ACTA TECHNICA

ACADEMIAE SCIENTIARUM HUNGARICAE

REDIGIT: M. MAJOR

TOMUS 90
FASCICULI 1–2

AKADÉMIAI KIADÓ, BUDAPEST 1980

ACTA TECHNICA

Az *Acta Technica* angol, francia, német és orosz nyelven közöl értekezéseket a műszaki tudományok köréből.

Az *Acta Technica* változó terjedelmű füzetekben jelenik meg, több füzet alkot egy kötetet.

A közlésre szánt kéziratok a következő címre küldendők:

Acta Technica
1051 Budapest, Münnich Ferenc u. 7.

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Megrendelhető a belföld számára az „Akadémiai Kiadó"-nál (1363 Budapest Pf. 24. Bankszámla 215-11448), a külföld számára pedig a „Kultura" Külkereskedelmi Vállalatnál (1389 Budapest 62, P.O.B. 149 Bankszámla: 218-10990) vagy annak külföldi képviseleteinél és bizományosainál.

Die *Acta Technica* veröffentlichen Abhandlungen aus dem Bereiche der technischen Wissenschaften in deutscher, englischer, französischer und russischer Sprache.

Die *Acta Technica* erscheinen in Heften wechselnden Umfanges. Vier Hefte bilden einen Band.

Die zur Veröffentlichung bestimmten Manuskripte sind an folgende Adresse zu senden:

Acta Technica
H-1051 Budapest
Münnich Ferenc u. 7.
Ungarn

An die gleiche Anschrift ist auch jede für die Schriftleitung und den Verlag bestimmte Korrespondenz zu richten.

Bestellbar bei »Kultura« Außenhandelsunternehmen (H-1389 Budapest 62, P.O.B. 149, Bankkonto Nr. 218-10990) oder seinen Auslandsvertretungen.

ACTA TECHNICA

ACADEMIAE SCIENTIARUM HUNGARICAE

REDIGIT: M. MAJOR

TOMUS 90

AKADÉMIAI KIADÓ, BUDAPEST 1980

ACTA TECHNICA

TOMUS 90, FASC. 1–4

INDEX

BOOK REVIEW

BUCHBESPRECHUNG

ACTA TECHNICA

ACADEMIAE SCIENTIARUM HUNGARICAE

REDIGIT: M. MAJOR

TOMUS 90

AKADÉMIAI KIADÓ, BUDAPEST 1980

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (1—2), pp. 3—14 (1980)

SÄTZE ÜBER DAS MINIMUM DER FORMÄNDERUNGSENERGIE IN DER ELASTOSTATIK

J. BARTA*

DOKTOR DER TECHNISCHEN WISSENSCHAFTEN

Eingegangen am 17. Juli 1979

Von den aktiven äußeren Kräften, für die gewisse Bedingungen vorgeschrieben sind, soll jene Gruppe herausgewählt werden, welche die kleinste Formänderungsenergie hervorruft. Beiträge zur Lösung dieser Aufgabe sind in den in der vorliegenden Abhandlung angeführten Sätzen enthalten. Ein numerisches Beispiel erläutert die Einzelheiten der Berechnung.

Der Zweck dieser Abhandlung ist die Sätze II—IV und VI—IX zu beweisen. Satz I folgt unmittelbar aus den bekannten Eigenschaften des Verschiebungsellipsoides [1] und wird hier nicht bewiesen. Satz II ist eine Verallgemeinerung des Satzes I. Der Satz V stammt von I. Ecsedi [2] und wird hier als Beweis des Satzes VI benützt. Satz VI dient zum Beweis des Satzes VII. Eine Auslegung der Sätze VIII und IX auf Grund der bekannten Eigenschaften der potentiellen Energie und Ergänzungsenergie ist wohl möglich, liegt aber außerhalb dieser Abhandlung.

Der Körper, auf den sich diese Abhandlung bezieht, ist ein Stab, ein Fachwerk, ein Rahmen, eine Platte etc., oder ein Kontinuum von einfach oder mehrfach zusammenhängendem endlichem Bereich (im Satz III ist der Körper ein Kontinuum). Übliche Voraussetzungen der Elastostatik werden benützt. Es werden also vorausgesetzt:

daß reibungsfreie oder starre Verbindungen und Stützen die Bewegung des Körpers verhindern, (im Satz III ist der Körper frei, d. h. er hat keine Verbindungen oder Stützen),
daß der Körper und seine Stützung statisch bestimmt oder statisch unbestimmt und frei von Anfangsspannungen ist,
daß nur statische Effekte hervorgerufen werden, also keine kinetische oder thermische Effekte entstehen,
daß die Formänderungen klein sind, daß das Elastizitätsgesetz linear ist, daß das Superpositionsgesetz gültig ist.

Das Material des Körpers ist homogen oder heterogen, und isotrop oder anisotrop (im Satz III ist es homogen isotrop).

* Prof. Dr. J. Barta, József körút 35, H-1085 Budapest, Ungarn

S a t z I: Sind für die aktive äußere Kraft ihre Größe F und ihr Angriffspunkt P vorgeschrieben, aber ihre Richtung nicht vorgeschrieben, so ist die aktive äußere Kraft, für die die Formänderungsenergie ein Minimum wird, die Kraft, deren Vektor $\mathbf{F}$ parallel zum Verschiebungsvektor $\mathbf{u}$ des Angriffspunktes P ist.

S a t z II: Weisen die aktiven äußeren Kräfte die Größen $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_k$ und die Angriffspunkte $P_1, P_2, \ldots P_k$ auf und sind ihre Richtungen nicht vorgeschrieben, so sind die aktiven äußeren Kräfte, für die die Formänderungsenergie ein Minimum wird, die Kräfte, deren Vektoren $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_k$ parallel zu den Verschiebungsvektoren $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ sind.

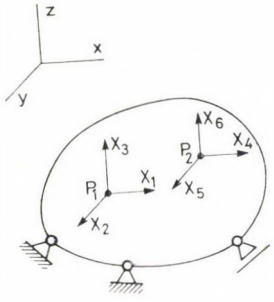


Bild 1

Beweis des Satzes II. Wir führen diesen Beweis für den Fall $k = 2$ an. (Bei anderen Werten von k kann man einen analogen Gedankengang verfolgen.) P_1 ist der Angriffspunkt von $\mathbf{F}_1$ und P_2 ist der Angriffspunkt von $\mathbf{F}_2$, (Bild 1). Ein rechtwinkliges Achsenkreuz xyz sei im Raume festgelegt. Die in die Richtungen x, y, z entfallenden Komponenten von $\mathbf{F}_1$ seien mit X_1, X_2, X_3, diese von $\mathbf{F}_2$ mit X_4, X_5, X_6, diese von $\mathbf{u}_1$ mit $\zeta_1, \zeta_2, \zeta_3$, diese von $\mathbf{u}_2$ mit $\zeta_4, \zeta_5, \zeta_6$ bezeichnet. Für die Formänderungsenergie des elastischen Körpers gilt die bekannte Formel

$$\begin{aligned} U = \frac{1}{2}\,[&X_1(a_{11}X_1 + a_{12}X_2 + \ldots + a_{16}X_6) \\ &+ X_2(a_{21}X_1 + a_{22}X_2 + \ldots + a_{26}X_6) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &+ X_6(a_{61}X_1 + a_{62}X_2 + \ldots + a_{66}X_6)]\,, \end{aligned} \tag{1}$$

wobei die Konstanten $a_{11}, a_{12}, \ldots a_{66}$ Einflußzahlen genannt werden. Gemäß der Prämisse des Satzes II bestehen die Gleichungen

$$\begin{aligned} \varphi_1 &\equiv X_1^2 + X_2^2 + X_3^2 = F_1^2\,, \\ \varphi_2 &\equiv X_4^2 + X_5^2 + X_6^2 = F_2^2\,. \end{aligned} \tag{2}$$

Gesucht sind also die Werte der sechs Veränderlichen

$$X_1, X_2, X_3, X_4, X_5, X_6 ,$$

welche die Funktion (1) auf ein Minimum reduzieren und zugleich die Nebenbedingungen (2) erfüllen. Nach der bekannten Regel der Berechnung des Minimums benützt man die Funktion $\Phi = U + \lambda_1\varphi_1 + \lambda_2\varphi_2$, wobei λ_1 und λ_2 konstante Multiplikatoren sind. Aus den acht Gleichungen

$$\frac{\partial\Phi}{\partial X_1} = 0 , \quad \frac{\partial\Phi}{\partial X_2} = 0 , \quad \frac{\partial\Phi}{\partial X_3} = 0 , \quad \frac{\partial\Phi}{\partial X_4} = 0 ,$$
$$\frac{\partial\Phi}{\partial X_5} = 0 , \quad \frac{\partial\Phi}{\partial X_6} = 0 , \quad \varphi_1 - F_1^2 = 0 , \quad \varphi_2 - F_2^2 = 0 \qquad (3)$$

kann man dann die acht Unbekannten

$$X_1, X_2, X_3, X_4 , X_5 , X_6, \lambda_1, \lambda_2$$

berechnen. Die ersten sechs Gleichungen (3) lauten

$$\begin{aligned}
a_{11}X_1 + \ldots + a_{16}X_6 + 2\lambda_1 X_1 &= 0,\\
a_{21}X_1 + \ldots + a_{26}X_6 + 2\lambda_1 X_2 &= 0,\\
a_{31}X_1 + \ldots + a_{36}X_6 + 2\lambda_1 X_3 &= 0,\\
a_{41}X_1 + \ldots + a_{46}X_6 + 2\lambda_2 X_4 &= 0,\\
a_{51}X_1 + \ldots + a_{56}X_6 + 2\lambda_2 X_5 &= 0,\\
a_{61}X_1 + \ldots + a_{66}X_6 + 2\lambda_2 X_6 &= 0. \qquad (4
\end{aligned}$$

Die Gleichungen

$$\begin{aligned}
a_{11}X_1 + \ldots + a_{16}X_6 &= \zeta_1,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
a_{61}X_1 + \ldots + a_{66}X_6 &= \zeta_6
\end{aligned}$$

drücken die fundamentale Eigenschaft der Einflußzahlen aus. Mithin lassen sich die Gleichungen (4) in der Form

$$\begin{aligned}
\zeta_1 + 2\zeta_1 X_1 &= 0,\\
\ldots\ldots\ldots\ldots\\
\zeta_6 + 2\lambda_2 X_6 &= 0
\end{aligned}$$

aufschreiben, woraus die Proportionalitäten

$$\frac{X_1}{\zeta_1} = \frac{X_2}{\zeta_2} = \frac{X_3}{\zeta_3}, \quad \frac{X_4}{\zeta_4} = \frac{X_5}{\zeta_5} = \frac{X_6}{\zeta_6} \tag{5}$$

folgen. Aus (5) geht die Richtigkeit des Satzes II hervor.

S a t z III: Der elastische Körper sei ein Kontinuum von endlichem Bereich und homogenem isotropem Material. Dann besteht von allen Gruppen der aktiven äußeren Kräfte, deren jede dieselbe gegebene Volumänderung hervorruft, die, für welche die Formänderungsenergie ein Minimum wird, aus einem an der ganzen Oberfläche wirkenden gleichmäßig verteilten Zug oder Druck (hydrostatischer Druck).

Beweis des Satzes III. E, G, μ sind die Elastizitätskonstanten. Die Spannungskomponenten seien mit σ_x, σ_y, σ_z, τ_{yz}, τ_{zx}, τ_{xy}, das Volumen des elastischen Körpers mit V bezeichnet. Bekanntlich lautet die Formel für die spezifische Formänderungsenergie

$$D = \frac{1}{2E}\left[\sigma_x^2 + \sigma_y^2 + \sigma_z^2 - 2\mu(\sigma_y\sigma_z + \sigma_z\sigma_x + \sigma_x\sigma_y)\right] + \\ + \frac{1}{2G}(\tau_{yz}^2 + \tau_{zx}^2 + \tau_{xy}^2) \tag{6}$$

und die für die spezifische Volumänderung

$$\vartheta = \frac{1 - 2\mu}{E}(\sigma_x + \sigma_y + \sigma_z)\,. \tag{7}$$

Der Wert der Funktion (6) ist, wie bekannt, nicht negativ. $\sigma_x, \ldots, \tau_{xy}$ als Funktionen von x, y, z sollen die Bedingung

$$\int_V D\,dV = \text{Minimum} \tag{8}$$

und die Nebenbedingung

$$\int_V \vartheta\,dV = C \tag{9}$$

erfüllen, wobei C die gegebene Volumänderung, somit eine gegebene Konstante ist. Die Regel der Berechnung des Minimums führt jetzt zu den sechs Gleichungen

$$\frac{\partial(D + \lambda\vartheta)}{\partial\sigma_x} = 0\,, \quad \ldots\,, \frac{\partial(D + \lambda\vartheta)}{\partial\tau_{xy}} = 0\,,$$

worin λ der konstante Multiplikator ist. Mit (6) und (7) erhalten sie die Form

$$\sigma_x - \mu\sigma_y - \mu\sigma + \lambda(1 - 2\mu) = 0, \ \ldots, \ \tau_{xy} = 0 ,$$

woraus $\sigma_x = \sigma_y = \sigma_z = -\lambda$, das heißt

$$\sigma_x = \sigma_y = \sigma_z = \text{konstant} \tag{10}$$

und

$$\tau_{yz} = \tau_{zx} = \tau_{xy} = 0 \tag{11}$$

folgen. (10) und (11) zeigen, daß überall in dem elastischen Körper ein in jeder Richtung gleichmäßiger Zug oder Druck (hydrostatischer Druck) herrscht. Dies ist also auch an der Oberfläche der Fall.

S a t z IV: A und B sind zwei gegebene Punkte des elastischen Körpers. $F_1, F_2, \ldots, F_n$ sind die aktiven äußeren Kräfte. Ihre Angriffspunkte, Größen und Richtungen sind nicht vorgeschrieben. c ist die durch $F_1, F_2, \ldots, F_n$ hervorgerufene Abstandsänderung der gegebenen Punkte A und B. Dann besteht von allen Gruppen von $F_1, F_2, \ldots, F_n$, deren jede dieselbe Abstandsänderung

$$c$$

der gegebenen Punkte A und B hervorruft, eine, für welche die Formänderungsenergie ein Minimum wird, nur aus zwei Kräften F_A und F_B, deren Angriffspunkte die Punkte A und B sind. Ihre Beträge sind gleich, ihre Richtungen sind entgegengesetzt, ihre Wirkungslinie ist die Verbindungslinie der Punkte A und B, (Bild 2 oder Bild 3).

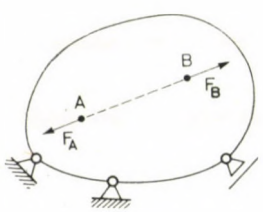


Bild 2

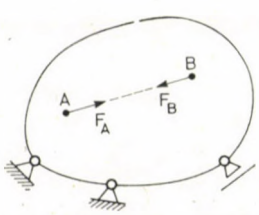


Bild 3

Beweis des Satzes IV. Stellen wir uns die Vektoren $\mathbf{F}_1, \mathbf{F}_2, \mathbf{F}_3, \ldots$ der aktiven äußeren Kräfte vor. Ein rechtwinkliges Achsenkreuz xyz sei im Raume so angenommen, daß der positive Ast der x-Achse vom Punkt A zum Punkt B führt. Dementsprechend bezeichnen wir die Komponenten von $\mathbf{F}_1, \mathbf{F}_2, \mathbf{F}_3, \ldots$ mit

$$X_1, X_2, X_3, \quad X_4, X_5, X_6, \quad X_7, X_8, X_9, \ldots$$

Zum Beispiel sei X_1 die Komponente von $\mathbf{F}_1$ in der Richtung x, und X_4 die Komponente von $\mathbf{F}_2$ in der Richtung x. Im weiteren verfahren wir so, wie beim Beweis des Satzes II. Dabei ist

$$U = \frac{1}{2}[X_1(a_{11}X_1 + a_{12}X_2 + \ldots) + \\ + X_2(a_{21}X_1 + a_{22}X_2 + \ldots) \\ \ldots\ldots\ldots\ldots\ldots\ldots]\,. \tag{12}$$

Beim Beweis dürfen wir annehmen, daß F_1 in A, F_2 in B angreift. So ist

$$\varphi \equiv -(a_{11}X_1 + a_{12}X_2 + \ldots) + a_{41}X_1 + a_{42}X_2 + \ldots = c, \tag{13}$$

wobei c die gegebene Längenänderung, also eine gegebene Konstante ist. Die Gleichungen

$$\frac{\partial(U + \lambda\varphi)}{\partial X_1} = 0\,, \quad \frac{\partial(U + \lambda\varphi)}{\partial X_2} = 0\,, \ldots$$

lauten

$$\begin{aligned} &a_{11}X_1 + a_{12}X_2 + \ldots + \lambda(-a_{11} + a_{41}) = 0\,, \\ &a_{21}X_1 + a_{22}X_2 + \ldots + \lambda(-a_{12} + a_{42}) = 0\,, \\ &a_{31}X_1 + a_{32}X_2 + \ldots + \lambda(-a_{13} + a_{43}) = 0\,, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned} \tag{14}$$

Aus den Gleichungen (13) und (14) berechnet man die Unbekannten λ, X_1, X_2, X_3, ... Man erhält die Werte $\lambda = -c/(a_{11} - 2a_{14} + a_{44})$ und

$$X_1 = -\,c/(a_{11} - 2a_{14} + a_{44})\,, \quad X_4 = c/(a_{11} - 2a_{14} + a_{44})\,, \tag{15}$$

$$X_2 = X_3 = X_5 = X_6 = X_7 = X_8 = X_9 = X_{10} = \ldots = 0\,. \tag{16}$$

Die Ergebnisse (15) und (16) bestätigen die Richtigkeit des Satzes IV.

Satz V
Satz VI
: $P_1, P_2, \ldots, P_k$ sind gegebene Punkte des elastischen Körpers. Diesen Punkten sind gegebene Richtungen $e_1, e_2, \ldots, e_k$ zugeordnet, (im Bild 4 ist $k = 3$). $F_1, F_2, \ldots, F_k, F_{k+1}, \ldots, F_n$ sind aktive äußere Kräfte, (im Bild

4 ist $n = 5$). $F_1, F_2, \ldots, F_k$ greifen in den gegebenen Punkten $P_1, P_2, \ldots, P_k$ an, ihre Größen sind nicht vorgeschrieben

und { ihre Richtungen sind die gegebenen Richtungen $e_1, e_2, \ldots, e_k$. / auch ihre Richtungen sind nicht vorgeschrieben, (Bild 4). }

Die Angriffspunkte, Größen und Richtungen von $F_{k+1}, \ldots, F_n$ sind nicht vorgeschrieben. $u_1, u_2, \ldots, u_k$ sind die durch $F_1, F_2, \ldots, F_k, F_{k+1}, \ldots, F_n$ hervorgerufenen, in die Richtungen $e_1, e_2, \ldots, e_k$ entfallenden Verschiebungskomponenten der Punkte $P_1, P_2, \ldots, P_k$. Dann besteht von allen Gruppen von $F_1, F_2, \ldots, F_k, F_{k+1}, \ldots, F_n$, deren jede dieselben gegebenen Verschiebungskomponenten

$$u_1, u_2, \ldots, u_k$$

hervorruft, eine, für welche die Formänderungsenergie ein Minimum wird, nur aus k Kräften, deren Angriffspunkte die Punkte $P_1, P_2, \ldots, P_k$ sind.

Beweis des Satzes VI. Wir führen den Beweis für den im Bild 4 veranschaulichten Fall

$$k = 3, \quad n = 5$$

an. (Bei anderen Werten von k und n kann man einen analogen Gedankengang verfolgen). Es handelt sich um die Kräfte $F_1, F_2, \ldots, F_5$ wie es das Bild 4 zeigt. Statt diesen fünf Kräften dürfen wir uns die acht Kräfte $F_1', F_2', F_3', F_4, F_5, F_6, F_7, F_8$ vorstellen wie in Bild 5 dargestellt: statt F_1 ihre Komponenten

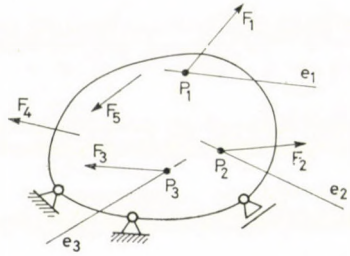


Bild 4

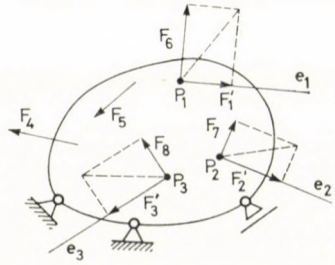


Bild 5

F_1' und F_6, statt F_2 ihre Komponenten F_3' und F_8, statt F_3 ihre Komponenten F_3' und F_8. Gemäß Bild 5 haben wir es also mit dem Fall

$$k = 3, \; n = 8$$

zu tun. Auf diesen Fall sei nun der Satz V angewendet. Aus dieser Betrachtung geht die Richtigkeit des Satzes VI hervor.

S a t z VII: $P_1, P_2, \ldots, P_g$ sind gegebene Punkte des elastischen Körpers. Für diese Punkte sind gegebene Verschiebungsvektoren $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_g$ vorgeschrieben. Die Anzahl, Angriffspunkte, Größen und Richtungen der aktiven äußeren Kräfte sind nicht vorgeschrieben. Dann besteht von allen Gruppen der aktiven äußeren Kräfte, deren jede dieselbe gegebene Gruppe von Verschiebungen

$$\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_g$$

hervorruft, eine, für welche die Formänderungsenergie ein Minimum wird, aus g Kräften, deren Angriffspunkte die Punkte $P_1, P_2, \ldots, P_g$ sind.

Beweis des Satzes VII. Um den Satz VI benützen zu können, sei ein rechtwinkliges Achsenkreuz xyz im Raum festgelegt. Die Komponenten der gegebenen Vektoren $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_g$ nach Richtungen x, y, z seien mit

$$\begin{gathered} u_1, u_2, u_3, \\ u_4, u_5, u_6, \\ \ldots\ldots\ldots \\ u_{k-2}, u_{k-1}, u_k \end{gathered} \tag{17}$$

bezeichnet, wobei $k = 3g$ ist. Wird der Satz VI auf den durch (17) gekennzeichneten Fall angewendet, so sieht man, daß jene Gruppe der aktiven äußeren Kräfte, für welche die Formänderungsenergie ein Minimum wird, aus g Kräften besteht, deren Angriffspunkte die Punkte $P_1, P_2, \ldots, P_g$ sind.

S a t z VIII: $P_1, P_2, \ldots, P_n$ sind gegebene Punkte des elastischen Körpers. Für diese Punkte sind gegebene Verschiebungsvektoren

$$\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$$

vorgeschrieben. $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n$ sind die Vektoren der aktiven äußeren Kräfte. $F_1, F_2, \ldots, F_n$ greifen in den gegebenen Punkten $P_1, P_2, \ldots, P_n$ an, ihre Größen und Richtungen sind nicht vorgeschrieben. C ist eine gegebene reelle Konstante. Dann, wird von allen Gruppen von $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n$, welche dieselbe gegebene virtuelle Arbeit C leisten, das heißt, die Bedingung

$$\mathbf{F}_1\mathbf{u}_1 + \mathbf{F}_2\mathbf{u}_2 + \ldots + \mathbf{F}_n\mathbf{u}_n = C \tag{18}$$

erfüllen, für eine, welche die gegebenen Verschiebungen $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ hervorruft, die Formänderungsenergie ein Minimum.

Beweis des Satzes VIII. Wir führen den Beweis für $n = 2$ an. (Bei anderen Werten von n kann man einen analogen Gedankengang verfolgen.) Bild 1, die Bezeichnungen $\mathbf{F}_1, \mathbf{F}_2, X_1, \ldots, X_6, \zeta_1, \ldots, \zeta_6$ und die Formel (1) werden jetzt ebenso benützt, wie sie schon beim Beweis des Satzes II benützt wurden. Beim Beweis nehmen wir an, daß die Gruppe $\mathbf{F}_1, \mathbf{F}_2$ die gegebenen Verschiebungsvektoren $\mathbf{u}_1, \mathbf{u}_2$ hervorruft. Eine andere Gruppe sei mit $\mathbf{F}_1', \mathbf{F}_2'$ bezeichnet. Es gilt jetzt also nicht nur die Formel (1), sondern auch die ähnlich aufgebaute Formel

$$U' = \frac{1}{2}[X_1'(a_{11}X_1' + \ldots + a_{16}X_6') \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad + X_6'(a_{61}X_1' + \ldots + a_{66}X_6')] \,. \tag{19}$$

Wird in (19)

$$X_1' = X_1 + Q_1, \ldots, X_6' = X_6 + Q_6$$

eingesetzt, so ist

$$\begin{aligned} U' = {} & \frac{1}{2}[X_1(a_{11}X_1 + \ldots + a_{16}X_6) + \ldots + X_6(a_{61}X_1 + \ldots + a_{66}X_6)] \\ & + \frac{1}{2}[Q_1(a_{11}Q_1 + \ldots + a_{16}Q_6) + \ldots + Q_6(a_{61}Q_1 + \ldots + a_{66}Q_6)] \\ & + Q_1a_{11}X_1 + \ldots + a_{16}X_6) + \ldots + Q_6(a_{61}X_1 + \ldots + a_{66}X_6) \,. \end{aligned} \tag{20}$$

Infolge von (18) dürfen wir beim Beweis die Gleichungen

$$(X_1 + Q_1)\zeta_1 + \ldots + (X_6 + Q_6)\zeta_6 = C,$$
$$X_1\zeta_1 + \ldots + X_6\zeta_6 = C$$

aufschreiben, woraus

$$Q_1\zeta_1 + \ldots + Q_6\zeta_6 = 0 \tag{21}$$

folgt. Fassen wir die rechte Seite der Gleichung (20) ins Auge. Die erste Reihe ist wie aus (1) ersichtlich, gleich U. Die zweite Reihe ist ein positiv-definiter Ausdruck. Die dritte Reihe ist zufolge der Gleichung (21) gleich Null. Daraus folgt die Beziehung $U' \geq U$. Damit ist der Satz VIII bewiesen. — Ein ähnlicher Beweis, dessen Mitteilung hier überflüssig zu sein scheint, liefert den

S a t z IX: $P_1, P_2, \ldots, P_n$ sind gegebene Punkte des elastischen Körpers. In diesen Punkten greifen aktive äußere Kräfte an, deren Vektoren

$$\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n$$

gegeben sind. $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ sind die Verschiebungsvektoren der Punkte $P_1, P_2, \ldots, P_n$, ihre Größen und Richtungen sind nicht vorgeschrieben. K ist eine gegebene reelle Konstante. Dann wird von allen Gruppen $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$, die alle dieselbe gegebene virtuelle Arbeit K liefern, das heißt, die Bedingung

$$\mathbf{F}_1\mathbf{u}_1 + \mathbf{F}_2\mathbf{u}_2 + \ldots + \mathbf{F}_n\mathbf{u}_n = K \tag{22}$$

erfüllen, für eine, welche von den gegebenen aktiven äußeren Kräften $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n$ hervorgerufen ist, die Formänderungsenergie ein Minimum.

Beispiel für die Anwendung des Satzes VII. Der Körper $ABCDE$ ist ein ebenes Gebilde, besteht aus vier gewichtslosen starren Stäben und aus den elastischen Gelenken B, C, D, und ist durch das elastische Gelenk E mit dem festen Untergrund verbunden (Bild 6). Jedes elastische Gelenk hat die Steifigkeit 132 000. Der Punkt T ist die Mitte des Stabes DC.

Was für aktive äußere Kräfte müssen an diesen Körper angebracht werden, wenn der Punkt A die horizontale Verschiebungskomponente 0,0001 nach rechts und die vertikale Verschiebungskomponente 0,0002 abwärts, der Punkt T die reine vertikale Verschiebung 0,00003 abwärts erfahren soll und die Formänderungsenergie des Körpers ein Minimum sein soll. Diese Frage beantworten wir wie folgt.

