c = 1\,\mathrm{kHz}$, and level 70 dB SPL. Modulation rate was equal $f_{\mathrm{mod}} = 2$, 5 and 10 Hz. Bands of noise had essentially rectangular spectral envelopes and cut-off frequencies: $f_l = 500\,\mathrm{Hz}$, $f_u = 800\,\mathrm{Hz}$ – low-band noise and $f_l = 1250\,\mathrm{Hz}$, $f_u = 2000\,\mathrm{Hz}$ – high-band noise. Spectrum level was equal to 49 and 45 dB (re. 20 μPa) for low- and high-band noise respectively. The parameters of these bands were chosen to give a crossing point of the excitation patterns [15] produced by the carrier signal and by each band of noise, 10 dB below maximum of the excitation pattern evoked by the carrier. Moreover, the maxima of the excitation patterns produced by bands of noise were approximately equal.

These parameters of the bands of noise allowed a little effect of low-band noise on the high-frequency side of the excitation pattern and *vice versa*. Noise used in this study was not so-called frozen noise but it was calculated individually for each stimulus by means of Inverse Fourier Transform. Bands of noise were not presented continuously as in MOORE and SĘK's [18] experiments but were gated synchronously with stimuli. Signals were generated via 16-bit digital-to-analogue converter (Tucker and Davis Technology) at a sampling rate of 50 kHz and presented monaurally by means of Sennheiser 414 headphones in a sound attenuating chamber. Three normal hearing subjects were used.

## 2.2. Results and discussion

As results of the experiment, probabilities of correct answers for each of 5 values of appropriate modulation indices ($m$ or $\beta$) were obtained. These probabilities were transformed into detectability $d'$ [7, 10]. Functions describing dependencies of $d'$ on $m$ or $\beta$, i.e. $d'(m)$ and $d'(\beta)$ were similar for all subjects and in general they could be approximated by linear functions of the modulation index square:

$$d'_{\mathrm{AM}} = K_{\mathrm{AM}} m^2 \,, \tag{4}$$

$$d'_{\mathrm{FM}} = K_{\mathrm{FM}} \beta^2 \,, \tag{5}$$

where $K_{\mathrm{AM}}$ and $K_{\mathrm{FM}}$ are constants expressing slopes of the best-fitting lines. Correlation coefficients of collected data were relatively high and no smaller than 0.92.
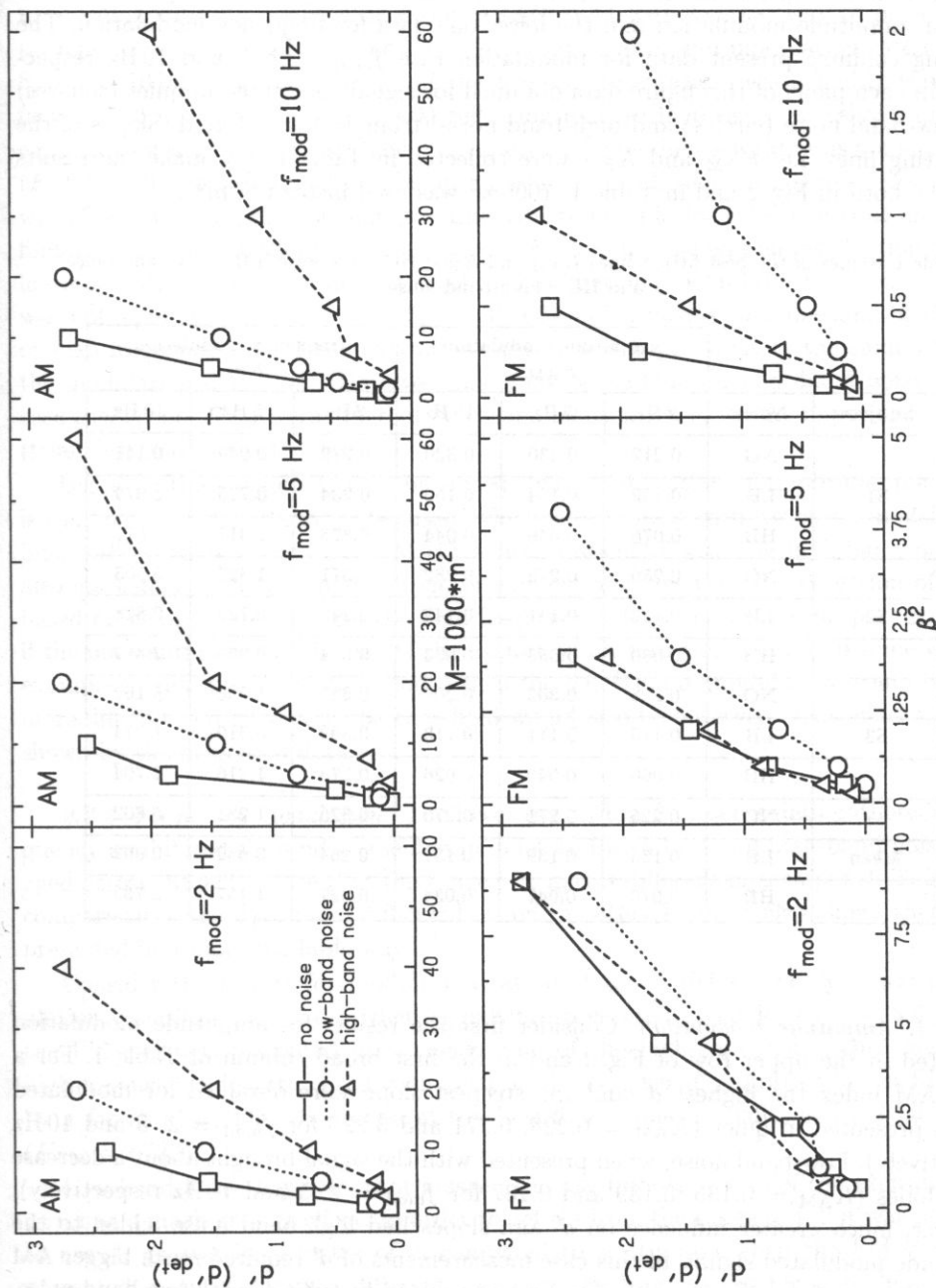
Fig. 2. Psychometric function for the detection of AM (upper row) and FM (lower row). The following columns show data for modulation rates $f_{mod} = 2$, 5 and 10 Hz for different ways of signal presentation: in quiet (squares), with low-band noise (circles) and with high-band noise (triangles).

The data averaged across the subjects are presented in Fig. 2. The upper row shows data for amplitude modulation and the lower one data for frequency modulation. The following columns present data for modulation rate $f_{mod}$ = 2, 5 and 10 Hz respectively. In each panel of this figure data obtained for signals presented in quiet (squares) with low-band noise (circles) and high-band noise (triangles) are plotted. Slopes of the best-fitting lines, i.e. $K_{AM}$ and $K_{FM}$ were collected in Table 1. (To make the results readable, both in Fig. 2 and in Table 1, $1000m^2$ was used instead of $m^2$).

**Table 1.** Slopes of the best-fitting lines $K_{AM}$ and $K_{FM}$: NO – no noise, LB – low-band noise and HB – high-band noise.

| Subject | Noise | Amplitude modulation $K_{AM}$ | | | Frequency modulation $K_{FM}$ | | |
|---------|-------|--------|--------|--------|--------|--------|--------|
|         |       | 2 Hz | 5 Hz | 10 Hz | 2 Hz | 5 Hz | 10 Hz |
|         | NO    | 0.212 | 0.230 | 0.320 | 0.279 | 0.949 | 6.141 |
| S1      | LB    | 0.159 | 0.164 | 0.156 | 0.234 | 0.725 | 1.074 |
|         | HB    | 0.070 | 0.049 | 0.044 | 0.376 | 1.317 | 3.077 |
|         | NO    | 0.259 | 0.279 | 0.282 | 0.371 | 1.327 | 5.503 |
| S2      | LB    | 0.132 | 0.141 | 0.142 | 0.248 | 0.720 | 0.828 |
|         | HB    | 0.080 | 0.053 | 0.033 | 0.324 | 0.939 | 2.419 |
|         | NO    | 0.295 | 0.303 | 0.209 | 0.325 | 1.568 | 5.162 |
| S3      | LB    | 0.115 | 0.111 | 0.113 | 0.309 | 0.510 | 1.074 |
|         | HB    | 0.060 | 0.040 | 0.026 | 0.274 | 1.216 | 2.704 |
|         | NO    | 0.225 | 0.271 | 0.270 | 0.325 | 1.281 | 5.602 |
| Mean    | LB    | 0.135 | 0.139 | 0.137 | 0.264 | 0.652 | 0.992 |
|         | HB    | 0.070 | 0.047 | 0.034 | 0.327 | 1.157 | 2.733 |

