

T E C H N I C A L R E P O R T S

Rauschbefreiung durch binaurales Hören Improvement of weak signal detection with binaural reception

von U. Schreiber DF1DM, Dahlmannstr. 22, D-3400 Göttingen

D.: Einleitung:

Es gibt wohl viele Unterschiede zwischen einem waschechten KW- Amateur und einem gestandenen VHF- (bzw. UHF-) DXer. Einer der wesentlichsten ist dabei wohl die hinsichtlich Strahlungsverlusten optimierten Stationen der VHF-Enthusiasten. Mit viel Aufwand wird hier besonders im Empfangszweig um jedes 1/10- dB gekämpft, denn eine empfindliche Empfangsstation entscheidet häufig über das Zustandekommen einer Verbindung.

Die nachfolgende Abhandlung soll zeigen, wie man durch Ausnutzung des Effektes des binauralen Hörens eine Absenkung der Mithörschwelle erreicht, was einer Verbesserung des Signal- Rauschverhältnisses um mehrere dB gleichkommt. Im Anschluß daran soll noch der prinzipielle Aufbau einer Schaltung erläutert werden, die eine Ausnutzung dieses Effektes zuläßt.

E.: Introduction:

Special operating modes, like EME for example, require the best possible equipment on the receiving part of the station because of the poor signal to noise ratio. The following article will show how the audio detection threshold can be lowered by means of binaural unmasking. Finally, a circuit to make use of this effect is introduced schematically.

D.: Zur Theorie des binauralen Hörens.

Den Vorteil, den man aus binauralem (zweiohrigen) Hören zieht, kann man am Besten an einem praktischen Beispiel erläutern. In einer stark lärmbelasteten Umgebung kann man sich auf einen Gesprächspartner konzentrieren, auch wenn dieser in einer Gruppe anderer Menschen steht, die ebenfalls gleichzeitig ein Gespräch führen. Offensichtlich ist man in der Lage die Störgeräusche gegenüber den erwünschten Signalen zu unterdrücken, was z.B. einer einseitig hörbehinderten Person nicht gelingt. Dieser Effekt wird in der Literatur als "Cocktailparty-Effekt" bezeichnet. Durch die Anwesenheit einer maskierten Geräuschquelle wird also die Hörschwelle für ein Signal (fortan sei damit ein CW- Sinuston von 1000 Hz gemeint) deutlich heraufgesetzt. Man hat also zwischen der absoluten Hörschwelle und der Mithörschwelle zu unterscheiden. Das Erste bezeichnet die Intensität eines angebotenen Testsignales, bei der der Reiz noch wahrgenommen wird. Dieser Reiz wird dabei in einer gänzlich ungestörten Umgebung ausgeübt. Unter der Mithörschwelle dagegen versteht man nun die entsprechende Intensität des Testsignales unter Anwesenheit einer maskierten Rauschquelle. Dabei liegt naturgemäß der erforderliche Pegel für die Mithörschwelle deutlich über dem der absoluten Hörschwelle. Durch binaurales Hören, wie am "Cocktailparty-Effekt" veranschaulicht, kann diese Mithörschwelle herabgesetzt werden. Dabei wird der Unterschied zwischen monauraler (einohriger) und binauraler Mithörschwelle ausgewertet. Diese Pegeldifferenz wird in der Literatur mit BMLD (Binaural Masking Level Difference) bezeichnet. Sie ist abhängig von der interauralen (d.h. der zwischen den Ohren auftretenden) Phasendifferenz des Maskierers und des Testsignales. Ein praktisches Beispiel soll dies weiter veranschaulichen. Man biete ein breitbandiges Rauschen und ein sinusförmiges Testsignal EINEM Ohr an und bestimme durch Leistungserniedrigung des Testtones den Pegel, an dem der Ton gerade noch wahrgenommen wird (Dabei wird man feststellen, daß man diese Schwelle nicht beliebig scharf festlegen kann). Bietet man nun noch zusätzlich dasselbe Rauschen dem anderen Ohr gleichzeitig an, so ist der Testton wieder deutlich hörbar, obwohl sich die abgegebene Rauschleistung verdoppelt hat. Der