Die Anzahl und Angriffspunkte der gesuchten aktiven äußeren Kräfte kennen wir im voraus nicht. Der Satz VII besagt aber, daß ihre Anzahl zwei, ihre Angriffspunkte die Punkte A und T sein sollen. Wir haben also nur die Komponenten X_A, Y_A, X_T, Y_T zu ermitteln. Ein Achsenkreuz xy sei nach Bild 6 angenommen. Die Drehmomente sind

$$\begin{aligned} M_B &= \phantom{1X_A + {}} 1Y_A \\ M_C &= 1X_A + 1Y_A \\ M_D &= 1X_A + 3Y_A \phantom{{}+ 1X_T} + 1Y_T \\ M_E &= 2X_A + 3Y_A + 1X_T + 1Y_T \end{aligned}$$

Die Drehwinkel sind

$$\begin{aligned} \vartheta_B &= M_B/132\,000, \\ \vartheta_C &= M_C/132\,000, \\ \vartheta_D &= M_D/132\,000, \\ \vartheta_E &= M_E/132\,000\,. \end{aligned}$$

Die Verschiebungskomponenten sind

$$\begin{aligned} \xi_A &= \phantom{1\vartheta_B + {}} 1\vartheta_C + 1\vartheta_D + 2\vartheta_E \\ \eta_A &= 1\vartheta_B + 1\vartheta_C + 3\vartheta_D + 3\vartheta_E \\ \xi_T &= \phantom{1\vartheta_B + 1\vartheta_C + 3\vartheta_D {}} 1\vartheta_E \\ \eta_T &= + 1\vartheta_D + 1\vartheta_E \end{aligned}$$

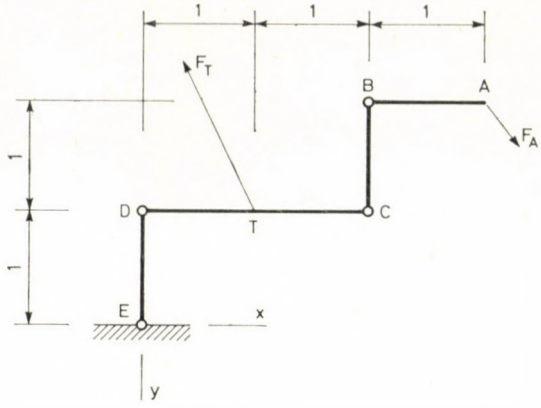


Bild 6

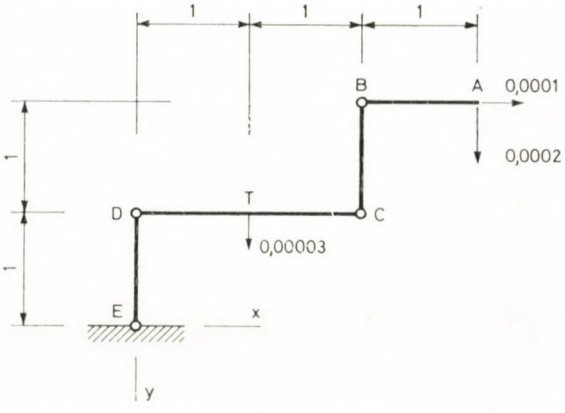


Bild 7

das heißt

$$\xi_A = (\ 6X_A + 10Y_A + 2X_T + 5Y_T)/132\,000\ ,$$
$$\eta_A = (10X_A + 20Y_A + 3X_T + 6Y_T)/132\,000\ ,$$
$$\xi_T = (\ 2X_A + 3Y_A + X_T + Y_T)/132\,000\ ,$$
$$\eta_T = (\ 3X_A + 6Y_A + X_T + 2Y_T)/132\,000\ .$$

Zur Berechnung von X_A, Y_A, X_T, Y_T dienen also die Gleichungen

$$\frac{1}{132000}(\ 6X_A + 10Y_A + 2X_T + 3Y_T) = 0{,}0001\ ,$$

$$\frac{1}{132000}(10X_A + 20Y_A + 3X_T + 6Y_T) = 0{,}0002\ ,$$

$$\frac{1}{132000}(\,2X_A + 3Y_A + X_T + Y_T) = 0\,,$$

$$\frac{1}{132000}(\,3X_A + 6Y_A + X_T + 2Y_T) = 0{,}00003$$

und so erhält man

$$X_A = 3{,}96\,,\ Y_A = 5{,}28\,,\ X_T = -7{,}92\,,\ Y_T = -15{,}84\,. \tag{23}$$

Die Vektoren $\mathbf{F}_A$ und $\mathbf{F}_T$ der gesuchten aktiven Kräfte sind im Bild 7 veranschaulicht und haben die Komponenten (23). Sie rufen die im Bild 6 eingezeichneten Verschiebungen hervor und reduzieren die Formänderungsenergie auf ein Minimum. Dieses Minimum ist

$$\frac{1}{2}(3{,}96 \cdot 0{,}0001 + 5{,}28 \cdot 0{,}0002 - 15{,}84 \cdot 0{,}00003) = 0{,}0004884\,. \tag{24}$$

SCHRIFTTUM

1. Encyklopaedie der mathematischen Wissenschaften, IV. 4, C 29 a, Theorie der Baukonstruktionen. B. G. Teubner, Leipzig 1907—1914, p. 493
2. Ecsedi, I.: A Theorem on the Strain Energy. *Acta Techn. Hung.* (im Druck)

Theorems on the Minimum of Strain Energy in Elastostatics. External forces are acting on an elastic body. Certain conditions are prescribed for their points of application, magnitudes, directions or for the displacements caused by these. Under the prescribed conditions, one raises the question: what are the characteristics of the active external forces, if the strain energy should be minimum. The theorems which are formulated and proved in this paper, yield some contributions to answering the raised question. A numerical example explains the application of the theorems.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (1—2), pp. 15—27 (1980)

EFFECT OF TITANIUM OXIDE CONTENT OF THE BAUXITE CHARGE ON THE SETTLING PROPERTIES OF THE REDMUD

SATYA PRAKASH*

M. OF SCI (MET. ENG.)

and

Z. HORVÁTH**

DR. OF TECHN. SCI.

(Manuscript received 22. January, 1979)

The settling characteristic of the redmud is effected by the type of impurity minerals in the bauxite and their particle size, crystal structure and mode of distribution. Titanium oxide content of the bauxite also effects the settling property. The settling ability of the redmud is decreased with increase in titanium oxide content of the bauxite charge. Addition of settling agent (flour) does not significantly improve the settling ability of the redmuds in laboratory scale investigation.

1. Literature

In the Bayer's process settling is one of the important steps in the separation of liquor and redmud. Settling of redmud is governed by factors like 1) Origin, geological age, chemical and mineral composition of the bauxite 2) Method of ore dressing 3) Digestion conditions 4) Conditions after digestion.

Sulphide minerals in the bauxite adversely effect the settling characteristics of the redmud and this has been linked to the observance of colloidal particles of iron sulphide when S content of the bauxite was above 2 per cent [1]. In general, hematite containing bauxite have better settling properties as compared to goethite type which gives rise to dispersed redmud [2]. On the contrary, some Ghanian bauxites having crystallised goethite have relatively higher settling rate as compared to some Jamaican bauxites having hematite in a highly dispersed form.

Titanium minerals present in bauxites have been found to effect the settling property of the redmuds. ZÁMBÓ and KOVÁCS [3] examined the settling characteristics of the redmuds obtained from the digestion of Indian bauxite. They found the settling rate and compacting to be unsatisfactory. Addition of CaO is amounts more than TiO_2 content of the bauxite led to improvement in settling rate, whereas lesser amount of CaO led to decrease in settling rate.

* Satya PRAKASH, Research Scholar
** Prof. Dr. Z. HORVÁTH, University of Heavy Industry, H-3515 Miskolc, Hungary

SCHEPERS [4] has also reported improvement in settling rate with addition of 2 per cent CaO to the bauxite charge and this has been attributed to crystallisation of sodalite.

KOVÁCS [5] in his investigation with the Indian bauxite from Phutkapahar concluded that with increase in Na_2O content, solid content, and decrease in molar ratio of the slurry proportionately decreases the settling properties of the redmud. They also found that the redmud from Indian bauxite has inferior settling characteristics as compared to one from the Hungarian bauxite and this has been explained to be due to higher TiO_2 content of the Indian bauxite.

KOVÁCS [6] has reported the effect of digestion temperature on the settling ability of the redmud from the Indian bauxite. The redmud from lower temperature digestion was found to have inferior settling properties as compared to the one obtained from higher temperature digestion. This is supported by the findings of BIELFELDT [7] and SCHEPERS [4] who have reported improvement in settling ability with higher temperature digestion.

KOVÁCS [6] has also compared the settling characteristics of the redmuds obtained from Phutkapahar and Amarkantak bauxites and found that the redmud from the latter has better settling property.

KOVÁCS and SOLYMÁR [8] investigated the effect of settling agent, namely, rye-flour and rice-flour on the settling characteristics of Amarkantak and Phutkapahar bauxites and could conclude that the effect of these two agents were similar and increase in the amount of the rice-flour did not further improve the settling rate. According to them the use of settling agent is not very promising.

Use of alkali free water was proposed for the preparation of flour solution [9]. GINSBERG [10] has reported that particle size effects the settling rate, the finer is the particle size, the worse is the settling property.

The data from settling studies in the tubes can be extrapolated to industrial settling equipments using methods proposed by authors [11, 12]. MIKLÓS et. al have developed a mathematical model for settling [13]. MÁRIÁSSY [14] has reported a relation between the specific surface of the bauxite and the heights of the redmud layers and has found these to be directly proportional to each other.

2. Experimental

The bauxite sample was obtained from Bharat Aluminium Co. Ltd, India. It was subjected to physical and chemical analysis. Chemical analysis of the bauxite is shown in Table 1. Derivatograph and X-ray diffractogram of the representative bauxite sample are shown in Figs 1 and 2, respectively. Table 2 shows the mineralogical composition of the bauxite. Thin microsection and electronmicro-probe analysis showed that bauxite is a laterite type having an oolitic and mixed structure. Titanium minerals are found to occur in a very

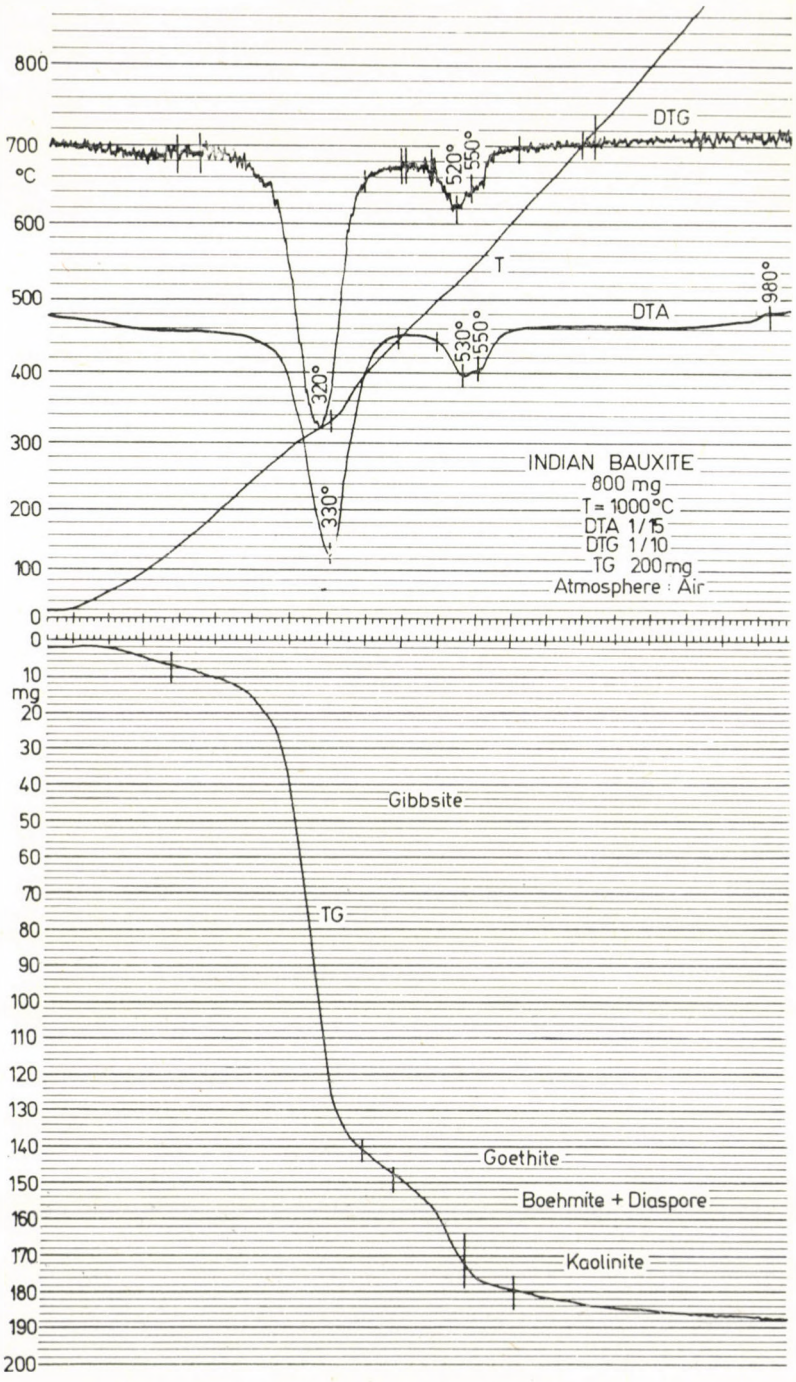


Fig. 1. Derivatograph of Indian bauxite

Table 1.

Chemical composition of the Indian bauxite

Component	Percentage	Component	Percentage
Al_2O_3	49,5	CaO	0,2
Fe_2O_3	15,9	MgO	0.04
SiO_2	4,0	P_2O_5	0.14
TiO_2	6,5	F	0,05

fine dispersed form as can be seem from the electron-reflection image in Fig. 3. The bauxite had 6,5 per cent TiO_2 of which 6,0 per cent was in anatase form and 0,5 per cent in rutile form.

To study of the effect of increasing TiO_2 content, TiO_2 in anatase form (p.a. Merck) was mixed with the original bauxite to obtain samples with TiO_2 content of 12 and 18 per cents, respectively. All three bauxites were digested in an autoclave under similar conditions. The digested slurry was diluted by

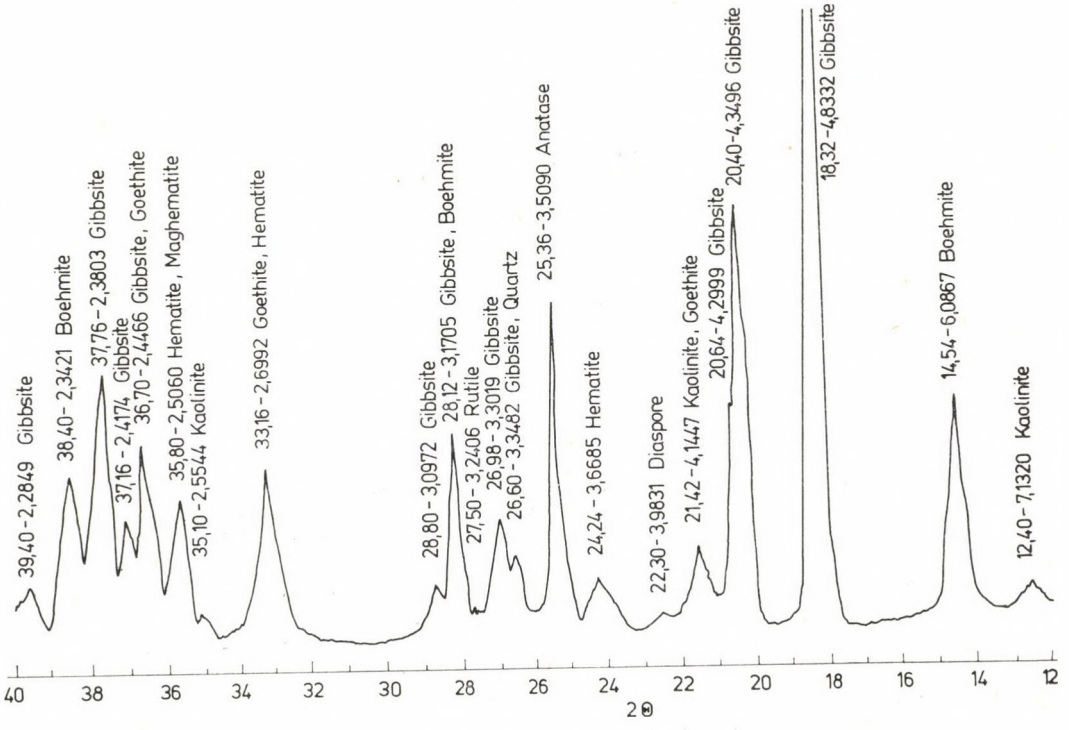


Fig. 2. X-ray diffractogram of the bauxite

Table 2

Mineralogical composition of the bauxite

Minerals	Percentage	Minerals	Percentage
Gibbsite	59,2	Hematite	10,7
Boemite	7,8	Anatase	6,0
Diaspore	1,2	Rutile	0,5
Goethite	6,2	Quartz	1,4
Kaolinite	5,6	Maghematite	0,8

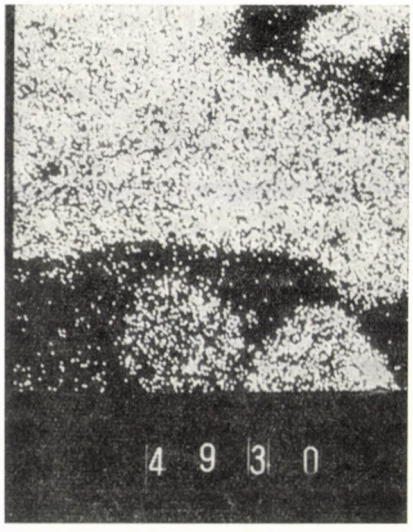


Fig. 3. Electron reflection image of the bauxite showing Ti-distribution. M: $\approx 600\times$

the addition of 100 g/l Na_2O solution to get Na_2O content of $\sim$150 g/l in slurry. The set-up used for settling is shown in Fig. 4. Settling tube consisted of glass tubes 43,1 mm inner diameter and 500 mm long which were immersed in a thermostat bath filled with water-glycerin mixture and fitted with a level controller. Tubes were lighted from behind.

Diluted slurry was poured into the settling tubes and stirred by a vertical movement of the stirrer. In the first tube slurry without flour and in the second and third settling tube 0,1 and 0,2% flour were added to the slurries. The thermostat was maintained at 98 °C temperature. The level of the redmud/clear solution interface were measured at intervals of 5 minutes up to 1 hour of settling period and the after 1 hour up to 4th hour and then only after 24 hours. The redmud after settling was dried at 105 °C for 8 hours and weighed.

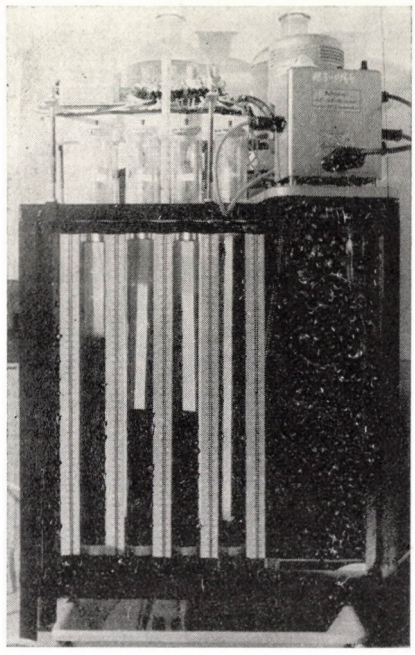


Fig. 4. Experimental set-up used for the settling studies

3. Results and discussions

Plots were drawn between the height of the redmud/liquor interface with respect to the period of settling as shown in Figs 5, 6 and 7 for the redmuds from the original bauxite and the bauxite mixtures having 12 and 18 per cent TiO_2,

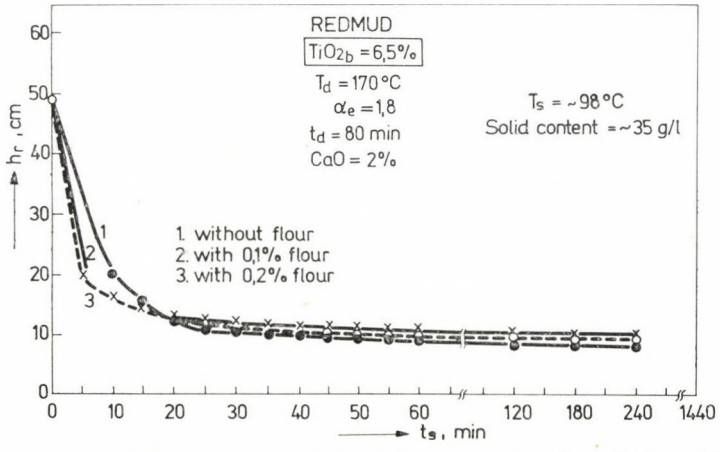


Fig. 5. Settling curves for the redmud obtained from the digestion of the original bauxite showing the effect of flour addition

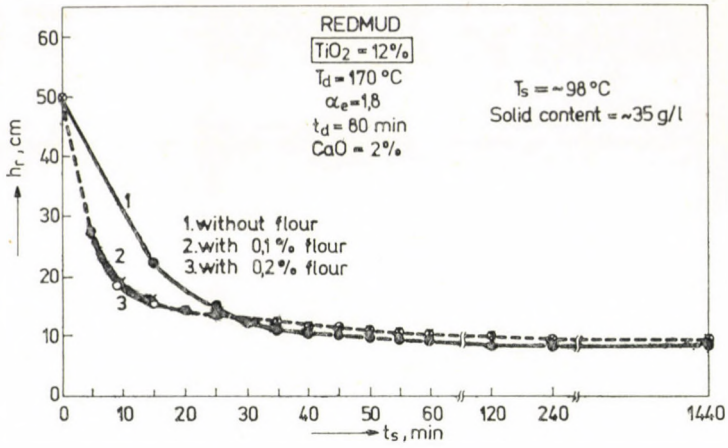


Fig. 6. Settling curves for the redmud obtained from the digestion of the bauxite mixture containing 12% TiO_2 showing the effect of flour addition

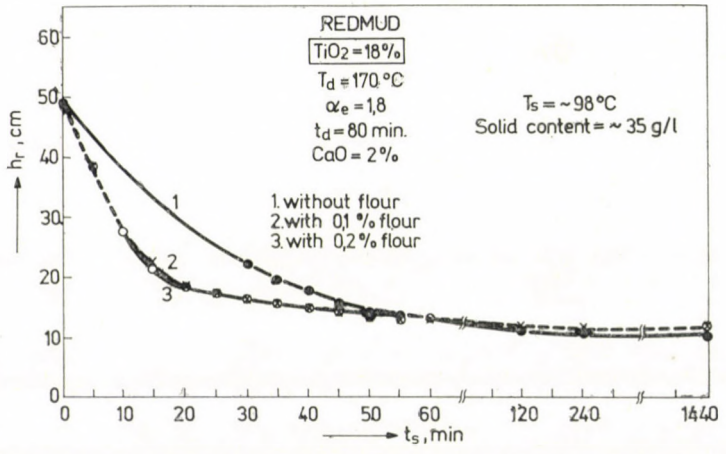


Fig. 7. Settling curves for the redmud obtained from the digestion of the bauxite mixture containing 18% TiO_2

respectively, without and with the additions of flour. It can be observed that the settling agent significantly increases the settling velocities only in the earlier periods of settling in all three cases. Increase in the amount of flour from 0,1 to 0,2 per cent does not markedly improve the settling velocities as can be observed from Table 3.

The settling velocities of the redmuds are decreased, with increase in TiO_2 content of the bauxite charge as can be seen from the settling plots in Figs 8, 9 and 10. The decrease in settling velocities with increase in TiO_2 contents was represented in percentage as compared to the settling velocities of the

Table 3

Settling velocities for the redmuds with a variation in time and amount of the settling agent $T_s = 99$ °C. Solid content = ~ 35 g/l slurry

S. N.	Bauxite TiO_2 content, %	Flour added wt. % redmud	Velocity of movement of redmud/liquor interface, V_s, cm/hr							
			t_s, minutes							
			5	10	20	30	60	120	240	1440
1	6,5	0	168,0	174,0	110,0	77,0	39,9	20,4	10,2	1,71
		0,1	336,0	195,0	108,0	73,4	38,3	19,4	9,77	1,31
		0,2	348,0	198,0	108,9	74,6	38,9	19,7	9,92	1,66
2	12	0	108,0	108,0	94,5	74,0	39,4	20,25	10,15	1,69
		0,1	258,0	179,4	103,5	72,8	38,6	19,17	9,95	1,65
		0,2	274,8	182,4	103,2	72,8	38,6	19,17	9,95	1,65
3	18	0	66,0	66,0	58,5	54,0	35,8	18,95	9,65	1,60
		0,1	128,4	124,8	91,8	65,4	35,7	18,77	9,5	1,58
		0,2	128,4	127,2	91,6	71,2	35,6	18,75	9,5	1,58

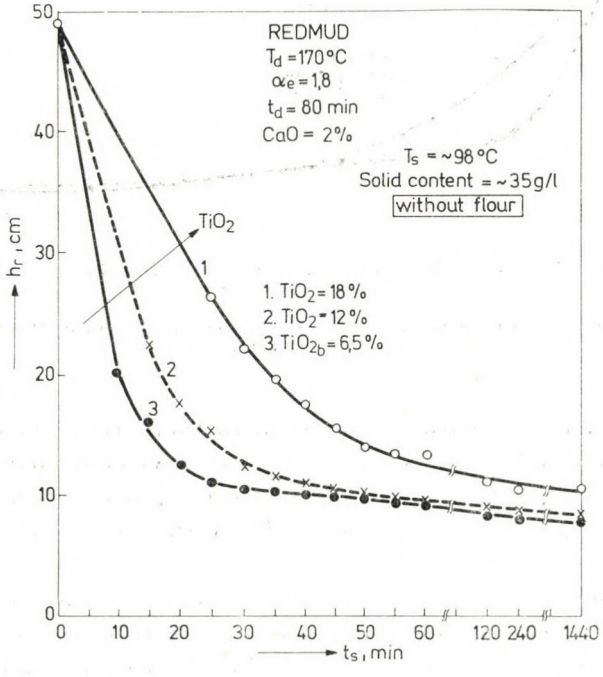


Fig. 8. Settling curves for the redmuds obtained from the digestion of the bauxites containing 6,5, 12 and 18% TiO_2 without addition of flour

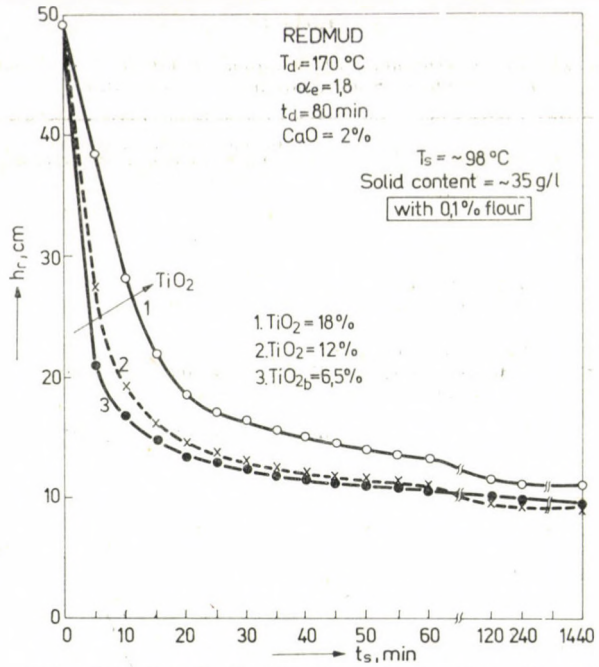


Fig. 9. Settling curves for the redmuds obtained from the digestion of bauxites containing 6,5, 12 and 18% TiO_2 with addition of 0,1% flour

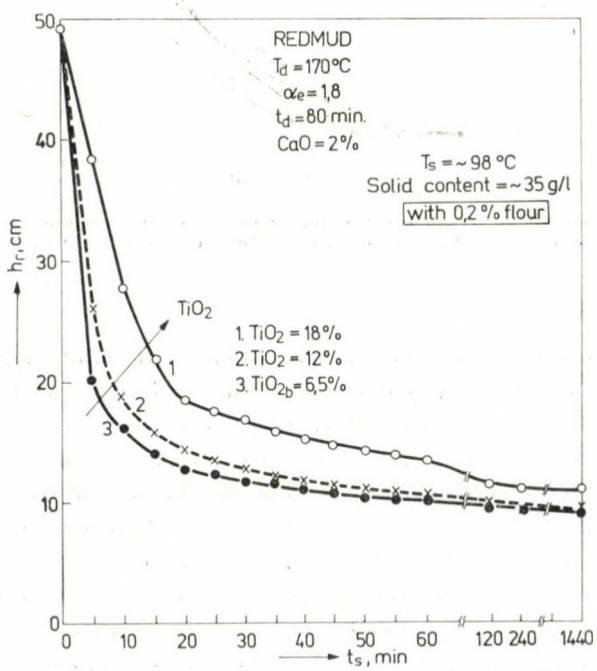


Fig. 10. Settling curves for the redmuds obtained from the digestion of bauxites containing 6,5, 12 and 18% TiO_2 with addition of 0,2% flour

Table 4

Change in settling velocities of redmuds with variations in TiO_2 content of the bauxite charge, $T_s = \sim 99$ °C, Solid content $= \sim 35$ g/l slurry

S. N.	TiO_2 content of the bauxite charge, %	Digestion conditions	Decrease in settling velocity as compared to redmud from the original bauxite containing 6,5% TiO_2				
			t_s, minutes				
			5	20	30	120	1440
1.	6,5	$T_d = 170$ °C $\alpha_e = 1{,}8$	—	—	—	—	—
2.	12,0	CaO = 2% of bauxite wt. $t_d = 80$ min	35,71	14,09	3,89	0,73	1,16
3.	18,0		60,71	46,81	29,87	6,37	6,43

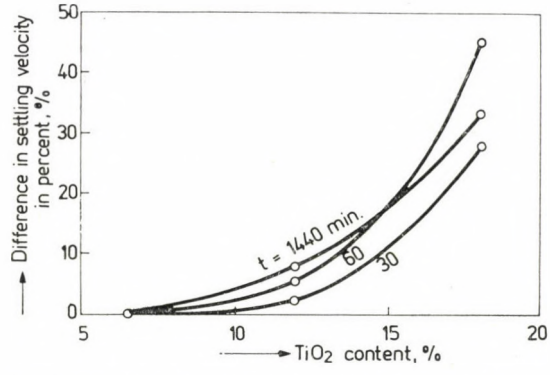


Fig. 11. Curves showing decrease in the settling velocities of the redmuds with varying TiO_2 content of the bauxite mixture for different settling times

redmud from the original bauxite are represented in Table 4 and plotted in Fig. 11. It can be observed that decrease in the settling velocity is appreciable when TiO_2 content increases from 12 to 18 per cent.

The decrease in settling velocities with increase in TiO_2 content of the bauxite charge may be attributed to the increase in the specific surface area of the redmuds as can be seen from the plot in Fig. 12. It is in agreement with the finding of SOLYMÁR et al. [15, 16] that there exists an inverse proportionality between the specific surface of the redmuds and its sedimentation rate.

