

**Zeitschrift:** Bulletin des Schweizerischen Elektrotechnischen Vereins  
**Herausgeber:** Schweizerischer Elektrotechnischer Verein ; Verband Schweizerischer Elektrizitätswerke  
**Band:** 62 (1971)  
**Heft:** 12

**Artikel:** Signalerkennung mit parametrischen und nichtparametrischen Methoden  
**Autor:** Schlaepfer, H.J.  
**DOI:** <https://doi.org/10.5169/seals-915828>

### **Nutzungsbedingungen**

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

### **Conditions d'utilisation**

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

### **Terms of use**

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

**Download PDF:** 17.02.2026

**ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>**

# BULLETIN

DES SCHWEIZERISCHEN ELEKTROTECHNISCHEN VEREINS

Gemeinsames Publikationsorgan des Schweizerischen Elektrotechnischen Vereins (SEV)  
und des Verbandes Schweizerischer Elektrizitätswerke (VSE)

## Signalerkennung mit parametrischen und nichtparametrischen Methoden

Von H. J. Schlaepfer, Zürich.

621.391.8

Die vorliegende Arbeit enthält in Anlehnung an [1]<sup>1)</sup> eine kurze Einführung in das Problem der Erkennung von Signalen mit bekanntem zeitlichen Verlauf in überlagertem Rauschen. Im besonderen werden nichtparametrische und parametrische Methoden einander gegenübergestellt und im Hinblick auf die Realisierungsmöglichkeit und die Leistungsfähigkeit miteinander verglichen. Es werden einige neue Ergebnisse von Berechnungen des Wirkungsgrades des (nichtparametrischen) Vorzeichen-Detektors bei endlichen Stichprobenzahlen und gaußischem Rauschen gegeben. Für das Verständnis der Arbeit ist die Kenntnis der grundlegenden Zusammenhänge der Wahrscheinlichkeitsrechnung unerlässlich.

L'exposé précité — basé sur [1]<sup>1)</sup> — renferme une courte introduction au problème de la perception des signaux à allure chronologique connue du bruit superposé. On confronte en particulier des méthodes paramétriques et non-paramétriques pour les comparer du point de vue de leurs possibilités de réalisation et de leur capacité. On fournit de nouveaux résultats des calculs de rendement du détecteur de signes (non-paramétrique) lors de sondages finis et de bruit de Gauss. La compréhension de ce travail présume la connaissance des relations fondamentales du calcul des probabilités.

### 1. Einleitung

#### 1.1 Allgemeines über die Signalerkennung in digitalen Übertragungssystemen

In der vorliegenden Arbeit wird das Problem untersucht, ein dem Empfänger eines digitalen Übertragungssystems zum voraus bekanntes Signal  $s_1(t)$ , welches vom Kanal mit Rauschen überlagert wurde, dann mit möglichst grosser Sicherheit zu erkennen, wenn es tatsächlich vorhanden ist. Das Ziel besteht in der funktionalen Definition eines Entscheidungsgliedes  $D$  («Detektor»), welches in bestimmter, noch zu definierender Weise optimale Eigenschaften aufweist. Diese Problemstellung ist in der Nachrichtentechnik charakteristisch für alle digitalen Übertragungssysteme, bei welchen der Sender in jedem Takt aus einem beschränkten Vorrat von Signalen jeweils eines auswählt und übermittelt. Als Beispiel sei die Übertragung der Wählinformation beim Telefon genannt, wo die 10 Zeichen 0...9 zehn verschiedenen (hier allerdings zeitlich ungleich langen) Signalformen entsprechen. Die abzählbare Menge möglicher Sendesignale eines digitalen Übertragungssystems (Fig. 1) heisst Signalalphabet. Den folgenden Ausführungen wird ein binäres Alphabet zu Grunde gelegt, welches nur die beiden Signale  $S \equiv 0$  und  $S = s_1(t)$  enthält. Das binäre System ist naturgemäss das einfachste, mit welchem überhaupt Information übertragen werden kann.

Die theoretischen Hilfsmittel zur Lösung der gestellten Aufgabe stellt die mathematische Statistik bereit; es handelt sich im vorliegenden Fall insbesondere um den Entscheid zwischen einer einfachen Hypothese und einer einzigen Alter-

nativen. Die Herleitung des optimalen Detektors geht von der Annahme aus, dass der Empfänger im zeitlichen Intervall  $0 \leq t < T$  (einem Takt entsprechend) das Signal  $y(t)$  beobachten kann, welches als lineare Überlagerung des Signales  $S$  und des Rauschens  $n(t)$  entstanden ist:

$$y(t) = n(t) \quad \text{falls } S \equiv 0$$
$$y(t) = n(t) + s(t) \quad \text{falls } S = s(t)$$

Im speziellen werden in diesem Intervall nur eine Anzahl  $N$  von Momentanwerten der beobachteten Funktion  $y(t)$  untersucht und als Vektor dargestellt, siehe Fig. 2:

$$[Y] = [y(t_1), y(t_2), \dots, y(t_N)]$$

Da sich die einzelnen Momentanwerte als Überlagerung des Signalvektors  $[S]$  und des Rauschvektors  $[N]$  ergeben, gilt die Vektorbeziehung:

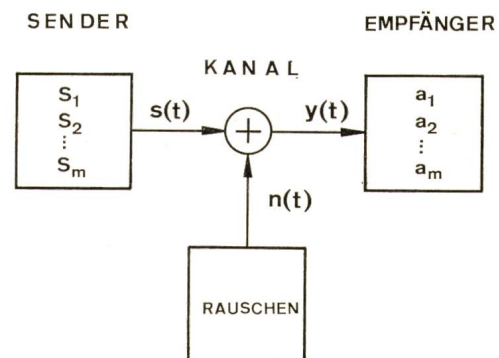


Fig. 1

Grundsätzliches Modell eines Digital-Übertragungssystems  
 $s_1, \dots, s_m$  verschiedene Signale;  $a_1, \dots, a_m$  Empfänger-Entscheide