Unterschied in der Mithörschwelle beträgt dabei etwa 6 dB. Es gibt eine ganze Reihe von Modellvorstellungen, die quantitative Aussagen über das Zustandekommen dieser Unterschiede in den Mithörschwellen machen. Eine sehr anschauliche und in der Literatur favorisierte Vorstellung ist das von DURLACH (1) vorgeschlagene EC-Modell (Equalisation and Cancellation Theory). Es geht davon aus, daß in einem ersten Verarbeitungsschritt die von beiden Ohren stammenden bandpaßgefilterten Signale einander angeglichen (equalisation) und im zweiten Schritt von einander subtrahiert (cancellation) werden. Abb. 1 zeigt im Blockdiagramm die grobe Struktur dieser Theorie. Es sollte in diesem Zusammenhang noch mal deutlich erwähnt werden, daß es sich dabei um ein physikalisches Modell handelt, welches die beobachteten Effekte gut wiedergibt, ohne daß dabei auf physiologische Strukturen Rücksicht genommen werden muß.

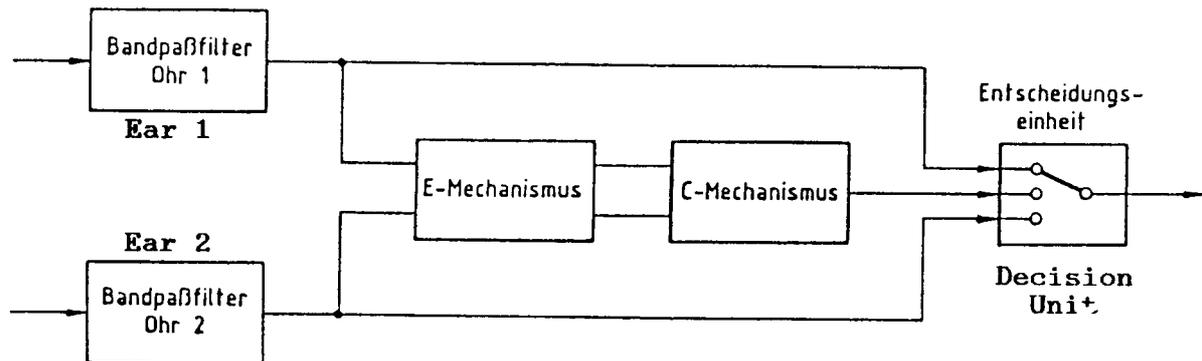


Abb. 1 Blockschaltbild des EC-Modells

Im Rahmen dieses Modells sei das obige Beispiel noch mal betrachtet. In der Gleichsetzungsstufe können durch eine geringfügige Pegel- und Zeitverzögerungsvariation die Signale beider Ohren angeglichen werden. In der nachfolgenden Stufe wird nun das Rauschen von gleichem Rauschen unterlegt und der Testton subtrahiert. Wäre dieses Modell beliebig exakt und die einzelnen Mechanismen fehlerfrei, so bliebe nur noch der Testton an der "Entscheidungseinheit", welche immer den Kanal mit dem besten Signal- Rauschverhältnis "durchschaltet", übrig.

Das käme aber einer Verbesserung von unendlich gleich, was nicht realistisch ist und experimentell auch nicht gefunden werden kann. Aus diesem Grunde wurde eine mit statistischen Fehlern behaftete Arbeitsweise des Gleichsetzungsapparates angenommen. Dies kann durch das in Abb. 2 dargestellte, modifizierte Modell erreicht werden. Dabei müssen keine zusätzlichen Annahmen über diese Verarbeitungsfehler gemacht werden.