The measurement of the specific surface area of the original bauxite and bauxite samples with 12 and 18 per cent TiO_2 showed that it does not greatly differ as can be seen in Table 5 so the manifold increase in the specific surface

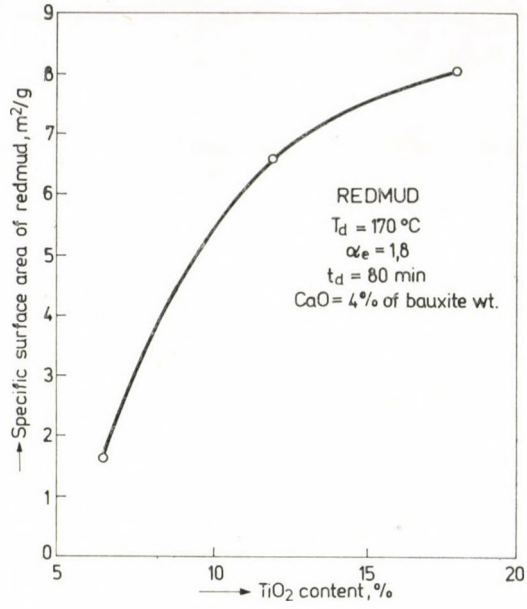


Fig. 12. Plot showing variation in the specific surface area of the redmud as a function of the TiO_2 content of the bauxite charge

Table 5

Specific surface areas for the bauxite samples with different TiO_2 content and the redmuds obtained after digestion

S. N.	Bauxite		Digestion conditions			Specific surface area of the redmud m^2/g
	TiO_2 content %	Specific surface area m^2/g	End molar ratio α_e	CaO added % of bauxite wt.	Temperature of digest. T_d	
1	6,5	13,0	2,3	2%	240 °C	1,60
2	6,5	13,0	1,8	4%	170 °C	1,63
3	12,0	11,3	1,8	4%	170 °C	6,54
4	18,0	11,8	1,8	4%	170 °C	8,04

area of the redmuds from higher TiO_2 containing bauxite may be due to a higher degree of transformation of TiO_2 content to amorphous sodium titanates during digestion.

4. Conclusions

Following conclusions may be drawn from this laboratory investigation:

1. Increase in TiO_2 content of the bauxite charge decreases the settling velocities of the redmuds.

2. The decrease in settling velocities is related to the increase in specific surface area of the redmuds with increase in TiO_2 content of the bauxite charge.

3. a) Use of settling agent (flour) does not siginificantly improve the settling velocities of the redmuds. The improvement is only in the earlier period of the settling.

b) Increase in the amount of the settling agent does not further improve the settling velocities of the redmuds.

Acknowledgements

We wish to express our thanks to the Research Institute for Non-ferrous Metals, Budapest, for the help which we received from them in course of this work in preparation of X-ray diffractogram and electronmicroprobe analysis, chemical analysis of the bauxite and redmud samples.

REFERENCES

1. Nepokryth, T. A.—Kuznetsov, S. F. and Fedyaev, F. F.: The Influence of Iron Sulphide Mineral Impurities in SUBR Bauxites on the Settling of Redmud Sludge during Production of Alumina by the Bayer Process. *Izv. Vyssh. Ucheb. Zaved. Tsvet. Metall*, **5** (1974), 42—44 (Russian)
2. Szabó, Z.—Solymár, K.: A bauxitok ásványos összetételének szerepe a Bayer technológiában c. témáról. (Role of Bauxite's Mineral Composition in the Bayer Technology) Fémipari Kutató Intézet, Budapest, Időszaki jelentés. Témaszám: 11—10
3. Zámbó, J.—Kovács, B.: Indiai bauxit timföldtechnológiai vizsgálata. (Alumina Production Technological Examination of the Indian Bauxite) *Fémipari Kutató Intézet Közleményei*, **7** (1964), 77—88
4. Schepers, B.: Aufschluß von Bauxiten (Digestion of Bauxites) *Erzmetall*, **29** (1976) 61—66
5. Kovács, B.: A phutkapahari bauxit nagy laboratóriumi vizsgálatairól. (Pilot Plant Examination with the Phutkapahar Bauxite) *Fémipari Kutató Intézet*, 511. sz. megbízása. Budapest 1965
6. Kovács, B.: „Bauxit Kataszter" c. téma keretében jelentés az indiai bauxitokkal végzett laboratóriumi előkísérletekről. ("Bauxite cadastral" Entitled Report on the Laboratory Investigations with the Indian Bauxite) *Fémipari Kutató Intézet*. 112. sz. megbízás. Budapest 1964
7. Bielfeldt, K.: Practical Experiences with Tube Digesters. *Jl. of Metals*, **20** (9) (1968), 48—54
8. Kovács, B.—Solymár, K.: Amarkantaki bauxit laboratóriumi vizsgálatairól (About Laboratory Investigations on the Amarkantak Bauxite) *Fémipari Kutató Intézet*, Budapest 511, Megbízásos téma (1969)
9. Bogárdi, E.: Adatok a timföldgyári lisztes vörösiszap ülepítés javítás technológiához (Data Regarding Technology for the Improvement in Settling of the Redmuds with Addition of Floor), *Kohászati Lapok*, **141** (1954), 172—177
10. Ginsberg, H.: Metallurgische Probleme der Tonerde Erzeugung (Metallurgical Problems in the Alumina Production) *Erzmetall*, **7** (1954), 97
11. Matula, M.—Tóth, L.—Bujdosó, E.: Vörösiszapos ülepítők tervezése, méréstechnikai és matematikai modell (The Designing of Measuring Technology and Mathematical Model for the Redmud Settlers), *Bányászati és Kohászati Lapok — Kohászat*, **109** (1976), 228—234

12. COUCHE, R. A.—GOLDEN, L. H.: The Design of Continuous Thickeners for Flocculated Materials. *Australian Inst. of Mining and Metallurgy*, **191** (1957), 117—135
13. SHANNON, P. T.—TORY, E. M.: *Ind. Engg. Chem.*, **57** (1965), 18
14. MÁRIÁSSY, M.: Adatok a vörösiszap ülepítésének kérdéséhez (Particulars Regarding Problems in the Settling of Redmuds), *Fémipari Kutató Intézet Közleményei* **3** (1959), 60
15. SOLYMÁR, K.: Recent Results in the Chemistry and Modelling of the Bayer Process. *Bauxite — Alumina — Aluminium*, **3**, Proceedings of the Second International Symposium of ICSOBA, Held in Budapest, Oct. 6—10, 1969. Edited by Research Institute of Nonferrous Metals, Budapest 1971, pp. 60
16. SOLYMÁR, K.: Methods Applied in Hungary for the Evaluation of Bauxites from the View-Point of Alumina Production., Mineralogical and Technological Evaluation of Bauxites, Edited by the Research Institute of Nonferrous Metals, Budapest 1975, pp. 227

Einfluß des Titanoxydgehalts der Bauxitcharge auf die Sedimentationseigenschaften des Rotschlammes. — Die Sedimentationseigenschaften des Rotschlammes werden vom Typ, der Korngröße, der Kristallstruktur und der Verteilung der mineralischen Verunreinigungen des Bauxits beeinflußt. Der TiO_2-Gehalt des Bauxits hat ebenfalls Einfluß auf die Sedimentation. Mit ansteigendem TiO_2-Gehalt sinkt die Sedimentationsfähigkeit des Rotschlammes. Es scheint, daß unter Laboratoriumsbedingungen die Zugabe eines Sedimentationsmittels das Absetzen nicht wesentlich verbessert.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (1—2), pp. 29—58 (1980)

ELECTROMAGNETIC WAVE PROPAGATION IN MOVING MEDIA WITH SPECIAL REGARD TO FREQUENCY-SHIFTS IN ASTRONOMY ("ANOMALOUS" FREQUENCY-SHIFTS IN ASTRONOMY)

PART II. THE RELATIVISTIC RAY TRACING METHOD

CS. FERENCZ*

[Manuscript received: 2. August, 1978]

In the first part after a comprehensive review of problems which appear to be fundamental in the study of electromagnetic wave propagation in moving media it is shown how to eliminate the observed discrepancy and the possible ways of analyzing the electromagnetic wave propagation in moving media are discussed. One of the discussed methods, developed under the name of relativistic ray tracing, is described in detail in the second part. It is shown how to determine the signal spectrum and the so-called basic effect for which expressions formulated in cases thought to be most important. The weighting functions and the proofs of their form are given.

Finally, in the third part results of earlier similar analyses are utilized for explaining "anomalous" frequency shifts and other associated wave propagation phenomena (line broadening, polarisation rotation, etc.) observed in the solar system. It is suggested that similar, apparently anomalous frequency shifts as well as other associated phenomena observed in galactic and extragalactic wave propagation could be interpreted in the same way.

Ray tracing is an early and useful method [50, 54] for the study of electromagnetic wave propagation in media which can be linearized or regarded as linear. One of its versions, the "principle of modified ray tracing" proved to be highly succesful in inhomogeneous media, among others, for the following of polarization conditions [20].

Besides our studies on application and interpretation of ray tracing in the case of moving inhomogeneous media [13—19] other attempts have been reported on its application to moving media [48].

Other methods, thought to be inadequate, have been critically discussed in part I. Now our review will be restricted to the "relativistic ray tracing" in accordance with the above conclusions.

* Dr. Cs. FERENCZ, Puskin u. 24, H-1088 Budapest, Hungary

II./1. The basic physical problem

As we have seen and understood in the foregoing the total equivalence of l_i and u_i along with their established role, it seems possible to solve any problem in infinitely extended homogeneous systems. However, the method which links l_i with u_i reveals, at the same time, that neither of them can be used in the case of non-homogeneous, non-plane arrangement (point I/3). Anyhow, actually, any even "homogeneous" and "plane" case is such that it can be approximated by a quasi-homogeneous and plane system, only. Our considerations have shown in addition that a correct result can be obtained only from a complete model, not from singled out parts.

It can readily be seen that if an inhomogeneous medium with parameters λ_{ikrs} or ϵ, μ, etc. moves in a laboratory system K at an inhomogeneously distributed velocity, we have actually several (an arbitrary number of) "co-moving" K_i systems. It is an open question which value of ω_i' represents the signal exp $j\omega_0 t$ in a K since these ω_i'-s are not independent of $\lambda_{ikrs}(x_i)$ nor of the other ω_i'-s nor of ω_0.

It is known that by transforming *one and the same physical phenomenon* from observer to observer, in each of the different inertial systems, and by returning on the so adopted way to the original observer in any of the inertial systems chosen by us *a completely closed cycle can be formed*, that is, the result of the transformation series is identical with the initial form. In other words, the selection of the observer does not change the objective phenomenon. For instance, the value of ω_i' varies from observer to observer but on returning through any Lorentz-transformation series to any inertial system, always the same value is obtained.

However, in the present considerations this is not the case. On the propagation in moving inhomogeneous media the sign passes from the transmitter to the receiver across the moving inhomogeneous medium and its interaction with the medium causes its parameters to change, to transform. Thus, the physical phenomenon changes, consequently the relation of the parameters in the initial to those in the resultant phenomenon is not a "path"-independent association. This is particularly so in cases when the model involves energy sources/absorbers of infinite capacity; e.g. most of the linear media, a mirror of infinite mass.

In inhomogeneous media the "propagation path" or "ray path" etc. is curved. In strictly plan-parallel layers when only the medium parameters are inhomogeneous but the whole (inhomogeneous) medium moves at the same $\bar{v}$ = constant velocity and generates a true plane wave, the ingoing and outgoing wave shows several invariant relations, e.g. in a given inertial system the frequency of the ingoing and outgoing wave does not change.

Generally, however, the structure of the inhomogeneity differs from that

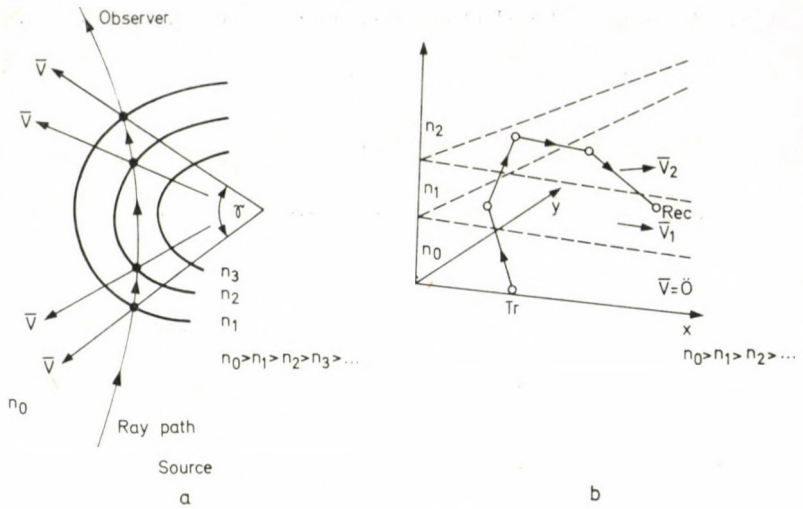


Fig. 3

specified above, the velocity field also changes along a coordinate even for a phenomenon stationary in space and time. Therefore, appreciable changes in the direction of propagation and in velocity have to be expected in most of the time invariant systems, too. These may be due to possible changes in the velocity field and/or to changes in ϵ, μ, etc. (Fig. 3). The resultant changes can be described in terms of a mirror system by use of geometrical-optical approximations. This is obvious in the case of the total "return" (total reflection) of the "ray path" and starting precisely from this, it can be understood in any case on considering Fig. 2. It can also be seen that the above mentioned special plan-parallel case can be represented by a mirror system which eventually "retransforms" the "distorted" input signal to the output signal required by the frequency invariance, etc. However, a moving mirror, mirror system or inhomogeneity causes a transient change in time. In order to obtain a stationary effect in time, other conditions have to be satisfied, too. Before discussing these conditions, it is thought useful to recall the effect of a single moving mirror or of any equivalent system.

The relations which hold for moving mirrors are known [29]. They are now recalled, because in a very simple inhomogeneous arrangement, they clearly show that signal parameters actually do change if the signals propagate in moving inhomogeneous media of which the moving mirror is though to be an extreme but still a case.

First of all, it follows trivially from (9.b) and (9.c) that e.g. a metallic mirror of conductivity ∞ ($\varepsilon' \doteq \infty$) in K' is observed in K still as an unchanged ideal mirror ($\lambda_{ikrs} \doteq \infty$). Let ϑ' and ϑ then be angles of incidence with ϑ

and ϑ_r being the angles of reflection, respectively. We know that in K'

$$\vartheta' = \vartheta'_r \tag{41}$$

and

$$\omega' = \omega'_r \,.$$

However, we also know — without going into the pertinent calculations — that in K

$$\vartheta \neq \vartheta_r \,, \tag{42}$$

but that

$$\frac{\operatorname{tg} \vartheta_r/2}{\operatorname{tg} \vartheta/2} = \frac{1 - v/c}{1 + v/c} \,.$$

Further

$$\omega \neq \omega_r \,,$$

$$\omega_r = \omega \varkappa^2 \left(1 + 2\frac{v}{c} \cos \vartheta + \frac{v^2}{c^2}\right) \tag{43}$$

and a similar relationship exists between incident and reflected energies. Alternatively, this result means that, as regards light, the mirror is an infinitely large source and absorber of energy because its mass is taken to be infinite and that the phenomenon which is observed in one of the inertial systems, as a change of impulse, is seen in the other as a change of both impulse and energy.

Actually, part of the signal is always reflected by the inhomogeneity while the other fraction continues to proceed, possibly so that part of it appears in another form (it dissipates, etc.), consequently the effects of these basic phenomena are necessarily due to be observed.

This fact is usually taken into account for inhomogeneities causing a transient effect in time (the most trivially equivalent case being a *single* mirror or mirror arrangement). Let us now see arrangements producing effects which are stationary in time.

It has to be noted here that now we do not try to find as described in point I/3 the exact solution in terms of wave theory but that we are studying a ray tracing procedure.

First, let us consider a simple interface at which the velocity field changes (Fig. 4) [19]. It is obvious from the figure that if a moving medium in ② is illuminated by a wave front stationary in the K system, it will appear in K', that is in ② as a moving excitation since individual parts of the medium *are subject to an excitation varying in time.* One cannot construct any other kind of a physically real model. It is important to realize that ***motion***, that is a ***moving medium*** cannot be observed and defined unless the above reasoning holds. If, for instance, a medium (in a model) is truly continuously homogeneous, its parts are undistinguishable and if it is infinitely extended, then its movement

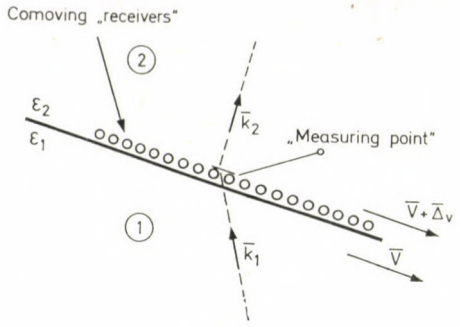


Fig. 4

in the direction of infinite extension cannot be interpreted. — It is precisely at this point that a number of considerations is based on a fallacy [4, 11, 12, 48, etc.]. — Thus, if we deal with a *moving medium*, then a system which is stationary, say, in K, will not be stationary, at least from the point of view of excitation, in K'. This is *an irrefutable, fundamental physical fact!*

It is important here to point out that excitation of media can also be treated in this way and that by making use of Huygens elementary wave approach as complementary to those discussed in I/3 we can try to find a solution *only* in terms of wave theory. If we want to evaluate every parameter by ray tracing, we have to use the same procedure. In purpose-oriented analyses a simplified method is available which will be described in the following.

One of the most simple arrangements which yields a stationary resultant is the moving set of mirrors shown in Fig. 5. If the size and spacing of mirrors is suitable, then the wave pattern received from a directed transmitter or by a receiver installed at a single point takes the form to be seen in Fig. 2, where the

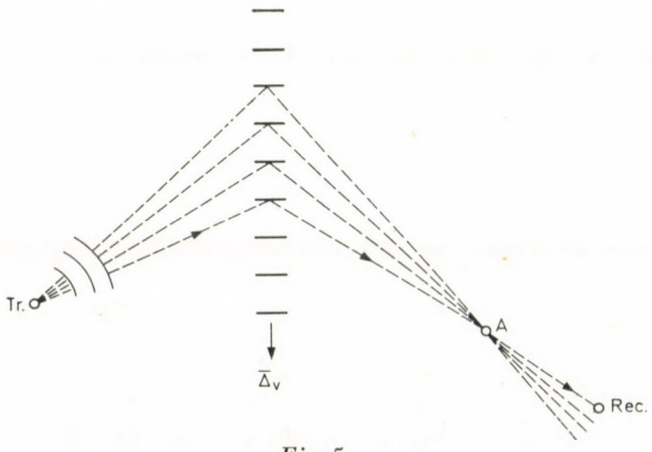


Fig. 5

signal is broken up by the set of mirrors to sections consisting of n complete waves. The spectrum formed in this case will be discussed later.

If we do not want to work with the simple mirror model and when Huygens' elementary wave approach is not yet needed, then we can proceed as follows. We advance along, with the signal, from layer to layer, where layer means ranges where $\bar{v}_i$ = constant, going in the direction of the main radiation taken to be a pattern representing a local plane-wave. The changes in the signal and its parameters are determined in the time intervals when the illuminated fraction of the medium is reached by the pattern representing the plane-wave taken to be the signal at maximum. Thus, going along, with the signal, from layer to layer we find that the resultant total change observed in the signal parameter under investigation is one of the determining parameters of the resultant spectrum form.

Two types of "interaction" are observed in this procedure. One is the reception of the signal by the moving medium from the previous layer (Fig. 4), the other is the penetration of the signal into the matter. The first "interaction" is described by the Lorentz transformation, the second by the law of refraction-reflection. On the principle of energy conservation it can readily be seen that any time a boundary between layers is crossed we have to apply first the Lorentz transformation, second the law of refraction-reflection.

It is now obvious that the interaction occurs with a single fraction (particle-group) K_i' of the medium in the manner described above only for an "instant", therefore, the results thus obtained have to be "weighted" by use of spectrum calculations or of other considerations.

Since the examples of application are frequency variation-oriented, it is useful to consider the following: On integrating the Doppler-effect over time from $-\infty$ to $+\infty$, the average frequency change $\widetilde{\Delta f}$ always goes to zero in the case of individual particles, etc. In the case of medium propagation this holds for any individual constituent, particle, of the medium. It follows from the resultant that for an isotropic transmitter antenna we get $\widetilde{\Delta f} = 0$ if we integrate the phenomenon at some spatial (receiver) point in definable relationship with the phenomenon over the whole space angle with respect to $\bar{r}$. (E.g. at point A in Fig. 5).

Moreover, it can be shown [19] that we have *always* $\widetilde{\Delta f} = 0$ *in homogeneous media with any propagation profile* for which Grad $\bar{v}$ holds, if the transmitter and receiver are jointly fixed in the same intertial system.

Similarly, it can be seen [19] that a useful informative result can be obtained for low velocities $(v/c)^i \cong 0$, $i \geq 2$ by using the following crude approximation. The dispersion relation assumed to hold for the medium [22, 23, 53] is solved together with the equation

$$\overline{\text{grad}}\ \omega = -\bar{k}\ \overline{\overline{\text{Grad}}}\ \bar{v}$$

or, for the notation $\bar{F} = \bar{F}_0 \exp j\,\Phi$, with the equation

$$\overline{\text{grad}}\,\frac{\partial \Phi}{\partial t} = \overline{\text{grad}}\,\Phi \cdot \overline{\overline{\text{Grad}}}\,\bar{v}$$

The medium parameters are involved with their assumed values in K' because of the low velocity.

II./2. The relativistic ray tracing method

It has already been seen in the foregoing that the form we try to find is not yet available in terms of the total wave theoretical - particle description (point I/3). On the other hand, it has been shown that individual parts of the moving medium are excited at different times. We also know that, in general, even for a stationary configuration the structure of the velocity field does not coincide with that of the medium parameter field in a moving inhomogeneous medium.

For this reason, the phenomenon must be decomposed in several ways into parts which can be regarded from some point of view as homogeneous (or analog) even if we apply the ray tracing method.

In the present method to the following decomposition will be used:

— We shall consider separately and follow over the whole system the "propagation" in (instantaneous) sections where the interactions of medium particles with electromagnetic wave can be regarded as stationary and complete, that is, as producing the maximum effect.

Keeping in mind that this interaction varies with time, we shall establish how the maximum effect should be weighed to approximate as far as possible the actually observed effect.

The maximum, ideal or basic effect will be determined as follows:

— The velocity vector field is approximated by a step function consisting of sections with constant velocity, that is, layers with "homogeneous velocity" will be formed with respect to the velocity field (Fig. 4).

The phenomenon can then be differentiated with respect to the velocity field. The phenomenon is stationary if the velocity at a given point of the field is constant in time. If the velocity is not constant, the phenomenon is regarded as quasi-stationary and as analysable in this manner if the velocity can be taken as constant during the transmission of the signal through the layer. If this is not the case, the phenomenon cannot be studied without further analysis and restrictions.

As to the medium parameters (ϵ, μ, etc.), the phenomenon is stationary, if because of propagation the arrangement of the "refraction index" does not change in either K' or K. It is quasi-stationary if the "refraction index"

arrangement does not change in K but in a strict sense it does in K_i', while the shift in the "refraction index" arrangement caused by $\overline{\Delta v}_i$ is negligible as compared to the transmission time of the signal. The phenomenon is non-stationary if neither of the above conditions is fulfilled, then a propagation in media inhomogeneous even in time, more precisely, of the form $\partial\lambda_{ikrs}/\partial t \neq 0$ has to be investigated also within the different layers in the K_i'-systems.

— Now for the application of the ray tracing procedure the layers of constant velocity, but of inhomogeneous medium parameters, defined above as stationary or as quasi-stationary will be decomposed to layers of homogeneous "refraction index" in K_i'.

It has already been shown that even slightly inhomogeneous media generally lead to scattering, etc. of the signal [20]. In the present paper these problems of ray tracing like the definition of plane parallelism, its conditions, consequences of the deviation, scattering, inhomogeneous medium parameters, etc. [20, 23, 26, 69] are not discussed as they are assumed to be known.

Similarly, it is assumed to be known, that it is possible to make energetically precise analyses concerning polarization etc. by using a ray tracing procedure of *modified* construction [20]. To follow the scattering pattern we have to use first a straight tracing and sometimes the straight and modified procedures are alternated in order to obtain an approximation of adequate accuracy [20]. Thus, the "basic effect" is determined according to the general block diagram to be seen in Fig. 6.

It has also been expounded in the foregoing and summarized in Fig. 6 how the "basic" or "ideal" effect is determined in terms of the "relativistic ray tracing method". However, we are aware that the so obtained characteristic is not the measurable parameter of the wave-pattern, but only one of its characteristic features. For instance, if we study the variation of frequency, the so obtained basic effect does not show the actual shift in frequency but it shows — remember — in Fig. 2 — the "frequency" within the different sections constituting the wave train in the determination of the total spectrum.

In the determination of the actual *measurable* effect we have also to take into account the already discussed phenomenon that the *spatial* wave pattern the interference pattern *in space* observable in the stationary case of our initially assumed system transforms to a space and *time dependent* wave pattern, interference pattern in space and *time*. The so formed total resultant spectrum, which can be characterized by the basic effect, determines the measurable, observable parameters. This analysis can be made — in principle — by the determination of the total resultant spectrum (Huygens' method, etc.). However, in practically important cases this seems to be too complicated for the time being and in most cases a good approximation is sufficient.

Thus, for a given "basic effect" the measurable effect produced in the total spectrum will be determined in the following as the phenomenon multi-

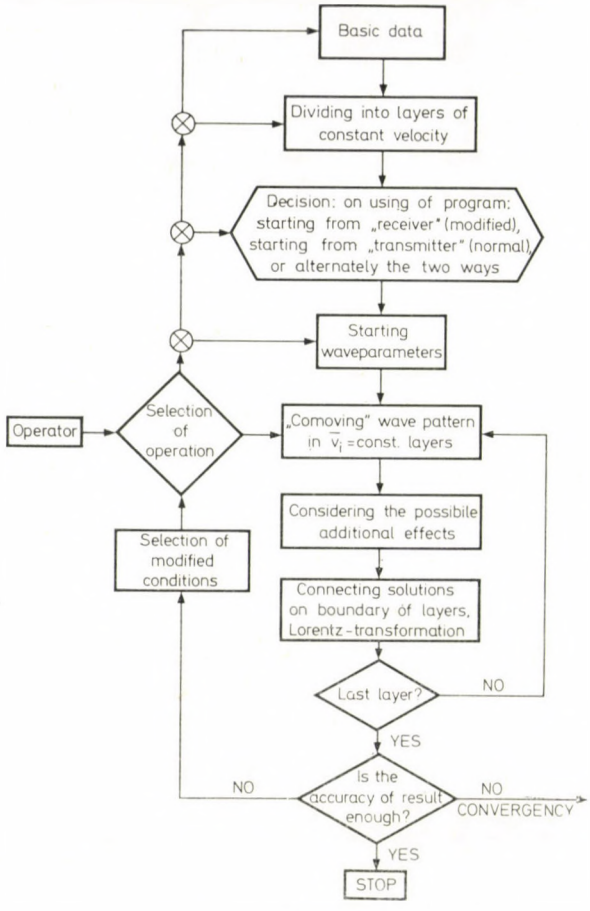


Fig. 6

plied by the associated weighting function obtained by making use of the ray tracing method. The weighting function can be associated with the parameter under investigation. Now let us see in detail, as part of the frequency shift analysis, the weighting function which can be associated with the frequency variation.

Finally, we shall utilize in the following the fact that for the determination of the weighting function in all cases studied to data [1, 13, 15, 16, 18] — the phenomenon the part of the signal propagating from the source to the receiver can be represented by an equivalent (as to reflection factor, velocity etc.) set(s) of moving mirrors. Since our aim is to analyze frequency shifts caused by moving inhomogeneous media, the spectrum analyses will be made in a signal spectrum similar to that generated by an equivalent set of moving mirrors. Now, it is expedient to delimit the scope of our analysis. The variations

of the electromagnetic wave spectrum in matter have already been studied in detail (line broadening due to thermal convection, shift in the spectrum of signals emitted from moving matter serving as signal generator, frequency dependent reflection, line broadening due to turbulence, scattering etc.). These effects the resultant of which also appears in a real phenomenon *will not be dealt with* in the following. Our considerations will be restricted to the *effect of moving media, only*, and even their range will be restricted to models which can be classified as *stationary* and which are thought to be important.

II./3. Analysis of spectra generated by a set of moving mirrors

Now let us consider the signal spectra generated by the set of moving mirrors or by equivalent models as shown in Fig. 5. The aim of this study is to determine the spectrum of a signal generated by a monochromatic continuous excitation function of the form

$$\bar{F} = \bar{F}_0 e^{j(\omega_0 t - \bar{k}\bar{r})} = \bar{F}(\bar{r}) e^{j\omega_0 t} \tag{44}$$

transmitted from a suitable transmitter in the K system to a receiver in the same system. Furthermore, we will try to find an expression for the measurable average frequency and to establish the relation between f_0 and f_{received}. (Since the case under investigation is taken to be stationary, we need not deal with problems mentioned in connection with the spectrum in [20].)

Since in the present case the aim is to identify a basic phenomenon seeing the number of misunderstandings which exist [4, 9—12, 21, 42—49, 55], we will not try now to construct a final refined approach to a problem and content ourselves with the study of the following signal form.

The arrangement to be seen in Fig. 5 and its equivalent versions cut the signal generated according to (44) as represented in Fig. 2, by a broken line into *sections* characterized by some altered frequency according to (43) in two possible ways. — Now only this substituted pattern will be analyzed.

First, we shall study the simpler case in which every section consists of n full waves, then we shall analyse the case in which though the lengths of individual sections are equal, the length is not associated with the period of the original signal. Finally, the average frequency and the crude change in line broadening will be studied in detail on the first (simpler) spectrum.

a) The spectrum of the signal consisting of sections of an integer number of complete waves

Let us determine the signal spectrum generated from the monochromatic and e.g. purely sinusoidal signal

$$\omega_0 = \frac{2\pi}{T_0} = 2\pi f_0 \tag{45}$$

by the set of mirrors in Fig. 5. if the signal is decomposed by the movement of the different elements in the set of mirrors to sections of frequency

$$\omega_\alpha = (1 + \alpha)\omega_0 \tag{46}$$

where $\alpha \lessgtr 0$ is the "basic effect". It is assumed, in addition, that the length of each section — as determined by the size, spacing of the mirrors and by the angle of incidence — is given as

$$T = nT_0 \tag{47}$$

where n is a positive, integer number. Furthermore, the arrangement of the set of mirrors is such, that no parts of the signal can "escape", that is, all the incident waves are reflected.

This is the simplest model of an excitation variable in time which is also made simple by the choice of n. The so formed reflected wave pattern is to be seen in Fig. 2. Since we have to deal with a signal of strict internal periodicity, the spectrum obtainable by Fourier transformation can be evaluated according to the rules of function theory. (The calculations involved are trivial.)