<sup>1)</sup> Siehe Literatur am Schluss des Aufsatzes.



$$[Y] = [S] + [N] \quad (1)$$

Das Ziel besteht nun in der Definition der optimalen Detektorfunktion  $D(Y)$ . Dazu wird eine bestimmte Fehlalarm-Wahrscheinlichkeit  $\alpha$  vorgegeben, das heisst eine gewisse Wahrscheinlichkeit dafür, dass sich der Detektor für die Hypothese «Signal anwesend» entscheidet, wenn tatsächlich  $S = 0$  war. Zusammen mit dieser Nebenbedingung wird nun derjenige Detektor gesucht, welcher eine maximale Erkennungswahrscheinlichkeit  $\beta$  aufweist, das heisst, dass die Wahrscheinlichkeit, ein tatsächlich vorhandenes Signal zu erkennen, möglichst gross wird. Diese Maximierung der Erkennungswahrscheinlichkeit bei gegebener Fehlalarmrate ist der Inhalt des Optimierungskriteriums von Neyman-Pearson. In der Literatur (z.B. [2; 3]) wird gezeigt, dass der im Sinne von Neyman-Pearson optimale Empfänger als Entscheidungsgrundlage den Quotienten zweier Wahrscheinlichkeiten<sup>2)</sup> verwendet, («likelihood ratio test»):

$$L(Y) = \frac{f_Y(Y | S = s_1(t))}{f_Y(Y | S = 0)} \quad (2)$$

Die Grössen  $f_Y(Y | S = s_1(t))$  und  $f_Y(Y | S = 0)$  stellen die bedingten Wahrscheinlichkeiten für die beobachteten Momentanwerte  $y_1 \dots y_N$  dar, unter der Voraussetzung, dass das Signal  $S = s_1(t)$  vorhanden war, bzw. dass  $S = 0$  war. Die Entscheidungsfunktion  $D(Y)$  des optimalen Detektors lautet daher

$$\begin{aligned} D(Y) &= 0 \text{ falls } L(Y) < L_0 \\ D(Y) &= 1 \text{ falls } L(Y) \geq L_0 \end{aligned} \quad (3)$$

Dieser Vorschrift entsprechend ist der Ausgang des Detektors  $D(Y) = 1$  («Signal  $S_1(t)$  anwesend»), falls der Quotient der bedingten Wahrscheinlichkeiten eine gewisse Schwelle  $L_0$  überschreitet und umgekehrt ist  $D(Y) = 0$  («kein Signal anwesend»), falls der Quotient der beiden bedingten Wahrscheinlichkeiten kleiner ist als  $L_0$ . Die Zuordnung des Gleichheitszeichens zu einem der Entscheide ist willkürlich, denn die Gleichheit tritt nur mit verschwindender Wahrscheinlichkeit ein.

## 1.2 Gegenüberstellung parametrischer und nichtparametrischer Methoden zur Signalerkennung

Bis hierher wurde ganz allgemein das Problem der Erkennung eines deterministischen Signales in überlagertem Rauschen formuliert; in der Art und Weise, wie sie dieses Ziel erreichen, unterscheiden sich jedoch die parametrischen von den nichtparametrischen Methoden. Damit eine explizite Form der Detektorfunktion gefunden werden kann, müssen noch gewisse Angaben über die Statistik der Rauschproben vorliegen.

Die parametrischen Detektoren gehen nun von der Annahme aus, dass die statistische Verteilungsfunktion  $F_n(n_i)$  der Momentanwerte  $n_i$  des Rauschens mit allen ihren Parametern bekannt sei. Dieser Weg wird in Abschnitt 2 besprochen werden, wo für die Rausch-Proben die Gaussverteilung vorausgesetzt wird. Damit gelingt es, einen einzigen,

<sup>2)</sup> Mit  $f_Y(x)$  wird in dieser Arbeit die Wahrscheinlichkeitsdichte der Zufallsgrössen  $y$  an der Stelle  $x$  bezeichnet, während mit

$$F_Y(x) = \int_{-\infty}^x f_Y(z) dz$$

die (kumulative) Verteilungsfunktion von  $y$  an der Stelle  $x$  dargestellt wird.

im Sinne von Neyman-Pearson optimalen Detektor zu definieren, welcher bei einer vorgegebenen Fehlalarm-Wahrscheinlichkeit die grösstmögliche Erkennungswahrscheinlichkeit erreicht. In den allereinfachsten Fällen ist dieser optimale Detektor auch noch realisierbar, doch werden die entsprechenden Strukturen bei etwas realistischeren Annahmen bald derart kompliziert, dass sie nur noch angenähert nachgebildet werden können.

In der Praxis taucht aber eine zusätzliche Schwierigkeit in dem Sinne auf, dass sich die statistischen Eigenschaften des Rauschens auf unbekannte Art und Weise zeitlich ändern können, oder dass sie im Extremfall zum voraus überhaupt nicht bekannt sind. Tatsächlich wird man in den wenigsten Fällen soviel über die Statistik der Rauschproben wissen, dass die Voraussetzungen der parametrischen Detektoren exakt erfüllt sind. Würde man dann — in Verletzung der Voraussetzungen — versuchen, die Erkennungsaufgabe auch bei nicht-gaussem Rauschen mit dem für gausches Rauschen optimalen Detektor zu lösen, so würde man finden, dass mit wesentlich einfacheren Detektor-Strukturen möglicherweise bessere Resultate erzielt werden können.