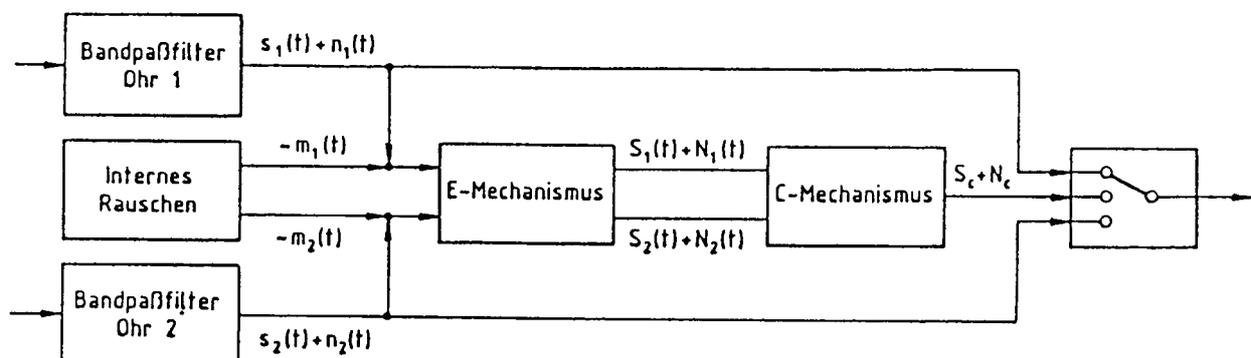


Abb. 2 Blockschaltbild des modifizierten EC-Modells

Ein zusätzliches "internes Rauschen" wird den Eingangssignalen im Gleichsetzungsmechanismus mit einer Amplitude aufaddiert, welche proportional zur effektiven Spannung des Eingangssignales ist. Auf diese Weise erhält man aus

dem Modell für das oben dargestellte Beispiel eine Schwellendifferenz von 6 dB, was experimentell auch gefunden wurde. Nun ist aber die Situation eines Funkamateurs nicht mit diesem Beispiel vergleichbar, denn hier kann das unerwünschte Rauschen nicht, wie im Experiment oben, von dem Sinussignal getrennt werden. Wie kommt man nun aber doch zu einer Verbesserung der Schwellendifferenz gegenüber dem monauralen Fall? Zu diesem Zweck wird das Lautsprechersignal des Empfängers unmittelbar auf eine Kopfhörermuschel gegeben. Das gleiche NF-Signal passiert noch eine Zeitverzögerungsschaltung mit einer Regelung der Zeitverzögerung von 0 bis ca. 7 Millisekunden (Abb. 3) bevor es auf die Kopfhörermuschel für das andere Ohr gelangt.

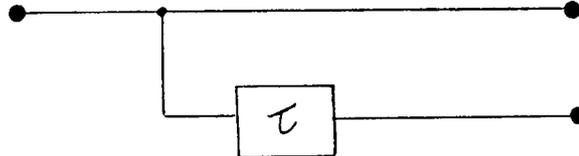


Abb. 3 Die erforderliche Baugruppe im Blockschaltbild

Die Wirkungsweise dieser Schaltung ergibt sich durch folgende Überlegung: Das interaural um einige Millisekunden verzögerte Rauschen kann im wesentlichen als unkorreliert angesehen werden, während das periodische Sinustonsignal stark korreliert bleibt und nach dem "Gleichsetzungsapparat" zwischen beiden Kanälen gegenphasig vorliegt. Daraus resultiert eine Absenkung der Mithörschwelle von ca. 3 dB gegen den rein monauralen Fall, was auch durch Labormessungen bestätigt wurde (Siehe Literatur). Im Anhang wird dies nochmals durch Berechnungen im Rahmen des EC-Modells ausgeführt.

E.: Theory of binaural reception.