If the time $t = 0$ is chosen so that the original signal is an odd function and if the decomposition according to (47) is associated symmetrically with this point, we get only components for an odd Fourier series and the value of these spectral lines is

$$B_k = (-1)^{n+1}\frac{2}{\pi}\sin\left(\frac{k\pi}{1+\alpha}\right)\cdot\frac{n(1+\alpha)}{n^2(1+\alpha)^2 - k^2} \tag{48}$$

where B_k is associated with the frequency

$$\omega_k = k\frac{2\pi}{T} = \frac{k}{n}\omega_0 .$$

By definition the amplitude of the signals is 1 in the range of the time under consideration. This spectrum can be analyzed by making use of (48). The average

frequency of the signal is given as

$$\tilde{f} = f_0 \frac{\sum_k \frac{k}{n} B_k^2}{\sum_k B_k^2}, \tag{49}$$

and the relative frequency shift (if any), as

$$\left(\frac{\widetilde{\Delta f}}{f}\right) = \frac{\tilde{f}}{f_0} - 1 . \tag{50}$$

The line broadening can be evaluated e.g. from the total number of lines needed for the determination of (49) and (50) to a given accuracy or by another method (from the value of $| B_k | / | B_{k\,\max} |$, etc.).

It is perhaps useful to note that if $\alpha = 0$, then $k \neq n$ $B_k \equiv 0$ and for $k = n$, we have $B_n = 1$.

It is again a special case if $(n\alpha)$ is an integer. Then B_k associated with $k = n(1 + \alpha)$ has to be evaluated separately from the expression

$$B_{n(1+\alpha)} = \frac{1}{1 + \alpha} .$$

Similarly, it can be seen that the interpretation range of (48) with respect to α is the open range $(-1, +\infty)$ in accordance with the physical phenomenon.

It is also apparent that the maximum of the spectrum lies somewhere close to $n(1 + \alpha)$ as a function of k.

b) General description of the spectrum of a signal reflected from a set of moving mirrors

If our analysis is made by considering distributions in Banach's space, then Heaviside's distribution offers a simple possibility to describe the time behaviour of an originally monochromatic signal, cut into any number of parts. This distribution, also called unit jump is defined [52] as

$$\langle \varphi, \Theta \rangle = \int_0^\infty \varphi(t)\, dt .$$

In the present case the original signal is the stationary signal existing for an infinite time on the transmitter side of the form

$$f_0(t) = A_0 \sin (\omega_0 t + \varphi_0)$$

where we can also take $\varphi_0 = 0$. The resultant signal after continuous chopping can be formed from this by modifying the frequency in each section according to (46). The chopping — since a stationary phenomenon is being studied — is also periodical, however, a period is given by

$$T = \nu T_0 , \quad \nu > 0 \tag{52}$$

where ν can be an integer, a rational or an irrational number. In the latter case the resultant phenomenon is aperiodical, while the former case is the same as that studied in point a). For simplicity's sake, the amplitude of the signal is taken by definition to be 1 in the covered range of time. (Because of this, for amplitude, etc. analysis with the so obtained spectrum, we have to multiply by the factor required by α, by the initial amplitude, etc.). By choosing the phase position of the chopping times relative to $t = 0$ so that the section including $t = 0$ should be symmetrical to $t = 0$, the general case has to be described by taking $\varphi_0 \neq 0$.

Using the above considerations, let us determine the spectrum of the following signal

$$\begin{aligned} f(t) = {} & A_0 \sin(\omega'_\alpha t + \varphi'_0) \cdot [1 - \Theta(t - T_{\alpha n-})] + \\ & + \sum_{l=n}^{m} A_1 \sin(\omega_\alpha t + \varphi_l) \cdot \Theta(t - T_{\alpha l-}) - \\ & - \sum_{l=n}^{m} A_1 \sin(\omega_\alpha t + \varphi_l) \cdot \Theta(t - T_{\alpha l+}) + \\ & + A_2 \sin(\omega''_\alpha t + \varphi''_0) \cdot \Theta(t - T_{\alpha m+}) . \end{aligned} \tag{53}$$

In (53) the value of ω_α is the value given by (46). The assumed model now requires the choice of either

$$\omega'_\alpha = \omega''_\alpha = \omega_\alpha = (1 + \alpha)\omega_0 ,$$

or

$$\omega'_\alpha = \omega''_\alpha = \omega_0 \quad \text{and} \quad \omega_\alpha = (1 + \alpha)\,\omega_0 .$$

It is useful and appropriate for our models if the values of the amplitudes are normalized to

$$A_0 = A_1 = A_2 = 1 .$$

Similarly, we take

$$\varphi'_0 = \varphi''_0 = \varphi_0 .$$

The detailed analysis of the decomposition to sections $T_{\alpha l\pm}$ and φ_l are obtained by taking also (52) into account as

$$\varphi_l = \varphi_0 - \alpha\omega_0 l\nu T_0 = \varphi_0 - 2\pi\alpha l\nu \tag{54}$$

and

$$T_{\alpha l\pm} = vlT_0 \pm \frac{vT_0}{2(1+\alpha)} .$$

It can be seen that (53) describes an aperiodical signal when v is an irrational number. We can also distinguish two periodical cases:

— The case described in point *a*) is obtained if $v = N$ which is a positive integer, when it is allowed to take $\varphi_0 = 0$. Then

$$\varphi_l = -2\pi\alpha lN$$

$$T_{\alpha l\pm} = NT_0\left[l \pm \frac{1}{2(1+\alpha)}\right]. \tag{55}$$

Our previous results are reproduced by taking the extreme cases

$$n \to -\infty , \; m \to +\infty \tag{56}$$

which lead to the stationary phenomenon.

— A periodical case, not discussed up till now, is obtained for the extreme cases (56) if $v = N/M$ is a rational number. (In a given case its particular analysis could be of interest.)

Finally, it is useful to note that in the case of $\alpha = 0$, $\varphi_0 = 0$, etc., the original signal and its spectrum are reproduced.

Let us symbolize the Fourier transform of the distribution (53) by

$$\mathfrak{F}[f(t)]$$

and let us interpret it by the usual definition [52] as

$$\langle \mathfrak{F}(\varphi), \mathfrak{F}(f)\rangle = 2\pi\langle \varphi, f\rangle .$$

In the present case we shall use, in the detailed calculations, the relation

$$\mathfrak{F}(f) = f_0(\sigma) = \langle e^{j\sigma t}, \overline{f(t)}\rangle$$

which holds for bounded, scalar distributions, where $\bar{f}$ is a conjugate of f.

To obtain the Fourier transform of (53), we have first to determine its formulating expression

$$\Theta(t - T)\sin(\omega t + \varphi) = f_l . \tag{57}$$

Making use of the above, omitting calculation details, we will get from

$$\mathfrak{F}(f_l) = \langle e^{j\sigma t}, \overline{\Theta(t - T)\sin(\omega t + \varphi)}\rangle$$

that

$$\mathcal{F}[\Theta(t-T)\sin(\omega t+\varphi)] = \\ = \frac{j}{2}\left\{e^{j[(\sigma-\omega)T-\varphi]}\left[\frac{j}{\sigma-\omega}+\pi\delta(\sigma-\omega)\right]- \right. \\ \left. - e^{j[(\sigma+\omega)T+\varphi]}\left[\frac{j}{\sigma+\omega}+\pi\delta(\sigma+\omega)\right]\right\} \tag{58}$$

By introducing the notation $\Phi = \omega T + \varphi$, we have

$$\mathcal{F}(f_l) = \frac{1}{2}\left[\frac{e^{j(\sigma T+\Phi)}}{\sigma+\omega} - \frac{e^{j(\sigma T-\Phi)}}{\sigma-\omega}\right] + \\ + \frac{j\pi}{2}\left[e^{j(\sigma T-\Phi)}\,\delta(\sigma-\omega) - e^{j(\sigma T+\Phi)}\,\delta(\sigma+\omega)\right]. \tag{59}$$

Now the total Fourier transform is to be understood as

$$\mathcal{F}(f) = \mathcal{F}(f_{\text{lover end}}) + \sum_{l=n}^{m}\left[\mathcal{F}(f_{l-}) - \mathcal{F}(f_{l+})\right] + \mathcal{F}(f_{\text{upper end}}). \tag{60}$$

In the calculations it is useful to make use of the relation

$$g(\sigma)\delta(\sigma-\sigma_0) = g(\sigma_0)\delta(\sigma-\sigma_0)$$

since this makes it possible to use the transformation

$$e^{j(\sigma T\pm\Phi)}\,\delta(\sigma\pm\omega) = e^{\pm j\varphi}\,\delta(\sigma\pm\omega).$$

If also (54) is taken into account, we get

$$\mathcal{F}[f(t)] = \mathcal{F}(f_{\text{lover end}}) + \mathcal{F}(f_{\text{upper end}}) + \\ + \frac{1}{2}\left\{\frac{1}{\omega_\alpha+\sigma}\sum_{l=n}^{m}\left[e^{j(\sigma T_{\alpha l-}+\Phi_{l-})} - e^{j(\sigma T_{\alpha l+}+\Phi_{l+})}\right]+ \right. \\ \left. + \frac{1}{\omega_\alpha-\sigma}\sum_{l=n}^{m}\left[e^{j(\sigma T_{\alpha l-}-\Phi_{l-})} - e^{j(\sigma T_{\alpha l+}-\Phi_{l+})}\right]\right. \tag{61}$$

where

$$\Phi_{l\pm} = \omega_\alpha T_{\alpha l\pm} + \varphi_l$$

in agreement with the notation in (59). In the stationary case the summation in (61) covers the range $-\infty < l < +\infty$ for all integers and the functions $f_{\text{lower end}}$ and $f_{\text{upper end}}$ automatically disappear.

(61) or some of its appropriate transforms will already permit some special parameters, e.g. the average frequency shift we are now interested in, to be determined. It can be expressed as

$$\tilde{\omega} = \frac{\int_0^\infty \sigma \mathcal{F}[f(t)]\,\overline{\mathcal{F}}[f(t)]\,d\sigma}{\int_0^\infty \mathcal{F}[f(t)\,\overline{\mathcal{F}}[f(t)]\,d\sigma} \tag{62}$$

It is also possible to define the band-with e.g. σ_{01} and σ_{02} are given according to the usual point -3 dB by

$$\frac{\mathcal{F}[f(t)]\,\overline{\mathcal{F}}[f(t)]}{\{\mathcal{F}[f(t)]\,\overline{\mathcal{F}}[f(t)]\}_{\max}} = \frac{1}{2}\,. \tag{63}$$

The associated ω_1 and ω_2 then give

$$\Delta f = \frac{\Delta\omega}{2\pi} = |\,\omega_1 - \omega_2\,|\,.$$

Of course, the band-width can be defined in other ways, too, — the convergence of $\tilde{\omega}$, etc.

If there are taken into account in (61) expression (55) and the total range $-\infty < l < +\infty$ the case described in point *a*) will be reproduced.

In a number of cases, it is useful to transform (61). On omitting the initial and closing — non-perturbed etc. — sections and considering the stationary case, we have

$$\mathcal{F}_{st}[f(t)] = j\left[e^{j\varphi_0}\,\frac{\sin \pi v\left(1 - \dfrac{1}{1+\alpha}\right)}{\omega_0(1+\alpha) - \sigma}\sum_{l=-\infty}^{+\infty} e^{j\left(\frac{\sigma}{\omega_0}-1\right)2\pi l v} - \right.$$

$$\left. - e^{j\varphi_0}\,\frac{\sin \pi v\left(1 + \dfrac{1}{1+\alpha}\right)}{\omega_0(1+\alpha) + \sigma}\sum_{l=-\infty}^{+\infty} e^{j\left(\frac{\sigma}{\omega_0}+1\right)2\pi l v}\right]. \tag{64}$$

This expression is already suitable for most of the analyses in question. It permits the spectrum to be considered as known to the extent required in this study.

c) Spectrum analysis of the signal consisting of sections with an integer number of complete waves

There is no difference in the phenomenal description of the two spectra considered in the foregoing which can be characterized by the "basic effect" given by frequency shift α. Thus, in the following it is taken as being sufficient for this study to analyze the spectrum considered in point *a*).

In the analysis the average frequency shift, as interpretable from (49) and (50), was determined in terms of the basic frequency shift α defined in (46) and in terms of n which, as interpretable from Fig. 2 and from (47), is the number of complete waves in each wave train and characterizes the medium-wave interaction. At this point the accuracy of the calculation is ε. In the figures showing the results, we find, in addition, that

$$\left(\frac{\Delta\tilde{f}}{f}\right) = g(\alpha, n) = \frac{\Delta\tilde{f}}{f_0} = \frac{\tilde{f} - f_0}{f_0} . \tag{65}$$

To characterize the band-width the number of all spectral lines — $N(\alpha, n)$ — needed to achieve the accuracy ε are taken into account as normalized to n in order to obtain an informative picture. Normalization necessarily follows from (48) and (49). Since the amplitude was taken in each case to be 1, the "total energy" accounted for is expressed as

$$P = \frac{4}{\pi^2} \sum_k B_k^2 \quad \text{and} \quad \Delta P_n = (P - 1)\ 100 . \tag{66}$$

This quantity is only informative.

Of course, the thus obtained band-width parameter $N(\alpha, n)/n$ is not the true band-width measurable with a real instrument — it contains more spectral lines than the true band — but it well characterizes the band-width agreeing well with both the nature of the measurable variations and the relative order of magnitude of the variations. It will be used only in this sense.

In Fig. 7 the functions $g(\alpha, n)/|\ \alpha\ |$ and $N(\alpha, n)/n$ are shown for $\alpha = +10^{-4}$ as plotted against the normalized calculation accuracy $\varepsilon/|\ \alpha\ |$. It is apparent that the convergence occurs for the values $\varepsilon/|\ \alpha\ | \leq 10^{-4}$. Problems of numerical convergence have to be taken into account if

$$\frac{1}{n} \leq \frac{\varepsilon}{|\ \alpha\ |} .$$

Moreover, the "total energy" parameters in (66), to be seen in Fig. 8, show a much earlier convergence. The total band measurable with real receivers

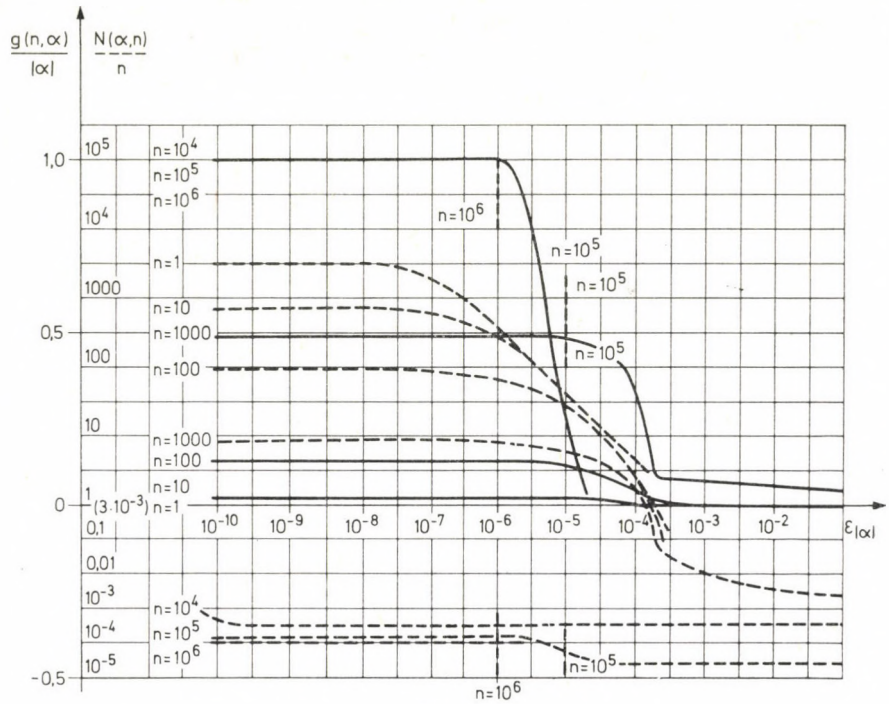


Fig. 7

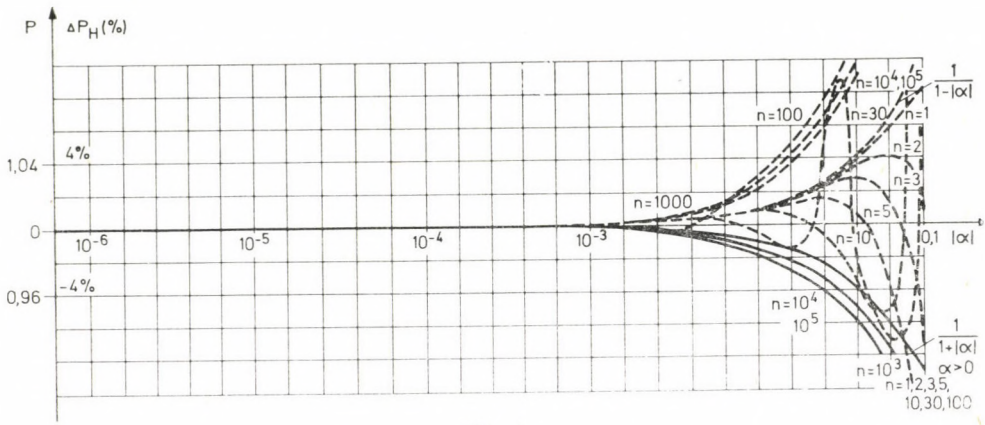


Fig. 8

is also limited. For this reason we have worked on the continued calculations to the accuracy

$$\frac{\varepsilon}{|\alpha|} = 10^{-4}. \tag{67}$$

(In a more detailed study a further accuracy analysis does not seem to be unnecessary.)

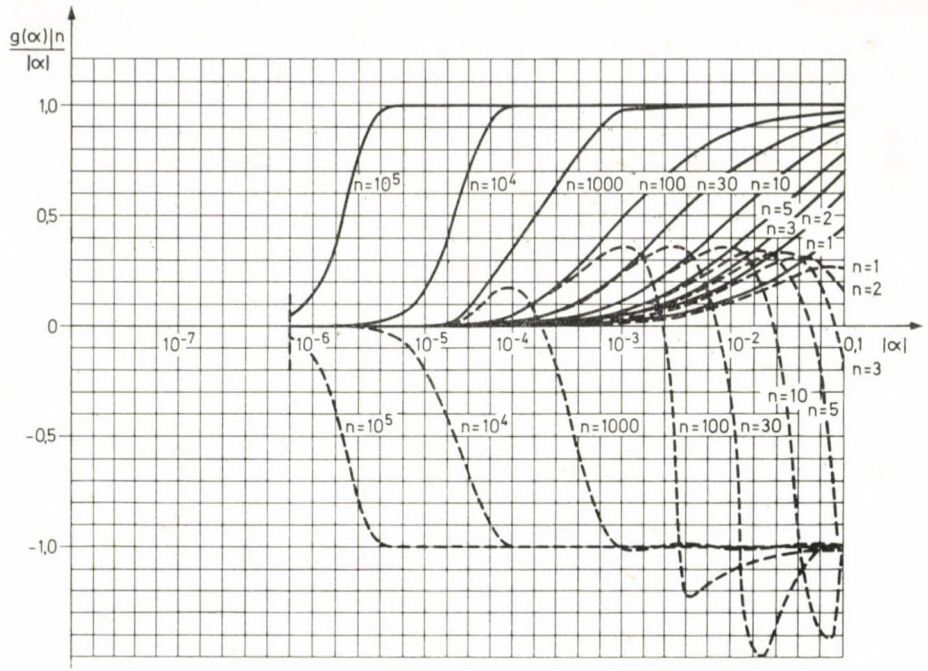

g(α)|n
|α|
1,0
0,5
0
-0,5
-1,0
10⁻⁷
10⁻⁶
10⁻⁵
10⁻⁴
10⁻³
10⁻²
0,1 |α|
n=10⁵
n=10⁴
n=1000
n=100
n=30
n=10
n=5
n=3
n=2
n=1
n=1
n=2
n=3
n=10
n=5
n=10⁵
n=10⁴
n=1000
n=100
n=30

Fig. 9

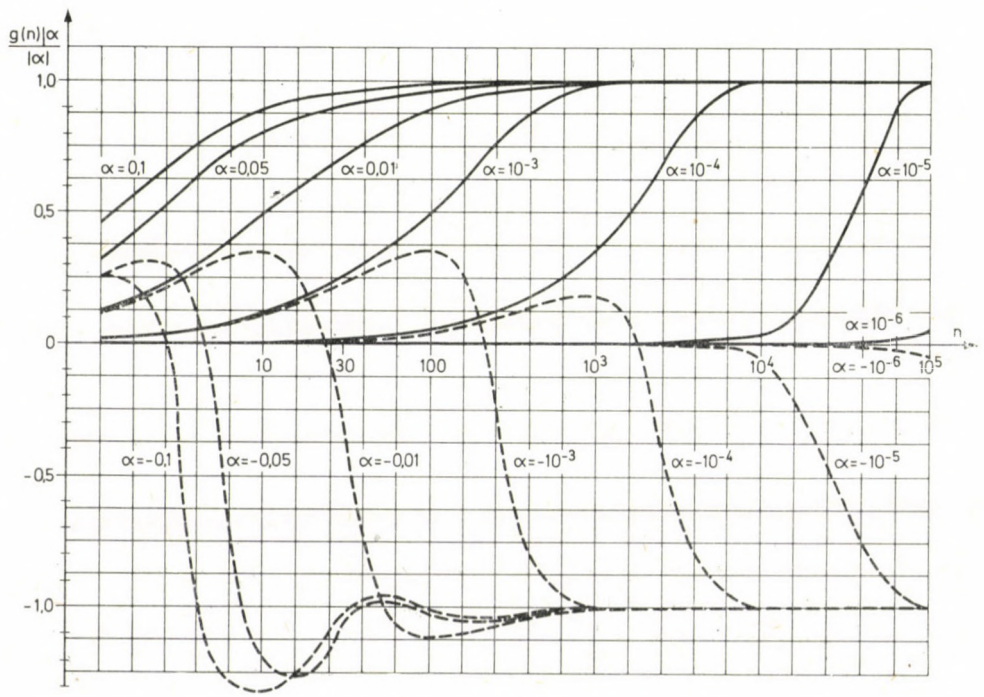

g(n)|α
|α|
1,0
0,5
0
-0,5
-1,0
α=0,1
α=0,05
α=0,01
α=10⁻³
α=10⁻⁴
α=10⁻⁵
α=10⁻⁶
α=-10⁻⁶
10
30
100
10³
10⁴
10⁵
n
α=-0,1
α=-0,05
α=-0,01
α=-10⁻³
α=-10⁻⁴
α=-10⁻⁵

Fig. 10

Fig. 9 shows the frequency shift normalized to the basic effect as a function of the basic effect $|\alpha|$ for both $\alpha > 0$ and $\alpha < 0$ using as parameter the "medium parameter" n.

Fig. 10 shows the frequency shift normalized to the basic effect as a function of the "medium parameter" n using as parameter the value of α.

It is an important result, for further use, that the curves are of

$$s(x) = [1 - e^{-f(x)}] \tag{68}$$

character and can well be approximated by $s(x)$, especially in phenomenon identifying analyses.

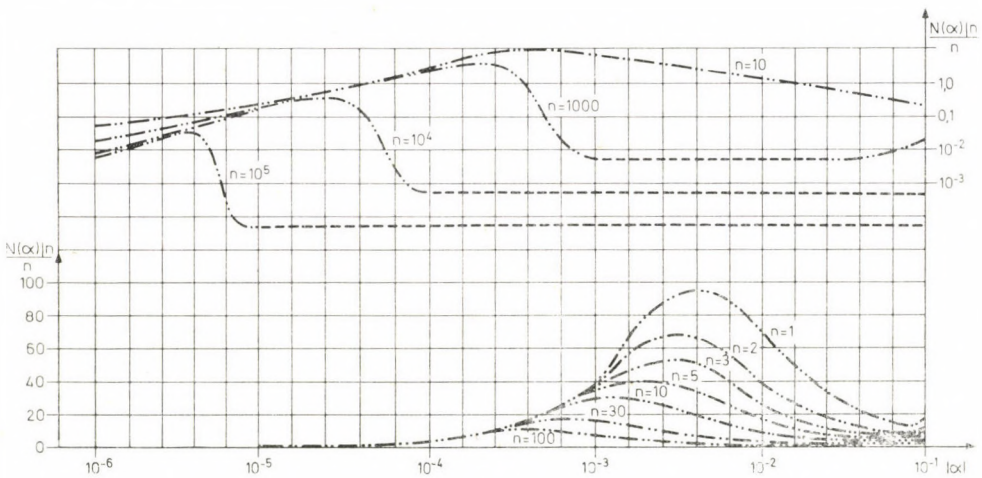


Fig. 11

Fig. 11 shows the values of N/n characteristic of the band-width as a function of $|\alpha|$ using n as parameter. In this case the more general result is important according to which in moving media line broadering occurs in the section coincident with the range of the transition of $s(x)$ in (68) from 0 to 1 if x corresponds to $|\alpha|$ and this increases the original band-width by a factor of $\sim 10 \lesssim 100 < 10^3$. N/n does not agree with a bandwidth measurable with *a given receiver system*. For this reason these results cannot be used to a higher accuracy than specified above. More precise analyses can be made by considering, in the light of the foregoing results, the original spectrum shape together with the receiver system used. For the present purpose the spectrum analysis can now be regarded as closed.

II./4. Determination of the average frequency shift

The "anomalous" or in other words the red shifts of non-cosmological and of non-gravitational origin have been considered until recently as an interesting problem and subject to rather peculiar interpretations [55, etc.]. For this reason it is thought to be of interest to see the change in signal frequency during propagation from transmitter to receiver, both fixed in K for cases stationary in K (non-stationary in K') in the above discussed sense.

For the time being, we are compelled to use the method of "relativistic ray tracing" which has been successfully applied for nearly a decade [13—18]. Thus, the measurable frequency shift can be expressed as the product of the "basic effect" multiplied by a "weighting function". Consequently

$$\frac{\widetilde{\Delta f}}{f} \cong \alpha[\bar{v}(\bar{r}),\ \epsilon(\bar{r}),\ \text{etc.}] \cdot s[\epsilon(\bar{r}),\ n,\ \text{etc.}]. \tag{69}$$

From the two factors in (69) α can be determined according to the procedure detailed in the foregoing, while the determination of s will be specifically dealt with later.

a) Calculation models for the determination of α

The value of α was analyzed in detail for a number of models. The analysis was not continued past an accuracy of v^2/c^2 which was thought to be sufficient for phenomenon identification by the usually formulated (non-modified [20]) ray tracing method. The results published up till now to illustrate the application of the relativistic ray tracing method, were evaluated to an accuracy of up to v^2/c^2. These calculations were reported in detail earlier [13—19] and now they will only be briefly summarized. As it has already been mentioned, at this point the considerations are restricted to the determination of the basic effect α on a signal propagating in an inhomogeneous moving medium of stationary configuration in K from a transmitter to a receiver both fixed in the K system. Other effects — due to gravitation, source movement, turbulences, etc. — are disregarded in the calculation models. They are taken into account only in some of the interpretations, to the extent to which the effect has been admitted and clarified to date.

a.1. Co-moving inhomogeneous media of plane parallel structure [19]: The medium is now described as

$$\bar{v} = \overline{\text{constant}} \qquad \text{and} \tag{70}$$

$$\bar{v}\ \overline{\text{grad}}\ \varepsilon_{ik} = 0$$

where ε' is the permittivity in K'; only in this manner can we have a picture which appears as stationary in K. Choosing the axes $x \,||\, x'$ to be parallel to $\bar{v}$, we find

$$f \cong f' \left(1 + \cos \alpha'_x \cdot n_\varepsilon \frac{v}{c} + \frac{1}{2} \frac{v^2}{c^2}\right),$$

$$\cos \alpha_x \cong \cos \alpha'_x + (1 - \cos^2 \alpha'_x)\left[\frac{1}{n_\varepsilon} \frac{v}{c} + \frac{1}{2} \cos^2 \alpha'_x \frac{v^2}{c^2}\left(1 - \frac{3}{n_\varepsilon^2}\right)\right], \tag{71}$$

where n_ε is the refraction index in K' — if $\boldsymbol{\epsilon}' = \varepsilon' \mathbf{1}$, $\boldsymbol{\mu}' = \mathbf{1}$, $\boldsymbol{\varkappa}' = \boldsymbol{\nu}' = \mathbf{0}$ and $v^i/c^i \cong 0$ for $i \geq 3$, then $n_\varepsilon = \sqrt{\varepsilon'}$; α_x is the measurable angle of the direction of propagation to the x-axis. The axes z and z' can be taken to lie without interruption parallel to the plane parallel layers.

In this simple case n_ε has the form corresponding to the law of refraction-reflection

$$\frac{n_{i+1\varepsilon}}{n_{i\varepsilon}} = \frac{\sin \vartheta_{i+1}}{\sin \vartheta_i}$$

where ϑ is the angle of incidence according to the usual definition.