Auf diesen Erkenntnissen beruhen die nichtparametrischen Detektoren, welche von wesentlich allgemeineren Voraussetzungen an die Statistik der Rauschproben ausgehen. Die in Abschnitt 3 besprochenen nichtparametrischen Methoden verlangen von den Dichtefunktionen  $f_n(n_i)$  der Rauschamplituden nur noch Symmetrie-Eigenschaften:

$$f_n(n_i) = f_n(-n_i) \quad (4)$$

Diese Klasse symmetrischer Dichtefunktionen enthält eine nicht abzählbare Menge möglicher Dichtefunktionen, insbesondere auch die Gaussverteilung mit dem Erwartungswert  $E(n_i) = 0$  und beliebiger Streuung. Ein nichtparametrischer Detektor ist nun gerade dadurch definiert, dass er eine konstante Fehlalarm-Wahrscheinlichkeit unabhängig von der tatsächlichen Form der Dichtefunktion der Rauschproben unter sehr allgemeinen Voraussetzungen (z.B. Symmetrie-Eigenschaften) aufweist [1]. Der nichtparametrische Detektor garantiert also gewisse minimale Eigenschaften für sehr zahlreiche mögliche Arbeitsbedingungen, verbunden mit — im allgemeinen — recht einfachen Entscheidungsvorschriften.

Zusammenfassend können die Eigenschaften der nichtparametrischen Detektoren denen der parametrischen folgendermassen gegenübergestellt werden:

- a) Sie weisen im allgemeinen relativ einfache Strukturen auf.
- b) Sie berücksichtigen einen nur sehr kleinen Bruchteil der möglicherweise zum voraus nicht bekannten statistischen Eigenschaften der Rauschproben.
- c) Sie sind sicher den optimalen parametrischen Detektoren bei bekannter Statistik des Rauschens unterlegen (suboptimale Detektoren).
- d) Sie sind aber parametrischen Detektoren möglicherweise weit überlegen, für Verhältnisse, bei denen die Voraussetzungen des optimalen parametrischen Detektors nicht mehr erfüllt sind.
- e) Sie garantieren eine bestimmte Fehlalarm-Wahrscheinlichkeit für eine sehr grosse Klasse von Dichtefunktionen der Rauschproben.

Die Leistungsfähigkeit eines nichtparametrischen Detektors wird beurteilt durch einen Vergleich mit dem optimalen parametrischen Detektor. Als Mass der Wirksamkeit wird das Verhältnis der erforderlichen Anzahl von Stichproben des nichtparametrischen ( $n_{np}$ ) gegenüber dem parametrischen ( $n_p$ ) Detektor bei wohldefinierten Bedingungen verwendet (definierte Verteilungsfunktion des Rauschens, gegebene Fehlalarm- und Erkennungswahrscheinlichkeiten). Als asymptotischer



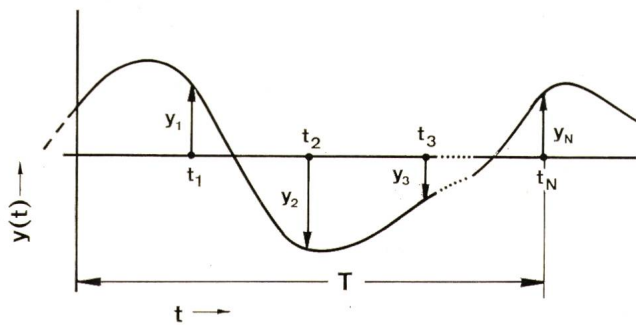


Fig. 2

Zur Vektordarstellung des Signals  $y(t)$

$t_1 \dots t_N$  Abtastmomente;  $t$  Zeit;  $y_1 \dots y_N$  Momentanwerte des Signals  $y(t)$  zu den Zeiten  $t_1 \dots t_N$

relativer Wirkungsgrad  $d_\infty$  des nichtparametrischen Entscheidungsgliedes wird der Quotient der beiden erforderlichen Stichprobenzahlen bezeichnet, wenn diese über alle Grenzen wachsen und gleichzeitig die Amplitude  $\lambda$  des Signales  $s_1(t)$  gegen Null strebt:

$$d_\infty = \lim_{\substack{n_p, n_{np} \rightarrow \infty \\ \lambda \rightarrow 0}} \frac{n_{np}(\alpha, \beta, \lambda)}{n_p(\alpha, \beta, \lambda)} \quad (5)$$

In Abschnitt 3 wird für verschiedene einfache nichtparametrische Detektorstrukturen der asymptotische Wirkungsgrad sowie der Wirkungsgrad bei endlichen Stichprobenzahlen für den Vorzeichen-Detektor bei nicht verschwindender Signalamplitude berechnet werden. Es wird sich zeigen, dass der asymptotische Wirkungsgrad auch für endliche Stichprobenzahlen eine sehr gute Annäherung darstellt.