The advantage which comes from binaural (two ears) listening is best explained with an example. In a very noisy location, such as a party, one is able to concentrate on someone talking in a background of other speech. This is called "the Cocktail Party-effect" and will not work with people having only one ear intact. The detection threshold of a tonal signal is raised when making with noise. Therefore, we have to distinguish between the absolute detection threshold which is reached when the power of an audio signal of let us say 1000 Hz is lowered, in the absence of any noise, until it is just inaudible; and a relative detection threshold, where the signal is masked with noise. The difference between these two thresholds is denoted by the MLD (Masking Level Difference). However, the incoming signals at two ears can have a phase difference between them, and the BMLD (Binaural- MLD) yields a gain, when compared to monaural signal detection (or, which is nearly the same, a two ears signal reception, with exactly the same stimuli on both ears). This can be demonstrated by the following experiment. Supply noise to one ear (via earphone) at a comfortable listening level; then add a signal consisting of a say 500 Hz tone interrupted every quarter of a second and adjusted in level until it is just inaudible. Now apply the same noise to the other earphone and the signal becomes clearly audible. The signal again disappears when it is added to the second earphone, thus making the sound in the two ears alike. There are quite a few models discussing these effects. The EC-model (Equalisation and Cancellation) of DURLACH (1) et. al. gives a good representation of the experimental data and has the advantage of being a pure physical model with no assumptions made on physiological structures. Fig. 1 shows how it works in principle. Small level adjustments of the amplitude and the phase of the bandpass filtered signals from the ears can be made. The outputs of this stage are then subtracted from each other in the cancellation mechanism. The decision unit finally switches to that channel, which has the best signal to noise ratio. Looking at the above example, we find the noise of the one ear set equal to and subtracted from the tone signal plus the same noise from the other ear. If this mechanism would work error free, we would have an unrealistic high gain in unmasking at the decision unit which is experimentally not found.

In fact we have to admit statistical errors in the equalization unit. Fig. 2

shows an improved version of this model. An "internal noise" is added to each input signal in such a way, that it is proportional to the effective voltage of the input signal. Mathematically we get a level difference of 6 dB in the above example, which is in good agreement to the experiment.

Unfortunately, the situation of a radio amateur is not comparable with this example, since noise and the tone signal cannot be separated. Now, how can we make use of binaural listening? Fig. 3 shows a block diagram of a device, which improves the signal to noise ratio by approximately 3 dB. The audiosignal is supplied to the left ear. The same signal is fed to a unit which delays the input by roughly 7 milliseconds. (The appropriate time to delay the signal can be adjusted to the particular needs) This delayed signal is then applied to the right ear. So the interaural delayed noise can be considered uncorrelated, while the masked periodic tonal signal remains correlated and after passing the equalisation unit, will be of opposite phase. The theoretically achievable gain is calculated in the appendix.

D.: Praktische Erfahrungen:

In der Anwendung zeigt sich zunächst ein wesentlich angenehmerer Rauschein- druck. Dies mag daran liegen, daß es dem auf binauralen Höreindrücken eingestellten Operator im Vergleich zu dem durch "normale" Kopfhörer realisierten monauralen Fall, vertrauter erscheint. Auch hier kann man sich von der Wirkung durch ein Experiment überzeugen. Wieder wird der Pegel des durch Rauschen maskierten Tones soweit herabgesetzt, daß er mit einem Kopfhörer, bei dem an beiden Hörmuscheln dasselbe Signal anliegt, soeben noch wahrgenommen werden kann. Schaltet man jetzt in einem Kanal eine Zeitverzögerungseinheit (von 0-7 Millisekunden regelbar) zu, welche vorher auf beste Verständlichkeit eingeregelt wurde, so ist der Ton wieder deutlich hörbar. Bei Versuchen mit dieser Anordnung zeigte sich, daß entgegen der Erfahrung im monauralen Fall, wo an der Mithörschwelle die Wahrnehmung des Testtones zu fluktuieren scheint, an der entsprechend niedrigen Mithör- schwelle, der einseitig zeitverzögernden Anordnung, der Ton entweder eindeutig zu hören war, oder nicht.

Abb. 4 zeigt im Blockdiagramm eine mögliche Realisierung dieser Schaltung. Es handelt sich im einfachsten Fall um einen Aufbau unter Verwendung des analogen Schieberegisters SAD 1024.