A detailed analysis of both the propagating and the reflected waves shows that independent of the magnitude and the direction of $\bar{v}$ and of the structure of ε' (disregarding the effect of plane parallelism) we find

$$\alpha \equiv 0 \quad \text{and thus} \quad \Delta \tilde{f} = 0 \,. \tag{72}$$

a.2. Non-comoving inhomogeneous media of plane-parallel structure [16, 19]:

The medium under consideration is taken as consisting of N layers with plane-parallel interfaces and their velocities in a direction parallel to the interface and differing from one another in magnitude, and possibly in their direction. For each of the layers it is assumed that within the layer

$$v_i = \text{constant} \quad \text{and} \quad n_{i\varepsilon} = \text{constant} \,.$$

Let us take the medium to be isotropic in the co-moving K'_i system with $\boldsymbol{\epsilon}'_i = \varepsilon'_i \mathbf{1}$, $\boldsymbol{\mu}'_i = \mathbf{1}$, $\boldsymbol{\varkappa}'_i = \boldsymbol{\nu}'_i = \mathbf{0}$. The sign of v_i may, of course, change arbitrarily from layer to layer. Then calculating in the same way as (71) taking ϑ to be the angle of incidence into the first layer in K, we get

$$\alpha = -\frac{n_{0\varepsilon}^2 \cos^2 \vartheta}{c^2} \sum_{i=1}^{N} \left\{ (v_{i-1}^2 - \bar{v}_i \, \bar{v}_{i-1}) \frac{n_{i-2\varepsilon}^2 - 1}{n_{i-2\varepsilon}^2} + \right.$$

$$\left. + (\bar{v}_{i-1} - \bar{v}_i) \sum_{k=1}^{i-2} \frac{n_{k-1\varepsilon}^2 - n_{k\varepsilon}^2}{n_{k\varepsilon}^2 \, n_{k-1\varepsilon}^2} \bar{v}_k \right\} \tag{73}$$

where v_i contains the sign, too. (73) gives

$$\alpha \sim \frac{v_i^2}{c^2}$$

which means than in many cases encountered in practice, we can calculate with $\alpha \cong 0$. For higher values of v_i (73) can give within the accuracy limits determined by (71)

$$\alpha \gtreqless 0 .$$

It can also be seen [19] that for v_i with an alternating sign — disregarding special cases — a red shift is observed which can be described as

$$\alpha \cong p_{\mathrm{II}}(n_\varepsilon, v_i) \frac{\tilde{v}^2}{c^2} < 0 . \tag{73a}$$

The form of $p_{\mathrm{II}}(n_\varepsilon, v_i) < 0$ and its value greatly depends on the total internal structure of the inhomogeneous medium.

a.3. Effect of expanding, radial flow [16, 19]:

The model is of importance in the study of stellar atmospheres, solar corona, galaxies of explosive nature, quasars, etc.

The analysis was made for the arrangement shown in Fig. 12. (Of course, a far more precise model can be constructed, if necessary, while it is sufficient to use the "equivalent-mirror arrangement" of Fig. 3.a for crude analyses.) The result contains coefficients of a highly complicated structure [19] taking the form

$$\alpha = p(n_{i\varepsilon}, \beta) \frac{v}{c} + q(n_{i\varepsilon}, \beta) \frac{v^2}{c^2} .$$

The analysis of the earlier specified [19] coefficients for the given model yields the values of p and q. For a given measurement the model permits a limit to be set at which the first jump ($n_{0\varepsilon} \to n_{1\varepsilon}$) is to be expected and from which it is possible to define the points of entrance and exit with the associated central

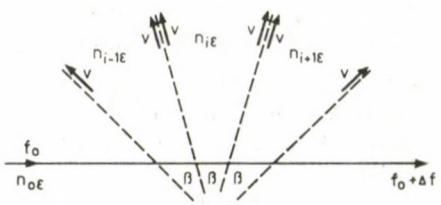


Fig. 12

angle γ (Fig. 3.a). Trivially, we get

$$\text{for } \gamma = 0^\circ \qquad p = 0,\ \alpha \sim 0\ ;$$

$$\text{for } n_{0\varepsilon} \equiv n_{i\varepsilon} \qquad \alpha \equiv 0\ .$$

In the cases of $n_{i\varepsilon} \neq n_{0\varepsilon}$, $\gamma \neq 0^\circ$ not only the values of $|\,p\,|$, $|\,q\,|$ are determined by the structure of the inhomogeneity, the properties of the medium (the medium is described by $n_\varepsilon \lesseqgtr 1$ or ϵ, μ, etc.) but also a γ dependent change of sign in p and q occur for some inhomogeneous mediums!

As in the following the theoretical results will be applied to the solar environment (Solar occultation $R \geqq 1\ R_\odot$, Solar limb effect etc.) let us see the values obtained from the analysis of this case [19] as

$$p(n_{i\varepsilon}, \beta) < 0\ ,\ q(n_{i\varepsilon}, \beta) > 0\ . \tag{75}$$

Calculating, in addition, with the refraction coefficients occurring for the corona and the solar surface [13, 56, 57], we find

$$|\,p\,| \sim |\,q\,| \sim 1 \tag{76}$$

Thus, for $v \ll c$

$$\alpha \cong -|\,p\,|\frac{v}{c} \quad \text{and} \quad \alpha_\odot \approx -\frac{v}{c}\ . \tag{77}$$

It is *important* that we get a *red shift* for radial expansion. (Of course, for radial *contraction* a blue shift of similar nature occurs.) It has also to be noted that p can change its sign after a $\gamma_{\max}$ in cases when γ varies with the structure over the $0 \to \pi$ range (e.g. galactic transillumination). In a plasma ($n_\varepsilon < 1$) the redshift for $\gamma < \gamma_{\max}$ can transform to a blueshift for $\gamma > \gamma_{\max}$ and also other effects can occur (total scattering, change of sign in $\widetilde{\Delta f}$ from layer to layer, etc.). However, it seems probable that we can calculate with (75) in a wide range of real phenomena.

a.4. Co-moving inhomogeneous medium of non-strictly plane parallel structure [17—19]:

Owing to the structure of the medium, in this case the wave-path is similar to that shown in Fig. 3.b seemingly stationary in the K-system. This leads, in non-strictly plane-parallel cases, always to the dispersion of the incident electromagnetic signal [20] whereas the phenomenon may still appear to be stationary in the laboratory system [1, 20]. If ϑ is taken to be an angle of the incident signal propagation vector to the medium velocity vector in the plane of the medium surface at the point of incidence, then we have for

$$\left(\frac{v}{c}\right)^i \cong 0\ , \quad i \geq 2$$

and for the "reflected" component of the signal

$$\alpha \cong -A \frac{v}{c} \cos \vartheta \tag{78}$$

where A varies with the geometry of the spatial arrangement and with the internal structure of the medium. Using the above notations and introducing $n_{0\varepsilon}$ to characterize the refraction index from the "exit" point of the signal, which on entering a vacuum is 1, the signal component "crossed" through the medium is given by

$$\alpha = \left\{ \frac{A[\epsilon'(\bar{r}'), \cos \vartheta, \text{etc}]}{n_{0\varepsilon}} - \cos \vartheta \right\} \frac{v}{c} . \tag{79}$$

In the case of the strict internal structure specified in a.1 the first term of the factor in the brackets [] of (79) gives cos ϑ and thereby (72) is reproduced.

a.5 Rotating medium of spherical symmetry [15].

A formally approximate relationship was formulated for the study of dense, strongly inhomogeneous planetary atmospheres rotating around their axis. The data of the signal source (space vehicle, etc.) are indexed by S, those of the observer (Earth, etc.) by E and those of the inhomogeneous medium by A, while "in" and "out" stand for the measurable data at the points of entrance and exit of the signal to and from the medium, respectively, with "rp" standing for points along the "ray path" which can be defined [50, etc.]. The angles are throughout the usual (optically) understood angles (of incidence, etc.), n_ε is the refraction index. For simplicity's sake the velocities due to the rotation around the axis are not converted to angular velocities for the arrangement similar to that of Fig. 3.a, but without expansion and contraction. Of course, the conversion is always possible. As laboratory system we have chosen an arrangement of spherical symmetry fixed at its centre (to the centre of rotation).

If the "atmosphere" rotates, we have

$$\alpha \cong \cos \vartheta_{\text{in}} \frac{|\bar{v}_S - \bar{v}_{A\text{in}}|}{c} + \cos \vartheta_{\text{out}} \frac{|\bar{v}_{A\text{out}} - \bar{v}_E|}{c} +$$

$$+ g[\cos \vartheta_{ln}, n_\varepsilon(\bar{r}_{rp})] \frac{|\bar{v}_{A\text{in}} - \bar{v}_{A\text{out}}|}{c} +$$

$$+ \frac{1}{c} \frac{\partial}{\partial t} \int_{\text{in}|rp}^{\text{out}} [n_\varepsilon(\bar{r}_{rp}) - 1] \, ds = \sum_{i=1}^{4} \alpha_i \tag{80}$$

where ds means integration over the arc length.

For "non-rotating" atmosphere, we write

$$\frac{\Delta f}{f_0} \cong \cos\vartheta_{\text{in}} \frac{|\bar{v}_s - \bar{v}_E|}{c} + \Delta(\cos\vartheta) \frac{|\bar{v}_s - \bar{v}_E|}{c} + \\ + \frac{1}{c}\frac{\partial}{\partial t}\int_{\text{in}|rp}^{\text{out}} [n_\varepsilon(\bar{r}_{rp}) - 1]\, \mathrm{d}s\,. \tag{81}$$

It can be seen that the second term in (81) arises from the effect of the "curved ray path" and the third term from the change in the phase path due to $n_\varepsilon \neq 1$ in the case of a "radio occultation" experiment. In the case of large atmospheric velocities, fast rotation, etc. expression (81) is of insufficient accuracy, as observed several times in practice [58, etc.]. In this case approximation (80) has to be used.

It is important to remember that it is the concrete arrangement, the actual model which determines whether or not the total expression or only the third term in (80) has to be multiplied according to (69) by any weighting function $s[\epsilon, n,$ etc.]. In most of the actual cases

$$\frac{\widetilde{\Delta f}}{f} \cong \alpha_1 + \alpha_2 + \alpha_3 s[\epsilon, n, \ldots] + \alpha_4\,. \tag{80a}$$

a.6 Now we have seen the results obtained for α in the cases of practical importance and elaborated up to present. It is worth mentioning that α is fundamentally dependent on the geometry of the "arrangement". It shows a frequency dependence through ϵ, μ, etc. where the values $\Delta\epsilon_i$, etc. are usually more important than the actual value (if it is not 1, 0 or ∞). As, according to Figs 9 and 10, the value of $\widetilde{\Delta f}/f \cong \alpha$ (or 0) for a wide range of given parameter values, the frequency dependence of $\widetilde{\Delta f}/f$ is in many cases very slight, even negligible. (Possible resonances may be responsible for "surprising" anomalies in this dependence). A detailed analysis of given cases is a further task, though the applicability of the approximation follows from the foregoing.

b) The weighting function

For applicability it is important to determine the weighting function. Two independent ways are available for the analysis of $s(\epsilon, n, \ldots)$: physical considerations and evaluation of calculated spectra from this point of view.

b.1 Physical considerations [13, 16, 19]

In this approximation part of the incident energy is taken to be continuously received and retranslated at a frequency $f_0(1 + \alpha)$ by the medium, while the other fraction of the incident energy passes through the medium

without interaction. If N is the characteristic density (neutral particle density, in a plasma the electron or charge density, etc.) of the medium, then the fraction of energy taking part in the interaction is given by

$$W_\alpha = W[1 - e^{-\int_{rp}\sigma N(s)ds}] ,$$

while the residual fraction of the total energy W, transmitted in the desired direction, is given by

$$W_0 = We^{-\int_{rp}\sigma N(s)ds}$$

Since the equations describing the medium, determining the values of ϵ, etc. as well as the expected values of α, are known, it seems justified to assume that over the total range of the generated spectrum $\sigma = \sigma_0 =$ constant. The determination of its value for specific media is a task of interpretation. Thus,

$$\frac{\widetilde{\Delta f}}{f_0} = \alpha[1 - e^{-\sigma_0\int_{rp}N(s)ds}] . \tag{82}$$

b.2 Results of spectrum analyses

The relation (68) already shows how to approximate the results we are interested in, which are shown in Fig. 10. n is a function of the structure of the "set of mirrors", it is the result of the interaction between the signal and the medium. According to spectrum analyses, we find

$$\frac{\widetilde{\Delta f}}{f_0} = \alpha[1 - e^{-\sigma_n n}] . \tag{83}$$

Thus, we get identical results from (82) and (83) for the shape of the weighting function. The fact that

$$n \sim \int_{rp} N(s)ds$$

is a justified and allowed approximation will not be discussed here in particular. It follows from the detailed analysis of individual cases, as it can be simply checked by any analyst.

The approximations (69) and (82) are thought to be useful for the determination of $\widetilde{\Delta f}/f_0$ until the availablity of the precise calculation procedures anticipated in chapter I or of the precise spectrum evaluations applying the Huygens principle or some analogous procedure described in chapter II.

It may be pointed out here that if for the total phenomenon

$$s(\varepsilon, n, \ldots) \sim [1 - e^{-\sigma_0\int_{rp}Nds}] \cong 1 \tag{84}$$

holds over the range of analysis with respect to independent variables q_i, then

$$\frac{\Delta f}{f_0} = \alpha(q_i) \tag{85}$$

and that this relation is generally frequency independent in a wide range of frequencies according to the analysis models in point II/4.a.

REFERENCES

1. W. P. Birkemeier, H. S. Merril, D. H. Sargeant, D. W. Thomson, C. W. Beamer, G. T. Bergmann: Observation of Wind-Produced Doppler-Shifts in Tropospheric Propagation; Radio Science; **3,** 309, 1968.
2. M. G. Adam: Interferometric Measurements of Solar Wavelength and an Investigation of the Einstein Gravitational Displacement; Mon. Not. Roy. Astron. Soc; **108,** 446, 1948.
3. L. A. Higgs: The Solar Red-Shift; Mon. Not. Roy. Astron. Soc; **121,** 421, 1960.
4. D. Sadeh, S. H. Knowles, B. S. Yaplee: Search for a Frequency Shift of the 21 centimeter Line from Taurus A. Near Occultation by the Sun; Science; **159,** 307, 1968.
5. R. M. Goldstein: Superior Conjunction of Pioneer 6; Science; **166,** 598, 1969.
6. J. A. Jacobs, T. Watanabe: Doppler Frequency Changes in Radio Waves Propagating Through a Moving Ionosphere, Radio Science, **1,** 257, 1966.
7. M. K. Bird: Coronal Tranzient Events Observed with S-Band Faraday Rotation Measurements during Solar Occultation; Space Res., **XVI,** 711, Akademie-Verlag, Berlin, 1976.
8. R. Woo, F. C. Yang, A. Ishimaru: Structure of Density Fluctuations Near the Sun Deduced from Pioneer-6 Spectral Broadening Measurements; The Astrophys. J; **210,** 593, 1976.
9. J. W. Kierein, B. M. Sharp: Compton Effect Interpretation of Solar Red-Shift; Solar Phys; **3,** 450, 1968.
10. P. Merat, J. C. Pecker, J. P. Vigier: Possible Interpretation at an Anomalous Redshift Observed at the 2292 MHz Line Emitted by Pioneer-6 in the Close Vicinity of the Solar Limb; Astron. Astrophys; **30,** 167, 1974.
11. A. A. Chastel, J. F. Heyvaerts: Perturbations of Pioneer 6 Telemetry Signal During Solar Occultation; Nature; **249,** 21, 1974.
12. A. A. Chastel, J. Heyvaerts: Broadening and Anomalous Shift of Pioneer VI Telemetry Line an Effect of Coronal Inhomogeneities Useful for Diagnostics; Astron. Astrophys; **51,** 171, 1976.
13. Cs. Ferencz, Gy. Tarcsai: A New Experimental Possibility of Investigating the Solar Corona: Frequency Measurements on Radio Sources When Occultated by the Sun; Planet. Space Sci; **18,** 1213, 1970.
14. Cs. Ferencz, Gy. Tarcsai: Redshift during Pioneer-6 Solar Occultation — Unexplained or Predicted; Nature; **252,** 615, 1974.
15. Cs. Ferencz, Gy. Tarcsai: Frequency Shift Effects due to Atmospheric Motions in Interplanetary Occultational Measurements; Space Res., **XVI.**, 705, Akademie-Verlag, Berlin, 1976.
16. Cs. Ferencz, Gy. Tarcsai: Theoretical Explanation of the Solar Limb Effect; Planet. Space. Sci; **19,** 659, 1971.
17. Cs. Ferencz, Gy. Tarcsai: Interaction of Gravitational and Electromagnetic Fields or Another Effect?: Nature; **233,** 404, 1971.
18. Cs. Ferencz, Gy. Tarcsai: Refraction Effects due to Moving Media in Doppler Measurement; Space Res; **XII,** 595, Akademie-Verlag, Berlin, 1972.
19. Cs. Ferencz: Electromagnetic Wave Propagation in Inhomogeneous Linear Media; Thesis submitted for the degree of Candidate of Sciences, Budapest, 1970, Library of the Hung. Ac. Sci. (in Hungarian).
20. Cs. Ferencz: Electromagnetic Wave Propagation in Inhomogeneous Media: The Analysis of the Rotation of Polarization, and the Application of the Principle of Modified Ray Tracing; Acta Techn. Ac. Sci. H., **86,** (1978), 363 and **87,** (1978), 263
21. D. Middleton: A Statistical Theory of Reverberation and Similar First-Order Scattered Fields-Part. III; IEEE Trans. on Inf. Theory; **IT-18,** 35, 1972.

22. Cs. FERENCZ: Electromagnetic Wave Propagation in Inhomogeneous Media: Strong and Weak Inhomogeneities: Acta Techn. Ac. Sci. H., **85,** 433, 1977.
23. Cs. FERENCZ: Electromagnetic Wave Propagation in Inhomogeneous Media: Method of Inhomogeneous Basic Modes; Acta Techn. Ac. Sci. H., **86,** 79, 1978.
24. Cs. FERENCZ: Electromagnetic Wave Propagation: The Analysis of the Group Velocity; Acta Techn. Ac. Sci. H., **86,** 169, 1978.
25. Cs. FERENCZ: A Geometric Resolution of the Contradiction between the Propagation of Electromagnetic Plane Waves in Moving Dielectrics and the Einsteinian Addition of Velocities; Acta Techn. Ac. Sci. H., **84,** 147, 1977.
26. Cs. FERENCZ: Permittivity of Inhomogeneous Media; Proc. of the Fourth Coll. on Micr. Com; **ET-10,** Akadémiai Kiadó, Budapest, 1970.
27. I. E. TAMM: Osznovi Teorii Elektritsestva, Izd. „Nauka", (in Russian) Moscow, 1966.
28. J. J. BRANDSTATTER: An Introduction to Waves, Rays and Radiation in Plasma Media; McGraw-Hill Book Co. Inc. New York, 1963.
29. K. NOVOBÁTZKY: The Theory of Relativity; Tankönyvkiadó (in Hungarian), Budapest, 1951.
30. A. EINSTEIN: Ann. der Physik; **17,** 891, 1905. („Válogatott Tanulmányok" — in Hungarian, Gondolat, Budapest, 1971).
31. J. L. SYNGE: Relativity, the Special Theory; North-Holland Publ. Co. Amsterdam, 1965.
32. M. V. LAUE: Die Relativitätstheorie, I; Vieweg und Sohn; Braunschweig, 1955.
33. H. OTT: Zum Energie-Impuls Tensor der Maxwell-Minkowskischen Elektrodynamik; Ann. der Physik, (6), **11,** 33, 1952.
34. F. BECK: Die Allgemeingültigkeit des Trägheitsgesätzes der Energie in der Planckschen Fassung; Z. für Physik; **134,** 136, 1953.
35. G. MARX: Relativistische Elektrodynamik der Magnete; Acta Phys. Ac. Sci. H., **II,** 67, 1952.
36. G. MARX: Das Elektromagnetische Feld in Bewegten Anisotropen Medien; Acta Phys. Ac. Sci. H., **III,** 75, 1953.
37. H. ARZELIES, de J. HENRY: Milieux Conducteurs on Polarisables en Mouvement; Travaux de l'Institut Scientifique Chérifien, Série Sciences Physiques; No. **5,** Rabat, 1959.
38. E. SCHMUTZER: Relativische Physik (Klassische Theorie); B. G. Teubner Verlagsgesellschaft, Leipzig, 1968.
39. R. P. FEYMANN, R. B. LEIGHTON, M. SANDS: The Feynmann Lectures on Physics, II; (in Hungarian) Műszaki Könyvkiadó, Budapest, 1970.
40. L. D. LANDAU, E. M. LIPSIC: Theoretitseskaia Fizika II; Teoria Polia; (in Hungarian) Tankönyvkiadó, Budapest, 1976.
41. W. P. ALLIS, S. J. BUCHSBAUM, A. BERS: Waves in Anisotropic Plasmas; M. I. T. Press, Cambridge, Mass., 1963.
42. I. V. LINDELL: On the Definiteness of the Constitutive Parameters of a Moving Anisotropic Medium; Proc. IEEE; **60,** 638, 1972.
43. J. A. ARNAUD, A. A. M. SALEH: Theorems for Bianisotropic Media; Proc. IEEE; **60,** 639 1972.
44. J. A. KONG, D. K. CHENG: Wave Reflections from a Conducting Surface with a Moving Uniaxial Sheath; IEEE Trans. on Ant. and Prop; **AP-16,** 577, 1968.
45. J. A. KONG, D. K. CHENG: Reflection and Refraction of Electromagnetic Waves by a Moving Uniaxially Anisotropic Slab; J. of Appl. Phys; **40,** 2206, 1969.
46. S. W. LEE, Y. T. LO: Radiation in a Moving Anisotropic Medium; Radio Science; **1,** 313, 1966.
47. J. A. KONG, D. K. CHENG: Modified Reciprocity Theorem for Bianisotropic Media; Proc. IEEE; **117,** 349, 1970.
48. D. CENSOR: Ray Tracing in Weakly Nonlinear Moving Media; J. Plasma Phys; **16,** 415, 1976.
49. M. E. KARPUK, W. G. TIEDERMAN: Effect of Finite-Size Probe Volume upon Laser Doppler Anemometer Measurements; AIAA Journal; **14,** 1099, 1976.
50. K. G. BUDDEN: Radio Waves in the Ionosphere; Cambridge at the Univ. Press; 1966.
51. L. B. FELSEN: Rays, Modes and Equivalent Networks; Proc. of the Fourth Coll. on Micr. Com; **ET-9,** Akadémiai Kiadó, Budapest, 1970.
52. R. CRISTESCU, G. MARINESCU: Introduction to the Theory of Distributions and its Applications; (in Hungarian) Műszaki Könyvkiadó, Budapest, 1969.
53. Cs. FERENCZ: Wave Propagation in Inhomogeneous Linear Media; Acta Techn. Ac. Sci. H; **68,** 215, 1970.
54. R. S. LAWRENCE, D. J. PASAKONY: A Digital Ray-Tracing Program for Ionospheric Research Space Res. **II**; 253, North-Holland Publ. Co; Amsterdam, 1961.
55. J. C. PECKER: Possible Explanations of Non Cosmological Redshifts; College de France Institute d'Astrophysique de Paris; (preprint) 1976.
56. A. ZIRIN: The Solar Atmosphere; Blaisdell, Waltham, Mass., 1966.

57. D. E. BILLINGS: A Guide to the Solar Corona; Academic Press, New York, 1966.
58. G. FJELDBO: Mariner 10 and Pioneer 10 and 11 Radio Occultation Measurements of Planetary Ionospheres; COSPAR XVIII. Plen. Meet; **VII. 4.5,** Varna, 1975.
59. E. H. SCHRÖTER: Zur Deutung der Rotverschiebung und der Mitte-Rand-Variation der Fraunhoferlinien bei Berücksichtigung der Temperaturschwankungen der Sonnenatmosphere; Zeitschrift für Astrophysik; **41,** 141, 1957.
60. E. N. PARKER: Coronal Expansion and Solar Corpuscular Radiation; Proc. of the Plasma Space Sci. Symp; 99, Reidel, Dortrecht, 1965.
61. I. A. BALL, D. F. DICKINSON, A. E. LILLEY, H. PENFIELD, T. I. SHAPIRO: Search for an Effect of the Sun on the Frequency of 18-Centimeter Radiation; Science; **167,** 1955, 1970.
62. V. R. ESHLEMAN: Jupiter's Atmosphere: Problem and Potential of Radio Occultation; Science, **189,** 876, 1975.
63. A. G. SOLOVJEV, A. D. KUZMIN, A. N. KAZANTSEV: "Pioneer 10" and Radio Astronomical Measurements of Jupiter; COSPAR XVIII. Plen. Meet; **VII 4.2,** Varna, 1975.
64. M. OHKUBO: Reflection and Transmission of Electromagnetic Waves by Anisotropic Plasma Half-Space Moving in Longitudinal Direction; IEEE Trans. on Ant. and Prop; **AP-19,** 569, 1971.
65. K. DAVIES: The Measurements of Ionospheric Drifts by Means of a Doppler Shift Technique; J. of Geophys. Res; **67,** 4909, 1962.
66. K. DAVIES: Doppler Studies of the Ionospheric Effects of Solar Flares; Proc. of the Int. Conf. on The Ionosphere, 1962; 76, Chapman and Hall Ltd; London, 1963.
67. M. G. ADAM, P. A. IBBETSON, A. D. PETFORD: The Solar Limb Effect: Observations of Line Contours and Line Shifts; Mon. Not. Roy. Astr. Soc., **177,** 687, 1976.
68. A. A. CHASTEL: A Critical Analysis of the Explanation of Red-Shifts by a New Field; Astron. Astrophys., **53,** 67, 1976.
69. G. L. JAMES: Geometrical Theory of Diffraction for Electromagnetic Waves; IEE Electromagnetic Waves Series 1; Peter Peregrinus Ltd. London, 1976.
70. L. F. BURLAGA: Magnetic Fields, Plasmas and Coronal Holes: The Inner Solar System; **SI.1.3.** XXI. COSPAR Plen. Meeting, Innsbruck, 1978.

Fortpflanzung von elektromagnetischen Wellen in bewegten Medien mit besonderer Berücksichtigung der Frequenzänderungen. (Anomalistische Frequenzverschiebungen in der Astronomie). II. Teil. Relativistisches Strahlenverfolgungsverfahren. — Im zweiten Teil der Arbeit wird die ausgearbeitete Methode, die »relativistische Strahlenverfolgung« eingehend diskutiert. Es wird das Verfahren für die Bestimmung des Signalspektrums gezeigt. Der Grundeffekt für die am wichtigsten scheinenden Fälle wird angegeben. Die Gewichtsfunktion wird angegeben und deren Form wird begründet.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (1—2), pp. 59—85 (1980)

HYDRAULISCHE GRUNDLAGEN DER GRUNDWASSERABSENKUNG DURCH VAKUUMBRUNNEN

G. ÖLLŐS*

DOKTOR DER TECHN. WISSENSCHAFTEN

[Eingegangen am 31. Januar 1979]

Die frühere Kennzeichnung der Vakuumbrunnenhydraulik beruhte in erster Linie auf der Hydraulik des Gravitationsbrunnens. Sowohl die Grundwasserabsenkung als physikalischer Vorgang, wie auch die Grenzbedingungen blieben eigentlich ungeklärt. Der Autor brachte die Gravitations- und die Vakuumhydraulik auf eine gemeinsame theoretische Basis und förderte die Deutung der Grenzbedingungen und der physikalischen Vorgänge. Er stellte eine neue, auch die Wasserführung des Kapillarraumes berücksichtigende Formel für die Ergiebigkeit des Vakuumfassungsstollens auf, woraus auch die Ergiebigkeit des Gravitationsfassungstollens, des einzelnen Vakuumbrunnens und des einzelnen Gravitationsbrunnens abgeleitet werden kann. Auch für eine Näherungsberechnung der Zonen der Wasser- und Luftströmungen abtrenennden Linie wird eine Formel abgeleitet.

1. Einleitung

Die Vakuumbrunnenhydraulik wurde durch die einschlägige Fachliteratur nur in einer ziemlich vereinfachten Form der theoretischen Grundlagen, der tatsächlichen Vorgänge, der Kinematik und Dynamik der Sickerung und der Randbedingungen entwickelt. Den Begriffen des Vakuumbrunnens, des Wasserspiegelabrisses, des Dreiphasenraumes, der Absenkungskurve und der Durchflußmenge sind in erster Linie Forschungsergebnisse zugeordnet, die eine weitere Entwicklung bedürfen. Offensichtlich ist daher, daß bei der Kennzeichnung der Physik, der Vorgänge und der Systeme der Gravitationssickerung eine einen günstigen Einfluß auf die Entwicklung ausübende analytische Anschauung in erhöhtem Grade zur Geltung gebracht werden sollte.

Demgemäß ist der Verfasser bestrebt, unter Berücksichtigung der Förderung der physikalischen Grundlagen und der Randbedingungen, die Gravitations- und Vakuumbrunnenhydraulik auf eine gemeinsame Basis zu bringen und zur Berechnung der die Zonen der Ergiebigkeit des Vakuumbrunnens und der Wasser-Luftströmungen trennenden Kurve theoretische Beziehungen zu erarbeiten.

* Prof. Dr. G. Öllős, Eszék u. 13-15, III. 4, H-1114 Budapest, Ungarn

2. Vergleich des Betriebes der Filter- und Vakuumbrunnen vom hydraulischen Standpunkt

Unsere brunnenhydraulische Anschauung schließt sich, in erster Linie, an die Hydraulik des Filterbrunnens an. Die Fachliteratur erschloß im allgemeinen deren Einzelheiten und ihre praktischen und theoretischen Zusammenhänge. Zunächst erscheint es als zweckmäßig, die aus den Konstruktions- und Betriebsabweichungen des *Filterbrunnens* und des *Vakuumbrunnens* stammenden grundlegenden hydraulischen Eigenarten *miteinander zu vergleichen.*

Der Vakuumbrunnen unterscheidet sich vom Filterbrunnen im wesentlichen dadurch, daß im Innenraum des ersteren ein niedrigerer Druck, als der atmosphärische über dem Brunnenwasserspiegel vorherrscht, was durch den *luftdichten* Abschluß des oberen Teils des Brunnens erreicht werden kann. Folglich befindet sich sowohl im Brunnenwasserraum, als auch in dem an den Brunnenfilter sich anschließenden Luftraum ein niedrigerer Druck, als der atmosphärische (p_0). Bei Ansaugen von Luft wirkt infolge der Luftströmung durch die Bodenporen in der Umgebung des Brunnens ein von oben nach unten gerichteter veränderlicher Luftdruck, d. h., nicht nur die Schwerekraft. Dies bedeutet, daß in solchen Fällen der auf die Identität des Potentials und des Wasserspiegels aufgebaute Gravitationsaspekt nicht ausschließlich angewandt werden kann.

Im folgenden wird der Unterschied zwischen den beiden Grundwasserabsenkungsmethoden in erster Näherung durch die Gegenüberstellung der vereinfachten Schemata des Filterbrunnens und des durch Luft ansaugenden Vakuumbrunnens dargestellt (Bild 1a und 1b).