## 2. Parametrische Signalerkennung

Das einfachste Beispiel einer Signalerkennung mit parametrischen Methoden stellt das folgende Entscheidungsproblem dar:

Gegeben sei der Vektor  $[Y] = [y_1, y_2, \dots, y_N]$  von Daten, welche gemäss Fig. 2 aufeinanderfolgende Momentanwerte einer vom Empfänger eines Übertragungssystems beobachtbaren Zeitfunktion  $y(t)$  darstellen. Von dieser Signalfunktion ist bekannt, dass sie als lineare Überlagerung von Gaußschem Rauschen mit dem Mittelwert  $\mu = 0$ <sup>3)</sup> und der Streuung (Leistung)  $\sigma^2$  und gegebenenfalls einem bekannten Signal  $s_1(t)$  entstanden ist. Das Entscheidungsproblem besteht nun darin, den nach *Neyman-Pearson* optimalen parametrischen Detektor  $D$  zu finden. In der Statistik wird die Aufgabe des Detektors als Entscheidung zwischen einer einfachen Hypothese  $H$  und einer einfachen Alternativen  $A$  beschrieben; nämlich

- $H_1$  Signal  $S$  anwesend: die Proben  $y(t_i)$  sind gaußverteilt mit dem Erwartungswert  $E[y(t_i)] = s_1(t_i)$  und der Streuung  $\sigma^2$ .
- $A_1$  Signal  $S = 0$ : die Proben  $y(t_i)$  sind gaußverteilt mit dem Erwartungswert  $E[y(t_i)] = 0$  und der Streuung  $\sigma^2$ .

Im folgenden soll im Interesse der Übersichtlichkeit angenommen werden, dass das Signal  $s_1(t)$  im untersuchten zeit-

<sup>3)</sup> Mit  $\mu = E(x)$  wird der Mittel- oder Erwartungswert der Zufallsgrösse  $x$  bezeichnet:

$$\mu = E(x) = \int_{-\infty}^{+\infty} x f_X(x) dx$$

Entsprechend gilt für die Streuung  $\sigma^2$  von  $x$ :

$$\sigma^2 = E[(x - \mu)^2] = \int_{-\infty}^{+\infty} (x - \mu)^2 f_X(x) dx$$

lichen Intervall  $0 \leq t < T$  konstant, nämlich  $s_1(t) = \lambda$  sei. Die interessanten Fälle zeitlich veränderlicher Signale werden in der Literatur ausführlich behandelt (siehe zum Beispiel [2]).

Die Verbundwahrscheinlichkeit  $f_N(y_1 \dots y_N | S = \lambda)$  für die Momentanwerte  $y_1 \dots y_N$  des beobachteten Signales ergibt sich, wenn statistische Unabhängigkeit der einzelnen Rauschproben vorausgesetzt wird, für den Fall, dass das Signal  $S = \lambda$  anwesend ist, zu

$$f_Y(y_1 \dots y_N | S = \lambda) = \prod_{i=1}^N f_Y(y_i | S = \lambda) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp - \left\{ \frac{\sum_{i=1}^N (y_i - \lambda)^2}{2\sigma^2} \right\} \quad (6)$$

Entsprechend findet man die Verbundwahrscheinlichkeit der Proben unter der Voraussetzung, dass der Vektor  $Y$  vom Rauschen allein stammt:

$$f_Y(y_1 \dots y_N | S = 0) = \prod_{i=1}^N f_Y(y_i | S = 0) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp - \left\{ \frac{\sum_{i=1}^N y_i^2}{2\sigma^2} \right\} \quad (7)$$

Nach den Ausführungen in Abschnitt 1 besteht die Funktion des optimalen Empfängers in der Bildung des Quotienten der bedingten Wahrscheinlichkeiten und im Vergleich dieser Grösse mit einer Schwelle  $L_0$ . Setzt man dazu die Ausdrücke (6) und (7) in die Gleichung (7) ein, so findet man für die Entscheidungsfunktion des optimalen Detektors

$$L(Y) = \exp \left[ \sum_{i=1}^N (2 y_i \lambda - \lambda^2) \right] \leq L_0 \quad (8)$$

Infolge der monotonen Abhängigkeit des Logarithmus von seinem Argument gilt aber auch

$$\ln[L(Y)] = 2\lambda \sum_{i=1}^N [y_i] - N\lambda^2 \geq \ln(L_0) \quad (9)$$

Verrechnet man konstante Summanden und Faktoren mit der Schwelle,  $L_0$ , so findet man schliesslich die allereinfachste Form des optimalen parametrischen Detektors, welcher, wegen seiner Funktion, auch «linearer Detektor» genannt wird:

$$z = \sum_{i=1}^N y_i \geq L_0' \quad (10)$$

Entsprechend dem Kriterium von *Neyman-Pearson* wird die Schwelle  $L_0'$  so bestimmt, dass die Fehlalarm-Wahrscheinlichkeit einen bestimmten Wert  $\alpha$  erreicht. Die Grösse  $z$  in Gl. (10) ist normalverteilt, weil sie eine Summe von normalverteilten Zufallsgrössen  $y_i$  ist. Ihr Erwartungswert ist  $E(z) = 0$ , sofern kein Signal vorhanden ist, und die Streuung  $\sigma_z^2$  ergibt sich als Summe der einzelnen Streuungen:

$$\sigma_z^2 = E(z^2) = E\left(\left[\sum_{i=1}^N y_i\right]^2\right) = N\sigma^2$$

Damit lautet die Dichtefunktion der Zufallsgrösse  $z$ , welche die entscheidende Grösse des Detektors ist, auf welcher seine Entscheidung beruht:

$$f_z(z) = \frac{s}{\sqrt{2\pi N\sigma^2}} \exp - \left\{ \frac{z^2}{2N\sigma^2} \right\} \quad (11)$$



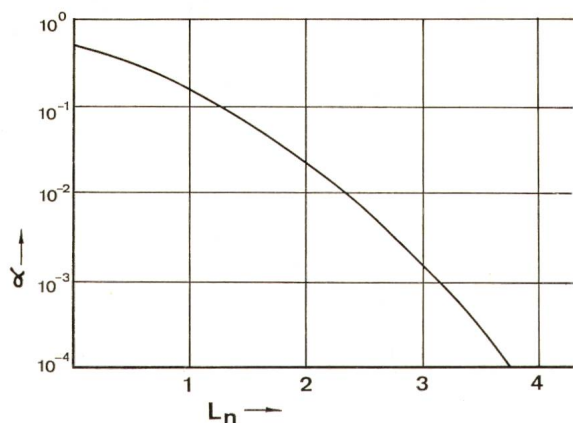


Fig. 3

Der Verlauf der normierten Schwelle  $L_n = L_0 / \sigma \sqrt{N}$  in Abhängigkeit der zulässigen Fehlalarm-Wahrscheinlichkeit  $\alpha$