*2.2.1. Amplitude modulation.* Consider first the results for amplitude modulation presented in the upper row of Fig. 2 and in the first broad column of Table 1. For a given AM index the highest $d'$ and the steepest slope were obtained for modulated signals presented in quiet ($K_{AM}$ = 0.225, 0.271 and 0.227 for $f_{mod}$ = 2, 5 and 10 Hz respectively). Low-band noise, when presented with the signal brought about a decrease in $d'$ slopes ($K_{AM}$ = 0.135, 0.139 and 0.137 for $f_{mod}$ = 2, 5 and 10 Hz respectively). However, much greater influence on $d'$ and slopes had high-band noise added to the amplitude modulated signal. In this case measurements of $d'$ required much bigger AM indices than in the case when the signal was presented in quiet or with low-band noise. Slopes are also shallower in this case: $K_{AM}$ = 0.070, 0.047 and 0.034 for $f_{mod}$ = 2, 5 and 10 Hz respectively. It should be noted that low-band noise had the same effect on $d'$ values for all modulation rates used. High-band noise, however, had a bigger effect for higher modulation rates.

Slopes of the best-fitting lines ([1]) for AM, as collected in Table 1, were subjected to a within-subjects analysis of variance (ANOVA) with factors: way of presentation (in quiet, with low-band noise and with high-band noise) and modulation rate ($f_{mod}$ = 2, 5 and 10 Hz). The main effect of the way of presentation was a highly significant factor ($F(2,4) = 142.1$, $p < 0.0001$) reflecting marked increase in threshold, observed especially for high-band noise. The main effect of the modulation rate was not significant ($F(2,4) = 0.1$, $p = 0.91$); $d'$s and slopes obtained for all modulation rates (for a given way of presentation) did not change markedly with modulation rate. This confirms earlier findings that the threshold for detecting of AM does not depend on modulation rate for low modulation rates. The interaction of the modulation rate and the way of presentation was not significant ($F(4,8) = 0.55$, $p = 0.702$) which means that masking of the low- or high-frequency side of the excitation pattern has approximately the same effect on the modulation detection. (However this is not quite true for high-band noise; in this case coefficients $K_{AM}$ were different across modulation rates, see the upper row of Fig. 2 (triangles) and Table 1).

Based on these results it may be stated that the detection of amplitude modulation is similar for all modulation rates used, which is consistent with Zwicker's model. Much higher $d'$s obtained when the AM signal was presented with the high-band noise are also consistent with this model. Zwicker's model assumes that the detection of AM is based on changes observed on the high-frequency side of the excitation pattern. Thus, if the information from this part of the excitation pattern were limited in some way (for example by masking, as in this study) then the detection of modulation would be much more difficult and growth in the threshold would also be observed. This effect is clearly shown by the presented data.

*2.2.2. Frequency modulation.* The lower row in Fig. 2 presents data gathered for frequency modulated sinusoid, presented in quiet and with low- and high-band noise, averaged across subjects. Note that the range of $x$ axis is different in each panel. This makes comparison of the data more difficult but, on the other hand, it also allows them to be presented in a more readable way.

Consider the results for modulation rate of $f_{mod} = 10$ Hz. The detection of modulation is easier (higher $d'$s) when signals are presented in quiet ($K_{FM} = 5.6$). When the FM signal is presented at a background of the high-band noise, that masks the high-frequency side of the excitation pattern, then a given $d'$ requires a bigger FM index by a factor of 2. Thus the slope of the best-fitting line is shallower by a factor of 2 ($K_{FM} = 2.733$).