Bei der Grundwasserabsenkung durch Filterbrunnen und mit Hilfe der im Bild 1 dargestellten Saugpumpe, ergibt sich als einem Absaugen s entsprechende Saughöhe (H')

$$H' = h + z + \Sigma h' \,, \tag{2.1}$$

$$s = h \,,$$

wo: h = den Höhenunterschied zwischen dem ursprünglichen Grundwasserspiegel und der Absenkungsfläche des Brunnenwassers;

z = den Höhenunterschied zwischen dem ursprünglichen Grundwasserspiegel und der Mittellinie der Saugpumpe; und

$\Sigma h'$ = die Summe der zwischen dem Brunnenwasserspiegel und der Mittellinie der Saugpumpe auftretenden durch Verlusthöhe ausgedrückten Strömungsverluste

bedeuten.

Im Falle von einer Grundwasserabsenkung durch *Vakuumbrunnen* (Bild 1b), wird in der Höhe des Brunnenwasserspiegels

$$p \neq p_0 \,,$$

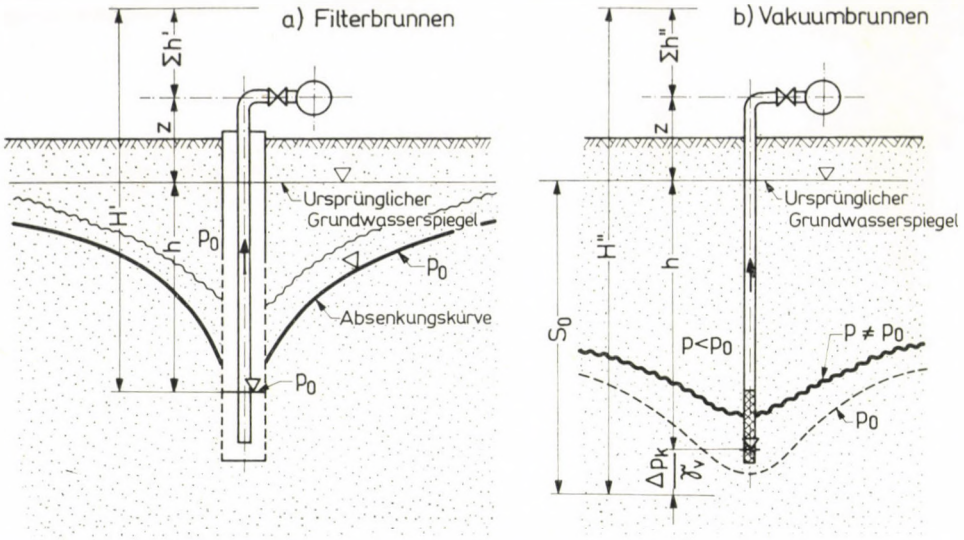


Bild 1. Gegenüberstellung der Betriebe des Filter- und der Gravitationsbrunnens. a) Filterbrunnen; b) Vakuumbrunnen

weshalb das »Absaugen« (s_0) als Summe der Brunnenwasserspiegeldifferenz (h) und der der Druckabnahme entsprechenden Druckhöhe

$$\frac{\Delta p_K}{\gamma_V}$$

zu berücksichtigen ist:

$$s_0 = h + \frac{\Delta p_K}{\gamma_V}, \tag{2.2}$$

wo $\Delta p_K = p_0 - p_t$.

Die Formel der mit Hilfe der Saugpumpe herzustellenden Gesamtdepression (unter alleiniger Berücksichtigung der im Saugrohr stattfindenden Wasserströmung, wodurch das Problem seitens der Sicherheit angenähert wird) lautet:

$$H'' = h + z + \frac{\Delta p_K}{\gamma_V} + \Sigma h'', \tag{2.3}$$

wo Δp_K die in der hydrostatischen Druckhöhe (WS) ausgedrückte Depression bedeutet. Praktisch genommen wird der Wasserspiegel im Brunnen durch das Niveau des bis an das untere Ende des Filterrohres reichenden Saugrohrendes (Bild 1b) bestimmt. Unter diesem Niveau befindet sich die die durch die Gleichung (2.2) definierte Absenkungshöhe s_0 ergebende hydraulische Drucklinie p_0 die im gegebenen Fall unter der Brunnensohle liegt.

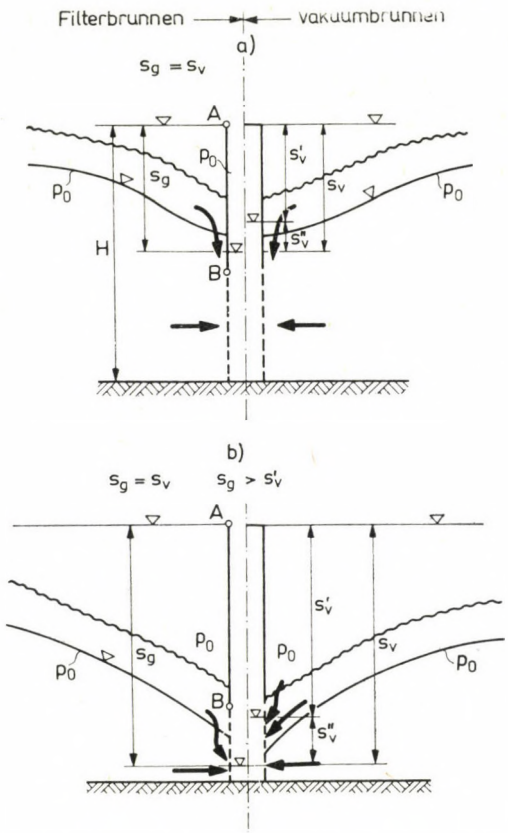


Bild 2. Gegenüberstellung der bodenhydraulischen Verhältnisse in der Umgebung des Filter- und des Vakuumbrunnens

Aus der Gegenüberstellung der Gleichungen (2.2) und (2.3) kann das folgende festgestellt werden: damit der Vakuumbrunnen wirkungsvoller als der Filterbrunnen sei, sollte der Wert s_0 der Absenkung wesentlich höher liegen als der Wert h, d. h., die Bedingung $\Delta p_K/\gamma_v$ sollte erfüllt werden.

Nach der Deutung der Gleichungen (2.1), (2.2) und (2.3) ist es zweckdienlich die Ausgangsfälle der in dem Filter- und Vakuumbrunnen umgebenden Boden auftretenden hydraulischen Verhältnisse zu vergleichen (Bild 2a bis 2d). Zur Gegenüberstellung wurde ein Brunnen bzw. Sickersystem gewählt, dessen konstruktiver Aufbau (d. h., die Filtermantelhöhe, der Durchmesser des Brunnens), die Ergiebigkeitsverhältnisse, die ursprüngliche Wasserschichtdicke und der Boden, der das Grundwasser leitenden Schicht für beide Brunnen gleich sind. Der Unterschied besteht nur darin, daß der Filterbrunnen oben offen, und der Vakuumbrunnen luftdicht abgeschlossen ist. Der Vergleich der beiden Grundfläche kann aufgrund der im Brunnen hervorgerufenen, mit der

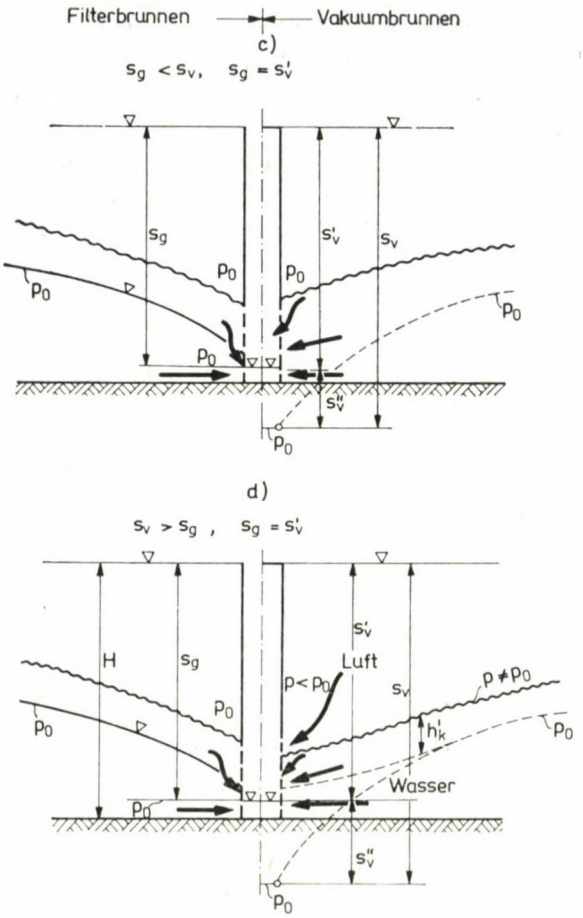


Bild 2 c—d

hydraulischen Druckhöhe (WS) ausgedrückten Absenkung durchgeführt werden.

a⟩ Fällt die Grundwasserabsenkung zwischen den ursprünglichen Grundwasserspiegel repräsentierenden Punkt A und die obere Kante des Filtermantels (Punkt B, Bild 2a), d. h. im Falle, wo die Bedingung s_g, bzw. $s_v < \overline{A-B}$ besteht, so schließt sich die Absenkungskurve an diesem Teil des Mantels an, nachdem zwischen den Punkten A und B der Mantel wasserdicht ist. Im Bereich unter der Absenkungskurve ist die Sickerung zweiphasig.

Über der Absenkungsoberfläche im Kapillarraum fließt das Sickerwasser bei beiden Brunnensystemen auf gleiche Weise, durch den unter der Absenkungsoberfläche befindlichen Raum in den Brunnen. Demzufolge, solange die Absenkungskurve über dem Punkt B sich dem Brunnenmantel anschließt, sind die im Boden vorherrschenden hydraulischen Verhältnisse sowohl bei dem

Filter- als auch bei dem Vakuumbrunnen identisch. Der Unterschied besteht bloß in der Differenz des Brunnenwasserspiegels beider Brunnentypen.

Wenn $s_g = s_v$ und $s_g > \overline{A-B}$, weiters $s_v'' > 0$ ist, (Bild 2b), dann ist die Grundwasserabsenkung bei beiden Systemen gleich. Jedoch gibt es den Unterschied, daß über dem Niveau s_0 sich ein Wasserraum befindet, worin ein Druck $p < p_0$ herrscht, der sich durch den Brunnenfilter dem Kapillarraum anschließt; somit kann das Wasser vom Kapillarraum unmittelbar in den Brunnen strömen. Dies vermindert natürlich auch das Maß des Abrisses, wodurch eine wirksamere Grundwasserabsenkung im Boden hervorgerufen wird.

Bild 2c stellt schon einen wirklichen Vakuumbrunnen dar, wobei $s_v > H$, folglich $s_v > s_g$ ist, jedoch findet hier kein Luftansaugen statt, weil der Kapillarraum über den Brunnenfilter hinaus reicht. Wie auf Bild ersichtlich ist;

— die bei den Filterbrunnen übliche Interpretation der Absenkungskurve p_0 hat keinen Sinn mehr, da diese Kurve unter die Brunnensohle sinkt. Deshalb kann diese Kurve auch nicht durch grobe Annäherung als eine den Sickerraum von oben abgrenzende Strömungslinie betrachtet werden, was sonst bei den Filterbrunnen üblich ist;

— die bei den Filterbrunnen gebräuchliche Näherungsmethode kann deshalb zur Bestimmung und zur Summierung der Wiederherstellung des Kapillar- und Gravitationsraumes nicht verwendet werden, da es sich um einen unmittelbaren Eintritt vom Kapillarraum in den Brunnen handelt;

— die für die Filterbrunnen entwickelten Deutungen können nicht unmittelbar zur Kennzeichnung des Abrisses angewandt werden.

Ein typischer, d. h. ein auch die Luft ansaugender Vakuumbrunnen ist auf Bild 2d dargestellt. Außer den obigen Feststellungen sollten hier in Zusammenhang mit dem Bild 2c die folgenden wesentlichen Eigenheiten erwähnt werden:

— über dem Wasserströmungsraum gibt es auch eine Luftströmung, daher kommt es, daß infolge deren Druckverluste der Wert des von oben wirkenden Drucks veränderlich ist;

— infolgedessen trennt sich die Absenkungskurve vom unteren Teil des Kapillarraumes;

— auf das Grundwasser wirkt außer der Schwerekraft auch eine durch den Luftdruckgradient ausgeübte Kraft;

— und als Ergebnis der gemeinsamen Wirkung der vorerwähnten Faktoren verfügt die den zweiphasigen Wassersickerraum oben abgrenzende Fläche, d. h., im eigentlichen Sinne »der Pumpenwasserspiegel des Vakuumbrunnens« auch über eine Inflexion.

3. Charakterisierung der Grenzbedingungen

Zur Berechnung der zur Lösung des Sickerungsproblems nötigen Charakteristiken (Geschwindigkeit, Druck, Ergiebigkeit), müssen außer der Kenntnis der grundlegenden kinematischen Beziehungen auch die *Grenzbedingungen* bestimmt werden. Bei Annahme einer stationären Strömung müssen die unterschiedlichen Sickersysteme in Abhängigkeit von der Absenkung erläutert werden. Dies erfolgt in Beziehung auf *einen einzigen Brunnen der Brunnenreihe.*

Im Falle eines Gravitationsbrunnens nach dem System der Absenkung im Brunnen über dem unteren Saugrohrende (Bild 3), kann das Wasser wenn auf die innere Fläche des Filterrohres noch der atmosphärische Druck p_0 wirkt, durch die wasserdichte Fläche C_1 nicht in den Sickerraum eintreten und auch nicht daraus entweichen. Deshalb kann der Geschwindigkeitsvektor keine auf eine derartige Fläche senkrechte Komponente haben, d. h., diese Linie ist im Planschnitt eine Stromlinie. Ihre analytische Kennzeichnung ist bei einer stationären Bewegung:

$$\psi = \text{konst.}$$

$$\frac{\partial \psi}{\partial s} = \frac{\partial \varphi}{\partial n} = 0 \,.$$

In den Punkten der den Oberwasserraum mit einer Stärke von H_1 — vom Sickerraum trennenden Grenzlinie C_2, ist das Geschwindigkeitspotential

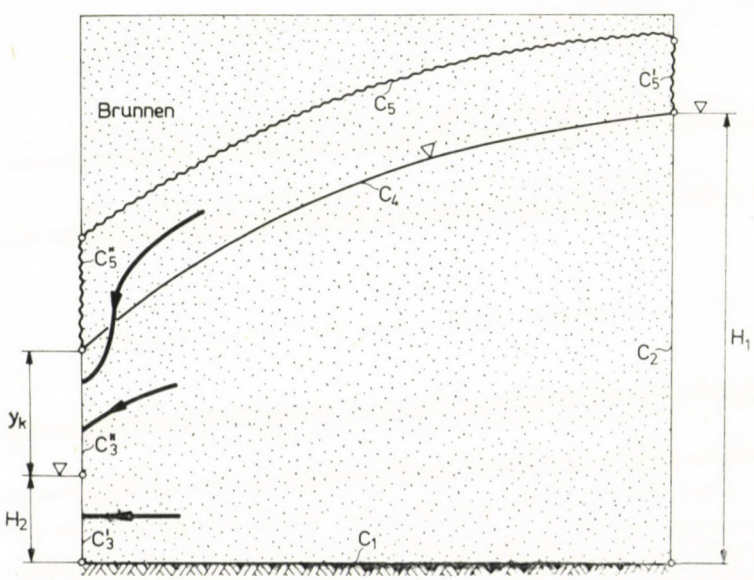


Bild 3. Die Sickerungsgrenzbedingungen bei dem Gravitationsbrunnen

konstant:

$$\varphi = k\left(h_1 + \frac{p_1}{\gamma}\right) + C = kH_1 + C\,. \tag{3.1}$$

Dafür kann auch folgende Gleichung aufgeschrieben werden:

$$\frac{\partial \varphi}{\partial s} = -\frac{\partial \psi}{\partial n} = 0\,. \tag{3.2}$$

Aus den Gleichungen (3.1) und (3.2) folgt, daß die Grenzlinie C_2 eine Potentiallinie ist, demzufolge die Stromlinien senkrecht zu dieser Grenzlinie verlaufen.

Die *Fassungfläche* repräsentiert eine bereits verwickeltere Grenzbedingung, als die Eintrittfläche ist.

Der Wert des Geschwindigkeitspotentials für die Grenzlinie C_3 wird:

$$\varphi = k\left(h_2 + \frac{p_2}{\gamma}\right) + C = kH_2 + C\,. \tag{3.3}$$

Für diese — die Potentiallinie repräsentierende — Grenzlinie ist auch die Beziehung (3.2) gültig.

Die Grenzlinie C_3'' ist eine der Erzeugenden der freien Aussickerungsfläche. Das Wasser strömt durch diese Fläche vom Sickerraum in den innerhalb des Filterrohres befindlichen Luftraum, in dem ein Druck p_0 vorhanden ist. An der inneren Fläche des Filterrohres wirkt der Druck in allen Punkten identisch gleich p_0. Folglich wird der Wert des dazugehörigen Geschwindigkeitspotentials

$$\varphi = ky_k\,, \tag{3.4}$$

woraus hervorgeht, daß die freie Aussickerungsfläche keine Potentialfläche ist. Sie ist auch keine Stromlinie, da ja das Wasser durch dieselbe in den Brunnenraum strömt.

Die freie Wasseroberfläche C_4, wenn die darauf wirkende Kapillaritätswirkung (und Oberflächenaktivität) vernachlässigt werden kann, und $p_{\text{atm}} = p_0$, dann ist dieselbe eine Stromlinie mit einem Druck p_0. Weshalb für sie die Beziehungen

$$\psi = \text{const}; \frac{\partial \varphi}{\partial s} = 0 \tag{3.5}$$

bzw.

$$\varphi = ky\,; \frac{\partial \varphi}{\partial s} = k\frac{\partial y}{\partial s} \tag{3.5a}$$

gültig sind. Jedoch, im Falle von einer Grundwasserabsenkung durch Vakuumverfahren, kann die Kapillaritätswirkung, nachdem die Böden feinkörniger

sind, nicht vernachlässigt werden. Somit kann in solchen Fällen die freie Wasseroberfläche C_4 keine Stromlinie sein, da zwischen dem Kapillarraum und dem Graviationssickerraum ein Wasseraustausch vorhanden ist, d. h., der Verlauf der Wasserpartikeln durchschneidet diese Oberfläche. Infolgedessen tritt die die Bedingung $\psi =$ konstant befriedigende obere abgrenzende Strömungslinie an der Grenzoberfläche C_5 des Kapillarraumes auf. An der Begrenzungsfläche C_5 ist der Wert des Überdrucks konstant, und h_k der Kapillarabsenkung entsprechend negativ:

$$p = -\gamma h_k \,, \tag{3.6}$$

und auf derselben ist die Bedingung

$$\varphi = k(y - h_k)$$

gültig. Dafür kann man auch die Beziehung

$$\frac{\partial \varphi}{\partial s} = k \frac{\partial y}{\partial s} \tag{3.7}$$

aufschreiben.

Zur oberen, mit C_5 bezeichneten, $\psi =$ konst. Grenzstromlinie gehören auch die Stromlinienabschnitte C_5' und C_5''. Für die Punkte derselben sind die Beziehungen

$$\begin{aligned} p &= -\gamma h_{k''} \\ \varphi &= k(y - h_k'') \end{aligned} \tag{3.8}$$

sowie (3.7) kennzeichnend. Für h_k' gilt die Bedingung

$$0 \leq h_k'' \leq h_k \,.$$

Nachdem längs des Grenzlinienabschnittes C_5'' an der inneren Seite des Filterrohres der atmosphärische Druck p_0 herrscht, weshalb das Kapillarwasser durch das Filterrohr *nicht unmittelbar* in den Fassungsraum des Brunnens *eindringen kann*, strömt es nur durch die Vermittlung des sich unter der Grenzoberfläche C_4 befindlichen Sickerraumes durch die Grenzoberfläche C_3'' in den Brunnen (Kovács, 1973).

Wenn das Grundwasser im Brunnen unterhalb des unteren Randes des Filterrohres sinkt, (s_0), weichen die hydraulischen Eigenschaften des den Brunnen umgebenden Sickersystems von den obenerwähnten Bedingungen in mehreren Beziehungen ab. Dieser Fall ist durch die folgenden zwei Grundfälle wiedergegeben.

Bild 4 stellt den *Vakuumbrunnen noch ohne Luftansaugen* dar. In diesem Falle wirkt im Sickerraum nur die Schwerekraft, die die Bewegung hervorruft. Vom Standpunkt der Grenzbedingungen um den Brunnen betrachtet können die Strömungsverhältnisse, wie folgt, gekennzeichnet werden.

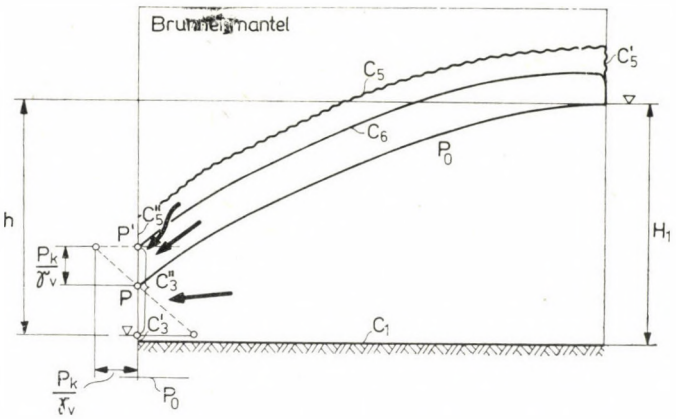


Bild 4. Die Sickerungsgrenzbedingungen bei einem Gravitationsbrunnen ohne Luftansaugen

Das Grundprinzip der Gravitationsbrunnen, daß das Wasser solange vom Boden unmittelbar in den Brunnen eintritt, bis der Wasserdruck im Boden höher als jener im Brunnen ist, können die den Brunnen entsprechenden Grenzbedingungen C_3', C_3'', und C_5 gemäß dem Bild erläutert werden. Daraus geht die folgende Eigenschaft des Vakuumbrunnens hervor:

— von einem Teil des ursprünglichen Kapillarraumes, dessen Höhe, aufgrund der Bezeichnung auf dem Bild p_K/γ_v, ist, tritt das Wasser unmittelbar in den Brunnen, und deshalb wird der Kapillarraum vom Standpunkt der hydraulischen Verhältnisse der unmittelbaren Umgebung des Brunnens aus betrachtet scheinbar zusammengedrückt.

Das Maß dieser Zusammendrückung ist:

$$\frac{h_k - p_K/\gamma_v}{h_k}.$$

Auch dies deutet darauf hin, daß die Vakuumbrunnen in den Böden mit höherem Kapillaranstieg verhältnismäßig wirksamer sind.

Als Folge der Absenkung p_K gelangt auch der dem Brunnen anschließende Abschnitt und die angrenzende Zone des Kapillarraumes C_5 auf ein niedrigeres Niveau.

Die Erzeugende der freien Sickerungsoberfläche ergibt sich dann aus der Summierung der Grenzlinienabschnitte:

$$C_3' + C_3'' .$$

Der Wert des Geschwindigkeitspotentials auf der freien Sickerungsoberfläche kann in ähnlicher Weise wie beim Filterbrunnen, aufgrund der zum Brunnenwasserspiegel in Verhältnis gestellten y_k ermittelt werden:

$$\varphi = k y_k - p_K/\gamma_v . \tag{3.9}$$

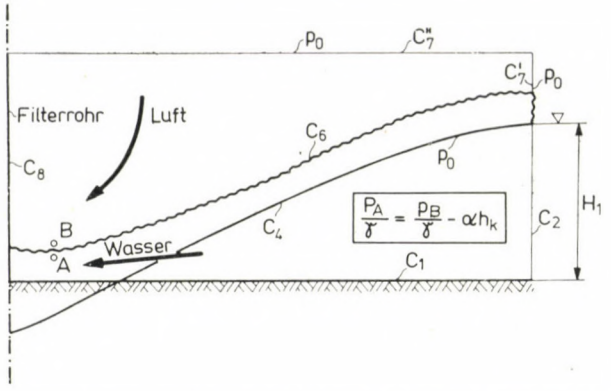


Bild 5. Die Sickerungsgrenzbedingungen bei einem luftansaugenden Vakuumbrunnen

Der Unterschied besteht nur darin, daß die Grenzlinie C_4 ihre frühere Funktion verloren hat.

Dagegen ist der Grenzlinienabschnitt C_6 ein neues Merkmal des Sickersystems. Die Abänderung des Potentials im Brunnen, die unter der Einwirkung des Potentials im Brunnen, die unter der Einwirkung des Vakuums eintritt, ist repräsentiert durch das Glied $-p_K/\gamma_v$.

Ein typisches Beispiel des *luftansaugenden Vakuumbrunnens* ist im Bild 5 dargestellt. Hier herrscht unter Einwirkung des innerhalb des Saugrohres herrschenden, schon erhöhten Vakuums an der ganzen Länge des Sickerraumes, der an der äußeren Fläche des Saugrohres für das Wasser zur Verfügung steht, ein Druck, der niedriger ist, als p_0. Dies bedeutet zugleich, daß die ursprüngliche Grenzoberfläche C_4 sich von dem Brunnenmantel entfernt und die Funktion der »Absenkungskurve« völlig verliert, da der Raum der Wasserströmung von oben durch eine andere Kurve (C_6) begrenzt ist, zu welcher die Grenzoberfläche C_4 nicht einmal parallel liegt. Wie das Bild 5 und die Laboruntersuchungen des Verfassers (Foto 1) beweisen, fließt die ganze in den Brunnen strömende Wassermenge *durch dieselbe* Kurve (C_4) dem Brunnen zu.

Die den Raum der Wasserbewegung von oben begrenzende Deckfläche C_6 ist zugleich die untere Grenzfläche des Raumes der Luftsickerung.

Für die an beiden Seiten befindlichen Punkte A und B (Bild 5) können (unter Berücksichtigung der Kapillaritätswirkung) die folgenden Gleichungen aufgeschrieben werden:

$$p_B < p_0\,, \tag{3.10}$$

$$p_A = p_B - \alpha h_k \gamma_v \tag{3.11}$$

und aus der letzteren Gleichung

$$\frac{p_A}{\gamma_V} = \frac{p_B}{\gamma_V} - \alpha h_k\,. \tag{3.12}$$

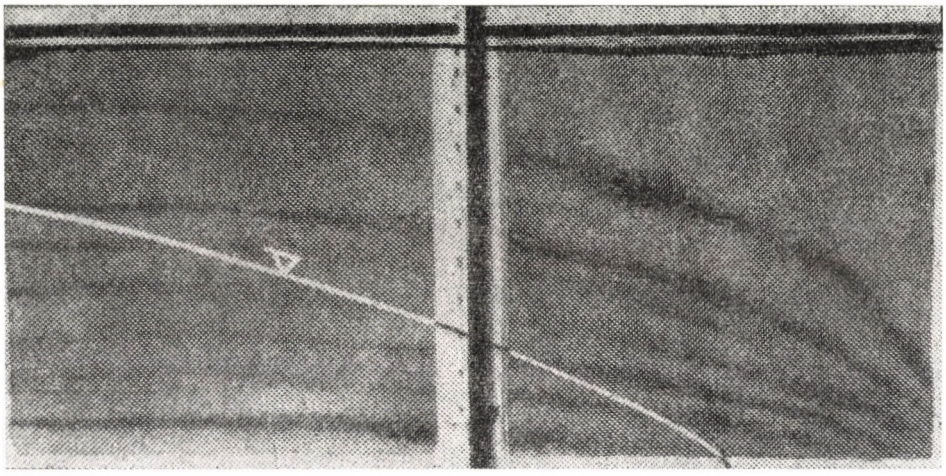


Foto 1. Die Sickerungsstromlinien schneiden die bei dem Vakuumbrunnen ihre Funktion verlorene Absenkungskurve p_0

Hierin ist

p_A der Druck im Punkt A unmittelbar unter der Grenzfläche C_6, und p_B der Druck im unmittelbar über der Grenzfläche C_6 befindlichen Punkt B.

Der Koeffizient $\alpha < 1{,}0$ dient zur Reduktion der ursprünglichen Kapillarschichtdicke. Das bedeutet, daß in irgendeinem Punkt der Grenzoberfläche der sich miteinander nicht vermischenden (Wasser-Luft-)Sickerräume, der durch den Luftstrom hervorgerufene Druck den Wasserdruck im untersuchten Punkt um einen Wert überschreitet, der dem tatsächlichen Kapillaranstieg $\alpha\, h_k$ entspricht.

Da p_B ein veränderlicher Wert ist, erscheint die Druckänderung als eine grundsätzliche Eigenschaft der Grenzfläche C_6. Folglich gibt es zwischen der Höhenlage eines Punktes und dem zugeordneten Potentialwert auf der obersten Stromlinie die folgende Beziehung:

$$\varphi = k(y_k - \alpha h_k + p_B/\gamma_v)\,. \tag{3.13}$$

Die Grenzflächen C_7' und C_7'' sind die Lufteintrittsflächen. Auf sie wirkt der atmosphärische Druck p_0. Die Grenzfläche (Wasser/Luft Abdichtungsfläche) C_8 ist eigentlich eine Stromlinie.

4. Rechnungsprinzip der kapillaren Ergiebigkeit

Aufgrund der Gleichung (3.1) kann der Potentialunterschied für den die ursprüngliche Wasserschichtdicke H_1 durchquerenden Vakuumbrunnen in folgender Form aufgeschrieben werden:

$$\Delta\varphi = k\left(H_1 + \frac{p_K}{\gamma_V}\right)\,. \tag{4.1}$$

Für den Gravitationsbrunnen wird der Potentialunterschied auf die Grundwasserdeckfläche bezogen:

$$\Delta\varphi = kH_1 \,. \tag{4.1a}$$

Es geht bereits aus der Gegenüberstellung dieser beiden Gleichungen hervor, daß die Ergiebigkeit des Vakuumbrunnens höher ist, als die des Gravitationsbrunnens.

Der die Ergiebigkeit beeinflußende Kapillarraum kommt im Sickersystem um die Vakuumbrunnenreihe in folgender Form in Erscheinung (Bild 6a bis 6c):

Im Gravitationssystem gibt es zwischen der Gravitations- und Kapillarzone einen Wasseraustausch wie im Bild 6a ersichtlich ist, und die gesamte Wassermenge strömt durch die Gravitationszone dem Brunnen zu.