Mit Hilfe von Tabellen der normierten kumulativen Gaussverteilung (z.B. [4]) findet man den in Fig. 3 gezeigten Verlauf der Schwelle  $L_0'$  in Abhängigkeit der zulässigen Fehlalarm-Wahrscheinlichkeit. Von speziellem Interesse ist, dass die Schwelle  $L_0'$  unabhängig von der zu erwartenden Signalamplitude  $\lambda$  ist, was bei allen nach dem Kriterium von *Neyman-Pearson* optimierten Detektoren zutrifft. Dagegen hängt natürlich die Erkennungswahrscheinlichkeit  $\beta$  sehr wohl von der Signalamplitude ab, wie es in Fig. 4 für verschiedene Verhältnisse von  $(N\lambda^2/\sigma^2)$  (Signal/Rauschverhältnis) dargestellt ist.

Im allgemeinen Fall eines zeitabhängigen Signales  $S = s_1(t)$  mit dem Momentanwert  $s_1(t_i)$  nimmt der lineare Detektor die Form

$$z = \sum_{i=1}^N s_1(t_i) y_i \geq L_0''$$

an, welche auch als «Korrelationsdetektor» bekannt ist. Der lineare Detektor entscheidet also in jedem Falle auf Grund einer Linearkombination der Stichproben. Beim inkohärenten Übertragungssystem mit unbekannter Trägerphase  $\varphi$  des Signales

$$s(t) = A(t) \sin(\omega_0 t + \varphi)$$

beruht der Entscheid auf einer Grösse  $e$ , der Umhüllenden der beobachteten Signalfunktion  $y(t)$  mit:

$$e = \sqrt{e_c^2 + e_s^2}$$

wobei

$$e_c = \sum_{i=1}^N y_i A(t_i) \cos(\omega_0 t_i)$$

$$e_s = \sum_{i=1}^N y_i A(t_i) \sin(\omega_0 t_i)$$

Dieser inkohärente Detektor heisst Enveloppendetektor. Seine Entscheidungssicherheit ist bei gleichem Signal/Rausch-Verhältnis natürlich kleiner als ein kohärenter Detektor nach Gl. (10), weil die unbekannte Trägerphase  $\varphi$  zusätzliche Unsicherheit für die Entscheidung mit sich bringt. Die Erkennung eines inkohärenten Signales in Rauschen entspricht im statistischen Sprachgebrauch einem Entscheid zwischen einer einfachen Hypothese («kein Signal da») und einer mehrfachen Alternativen («Signal  $S$  vorhanden», Phase  $\varphi$  unbekannt). Lässt man mehrere mögliche Signale zu, von denen der Sender in jedem Takt je eines übermitteln kann, so hat der Empfänger

einen Entscheid zwischen einer Hypothese («kein Signal vorhanden») und mehreren (möglicherweise wieder mehrfachen) Alternativen («Signal  $S_1$  vorhanden, Phase  $\varphi_1$  unbekannt») oder («Signal  $S_2$  vorhanden», Phase  $\varphi_2$  unbekannt usw.) zu entscheiden. Als Beispiel sei ein Übertragungssystem im Mehrfrequenzcode erwähnt, bei welchem in jedem Takt einer von  $m$  möglichen Sinustönen mit zufälligen Phasen übermittelt wird. Der Empfänger entscheidet sich nun für den wahrscheinlichsten Kanal [2]. Die dann optimalen Entscheidungsfunktionen sind aber im allgemeinen nur noch angenähert realisierbar.

### 3. Nichtparametrische Signalerkennung

Das in Abschnitt 2 beschriebene Erkennungsproblem bezog sich auf ein zeitlich konstantes Signal  $\lambda$  in additivem Rauschen. Die entsprechende nichtparametrische Erkennungsaufgabe geht nun aber im Gegensatz dazu nicht von einer exakt bekannten Dichtefunktion  $f_n(n_i)$  der Rauschproben  $n_i$  aus, sondern setzt wesentlich allgemeinere Eigenschaften voraus. In dieser Arbeit soll angenommen werden, dass die Dichtefunktion  $f_n$  eine gerade (symmetrische) Funktion sei:

$$f_n(n_i) = f_n(-n_i)$$

Der Meridian  $x_m$  einer Verteilung  $F_x$  ist definiert durch:

$$\int_{-\infty}^{x_m} f_x(x) dx = F_x(x_m) = 0,5 \quad (12)$$

Er fällt bei einer geraden Dichtefunktion mit dem Nullpunkt zusammen ( $x_m = 0$ ). Ein positives, additiv überlagertes Signal ( $\lambda > 0$ ) verschiebt nun den Meridian der Dichtefunktion  $f_y(y_i)$  der Proben  $y_i$  zu positiven Werten hin ( $x_m > 0$ ). Bezeichnet man mit  $p$  die Wahrscheinlichkeit, dass die Probe  $y_i$  ein positives Vorzeichen aufweist, ferner mit  $f_y^+(y_i)$  die Dichtefunktion der Probe  $y_i$  unter der Voraussetzung, dass sie positiv ist [ $f_y^+(y_i) = f_y(y_i | y_i \geq 0)$ ], sowie mit  $f_y^-(y_i)$  die entsprechende Dichte für den Fall, dass  $y_i < 0$  ist [ $f_y^-(y_i) = f_y(y_i | y_i < 0)$ ], so findet man mit Hilfe des Satzes über die bedingte Wahrscheinlichkeit (z.B. [5]) für die Dichtefunktion  $f_y(y_i)$  der Probe  $y_i$  den Zusammenhang:

$$f_y(y_i) = p f_y^+(y_i) + (1 - p) f_y^-(y_i) \quad (13)$$

Bei fehlendem Signal ( $\lambda = 0$ ) wird infolge der Koinzidenz des Meridianes mit dem Nullpunkt die Wahrscheinlichkeit  $p$