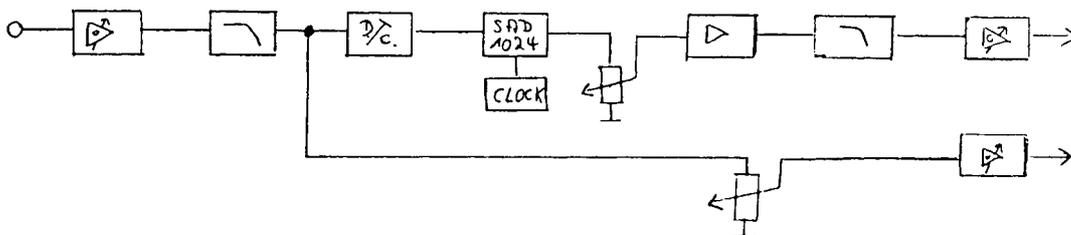


Abb. 4: Blockschaltbild der Verzögerungsschaltung

Das Eingangssignal wird verstärkt und passiert dann eine Tiefpaßfilterstufe. Anschließend wird ein Gleichspannungsanteil dem verrauschten Signal zuge- mischt, da der Eimerkettenpeicher- IS nur Spannungen, die größer als 0V sind, verarbeiten kann. Das Signal am Ausgang des SAD 1024 wird weiter verstärkt und passiert dann wieder ein Tiefpaßfilter. Dies ist notwendig, um das Taktsignal des "Schieberegisters", welches im Bereich der Audiofrequenzen liegen kann, auszufiltern. Eine nachgeschaltete Verstärkerstufe bringt das verzögerte Eingangssignal auf den nötigen Pegel, um einen Kopfhörer anzusteuern. Es ist sicherlich sinnvoll, auch das unverzögerte Signal entsprechend zu filtern und zu verstärken, denn der Effekt der Rauschbefreiung ist bei der Verzögerung möglichst ähnlicher Signale am größten.

Diese Schaltung stellt sicherlich nicht das erzielbare Optimum dar. Sie entstand in Anlehnung an einen Artikel aus der Zeitschrift Elektor (Heft 10/1978: Elektronischer Hall) und wurde entsprechend für die obige

Anwendung modifiziert. Da dieser Artikel nicht jedermann zugänglich sein wird, stelle ich ihn auf Anfrage gerne zur Verfügung.

E.: Operating Experience (Experimental setup)

When operating an unit as shown in Fig. 3, a more pleasant noise impression is obvious, when compared to "regular" earphones (monaural case). This may result out of the operator's being used to binaural reception in everyday life. An experiment will show the effectivity of the unit. Again the intensity of a tone signal applied to both ears, masked by noise at a comfortable listening level, is lowered until it just cannot be distinguished any more. When switching the time-delay-unit into the line of one ear, the signal again can be clearly heard.

Fig. 4 shows a more detailed block diagram. An analog shifting-register SAD 1024 is used. The first stage amplifies and lowpassfilters the incoming signal. Then a DC voltage is added to the signal, so that no negative voltages appear on the input of the SAD 1024. The delayed signal is once again amplified and lowpass filtered. This is to eliminate the variable clock frequency, which may reach down to the audio frequency range. A final amplifier adjusts the signal level so that earphones can be driven. The undelayed signal passes the low pass filter and amplifier stages, too, since it can be assumed that the unmasking effect works best on signals which are basically alike.

This circuit is surely not the ultimate realisation. It was taken from the electronic journal "Elektor" (issue 10/78 "Elektronischer Hall") (=E.: electronic reverberation) and was modified to the needs of this article. Since the description of the circuit will not be available to everyone, I can supply further information when asked for.