In einem mit einer Vakuumsystem verhältnismäßig noch niedrigere Absaugfähigkeit (Bild 6b) bleibt der bei den Filterbrunnen beobachtete Charakter des Wasseraustausches bestehen. Jedoch wirkt bereits am Brunnenmantel der unmittelbare Saugeffekt des Brunnens auf den Kapillarraum. Im Anfang kann das Wasser nur vom unteren Teil des Kapillarraumes, und dann, im Falle einer stärkeren Depression von dem ganzen Kapillarraum unmittelbar in den Brunnen strömen. Auf diese Weise nimmt die Menge und Proportion des Wassers, das durch den Kapillarraum strömt, allmählich zu.

Infolge der weiteren Steigerung der Grundwasserabsenkung kann sich auch der im Bild 6c dargestellte Kapillarraum entwickeln. In solchen Fällen reicht der untere Teil des Kapillarraums bereits in die undurchlässige Schicht hinein, was von dem obenerwähnten hydraulischen Benehmen und der Funktion des Kapillarraums abweicht, wobei man im üblichen Sinne von einer kapillaren Ergiebigkeit nicht mehr reden kann.

Infolge der von der Grundwasserabsenkung abhängigen hydraulischen Eigenheiten des Kapillarraumes sollten gewisse hydraulische Vereinfachungen — wenn man die kapillare Ergiebigkeit in jedem Fall als einen Begriff behalten will — eingeführt werden.

Die Hydraulik der Fassungsstollen behandelt den Kapillarraum als ein unabhängiges Sickersystem und betrachtet denselben als ein geschlossenes Rohr (Kovács, 1973). Folglich kann die Ergiebigkeit des Kapillarraumes auf der Einheitsbreite des Gravitationsfassungsstollens bei Kenntnis des konstanten Sickerquerschnitts ($F = h'_k$), der Länge L des Kapillarraumes, der an dessen beiden Enden wirkenden Potentialwerte, sowie des für den Kapillarraum als maßgebend betrachteten Filtrationskoeffizienten k, angenähert berechnet werden (Bild 6a):

$$q = Fv = kh'_k J = kh'_k \frac{H_1 - H_3}{L} \,. \tag{4.3}$$

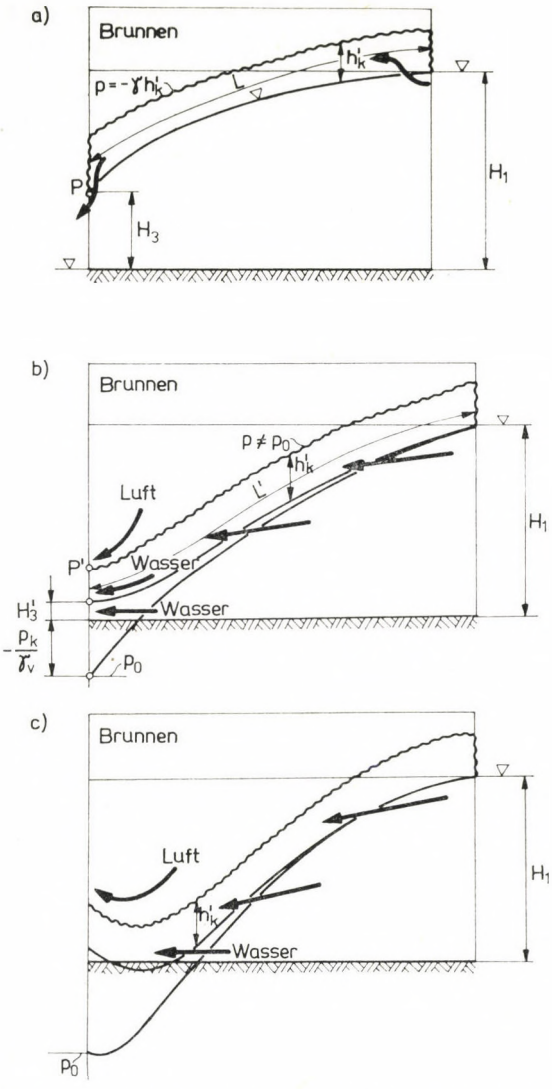


Bild 6a bis c. Erläuterung des Kapillarraumes im Falle von unterschiedlichen Absenkungen

Für den die Vakuumbrunnenreihe substituierende Vakuumfassungsstollen würde sich die Beziehung (Bild 6b)

$$q = Fv = kh'_k J = kh'_k \frac{H_1 - (H'_3 - p_K/\gamma_v)}{L'} \tag{4.4}$$

ergeben, jedoch darf schon, wie auch im Bild ersichtlich ist, im vorliegenden Fall die Behandlung des Kapillarraumes als ein geschlossenes Rohr nicht gestaltet werden.

Die Höhe H_3' des unteren Niveaus des sich am Brunnenmantel anschließenden Kapillarraumes über der Brunnensohle p_K/γ_v ist das unter dem Rande des Saugrohres befindliche, durch den hydrostatischen Druck gekennzeichnete Vakuum.

Zur Ermittlung der Gesamtergiebigkeit des Vakuumsickersystems, sollte — da der Kapillar- und Gravitationssickerraum vom dynamischen Standpunkt aus betrachtet ein *einheitliches System bildet* —, die gleichzeitige Berechnung der Gesamtergiebigkeit der beiden Systeme ermöglicht werden.

Das Grundprinzip der Lösung ist, daß einerseits die Ergiebigkeit eines einheitlichen Sickersystems berechnet, andererseits, bei den Grenzbedingungen der Kapillarraum mit einer Höhe von h'_k berücksichtigt werden sollte, d. h., es sollte ein solcher stellvertretender Kapillarsickerquerschnitt gewählt werden, für welchen die für den Zweiphasenraum bezogene Bedingung k = konst. als gültig betrachtet werden kann. Auf diese Weise kann die Veränderlichkeit des Filtrationskoeffizienten längs der Kapillarraumhöhe außer acht gelassen werden. Letzten Endes wird also eine kleinere, als die tatsächliche Höhe h_k, *die sog. Ersatzhöhe* (h_k') für den Kapillarraum eingeführt, wodurch die ursprüngliche Dicke des Zweiphasensickerraumes nach oben durch h_k' erhöht wird.

Unter Berücksichtigung der Veränderung $k(h)$ des Koeffizienten k in der Richtung h, wird die Ergiebigkeit des Kapillarraumes durch die Gleichung

$$q_x = J \int_0^{h_k} k(h)\, dh \tag{4.5}$$

bestimmt.

Dieselbe Ergiebigkeit wird unter Berücksichtigung der Ersatzhöhe h_k' und k = konst. auf eine Elementarlänge des Kapillarraumes durch die Beziehung

$$q_x = kJh_k' \tag{4.6}$$

erhalten.

Da die Ergiebigkeiten der beiden Gleichungen (4.5) und (4.6) gleich sind, ergibt sich die Beziehung

$$kJh_k' = J \int_0^{h_k} k(h)\, dh\,, \tag{4.7}$$

woraus die Ersatzhöhe

$$h_k' = \frac{\int_0^{h_k} k(h)\, dh}{k} \tag{4.8}$$

ermittelt werden kann.

Die Höhe h_k' des Kapillarraumes für eine bestimmte Bodenart kann aufgrund der Kurve pF geschätzt werden. Die bei der Grundwasserabsenkung nach dem Vakuumbrunnenverfahren vorkommenden Böden besitzen im allgemeinen einen bestimmten abgeschlossenen Kapillarbereich, und die *van der*

Waals—Londonsche Kraft folgt nur einem verhältnismäßig kleineren Teil des im Hohlraumgehalt befindlichen Wassers weshalb für solche Bodenarten der Bereich

$$h_{k2} > \psi(n)$$

als charakteristisch betrachtet werden kann,

wo:

h_{k2} = die dem höchsten Punkt der horizontalen Strecke der Kurve pF zugeordnete hydrostatische Druckhöhe; und

ψ = die auf der Oberfläche der Wassermembrane gemessene Saugspannung;

n = den Hohlraumgehalt

bedeuten

Ausführliche Berechnungen können nach KOVÁCS (1972), (S. 70, Bilder 1.3 bis 2b), unter Berücksichtigung der Tatsache durchgeführt werden, daß der Sättigungsverminderung des Bodens die Wasserdurchlässigkeit desselben stark abnimmt (KÉZDI 1969, 1976).

Im folgenden wird die vollständige theoretische Ableitung der auch die Kapillarwirkungen berücksichtigenden Ergiebigkeitsformel des Vakuumsickersystems vorgeführt.

5. Die Ergiebigkeitsformel des Vakuumsickersystems

Zur theoretischen Berechnung der Ergiebigkeit des Gravitationsfassungsstollens mit freier Oberfläche wird das wohlbekannte Dupuitsche Verfahren angewandt. Nach dem Ausgangsgrund desselben, wenn man das Beisein des Kapillarraumes außer acht läßt, erfüllt die Grundwasserabsenkungskurve $p = 0$ gleichzeitig auch die Funktion der höchsten Grenzstromlinie des Sickerfelds. Wenn man aber auch die Wirkung der Ergiebigkeitserhöhung des Kapillarraumes in Betracht zieht, verläuft die höchste Stromlinie des Sickersystems über der Absenkungskurve in einer gewissen Höhe parallel zu dieser letzteren.

Die zur theoretischen Berechnungen der Ergiebigkeit gehörenden hydraulischen Umstände der Umgebung der *Vakuumbrunnenreihe, die gleichzeitig Wasser und Luft ansaugt* (und die im vorliegenden Fall durch eine Vakuumfassungsstollenreihe ersetzt ist) weichen von den obenerwähnten Bedingungen ab (Bild 7):

a) Die Kurve des Druckes $p = 0$ ist keine Grundwasserabsenkungskurve mehr, da nämlich der zwischen den Gravitations- und Vakuumsickersystemen bestehende grundlegende Unterschied beachtet werden muß; bei dem Gravitationsverfahren gibt die die obere Abgrenzung des Strömungsbereichs repräsentierende Absenkungskurve (welche zugleich auch die äußere Stromlinie ist) eindeutig das irgendeinem Querschnitt x des Sickersystems zugeordnete Energieniveau ($\varphi_x = kH_x + C$) an. Im Falle von einem luftansaugenden Vakuumbrunnen verändert sich in der Konstante C der bisher stillschweigend berück-

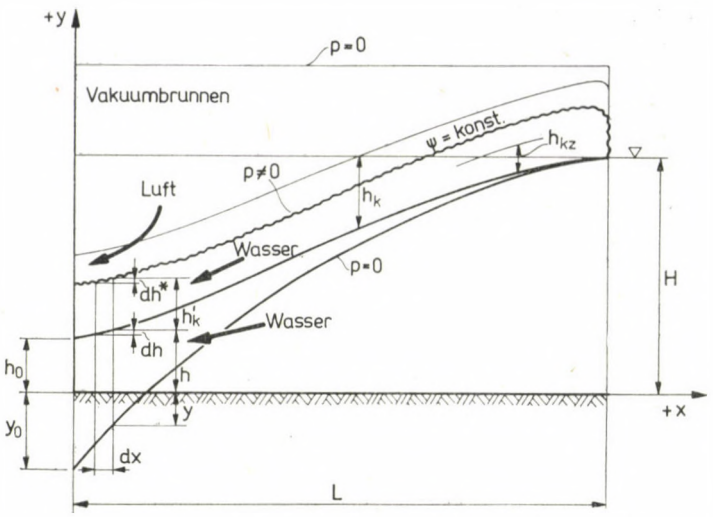


Bild 7. Erläuterung der zur Ableitung der Ergiebigkeitsformel für den Vakuumbrunnen angewandten Bezeichnungen

sichtigte, über dem Sickerraum herrschende Luftdruck p_0. Infolgedessen sollte die Luftdruckänderung besonders berücksichtigt werden (1. Erkenntnis)

$$\varphi_x = k\left(H_x + \frac{p_{lx}}{\gamma}\right) + C\,, \tag{5.1}$$

bzw. in einer mehr angebrachten Form

$$\varphi_x = k\left(H_x - \frac{p_0 - p_{lx}}{\gamma}\right) + C\,. \tag{5.2}$$

Deshalb kann die Kurve p_0 als eine Energiekennlinie gedeutet werden.

b) Die höchstliegende Stromlinie des Sickerbereiches, die zugleich auch die niedrigste Stromlinie des Luftsickerraumes ist, liegt zwar auch jetzt über der Kurve $p = 0$, jedoch verläuft sie nicht parallel zur letzteren, denn der Luftdruck über der Wassersickerzone ist konstant.

c) Unter dieser Stromlinie verläuft in einem Abstand von h_k' eine parallele Kurve, die als die untere Grenzlinie des Kapillarraumes definiert werden kann (die jedoch keine Stromlinie ist). Für h_k' sind folgende Bedingungen gültig

$$h_k > h_k' > h_{kz} > 0\,,$$

wo:

h_k = die kapillare Anstieghöhe;
h_{kz} = die Dicke des geschlossenen Kapillarraumes

bedeuten.

Aufgrund dieser Bedingungen kann man mit der Benutzung der im Bild angewandten Beziehungen schreiben:

$$\frac{dh^*}{dx} = \frac{dh}{dx} . \tag{5.3}$$

Die die Wasser- und Luftströmungen berücksichtigende Druckänderung wird:

$$\frac{dp}{dz} = \gamma \frac{dh}{dx} + \frac{dp_l}{dx} = \gamma \frac{dy}{dx} . \tag{5.4}$$

Durch Anwendung der von DUPUIT eingeführten, bekannten Bedingungen erhält man für die auf die einheitliche Zonenbreite entfallende spezifische Ergiebigkeit

$$q = vF , \tag{5.5}$$

wo

$$F = h + h'_k \tag{5.6}$$

ist.

Unter Berücksichtigung der Geschwindigkeit

$$v = k \frac{dp}{dx} = k \frac{dy}{dx} , \tag{5.7}$$

und nach Durchführung der Substituierungen erhält man

$$q = k \left(\frac{dh}{dx} + \frac{dp_l}{\gamma dx} \right) (h + h'_k) , \tag{5.8}$$

die die auch den *Kapillarraum berücksichtigende allgemeine Differentialgleichung des Vakuumfassungsstollens ist.*

Diese Gleichung kann, unter Berücksichtigung der Gleichung (5.4) auch, wie folgt, aufgeschrieben werden:

$$q = k \frac{dy}{dx} (h + h'_k) . \tag{5.9}$$

Trennung der Veränderlichen und Durchführung der Integration ergeben

$$\frac{q}{k} \int dx = \int (h + h'_k)\, dy , \tag{5.10}$$

wo

$$h = y + \frac{p_0 - p_{lx}}{\gamma} \tag{5.11}$$

und somit

$$\frac{q}{k}\int dx = \int \left(y + \frac{p_0 - p_{lx}}{\gamma} + h'_k\right) dy \,. \tag{5.12}$$

Führt man die Beziehung

$$\frac{p_0 - p_{lx}}{\gamma} = f(x) \tag{5.13}$$

ein, die diejenige von x abhängige Druckdifferenz bedeutet, welche auf der Wasser- und Luftströmungsräume abtrennenden Stromlinie interpretiert und mit p_0 in Zusammenhang gebracht wird (wo p_{lx} den Luftdruck im Punkte der Stromlinie in einer Entfernung von dem Fassungsstollen bedeutet), erhält man

$$\frac{q}{k}\int dx = \int \{y + f(x) + h'_k\}\, dy \,. \tag{5.14}$$

Die Funktion $f(x)$ kann theoretisch nicht exakt ermittelt werden, denn man könnte sie nur vom Strombild ableiten, welches die gemeinsame Strömung der Luft und des Wassers repräsentiert. Jedoch, mit Rücksicht auf

a) die Bedingung $\gamma_l \ll \gamma_v$, weiters

b) auf die Ergebnisse der Laborversuche (Bild 8), könnte man für die in der Praxis vorkommenden Einzelfälle annehmen, daß die Höhenänderung der am höchsten liegenden Stromlinie des Wassersickerraumes auf den an einer gegebenen Stelle x herrschenden Luftdruck keine Rückwirkung ausübt (2. Erkenntnis), d. h.,

$$f(x) = \frac{p_0 - p_{lx}}{\gamma} \neq f(y) \tag{5.15}$$

wovon man nach Durchführung der Integration die Gleichung in der Form erhält:

$$\frac{q}{k}x = \frac{y^2}{2} + yf(x) + yh'_k + C \,. \tag{5.16}$$

Wenn $x = 0$, dann $f(x_0) = h_0 - y_0$ (3. Erkenntnis), somit

$$\frac{q}{k}0 = \frac{y_0^2}{2} + y_0(h_0 - y_0) + y_0 h'_k + C \,, \tag{5.17}$$

und daraus ergibt sich die Gleichung

$$C = \frac{y_0^2}{2} - y_0 h_0 - y_0 h'_k \,. \tag{5.18}$$

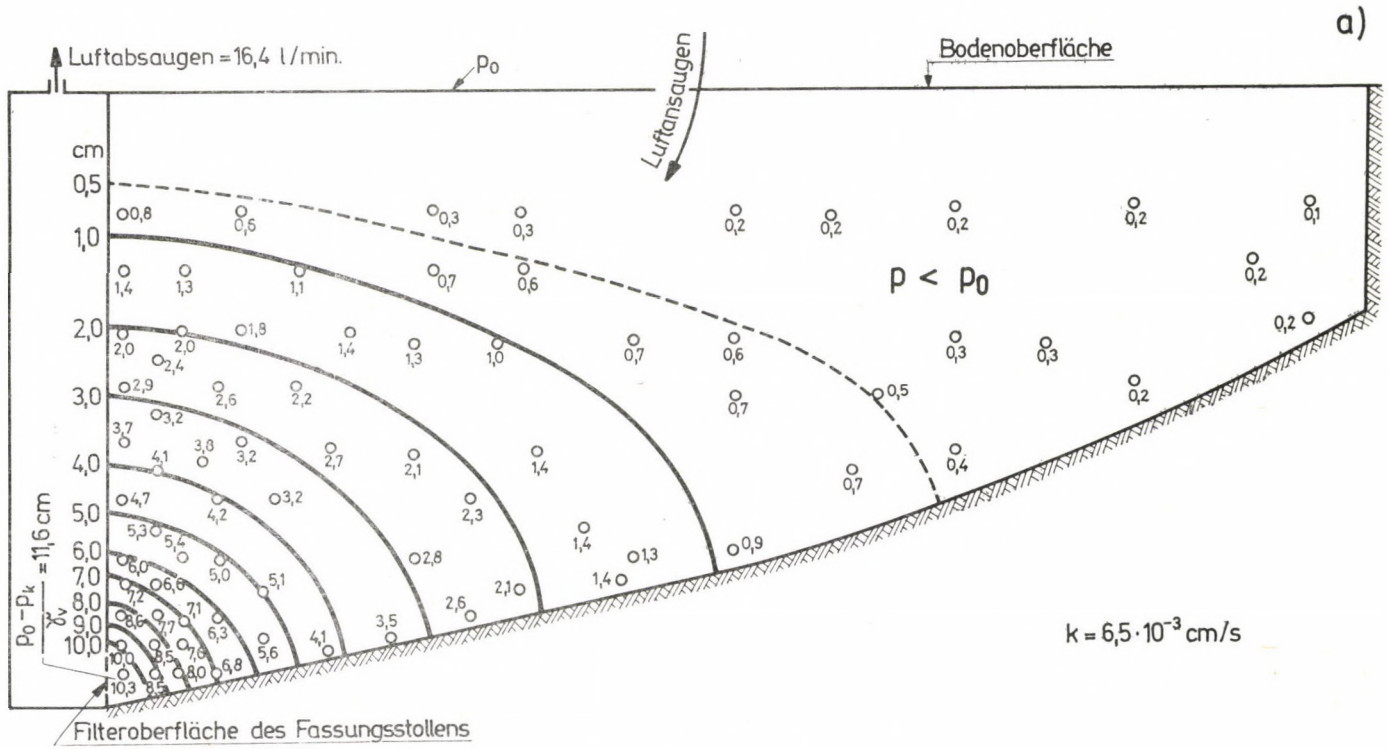


Bild 8. Eine typische Luftdruckverteilung

Nach der Interpretation des Bildes 7, ist h_0 die Höhe der unteren Linie des Kapillarraumes über der Brunnensohle. Hiezu sei bemerkt, daß h_0 auch als ein negativer Wert vorkommen kann.

Sei $x = L$, dann $f(x_L) = 0$, $y = H$ und

$$\frac{q}{k} L = \frac{H^2}{2} + Hh'_k + \frac{y_0^2}{2} - y_0 h_0 - y_0 h'_k , \tag{5.19}$$

wovon man die Ergiebigkeit des Kapillarraumes berücksichtigende *Formel der Ergiebigkeit des Vakuumfassungsstollens* ⟨*der Vakuumbrunnenreihe*⟩ erhält:

$$\frac{q}{k} 2L = H^2 + y_0^2 - 2y_0 h_0 + 2h'_k(H - y_0) . \tag{5.20}$$

Die Formel (5.20) kann als allgemein gültig betrachtet werden und demgemäß geht nach Durchführung der den üblichen Vernachlässigungen entsprechenden Substituierungen in die bekannte Formeln der Ergiebigkeit über.

So, z. B. vereinfacht sich die Gleichung (5.20) für den Graviationsfassungsstollen, mit Rücksicht auf die Ergiebigkeit des Kapillarraumes, im Falle der Gültigkeit der Voraussetzung $h_0 = y_0$ auf folgende Form:

$$\frac{q}{k} 2L = H^2 - h_0^2 + 2h'_k(H - h_0) . \tag{5.21}$$

Vernachlässigt man die erhöhende Wirkung des Kapillarraumes auf die Wassergiebigkeit, d. h., $h'_k = 0$, so nimmt die Formel die bekannte Form

$$\frac{q}{k} 2L = H_2 - h_0^2 \tag{5.22}$$

an.

Untersucht man nur einen einzigen Brunnen (anstatt eines Fassungsstollens), so erhält man auf gleiche Weise wie durch die oben geschilderte Ableitung die Formel der Ergiebigkeit für einen *einzelnen Vakuumbrunnen*:

$$\frac{Q}{\pi k} \ln \frac{R}{r} = H^2 + y_0^2 - 2y_0 h_0 + 2h'_k(H - y_0) . \tag{5.23}$$

Im Gravitationssystem geht auch diese, auf den einzelnen Vakuumbrunnen bezogene Formel nach Durchführung der entsprechenden Substituierungen in die schon bekannte Form über. Auf diese Weise erhält man für den auch die Wirkung des Kapillarraumes berücksichtigenden Einzelgravitationsbrunnen (d. h., im Falle der Erfüllung der Bedingung $y_0 = h_0$) die Formel:

$$\frac{Q}{\pi k} \ln \frac{R}{r} = H^2 - h_0^2 - 2h'_k(H - h_0) . \tag{5.24}$$

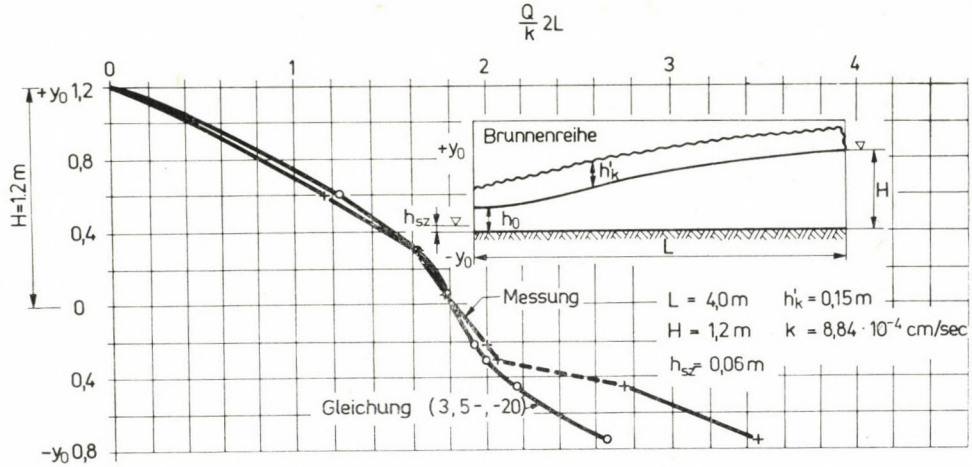


Bild 9. Laborkontrolle der Ergiebigkeitsformel

Vernachlässigt man aber die ergiebigkeitssteigernde Wirkung des Kapillarraumes (d. h., $y = h_0$ und $h_k = 0$), so ergibt sich für den einzelnen Gravitationsbrunnen die Formel

$$\frac{Q}{\pi k} \ln \frac{R}{r} = H^2 - h_0^2. \tag{5.25}$$

Die im Laboratorium erhaltenen und die Gültigkeit der Formel (5.20) beweisenden Angaben sind im Bild 9 ersichtlich.

6. Berechnung der oberen Grenzlinie des Zweiphasensickerraumes

Zweck der Grundwasserabsenkungssysteme ist die Reduktion der Höhe der Zweiphasenraumdeckfläche, d. h., das Herabsetzen des Niveaus des gesättigten Sickerraumes. Die obere (der Höhe h'_k zugeordnete Grenzstromlinie, die bei der theoretischen Berechnung der Ergiebigkeit, für das physikalische Modell eingeführt wurde) kann mit guter Annäherung als obere Grenzlinie des entwässerten Raumes betrachtet werden. Die diese Linie bestimmende Differentialgleichung wurde für den Vakuumfassungsstollen aufgeschrieben und erläutert. Mit der in Zusammenhang mit der Ermittlung der Ergiebigkeit angewandte Ableitung wurde die Bezeichnung $(p_0 - p_{lx})/\gamma = f(x)$ eingeführt. Aufgrund des physikalischen Modells und der Beziehung (5.14) ist einzusehen, daß zur Ermittlung der gesuchten Stromlinie die Kenntnis der Funktion $f(x)$ unerläßlich ist. (Dieser Umstand verursacht in der Ableitung der Ergiebigkeitsformel keine Schwierigkeiten, da in Zusammenhang mit den in der Praxis vorkommen-

den Fällen rechtmäßig vorausgesetzt werden kann, daß die Funktion $f(x)$ unabhängig von y ist und infolgedessen aus der Ableitung der Ergiebigkeitsformel entfällt.

Im vorliegenden Fall ist jedoch die Grundbedingung der Lösung die Kenntnis des Luftdrucks in Abhängigkeit von der Entfernung von der Fassungsoberfläche.

Aufgrund der Analyse der zur Verfügung stehenden Fachliteratur (GRIGOREW, 1973) und der Auswertung der von dem Autor durchgeführten Laborversuche konnte die Folgerung gezogen werden, daß im Fall von Fassungsstollen durch die Anwendung der folgenden Beziehung eine gute Näherung erzielt wird

$$f(x) = \frac{p_0 - p_{lx,0}}{\gamma} \frac{(L-x)^4}{L^4}, \tag{6.1}$$

wo

$$\frac{p_0 - p_{lx,0}}{\gamma} = h_0 - y_0 \tag{6.2}$$

bedeutet.

Als Beweis dient das Bild 10. In der obigen Formel bedeuten:

p_{lx} = den Luftdruck im Boden in der unmittelbaren Nähe vom Fassungsstollen;
γ = die Wichte des Wassers;
L = die Reichweite des Fassungsstollens;
x = die Entfernung der untersuchten Stelle von dem Fassungsstollen.