Low-band noise, however, that masked the low-frequency side of the excitation pattern had a much greater effect. In order to get $d'$s from the range of 1–2 it was necessary to use $\beta_s$ greater by a factor of 3, compared to a situation when the FM signal was presented in quiet. Quite a different situation can by observed for a modulation rate of $f_{mod} = 2$ Hz. The presentation of FM stimuli in quiet or with low-band noise that

---

([1]) Since the data for different modulation rates were gathered for markedly different ranges of an appropriate modulation index, raw data could not be subjected to this analysis. Such analysis would require many so-called "missing values" that could markedly influence the final results of the analysis.

masks the low-frequency side of the excitation pattern or with high-band noise that masks the upper side of the excitation pattern gives similar slopes of the best-fitting lines describing $d'$ as a function of FM index. The detection of frequency modulation in this case did not depend on presence of masking bands of noise, but took place for bigger values of $\beta$ than for $f_{\mod} = 10\,\text{Hz}$. For modulation rate of 5 Hz a sort of "half way through" situation, between these for $f_{\mod} = 2\,\text{Hz}$ and $f_{\mod} = 10\,\text{Hz}$, is observed (see Fig. 3).

The highest $d'$s were obtained for FM tone presented in quiet ($K_{\text{FM}} = 1.281$). High-band noise that masks the high-frequency side of the excitation pattern did not affect the results. However, the low-band noise markedly influenced the slope ($K_{\text{FM}} = 1.157$). Thus decreasing in modulation rate required much bigger FM indices to reach the detection threshold.

Slopes of the best-fitting lines for FM, presented in the second broad column of Table 1 were subjected to within-subjects analysis of variance (ANOVA) with factors: way of presentation (in quiet, with low-band noise and with high-band noise) and modulation rate ($f_{\mod} = 2$, 5 and 10 Hz). The main effect of the modulation rate was highly significant ($F(2,4) = 210.30, p < 0.0001$) which confirms earlier findings that if the threshold is expressed in units of $\beta$ then it decreases with increase in the modulation rate [25, 31]. The main effect of the way of presentation was also statistically significant ($F(2,4) = 301.72$, $p < 0.0001$). It emphasised the effect of masking, particularly for $f_{\mod} = 10\,\text{Hz}$. However, the most important result of this analysis is that the interaction of the modulation rate and the way of presentation was also highly significant ($F(4,8) = 79.44, p < 0.0001$). This interaction means that the masking effect was markedly different across modulation rates.

The results for FM stimuli are consistent with Zwicker's model but only for a modulation rate of 10 Hz. Only for this modulation rate the effect of the low-band noise was observed. As Zwicker's model suggests, the detection of frequency modulation is based on the low-frequency part of the excitation pattern. Thus if the information from this part of the excitation pattern is limited in some way (for example by masking, as in this study) the detection of FM would be much more difficult and the threshold much bigger. However, this was observed for modulation rate of $f_{\mod} = 10\,\text{Hz}$ only. If Zwicker's model had been correct for the whole range of modulation rate, then the same influence of the low-band noise should have been observed for modulation rate of $f_{\mod} = 2\,\text{Hz}$. Such an influence was not observed in this study.

2.2.3. *Interim summary.* In summary of this part of the results it can be stated that the effect of the high-band noise that masked the high-frequency side of the excitation pattern on $d'$s was bigger for amplitude modulated signals. The effect of this band as well as the effect of the low-band noise was approximately the same across all modulation rates used. The detection of frequency modulation at a modulation rate of $f_{\mod} = 2\,\text{Hz}$ did not depend on presence of any band of noise. However, an increase in the modulation rate brought about that the effect of the noise was bigger, particularly for the low-band noise that masked the low-frequency side of the excitation pattern.

## 3. Experiment 2. Effect of modulation rate on the discrimination of modulation type

As mentioned earlier, one of the possible ways of modulation type discrimination may be a comparison of magnitudes and phases of excitation level changes on the low- and high-frequency sides: changes on the two sides in phase suggest that amplitude modulation as a stimulating signal is present; if the changes are not in phase then this may suggest that the FM signal is present. If the auditory system indeed makes such a comparison of two sides of the excitation pattern then a selective masking of any of these two sides, that limits the amount of the information available from masked part of the excitation pattern should bring about noticeable decrease in probability of modulation type discrimination.

Thus in the next experiment the discrimination of modulation type (AM from FM) was analysed. This was done for AM and FM signals presented in quiet and with bands of noise that selectively masked the low- or high-frequency side of the excitation pattern.