D.: Schlußbemerkung:

Es liegt nahe, dieses Konzept nicht nur auf maskierte streng periodische Signale anzuwenden, sondern auch auf stark verrauschte Sprachübertragungen. Entsprechende Versuche im SSB-Bereich schienen das auch zu bestätigen. Exakte Untersuchungen in der Form einiger Silbentests unter definierten Laborbedingungen zeigten aber, daß für Sprachübertragungen offensichtlich zwei getrennte Mechanismen zur Anwendung kommen. Dabei handelt es sich um einen "Spracherkennungssapparat" und um einen "Sprachanalyseapparat". Durch die oben dargestellte Schaltung wird die Spracherkennung zwar erheblich verbessert (man hat unmittelbar den Eindruck, daß die Anordnung zu einem erheblich verbesserten Empfang führt), es ließ sich aber keine Steigerung in der Richtigkeit der übertragenen Informationen nachweisen. Eine Verbesserung bei der Übertragung von Sprache, wäre auch im Rahmen der EC-Theorie nicht zu erklären gewesen. Abschließend möchte ich noch bemerken, daß ich über weiterreichende Erfahrungen mit dem hier vorgestellten Verfahren dankbar wäre.

E.: Final remarks:

It sounds reasonable, to apply this concept to heavily masked speech transmissions, too. Experiments in the 2m SSB-Band seemed to yield good results. But exact investigations under laboratory conditions did not confirm it. It was found, that there are obviously (as reported in the literature) two different mechanisms on the perception of speech. There is a "speech recognition unit" and a "speech analysing unit". The circuit, introduced above, definitely improves the recognition of speech, but could not improve the number of correctly copied syllables in appropriate tests. The EC-model itself does not suggest any improvement with the unmasking of speech. Finally, I would appreciate reports on experiences gained with this method.

Literature:

- 1) Foundations of Modern Auditory Theory, Vol II
J. V. Tobias, Academic Press 1972
(Binaural Phenomena) Handbook of Perception IV
Colburn - Durlach, Academic Press

Anhang:
Appendix:

D.: Zur Berechnung der BMLD des durch Rauschen maskierten Sinus- Tonsignales auf einem Ohr und dem um τ verzögerten gleichen Signales auf dem anderen Ohr nach dem in Abb. 2 skizzierten Modell sei:

E.: To calculate the BMLD of a periodic, sinusoidal signal ($f=500..1000\text{Hz}$) masked by noise, interaurally delayed by τ , and treated by means of the model shown in Fig. 2, it is assumed:

$$s_1(t) = \sqrt{2A} \sin \omega t ; n_1(t) = \sqrt{N} n(t)$$

D.: Dabei zeichnet A die mittlere Leistung und ω die Frequenz des Tonsignales $S_1(t)$ am Ohr 1 und $n(t)$ einen Rauschprozeß der Leistung 1, der zusammen mit der Leistung N den Rauschprozeß an diesem Ohr wiedergibt. Entsprechend ergibt sich für das andere Ohr:

E.: where A denotes the mean power and ω the frequency of the tonal signal applied to ear 1. $n(t)$ is meant to be a noise process having the a power level normalized to 1, which together with the power N representing the total noise is applied to ear 1. For ear 2 we assume accordingly:

$$s_2(t) = \sqrt{2A} \sin \omega(t+\tau) ; n_2(t) = \sqrt{N} n(t+\tau)$$

D.: Am Ausgang des E-Mechanismus erhält man nun die Signale $S_1(t)$ und $S_2(t)$, die aus den vorherigen Signalen $s_1(t)$ und $s_2(t)$ durch Addition des "internen Rauschens" hervorgehen. Dabei sind $m_1(t)$ und $m_2(t)$ voneinander unabhängig und von der Leistung 1. Sie werden mit der Wurzel der Eingangsleistung multipliziert, so daß das interne Rauschen proportional zu der effektiven Amplitude des Eingangssignales ist. Weiterhin wird hier von der Fähigkeit des E-Mechanismus, im eng begrenzten Umfange die Signale gegenseitig zu verzögern, Gebrauch gemacht. Auf diese Weise wird mit $\tau+d\tau=\tau'+n\pi/\omega$ (ω ist die Signaltonkreisfrequenz, n =ungerade) die interaurale Gegenphasigkeit des Signaltones erreicht:

E.: The output of the E-mechanism will provide the two signals $S_1(t)$ and $S_2(t)$, which were obtained by adding an "internal noise" $m_1(t)$ and $m_2(t)$, each with a power of 1, to the input signals. They are multiplied by the square root of the input power, so that the internal noise is proportional of the effective amplitude of the E-mechanism of the input signal. Furthermore advantage of the E-mechanism's ability to manipulate the interaural delay slightly is taken. So, using $\tau+d\tau=\tau'+n\pi/\omega$ (ω the angular frequency and n a natural odd number) there are two tonal signals of opposite phase, thus giving:

$$\begin{aligned} S_1(t) &= \sqrt{2A} \sin \omega(t+\tau) + r\sqrt{A} m_1(t) \\ &= \sqrt{2A} \sin \omega t + r\sqrt{A} m_1(t) \end{aligned}$$

Entsprechend gilt:/ and:

$$S_2(t) = \sqrt{2A} \sin \omega t + \sqrt{A} m_2(t)$$

Im Cancellation- Schritt erhält man:/ The cancellation step yields:

$$S_c(t) = S_2(t) - S_1(t) = 2\sqrt{2A} \sin \omega t + r\sqrt{A} (m_2(t)-m_1(t))$$

Dann berechnet sich die mittlere Leistung unter der Annahme, daß $m_1(t)$ und $m_2(t)$ voneinander unabhängige Rauschprozesse sind zu:

Assuming that $m_1(t)$ and $m_2(t)$ do not depend on each other, we get for the mean power:

$$S_c^2(t) = 4A (1 - \cos 2\omega t) + 2Ar^2 (1 - m_1(t) m_2(t)) + 4Ar\sqrt{2} (m_2(t) - m_1(t)) \sin \omega t$$

Und nach der Integration über t: / Integrating over t yields:

$$S_c^2 = 4A + 2Ar^2 = 2A (2 + r^2)$$

In entsprechender Weise geht man für das Rauschen vor: / These steps are performed for the noise, too:

$$N_1(t) = n_1(t) + r\sqrt{N} m_1(t)$$

$$N_2(t) = n_1(t+\tau') + r\sqrt{N} m_2(t)$$

Dann wird: / So now is:

$$N_c(t) = (n_1(t+\tau') - n_1(t)) + r\sqrt{N} (m_2(t) - m_1(t))$$

Bei der Berechnung der mittleren Leistung geht mit ein, daß $n_1(t+\tau')$ und $n_1(t)$ nunmehr als unkorreliert anzusehen sind: / The mean power is calculated, keeping in mind that $n_1(t+\tau')$ and $n_1(t)$ are no longer correlated to each other:

$$\surd N_c^2 = 2N + 2Nr^2 = 2N (1+r^2)$$

Somit ergibt sich an der "Entscheidungseinheit" das Nutzsinal/ Rauschverhältnis f_c zu: / So the signal/ noise ratio f_c at the input of the decision unit is:

$$f_c = \frac{S_c^2}{N_c^2} = \frac{2A (2 + r^2)}{2N (1 + r^2)}$$

Um die Verbesserung im Nutzsinal/ Rauschverhältnis gegenüber dem monauralen Fall als Bezugsgröße zu berechnen, welche nach Voraussetzung gegeben ist zu $f_m = A/N$, bildet man: / The improvement of this method is seen, when compared to the monaural case ($f_m = A/N$):

$$f = \frac{f_c}{f_m} = \frac{2A (2 + r^2)}{2N (1 + r^2)} \bigg/ \frac{A}{N} = \frac{2 + r^2}{1 + r^2}$$

Also erhält man für BMLD in diesem Fall: / Since the BLMD is defined to:

$$L = 10 \log (f)$$

Aus Experimenten ist für die Kopplungskonstante r ein Wert von 0.214 ermittelt worden. / and the coupling constant r is found experimentally to be 0.214:

$$L \approx 2.9 \text{ dB}$$

Damit ergibt sich eine zu erwartende Absenkung der Mithörschwelle von nahezu 3dB. / That means, that we have to expect an total improvement roughly 3 dB.