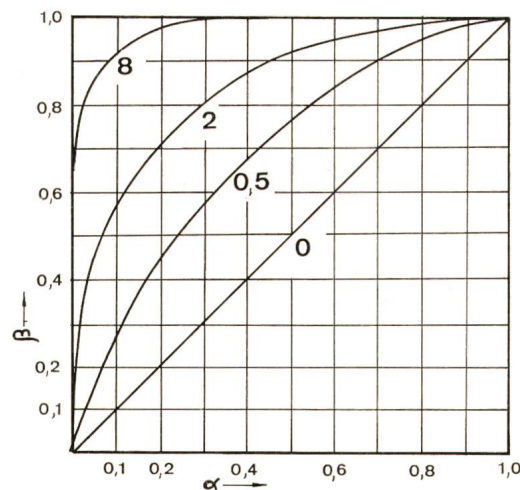


Fig. 4

Fehlalarm- und Erkennungswahrscheinlichkeiten  $\alpha, \beta$  des optimalen linearen Detektors mit dem Signal/Rauschverhältnis als Parameter



für ein  $y_i \geq 0$  gerade 0,5; die Dichtefunktion der Probe  $y_i$  lautet daher unter diesen Voraussetzungen:

$$f_y(y_i) = \frac{1}{2} [f_{y^+}(y_i) + f_{y^-}(y_i)] \quad (14)$$

Wie bei den parametrischen Detektoren führt hier die Anwendung des Optimierungskriteriums von *Neyman-Pearson* auf einen Quotienten bedingter Wahrscheinlichkeiten. Setzt man auch wieder statistische Unabhängigkeit der einzelnen Proben  $y_i$  voraus, so findet man für das Wahrscheinlichkeits-Verhältnis  $L(Y)$ :

$$L(Y) = \prod_{i=1}^N \frac{p f_{y^+}(y_i) + (1-p) f_{y^-}(y_i)}{\frac{1}{2} [f_{y^+}(y_i) + f_{y^-}(y_i)]} \quad (15)$$

Die einzelnen Multiplikatoren weisen dabei den Wert

$$\frac{p f_{y^+}(y_i)}{\frac{1}{2} [f_{y^+}(y_i) + f_{y^-}(y_i)]} = 2p \quad (16)$$

falls  $y_i \geq 0$  und

$$\frac{(1-p) f_{y^-}(y_i)}{\frac{1}{2} [f_{y^+}(y_i) + f_{y^-}(y_i)]} = 2(1-p) \quad (17)$$

auf, falls  $y_i < 0$ . Das Wahrscheinlichkeitsverhältnis in Gl. (15) nimmt daher die Form an:

$$L(Y) = 2^N p^{\sum_{i=1}^N u(y_i)} (1-p)^{N - \sum_{i=1}^N u(y_i)} \quad (18)$$

worin  $u(x)$  die Einheits-Schrittfunktion ist:

$$u(x) = 1 \text{ falls } x \geq 0 \\ u(x) = 0 \text{ falls } x < 0$$

Da auch hier der Ausdruck  $L(Y)$  eine monotone Funktion des Argumentes  $\sum_{i=1}^N u(y_i)$  ist, kann man ebensogut das Argument selbst mit einer geeigneten Schwelle  $L'''$  vergleichen. Damit lautet die optimale nichtparametrische Detektorfunktion:

$$k = L(Y) = \sum_{i=1}^N u(y_i) \geq L''' \quad (19)$$

Diesen Detektor nennt man «Vorzeichen-Detektor», weil er von den einzelnen Proben  $y_i$  nur noch die Vorzeichen berücksichtigt. Dementsprechend ist er einfacher zu realisieren: man braucht nur noch die Stichproben mit positivem Vorzeichen zu zählen und mit einer Schwelle zu vergleichen. Diese wird auch hier nach dem Kriterium von *Neyman-Pearson* durch Angabe einer bestimmten Fehlalarmwahrscheinlichkeit bestimmt.

Die Grösse  $k$  in Gl. (19) ist binomialverteilt, sie stellt die Anzahl der Proben mit positivem Vorzeichen dar. Unter der Voraussetzung, dass kein Signal vorhanden ist, lautet die Dichtefunktion von  $k$ , weil ja mit gleicher Wahrscheinlichkeit positive und negative Vorzeichen erscheinen:

$$f_k(k) = \binom{N}{K} \frac{1}{2^N} \quad 0 \leq k \leq N \quad (20)$$

während im Falle eines tatsächlich vorhandenen Signales die Wahrscheinlichkeit  $p$  für eine positive Probe grösser als 0.5 ist;

$$f_k(k) = \binom{N}{K} p^k (1-p)^{N-k} \quad 0 \leq k \leq N \quad (21)$$

Die Schwelle  $L'''$  ergibt sich unter Berücksichtigung von Gl. (20) und (19) aus:

$$\frac{1}{2^N} \sum_{k=L'''}^N \binom{N}{k} \geq \alpha \quad (21)$$

Die tatsächliche Fehlalarm-Wahrscheinlichkeit weist (infolge der hier diskreten Schwelle) einen Wert auf, der gerade grösser oder gleich dem vorgegebenen  $\alpha$  ist. Entsprechend findet man für die Erkennungswahrscheinlichkeit:

$$\beta = \sum_{k=L'''}^N \binom{N}{k} p^k (1-p)^{N-k} \quad (22)$$

Im folgenden sei die Leistungsfähigkeit des Vorzeichen-Detektors im Vergleich zum linearen Detektor nach Gl. (10) untersucht. Es kann gezeigt werden, dass für den asymptotischen Wirkungsgrad  $d_\infty$  angenähert gilt ([1]):

$$d_\infty \approx 4 \sigma^2 f_n^2(0) \quad (23)$$

Hierzu wurden die Voraussetzungen für die Ableitung des asymptotischen Wirkungsgrades von Abschnitt 1.2 berücksichtigt. Wenn nun die Rauschproben normal verteilt sind, so wird

$$f_n(0) = 1/\sqrt{2\pi\sigma^2}$$

und der asymptotische Wirkungsgrad für sehr grosse Stichprobenzahlen erreicht den Wert

$$d_{\infty \text{ gauss}} = \frac{2}{\pi} \approx 64\% \quad (24)$$

Falls aber die Rauschproben exponentiell verteilt sind nach

$$f_n(n_i) = \frac{\alpha}{2} \exp -\alpha |n_i| \quad (25)$$

so strebt der asymptotische Wirkungsgrad gegen:

$$d_{\infty \text{ exp}} = 200\%$$

Es lässt sich eine ganze Klasse von Dichtefunktionen  $f_n$  definieren, welche beim Vorzeichen-Detektor zu Wirkungsgraden über 100% führen [1].

Für die Anwendung interessant ist insbesondere die Frage wie weit die Beurteilung der Wirksamkeit durch den asymptotischen Wirkungsgrad schon bei endlichen Stichprobenzahlen gültig ist, und wie sich der Wirkungsgrad bei nicht verschwindender Signalamplitude  $\lambda$  verhält. Zu diesem Zweck wurde der Wirkungsgrad des Vorzeichendetektors bei gaußschem Rauschen für die Anzahl  $N$  der Stichproben im Bereich  $1 \leq N \leq 40$  numerisch berechnet (Fig. 5 und 6). Es zeigt sich, dass der asymptotische Wirkungsgrad auch für die Verhältnisse mit endlichen Stichprobenzahlen und nicht verschwindender Signalamplitude eine brauchbare Näherung darstellt. Immerhin ist dabei eine leichte Abhängigkeit des Wirkungsgrades von der zulässigen Fehlalarm-Wahrscheinlichkeit festzustellen, eine Abhängigkeit, welche für  $N \rightarrow \infty$  verschwindet. Die Oszillation der Kurven kommt durch die Tatsache zustande, dass infolge der diskreten möglichen Schwellen des Vorzeichendetektors die tatsächlichen Fehlalarm-Wahrscheinlichkeiten nicht immer gleich sind, wodurch eben auch der Wirkungsgrad schwankt.

Während die Entscheidungssicherheit des linearen Detektors bei gaußschem Rauschen von keiner anderen Funktion



übertroffen werden kann, weil er für diese spezielle Statistik der Störampplituden optimiert wurde, kann der Wirkungsgrad des Vorzeichendetektors bei anderen Dichtefunktionen der Rauschamplituden beträchtlich über 100 % liegen, nämlich z.B. 200 % beim exponentiellen Rauschen nach Gl. (25). Immerhin sind auch Fälle möglich, bei denen der asymptotische Wirkungsgrad wesentlich geringer ist; dies tritt immer dann ein, wenn die Dichtefunktion mit zunehmender Amplitude rasch genug abnimmt [1].

Nun gibt es aber erstaunlicherweise auch nichtparametrische Detektoren, welche einen bestimmten minimalen Wirkungsgrad garantieren. Unter ihnen ist der Wilcoxon-Detektor zu nennen, dessen Wirkungsgrad immer über 86 % liegt [6]. Der Wilcoxon-Detektor verwendet von den Stichproben mehr Information als nur ihr Vorzeichen, nämlich zusätzlich noch ein Mass für ihren Abstand vom Nullpunkt. Ordnet man die  $N$  Stichproben

$$y_1, y_2, \dots, y_N$$

in der Reihenfolge ihrer Beträge

$$|y_{g1}| < |y_{g2}| \dots |y_{gN}|$$

so ist die Entscheidungsfunktion des Wilcoxon-Detektors die Summe der Ordnungen:

$$L(Y) = \sum_{i=1}^N d_i \geq L_0''' \quad (26)$$

definiert, wobei

$$d_i = \begin{cases} 0 & \text{falls } y_{gi} \leq 0 \\ i & \text{falls } y_{gi} > 0 \end{cases}$$

In [1] wird gezeigt, dass der asymptotische Wirkungsgrad des Wilcoxon-Detektors gegenüber dem linearen Detektor bei beliebiger symmetrischer Dichtefunktion  $f_n(n_i)$  der Rauschproben  $n_i$  gegeben ist durch:

$$d_w = 12 \sigma^2 \left[ \int_{-\infty}^{\infty} f_n^2(n_i) dn \right]^2 \quad (27)$$