### 3.1. Method

An ability of discrimination of modulation type (AM from FM) was analysed for a sinusoidal carrier signal at a frequency of 1 kHz with level of 70 dB SPL, amplitude or frequency modulated by a sinusoid at a frequency of 2, 5, and 10 Hz. The same masking bands of noise were used as in Experiment 1 (see Sec. 2.1). A two-alternative forced-choice (2AFC) method was used. Two successive signals in a random order were presented on each trial. One of them was amplitude modulated, whereas the other one was frequency modulated. The subject's task was to indicate an order of the signals in each trial, i.e. AM was first then FM, or *vice versa*.

Values of appropriate modulation indices i.e. $m$ and $\beta$ for AM and FM respectively within each trial were chosen to produce equally detectable amount of each type of modulation. Six different values of $d'$ were chosen namely: $d' = 0.66, 1.16, 1.66, 2.16, 2.66$ and $3.16$. For these $d'$s six pairs of $(m_{d'}, \beta_{d'})$, i.e. $\{(m_{0.66}, \beta_{0.66}), (m_{1.16}, \beta_{1.16}), ..., (m_{3.16}, \beta_{3.16})\}$ were calculated based on equations (4) and (5) (see Sec. 2.2). Pairs of $(m_{d'}, \beta_{d'})$ were determined separately for each subject (based on individual data for each subject) and for modulation presented in quiet and with bands of noise.

A single experimental run consisted of 65 trials that were presented in random order. In the first five trials modulation indices $m$ and $\beta$ were the biggest i.e. $m = m_{3.16}$ and $\beta = \beta_{3.16}$. Twenty blocks of trials were run for each way of signal presentation and for each subject. This gave at least 200 judgements for each point on the psychometric function. Signals were presented in the same way as in Experiment 1. The same normal hearing subjects took part in this experiment.

### 3.2. Result and discussion

Probabilities of correct response, as raw experimental data for all pairs of $(m_{d'}, \beta_{d'})$ were transformed into $d'$ domain and presented in Fig. 3 as a function of detectability $d'$ for detection of AM or FM. Since $d'$s obtained for three subjects were similar,

Fig. 3 presents mean data. Top, middle and bottom panels show data for modulation rates $f_{mod} = 2$, 5 and 10 Hz respectively. Each panel shows the results for different ways of signal presentation: in quiet (squares), with low-band noise (circles) and with high-band noise (triangles). Solid lines with no data points show hypothetical situations where data would fall if probability of modulation detection were equal to probability of modulation type discrimination ($d'_{det} = d'_{dis}$).
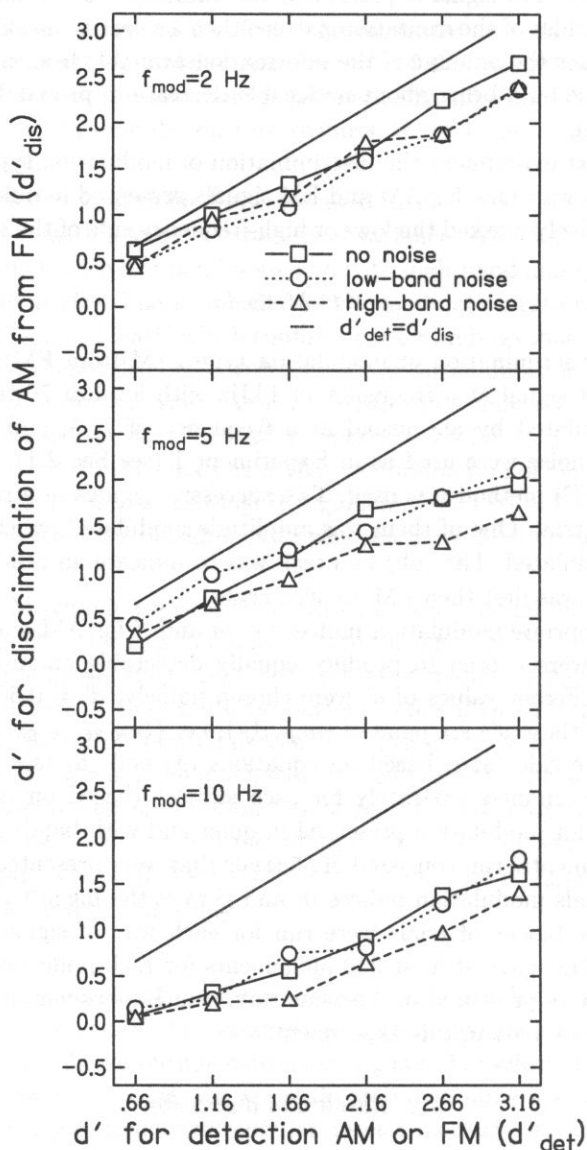


Fig. 3. Psychometric function for the discrimination of modulation type AM or FM as a function of $d'$ for detection of AM or FM. Each panel shows data for one modulation rate and for three ways of signal presentation: in quiet (squares), with low-band noise (circles) and with high-band noise (triangles).

In general it can be stated that an increase in detectability $d'$ for detection AM or FM brings about an increase in detectability $d'$ for the discrimination of modulation type. For $f_{\text{mod}} = 2\,\text{Hz}$ (top panel of Fig. 3) the discrimination (or identification) of modulation type is nearly as good as the detection of modulation. This finding is consistent with that of DEMANY and SEMAL [4]. Thus, for this modulation rate the discrimination of modulation type is quite possible at the detection threshold of modulation. For slightly higher $d'$s for detection of AM or FM the discrimination of modulation type (AM from FM) worsens. It is worth to add, that the way of modulation presentation (i.e. in quiet or with bands of noise) did not influence obtained data: these obtained for signals presented in quiet are very close to those when the signals were presented with low- or high-band noise. This means that the presence of masker, that selectively masked either the low- or high-frequency side of the excitation pattern did not impair the discrimination of modulation type. The middle panel of Fig. 3 shows data for modulation rate of $f_{\text{mod}} = 5\,\text{Hz}$. $d'$s presented in there are, in general, smaller than those for $f_{\text{mod}} = 2\,\text{Hz}$. They increase more slowly as $d'$ for detection of AM or FM increases. This means that the discrimination of modulation type (identification) in this case is more difficult. The way of stimuli presentation does not seem to be important. A pattern of the data obtained for $f_{\text{mod}} = 10\,\text{Hz}$ (bottom panel of Fig. 3) is similar to those for $f_{\text{mod}} = 2\,\text{Hz}$ and $5\,\text{Hz}$. An increase in $d'$ for discrimination is smaller than in the two first cases. No effect of the way of signal presentation was also observed. Thus, for a near-threshold modulation indices an increase in modulation rate makes the discrimination of modulation type progressively more difficult: the discrimination of AM from FM seems to be easy for $f_{\text{mod}} = 2\,\text{Hz}$, more difficult for $f_{\text{mod}} = 5\,\text{Hz}$ and almost impossible for $f_{\text{mod}} = 10\,\text{Hz}$.

The results were subjected to analysis of variance (ANOVA) with the following factors: modulation rate ($f_{\text{mod}} = 2$, 5 and $10\,\text{Hz}$), way of signal presentation (in quiet and with bands of noise) and detectability $d'$ for detection of AM or FM (6 values). As expected, the main effect of the detectability $d'$ for detection of AM or FM was highly significant ($F(2,4) = 210.34$, $p < 0.001$); it is clear that an increase in appropriate modulation indices largely improves performance. The main effect of the modulation rate was also highly significant ($F(2,4) = 15.90$, $p = 0.012$). That means, that the discrimination of modulation type depends on modulation rate. The main effect of the way of signal presentation was not significant ($F(2,4) = 2.61$, $p = 0.188$). The interaction of the modulation rate and the detectability for detection of AM or FM was also significant which indicates that an increase in $d'$ for detection of AM or FM brought about a different increase in $d'$ for discrimination for different modulation rates.

The results of this experiment suggest that when trying to identify modulation type the auditory system does not use information connected with an envelope of the phase of the excitation pattern changes in different frequency regions. However, two factors that affected the results should be borne in mind. First, the modulation depths used (for a given $d'$ for detection of AM or FM) were markedly greater where noise bands were presented than when they were not presented. Secondly bands of noise were not completely effective in masking one side of the excitation pattern only: a small area adjacent to the maximum of the excitation pattern always remained unmasked.

While trying to identify modulation type the auditory system does not use the phase of the excitation pattern changes on the low- and high-frequency side of the pattern. If this information had been used then the discrimination of modulation type would have worsened when AM or FM signals were presented with bands of noise. Subjects can perform the discrimination task even when noise band masks either the low- or high-frequency side of the excitation pattern and $d'$s for discrimination are nearly as good as $d'$s for detection of AM or FM for the lowest modulation rates.