Bei gaußischem Rauschen wird der Wirkungsgrad des Wilcoxon-Detektors

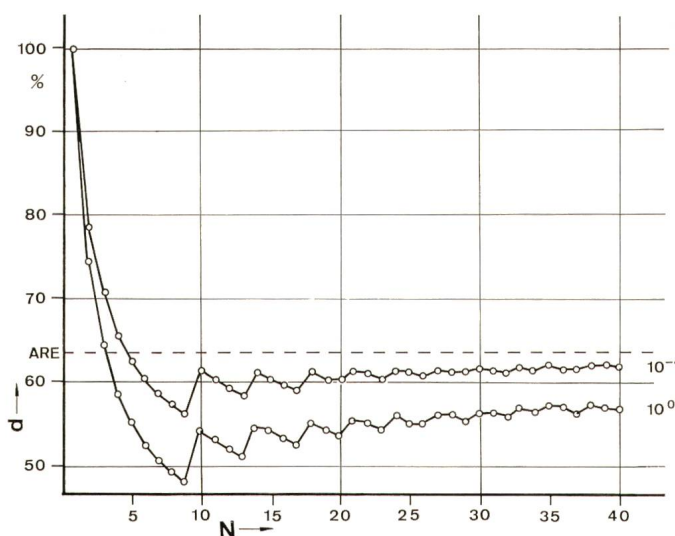


Fig. 5

Der Wirkungsgrad  $d$  des Vorzeichendetektors bei gaußischem Rauschen und endlichen Stichprobenzahlen  $N$  bei einer Fehlalarmwahrscheinlichkeit von  $\alpha = 0,001$

Signal/Rausch-Verhältnis  $\lambda^2/\sigma^2 = 10^{-4}$  bzw.  $10^0$

ARE: asymptotischer Wirkungsgrad

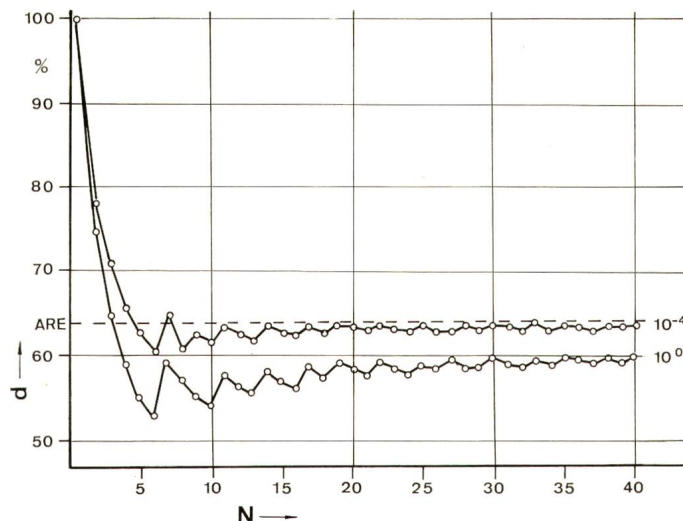


Fig. 6

Wie Fig. 5, aber mit der Fehlalarmwahrscheinlichkeit  $\alpha = 0,01$   
ARE: asymptotischer Wirkungsgrad

$$d_{w \text{ gauss}} = \frac{3}{\pi} \approx 95 \% \quad (28)$$

bei exponentieller Verteilung der Rauschproben wird dagegen

$$d_{w \text{ exp}} = \frac{3}{2} = 150 \% \quad (29)$$

In [7] wird gezeigt, dass der Wirkungsgrad des Wilcoxon-Detektors niemals unter 86 % sein kann, was die Dichtefunktion  $f_n(n_i)$  der Störampplituden auch immer für eine Form aufweise. Selbstverständlich erkaufte man diese Sicherheit mit erhöhtem Aufwand bei der Realisierung: während der Vorzeichendetektor mit einem Schwellwertschalter und einem Zähler auskommt, erfordert der Wilcoxon-Detektor unter anderem einen Speicher für alle  $N$ -Stichproben.

Das Repertoire nichtparametrischer Entscheidungsfunktionen ist mit den beiden erwähnten Beispielen durchaus nicht erschöpft. Sie stellen im Gegenteil die beiden einfachsten Fälle einer ausserordentlich grossen Zahl nichtparametrischer Entscheidungsregeln dar, [7; 8]. Dazu kommen neben den besprochenen Detektoren mit einem einzigen Signaleingang solche mit mehreren Eingängen, z.B. Korrelatoren (Polaritäts-Koinzidenz-Korrelatoren), ferner nichtparametrische Detektoren für zeitlich veränderliche Signale, usw.

## Literatur

- [1] J. B. Thomas: Nonparametric detection. Proc. IEEE 58(1970)5, p. 623...631.
- [2] C. L. Weber: Elements of detection and signal design. New York, McGraw-Hill, 1968.
- [3] D. Middleton: An introduction to statistical communication theory. New York, McGraw-Hill, 1960.
- [4] M. Abramowitz and I. A. Stegun: Handbook of mathematical function with formulas, graphs and mathematical tables. National Bureau of Standards, Applied Mathematics Series No. 55; Washington, US Department of Commerce, Government Printing Office, 1965.
- [5] A. Papoulis: Probability, random variables and stochastic processes. New York, McGraw-Hill, 1965.
- [6] J. L. Hodges and E. L. Lehmann: The efficiency of some nonparametric competitors to the t-test. Annals of Mathematics Statistics 27(1956), p. 324...335.
- [7] J. V. Bradley: Distribution-free statistical tests. Englewood/Cliffs, N. J. Prentice-Hall, 1968.
- [8] J. E. Walsh: Handbook of nonparametric statistics. Vol. 1/3, Princeton, N. J., Van Nostrand, 1962/1968.
- [9] A. V. Balakrishnan: Communication Theory. New York, McGraw-Hill, 1968.

## Adresse des Autors:

Dr. H. Schlaepfer, Institut für Fernmeldetechnik ETHZ, Sternwartstrasse 7, 8006 Zürich.