## 4. Concluding remarks

The results of the presented investigations suggest that the discrimination of modulation type at the modulation detection threshold is possible for the lowest modulation rates; this finding is not consistent with Zwicker's model. They also suggest that while trying to identify modulation type the auditory system does not use information connected with the excitation pattern changes on the low- and high-frequency side of the pattern. Information about relative phases of the excitation pattern changes is not crucial for the discrimination of modulation type. Changes in the excitation pattern seem likely to be more important for amplitude modulation because low- and high-band of noise impaired AM detection to a larger extent. Thus, if the information contained in changes in the excitation pattern do not play the most important role then an additional mechanism (or mechanisms) must exist in the auditory system that enables fine discrimination of modulation type at the detection threshold for AM and FM. This mechanism does not have to be connected with both types of modulation and it may be connected with either amplitude or frequency changes only. Since the detection of frequency modulation depends less on masking bands of noise it seems that this mechanism may be connected with frequency changes only. However, this mechanism operates more effectively for very low modulation rates. It enhances a sensitivity of the auditory system to frequency changes. This mechanism may be based on information available in phase locking. It seems likely that the auditory system is capable of analysing the time intervals between neural spikes, for relatively long time samples of the signal. If this is the case, the accuracy and efficiency of this analysis would depend on the absolute number of the impulses characterised by a constant time interval between them. The number of spikes is roughly proportional to the duration of the signal, if the signal level is constant. Increasing the duration of the signal with the constant frequency introduces more equal time intervals between spikes. This provides more precise information about the frequency of the signal and, as a consequence, better performance in frequency modulation detection. A strong argument supporting this point of view are the results of experiments, when the task was to detect a difference between two successive sinusoidal tones: these threshold are largely lower that those for modulation detection [28].

In summary it may be stated that there are two mechanisms for the detection of amplitude and frequency changes of a signal at low rates. One of them is a mechanism based entirely on changes in excitation level, proposed by Zwicker. This mechanism operates for amplitude and frequency changes for all range of carrier frequencies and modula-

tion rates, being most effective for modulation rates higher than 5 Hz. However, for very low modulation frequencies, less than 5 Hz, an additional mechanism enhances the auditory system's sensitivity to frequency changes. The mechanism provides additional information about frequency changes and brings about markedly higher detectability for the detection and discrimination of the frequency changes. The mechanism is probably based on the analysis of the phase-locking to samples of the signal.

The concept of the two independent mechanisms for detection amplitude and frequency changes has already been presented by FETH [5] and CONINX [2, 3]. However, they assumed that differences in the detection of amplitude and frequency changes occur over the whole range of modulation rates. The model suggested in this paper assumes that detection of changes in amplitude or frequency of a signal is based primarily on the changes in excitation level over the active region of the excitation pattern. However, for very low modulation rates, when the changes in the signal's parameters are very slow, the information conveyed by inter-spike intervals is used to evaluate the frequency of the signal.

## References

[1] J.T. ALLANSON and A.F. NEWELL, *Subjective responses to tones modulated simultaneously in both amplitude and frequency*, J. Sound. Vib., **3**, 135–146 (1966).

[2] F. CONINX, *The detection of combined differences in frequency and intensity*, Acustica, **39**, 137–150 (1978).

[3] F. CONINX, *The perception of combined frequency and amplitude modulation with clearly audible modulation depths*, Acustica, **39**, 151–154 (1978).

[4] L. DEMANY and C. SEMAL, *On the detection of amplitude modulation and frequency modulation at low modulation frequencies*, Acustica, **61**, 243–255 (1986).

[5] L.L. FETH, *Combinations of amplitude and frequency differences in auditory discrimination*, Acustica, **26**, 67–77 (1972).

[6] M. FLORENTINE and S. BUUS, *An excitation-pattern model for intensity discrimination*, J. Acoust. Soc. Am., **70**, 1646–1654 (1981).

[7] D.M. GREEN and J.A. SWETS, *Signal detection theory and psychophysics*, Krieger, New York 1974.

[8] W.M. HARTMANN and G.M. HNATH, *Detection of mixed modulation*, Acustica, **50**, 297–312 (1982).

[9] W.M. HARTMANN and M.A. KLEIN, *Theory of modulation detection for low modulation frequencies*, J. Acoust. Soc. Am., **67**, 935–946 (1980).

[10] N.A. MACMILLAN and C.D. CREELMAN, *Detection theory: A User's guide*, Cambridge University Press, Cambridge, England 1991.

[11] D. MAIWALD, *Die Berechnung von Modulationsschwellen mit Hilfe eines Funktionsschemas (calculation of differential thresholds by means of a functional model)*, Acustica, **18**, 193–207 (1967).

[12] D. MAIWALD, *Ein Funktionsschema des Gehörs zur Beschreibung der Erkennbarkeit kleiner Frequenz- und Amplitudenänderungen*, Acustica, **18**, 81–92 (1967).

[13] B.C.J. MOORE and B.R. GLASBERG, *Formulae describing frequency selectivity as a function of frequency and level and their use in calculating excitation patterns*, Hear. Res., **28**, 209–225 (1987).

[14] B.C.J. MOORE and B.R. GLASBERG, *Mechanisms underlying the frequency discrimination of pulsed tones and the detection of frequency modulation*, J. Acoust. Soc. Am., **86**, 1722–1732 (1989).

[15]  B.C.J. Moore and B.R. Glasberg, *Suggested formuale for calculating auditory-filter bandwidth and excitation patterns*, J. Acoust. Soc. Am., **74**, 750–753 (1983).

[16]  B.C.J. Moore and A. Sęk, *Detection of combined frequency and amplitude modulation*, J. Acoust. Soc. Am., **92**, 3119–3131 (1992).

[17]  B.C.J. Moore and A. Sęk, *Discrimination of modulation type (AM or FM) with and without background noise*, J. Acoust. Soc. Am., **96**, 726–732 (1994).

[18]  B.C.J. Moore and A. Sęk, *Effects of carrier frequency and background noise on the detection of mixed modulation*, J. Acoust. Soc. Am., **96**, 741–751 (1994).

[19]  B.C.J. Moore and A. Sęk, *Effects of carrier frequency, modulation rate and modulation waveform on the detection of modulation and the discrimination of modulation type (AM vs FM)*, J. Acoust. Soc. Am., **97**, 2468–2478 (1995).

[20]  E. Ozimek and A. Sęk, *AM and FM difference limens and their reference to amplitude-frequency changes of a sound in a room*, Acta Acustica, **82**, 114–122 (1996).

[21]  E. Ozimek and A. Sęk, *Perception of amplitude and frequency modulated signals (mixed modulation)*, J. Acoust. Soc. Am., **82**, 1598–1603 (1987).

[22]  J.O. Pickles, *An introduction to the physiology of hearing*, 2nd Ed., Academic Press, London 1988.

[23]  M.A. Ruggero and N.C. Rich, *Application of a commercially-manufactured doppler-shift laser velocimeter to the measurement of basilar-membrane motion*, Hear. Res., **51**, 215–230 (1991).

[24]  M.A. Ruggero and N.C. Rich, *Furosemide alters organ of Corti mechanics: Evidence for feedback of outer hair cells upon the basilar membrane*, J. Neurosci., **11**, 1057–1067 (1991).

[25]  A. Sęk, *Critical modulation frequency and critical band based on random amplitude and frequency changes*, Archives of Acoustics, **19**, 59–74 (1994).

[26]  A. Sęk and B.C.J. Moore, *Detection of mixed modulation using correlated and uncorrelated noise modulators*, J. Acoust. Soc. Am., **95**, 3511–3518 (1994).

[27]  A. Sęk and B.C.J. Moore, *The detection of modulation at low modulation rates*, [In:] Auditory Perception: Some Principles and Applications, U. Jorasz, Adam Mickiewicz University Press, Poznań 1996.

[28]  A. Sęk and B.C.J. Moore, *Frequency discrimination as a function of frequency, measured in several ways*, J. Acoust. Soc. Am., **97**, 2479–2486 (1995).

[29]  A. Sęk and E. Ozimek, *AM and FM difference limens*, Archives of Acoustics, **19**, 37–58 (1994).

[30]  E. Zwicker, *Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs*, Acustica, **6**, 356–381 (1956).

[31]  E. Zwicker, *Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones*, Acustica, **2**, 125–133 (1952).

[32]  E. Zwicker, *Temporal effects in simultaneous masking and loudness*, J. Acoust. Soc. Am., **38**, 132–141 (1965).