Unter Berücksichtigung der obigen Feststellungen kann der Wert von $f(x)$ aufgrund der Gleichung (6.1) angenommen werden, demzufolge erhält man

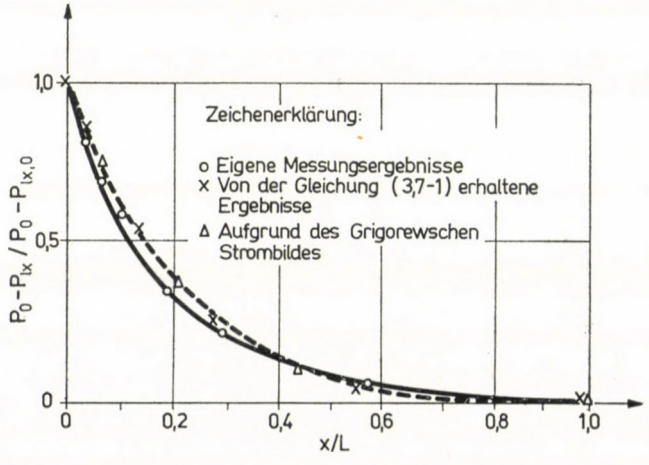


Bild 10. Die Anwendbarkeit der Gleichung (6.1) demonstrierende Versuchsergebnisse.

nach der Substituierung

$$\frac{q}{k}\int dx = \int\left[y + (h_0 - y_0)\frac{(L-x)^4}{L^4} + h'_k\right]dy\,. \tag{6.3}$$

Nach der Integration erhält man

$$\frac{q}{k}x = \frac{y^2}{2} + y(h_0 - y_0)\left(\frac{L-x}{L}\right)^4 + yh'_k + C. \tag{6.4}$$

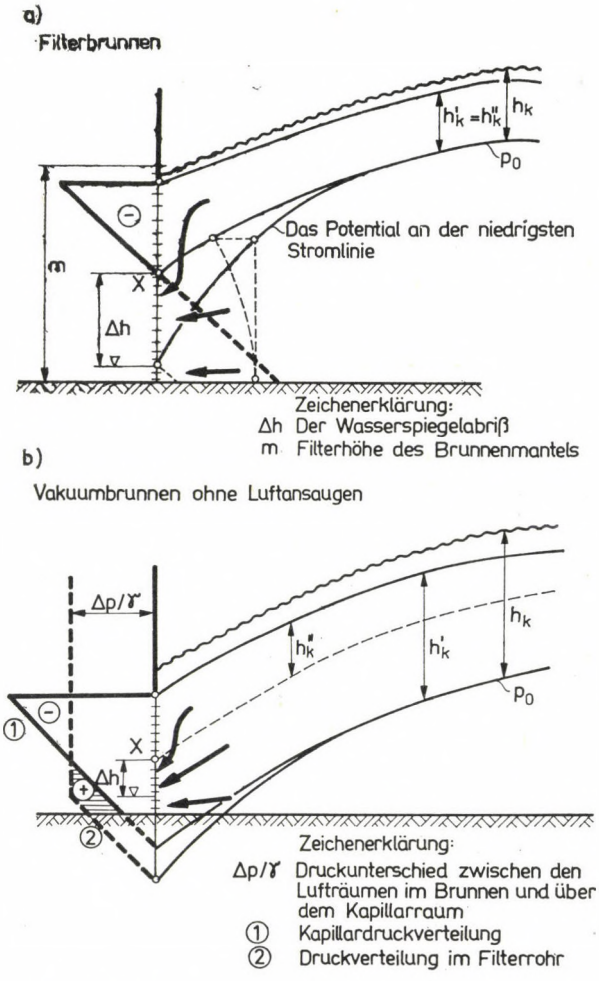


Bild 11a bis d. Deutung des Wasserspiegelabrisses Δh bei einem Vakuumbrunnen. a) Filterbrunnen. Zeichenerklärung: Δh Der Wasserspiegelabriß; m Filterhöhe des Brunnenmantels; b) Vakuumbrunnen ohne Luftansaugen; $\Delta p/\gamma$ Druckunterschied zwischen den Lufträumen im Brunnen und über dem Kapillarraum; ①Kapillardruckverteilung; ②Druckverteilung im Filterrohr; *c)* Vakuumbrunnen ohne Luftansaugen; *d)* Luftansaugender Vakuumbrunnen ($h'_k \geqq m$)

Wenn $x = L$, dann $y = H$. In diesem Fall

$$\left(\frac{L - x}{L}\right)^4 = 0$$

und so

$$y(h_0 - y_0)\left(\frac{L - x}{L}\right)^4 = 0. \tag{6.5}$$

Nach der Substituierung wird die Gleichung *der oberen Grenzstromlinie des* dem Vakuumfassungsstollen sich anschließenden, als zweiphasig betrachteten *Wassersickerraumes*

$$\frac{q}{k}x = \frac{y^2}{2} + y(h_0 - y_0)\left(\frac{L - x}{L}\right)^4 + yh'_k + \frac{y_0^2}{2} - y_0 h_0 - y_0 h'_k . \tag{6.6}$$

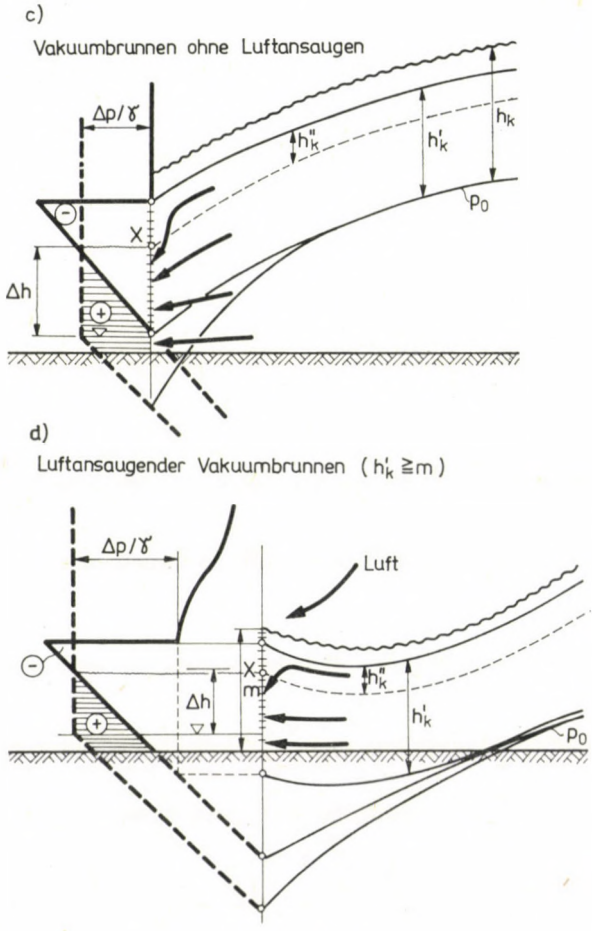


Bild 11.

7. Deutung des Wasserspiegelabrisses bei den Vakuumbrunnen

Der physikalische Verlauf des Wasserspiegelabrisses (Δh) wird durch zwe Hauptfaktoren beeinflußt:

— am Brunnenmantel, in der Zone der größten Sickergeschwindigkeiten nimmt der Widerstand zu;

— das Wasser tritt auch von oben vom Kapillarraum in den Brunnen ein.

Die Deutung der Druckverhältnisse an beiden Seiten des Brunnenmantels und zugleich des Abrisses Δh als Verlaufes ist in den Bildern 11a bis 11d wiedergegeben.

Wie aus diesen Bildern hervorgeht, ist infolge des Vorhandenseins des Druckunterschieds Δp, der unmittelbare Eintritt des Wassers in den Brunnen von einem gewissen Teil des Kapillarraumes von der $\Delta p/\gamma$ *hohen unteren Zone* möglich. Das kann auch anders ausgedrückt werden, und zwar: vom Gesichtspunkt des Wasserspiegelabrisses Δh wird die ursprüngliche h'_k Höhe des Kapillarraumes auf die Höhe $h''_k = h'_k - \Delta p/\gamma$ *reduziert*, d. h., von dieser h''_k starken Zone kann kein Wasser durch die untere Zone in den Brunnen strömen. Somit ist der Verlauf des Abrisses analog zu dem, der bei den Filterbrunnen vorkommt.

SCHRIFTTUM

JESENAK, J.: Vakuumentwässerung von feinkörnigen Böden (Kandidátska dizertacná práca) Bratislava 1959. (Auf slowakisch)

KARÁDI, G.: Hydraulik der linearen Grundwasserabsenkungssysteme. Doktorarbeit, Khartoum 1963*

KÉZDI, Á.: Bodenmechanik I. Tankönyvkiadó, Budapest 1969*

KÉZDI, Á.: Grundlagen einer neuen Bodenphysik. Akadémiai Kiadó, Budapest 1976*

KÉZDI, Á.: Fragen der Bodenphysik. Akadémiai Kiadó, Budapest 1976*

KOVÁCS, GY.: Theoretische Untersuchung der Mikrosickerung. *Hidrológiai Közlöny*, (1957), Nr. 3.*

KOVÁCS, GY.: Deutung und Ermittlung des Filtrationskoeffizienten, die Sickerung beeinflußenden bodenphysikalischen Kennwerte. *VITUKI. Tanulmányok és Kutatási Eredmények.* 1972, Nr. 35*

KOVÁCS, GY.: Sickerungshydraulik. Akadémiai Kiadó. Budapest 1972*

KOVÁCS, GY.: Evaluation of Investigation Related to Seepage through Earth Dams. II. Characterization of Steady Seepage through Homogeneous Earth Dams with Vertical Faces. *VITUKI Publ. in Foreign Languages*, Budapest 1973

MATSCHAK, H.: *Bergakademie* (1961) Nr. 3.

MATSCHAK, H.—E. STARKE, HOFMEISTER, I.—DRESCH, H.: Betriebsversuche zur Steigerung der Ergiebigkeit und Entwässerungswirkung von Filterbrunnen und Steckfilter durch Anwendung vom Vakuum. *Bergbautechnik* **12** (1962) Nr.7.

MOLNÁR, L.: Luftströmung zu den Vakuumbrunnen. *Hidrológiai Közlöny* (1966) Nr. 4.*

MÖLLER, B.: Das Vakuumverfahren und die Grundwasserabsenkung nach der Wellpoint Methode. *Baumaschine und Bautechnik.* **4** (1957), Nr. 1.

OELBERG, G.—MITÓK B.: Einfluß der Entwässerung auf die Stabilität der Baugrube. *MÉLYÉPTERV.* Műszaki Fejlesztési Osztály, (1970) Nr. 5, Budapest*

ÖLLŐS G.—DELI, M.—SZOLNOKY CS.: Ergebnisse von Modellversuchen über die Grundwasserabsenkung durch Vakuumbrunnen. *Acta Techn. Hung.* **49.** (1964)

ÖLLŐS G.: Fundamental Problems of Vacuum-Well Hydraulics. *Építőipari és Közlekedési Műszaki Egyetem Tudományos Közleményei*, X (1964) 4

* Auf ungarisch

Öllős G.: Einige Eigenartigkeiten der Luftströmung in körnigen Mitteln. *Építőipari és Közlekedési Műszaki Egyetem Tudományos Közleményei* **XI** Budapest. (1965)* Nr. 3/4

Öllős G.: Hydraulische Grundlagen der Grundwasserabsenkung durch Anwendung von Vakuumbrunnen. Doktorarbeit. Budapest, Juli. 1977*

Rétháti L.: Ingenieurtechnische Beziehungen der Bodenkapillarität. *Vízügyi Közlemények.** (1960) Nr. 1

Rückert, H.: Grundwasserabsenkung nach dem Vakuumverfahren. *Institut für Ingenieur- und Tiefbau. Leipzig. Mitteilungen* 1964. **65** (1964), H.3

Széchy K.: Beitrag zur Theorie der Grundwasserabsenkungen. *Bautechnik.* (1959), Febr.

Vastagh G.: Grundwasserabsenkung nach dem Vakuumverfahren. *Mélyépítéstudományi Szemle.* (1952).* Okt.

Young, E. G.: Horizontal Seepage through Unconfined Aquifers Taking into Account Flow in Capillary Fringe. Water in the Unsaturated Zone. IASH/AIHS-Unesco. *Proc. of the Wageningen Symposium.* Band 2. 1969, S. 897 bis 905.

* ungarisch

Hydraulic Bases of the Ground-Water Lowering with Vacuum Wells. — The earlier methods of treating the vacuum-well hydraulics were based mainly on the aspect of the gravitational well hydraulics. The lowering of the ground-water level as a physical process itself and the boundary conditions have not yet been cleared up. The author brought the gravitational and vacuum-well hydraulics to a common theoretical basis, and continued to develop the interpretation of the boundary conditions and physical processes. He established a new general formula for the yielding capacity of the filter galery which also accounts for the water conveyance of the capillary fringe, and from which also the yielding capacity of the gravitational filter galery, the single vacuum and gravitational well may be derived. Also the formula for the approximate calculation of the curve separating the zones of the air and water streams has been established.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (1—2), pp. 87—90 (1980)

A NOTE IN CONNECTION WITH THE DEFORMATION ENERGY

I. ECSEDI*

(Manuscript received February 13, 1979)

This paper deals with bodies made of linearly elastic material. It verifies an inequality relation in connection with the deformation energy, by applying the usual assumption of elastostatics.

Let us consider the body of linearly elastic material represented in Fig. 1. The load applied to the body is the volume load forming a system of forces in equilibrium, $q = q(r)$ and the surface load $p = p(r)$. The usual assumptions of elastostatics are applied, however, it should be noted that the material of the body may also be inhomogeneous, anisotropic.

In case of the simultaneous application of both the volumetric and surface load ($q_i = q_i(r)$ and $p_i = p_i(r)$ respectively) the deformation energy of the body

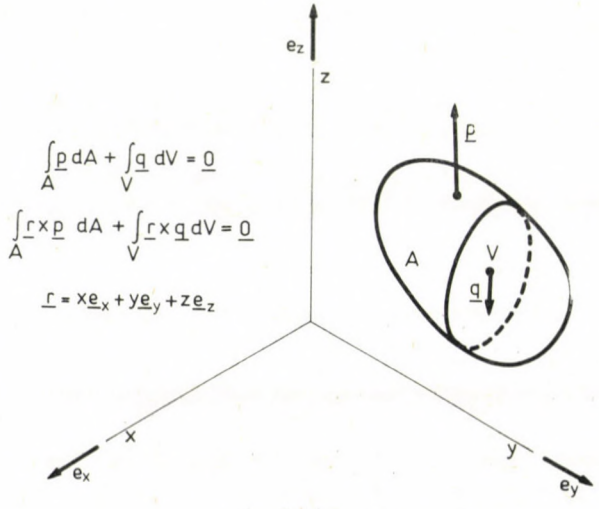


Fig. 1. Body of linearly elastic material

* Dr. István Ecsedi, Vászonfehérítő u. 24 IV/1. H-3531 Miskolc, Hungary

U_{ii}; and in case of the application of the volumetric load

$$q = \sum_{i=1}^{n} q_i(r)$$

and of the surface load

$$p = \sum_{i=1}^{n} p_i(r)$$

the deformation energy of the body be U.

The paper is intended to prove that in connection with the deformation energy U of the linearly elastic body the inequality relation

$$U \leq n(U_{11} + U_{22} + U_{33} + \ldots + U_{nn}) \tag{1}$$

holds true.

Proving of the inequality relation (1)

The verification is only carried out for case $n = 3$ (which, as a matter of fact, does not affect the generality), i.e., it is pointed out that the inequality relation

$$U \leq 3(U_{11} + U_{22} + U_{33}) \tag{2}$$

is valid wherein U denotes the deformation energy associated with the volumetric load $q = q_1 + q_2 + q_3$ and the surface load $p = p_1 + p_2 + p_3$.

It is known that

$$\begin{aligned} U = U_{11} + U_{12} + U_{13} \\ + U_{22} + U_{23} \\ + U_{33} . \end{aligned} \tag{3}$$

In Eq. (3) expressions U_{12}, U_{23}, U_{13} designate the extraneous deformation energies.

In case of the volumetric load $q_{12} = q_1 - q_2$ and of the surface load $p_{12} = p_1 - p_2$ the value of the deformation energy of the elastic body is

$$U' = U_{11} + U_{22} - U_{12} . \tag{4}$$

The deformation energy U' being non-negative, on the basis of Eq. (4) may be written as

$$U_{11} + U_{22} \geq U_{12} . \tag{5}$$

By similar reasoning it can be pointed out that

$$U_{11} + U_{33} \geq U_{13} , \tag{6}$$

$$U_{22} + U_{33} \geq U_{23} . \tag{7}$$

With the aid of combination of Eq. (3) and inequality relations (5), (6), (7) the inequality relation

$$U \leq 3(U_{11} + U_{22} + U_{33}) \tag{8}$$

to be proved is obtained.

An application

The elements of the structure shown in Fig. 2 are elastic and rigid bodies. Between the elements of the structure such connections and support conditions are assumed to achieve that stage in which the elements and the whole structure

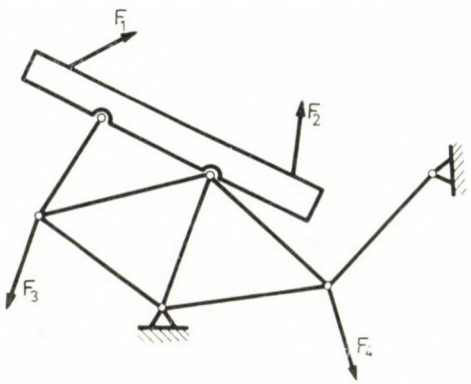


Fig. 2. Structure built up from elastic and rigid bodies

could not carry out finite, rigid-body-like movement at any load coming in question. The deformation energy of the structure is the function of the external forces $F_1, F_2, \ldots, F_n$. By making use of Maxwell's influence numbers c_{ij}, the deformation energy may be written in the form

$$U = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} F_i F_j . \tag{9}$$

Obviously, in the case in question

$$U_{ii} = \frac{1}{2} c_{ii} F_i^2 . \tag{10}$$

On the basis of the verified inequality relation (1) the following relation is obtained:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} F_i F_j \leq n \sum_{i=1}^{n} c_{ii} F_2 . \tag{11}$$

The inequality relation (11) holds true for any of the variables $F_1, F_2, \ldots, F_n$. Using the values $F_i = 1$ $(i = 1, 2, \ldots, n)$, and the inequality relation, the relation

$$\sum_{\substack{i,j=1 \\ i \neq j}}^{n} c_{ij} \leq (n-1) \sum_{i=1}^{n} c_{ii} \tag{12}$$

is obtained in connection with Maxwell's influence numbers c_{ij}.

Ein Satz im Zusammenhang mit der Formänderungsenergie. — Gegenstand der Studie ist Beweis einer Ungleichheitsbeziehung im Zusammenhang mit der Formänderungsenergie des Körpers aus linear elastischem Material.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (1—2), pp. 91—108 (1980)

TORSIONAL BUCKLING OF A CANTILEVER SUBJECTED TO DISTRIBUTED NORMAL LOADS

K. ZALKA*

[Manuscript received November 21, 1978]

The paper deals with the spatial stability analysis of a bar with a built-in lower and a free upper end, subjected to distributed normal loads along its axis. The simultaneous differential equations of combined torsional and flexural buckling are set up. The exact and approximate computation of the critical load of a bar with double-symmetric cross section is dealt with in detail. Closed formulae are given for the approximate analysis of the critical load of a bar with mono-symmetric cross section.

1. Introduction

According to the elementary stress analysis the elements of a bar exhibit three kinds of deformations: buckling in the two principal planes and pure torsional buckling. During torsion the two adjacent cross sections of the element have a relative rotation (twist) with respect to an axis through the shear center. Due to the three kinds of deformations the axis of the bar generally becomes a three-dimensional curve. Taking these possibilities into consideration we can speak about the phenomena of combined torsional and flexural buckling: an axially compressed bar with initially straight axis loses its stability at a critical load P_{cr} and buckling occurs by a combination of twist of the cross sections and bending of the axis of the bar. This critical load is generally smaller or at most equal to that calculated with the assumption of planar buckling.

In the case of bars with thin-walled open cross section the phenomena of torsional-flexural buckling is of great importance because the critical load of the pure torsional buckling can be considerably smaller than that of the planar buckling and so the critical load of the combined torsional-flexural buckling is small as well.

The general problems of the spatial stability analysis of a bar subjected to a dead vertical load have been analysed by several authors [2, 3, 4, 5, 8]. The value of the critical load equals the smallest roof of a tertiary algebraical equation deduced from the expansion of a third order determinant.

A bar with simple (hinged) supports has been analysed in detail by TIMOSHENKO assuming that the cross section has one axis of symmetry. For the

* K. ZALKA, Nyár u. 19. I/4, H-1043 Budapest, Hungary

calculation of the critical load a set of diagrams was given in function of different ratios of rigidity [8].

In practice, however, bars having a built-in lower and a free upper end, subjected to distributed normal loads (i.e. gravity loads) can be frequently encountered. The aim of this paper is to derive formulae of closed form for the calculation of the critical load of such bars.

The analysis of the critical load of a bar with double-symmetric cross section will be discussed in detail. A closed formula will be given for the approximate and a table and diagrams for the exact calculation of the critical load for pure torsion.

Closed, approximate formulas will be given for the calculation of the critical load of a bar with mono-symmetric and unsymmetric cross section. The exact calculation of the critical load in such cases will be presented in a following paper.

The bifurcation load of the bar will be investigated by assuming that the material of the bar is perfectly elastic and its cross section doesn't vary along the height. Deformation caused by shear forces will be neglected.

2. The simultaneous differential equations of the combined torsional-flexural buckling

Let us consider a bar with a lower built-in and an upper free end having an unsymmetrical cross section shown in Fig 1. The bar is subjected to uniformly distributed normal loads of intensity q (i.e. gravity loads). Let us set the coordinate system such that the axis z coincides with the centroidal axis of the

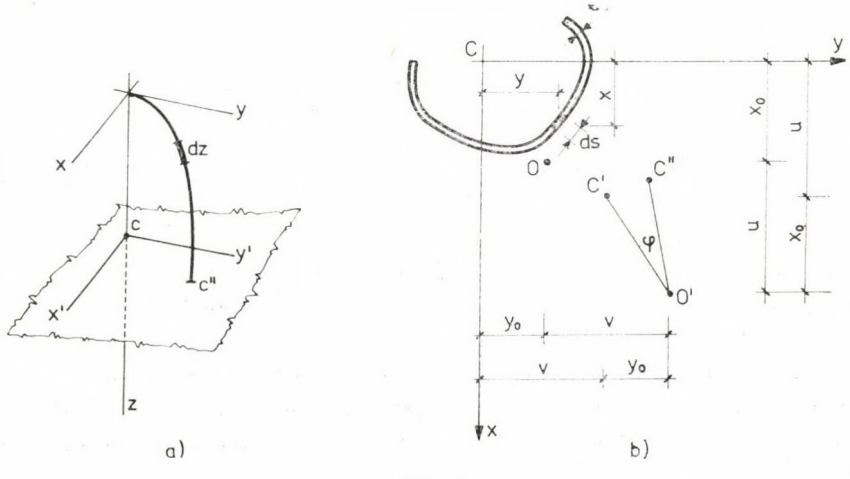


Fig. 1

bar and the axes x and y are the principal axes of inertia of the free upper cross section of the bar. The axes x and y are fixed to the upper cross section therefore they remain principal axes of intertia during buckling as well. C denotes the centroid and x_0, y_0 are the coordinates of the shear center O.

During buckling cross sections undergo translation and rotation. The translation is defined by the deflections u and v in the x and y direction, respectively, of the shear center O. Thus, during translation of the cross section point O moves to O' and point C to C'. The rotation of the cross section about the shear center is denoted by the angle φ and the final position of the centroid is C'' (Fig. 1). Therefore the final deflections of the centroid C during buckling are:

$$u + y_0\varphi \,, \tag{2.1}$$

$$v - x_0\varphi \,. \tag{2.2}$$

As a next move on the basis of the above statement the differential equation of the buckling of the shear-center axis in the direction x will be set up. The differential equation of the buckling in the direction y will be given disregarding details.

Let us consider the elementary section of the shear-center axis (Fig. 2). The following equations of equilibrium can be written.
Equation of moments:

$$N(du + y_0 d\varphi) - dM + T dz = 0 \,. \tag{2.3}$$

Hence:

$$T = - N(u' + y_0\varphi') + M' \,. \tag{2.4}$$

Equation of the horizontal projections:

$$dT = 0 \,,$$

or

$$T' = 0 \,. \tag{2.5}$$

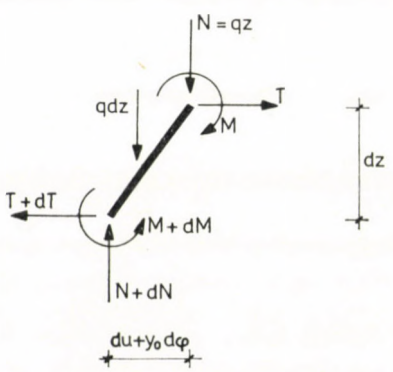


Fig. 8

Equation of the vertical projections:

$$qdz - dN = 0 . \tag{2.6}$$

Hence:

$$N' = q . \tag{2.7}$$

Let us differentiate Eq. (2.4) once with respect to z.
Using Eqs (2.5) and (2.7) we obtain the following equation:

$$-q(u' + y_0\varphi') - qz(u'' + y_0\varphi'') + M'' = 0 . \tag{2.8}$$

According to the elementary stress analysis

$$M = -EIu'' ,$$

so the differential equation for the deflected shape of the shear-center axis becomes

$$u'''' + \frac{N}{EI_y}(u'' + y_0\varphi'') + \frac{q}{EI_y}(u' + y_0\varphi') = 0 . \tag{2.9}$$

Proceeding as before the differential equation for the deflected shape in the direction y can be obtained:

$$v'''' + \frac{N}{EI_x}(v'' - x_0\varphi'') + \frac{q}{EI_x}(v' - x_0\varphi') = 0 . \tag{2.10}$$

Since the point of origin of the coordinate system has been fixed on the top of the column, Eqs (2.9) and (2.10) have the following boundary conditions.

a) At the top of the column the displacement in the directions x and y equals zero, respectively:

$$v(O) = u(O) = 0 . \tag{2.11}$$

b) At the bottom of the column the slope of the deflection in the directions x and y, respectively, is parallel with the axis z:

$$v'(H) = u'(H) = 0 . \tag{2.12}$$

c) At the top of the column bending moments equal zero:

$$v''(O) = u''(O) = 0 . \tag{2.13}$$

d) At the top of the column the value of the shear forces equals zero:

$$v'''(O) = u'''(O) = 0 . \tag{2.14}$$

The two differential equations for bending of the bar (2.9), (2.10) contain u, v and φ as unknown quantities. The required third equation is found by considering the torsion of the bar. As a next move this equation will be presented.

Let us take a longitudinal strip of cross section $t\,ds$ defined by coordinates x, y in the plane of the cross section (Fig. 1, Fig. 3). The components of its deflection in the directions x and y during buckling are, respectively,

$$u + (y_0 - y)\varphi\,, \tag{2.15}$$

and

$$v - (x_0 - x)\varphi\,. \tag{2.16}$$

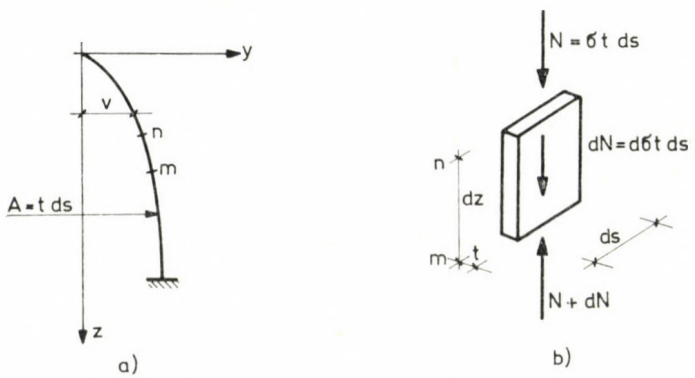


Fig. 3

The elementary strip is subjected to a compressive unit stress:

$$\sigma = \frac{zq}{A}\,, \tag{2.17}$$

wherein A denotes the cross-sectional area of the bar. The compressive force acting on the slightly rotated end of the element defined by the coordinate z is (Fig. 3):

$$N = \sigma t\,ds = zq\frac{t\,ds}{A}\,. \tag{2.18}$$

Considering the equilibrium of an element of length dz on the xz and yz planes of projections the following equations can be written.

Plane yz

Equation of moments:

$$N\{dv - (x_0 - x)\,d\varphi\} - dM + T dz = 0\,.$$

Hence:

$$T = M' - N\{v' - (x - x_0)\varphi'\} . \tag{2.19}$$

Equation of the horizontal projections:

$$dT = 0 ,$$

or

$$T' = 0 . \tag{2.20}$$

Equation of the vertical projections:

$$q\,dz\,\frac{t\,ds}{A} - dN = 0 .$$

Hence:

$$N' = q\,\frac{t\,ds}{A} . \tag{2.21}$$

Differentiating Eq. (2.19) once and substituting it into Eq. (2.20) using Eq. (2.21) it is found that

$$M'' - zq\frac{t\,ds}{A}\{v'' - (x_0 - x)\varphi''\} - q\frac{t\,ds}{A}\{v' - (x_0 - x)\varphi'\} = 0 . \tag{2.22}$$

Since

$$M = -EIv'' ,$$

the differential equation for the deflected shape of the elementary strip in the plane yz is as follows:

$$EI_x\,v'''' = -qz\frac{t\,ds}{A}\{v'' - (x_0 - x)\varphi''\} - q\,\frac{t\,ds}{A}\{v' - (x_0 - x)\varphi'\} . \tag{2.23}$$

We can now realize that Eq. (2.23) is the differential equation of an ordinary beam subjected to bending. In this case the right side of the equation can be understood as a fictitious load intensity in the plane yz:

$$-qz\frac{t\,ds}{A}\{v'' - (x_0 - x)\varphi''\} - q\frac{t\,ds}{A}\{v' - (x_0 - x)\varphi'\} . \tag{2.24}$$

The fictitious load intensity in the plane xz can be obtained in a similar manner:

$$-qz\frac{t\,ds}{A}\{u'' + (y_0 - y)\varphi''\} - q\,\frac{t\,ds}{A}\{u' + (y_0 - y)\varphi'\} . \tag{2.25}$$

These fictitious loads give rise to twisting. Taking the moment about the shear-center axis of the above forces for one longitudinal strip the following torque per unit length of the bar is obtained:

$$dm_z = +qz\frac{t\,ds}{A}(x_0 - x)\{v'' - (x_0 - x)\varphi''\} + q\frac{t\,ds}{A}(x_0 - x)\times$$

$$\times\{v' - (x_0 - x)\varphi'\} - qz\frac{t\,ds}{A}(y_0 - y)\{u'' + (y_0 - y)\varphi''\} -$$

$$- q\frac{t\,ds}{A}(y_0 - y)\{u' + (y_0 - y)\varphi'\}\,. \tag{2.26}$$

Integrating over the entire cross-sectional area and observing that

$$\int_A t\,ds = \mathrm{A}\,,$$

$$\int_A x^2 t\,ds = I_y\,,\ \int_A y^2 t\,ds = I_x\,,$$

$$I_0 = I_x + I_y + A(x_0^2 + y_0^2)\,,$$

$$\int_A x\,t\,ds = \int_A y\,t\,ds = 0\,,$$

we obtain

$$m_z = \int_A dm_z = qz(x_0 v'' - y_0 u'') + q(x_0 v' - y_0 u') - qz\frac{I_0}{A}\varphi'' - q\frac{I_0}{A}\varphi'\,. \tag{2.27}$$

In these expressions I_x and I_y are the principal centroidal moments of inertia of the cross section and I_0 is the polar moment of inertia about the shear center O.

The torque per unit length of the internal forces for nonuniform torsion is given by the following expression:

$$\frac{dM_t}{dz} = GI_t\,\varphi'' - EI_\omega\,\varphi'''\,, \tag{2.28}$$

where the following notations are used:

G shearing modulus of elasticity,
I_t torsinal constant,
GI_t torsional rigidity,
I warping constant,
EI_ω warping rigidity.

The positive direction of M_t and m_z are given by the right-hand rule (Fig. 4):

$$m_z = -\frac{dM_t}{dz}\,,$$

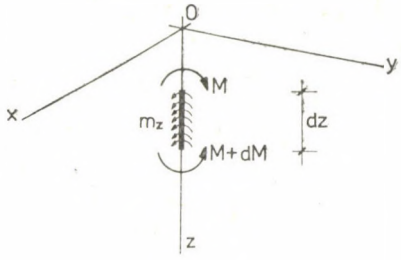


Fig. 4

and thus, using Eqs (2.27) and (2.28) we obtain a fourth order homogeneous differential equation of variable co-efficient for torsion:

$$EI_\omega \varphi'''' - \left(GI_t - qz\frac{I_0}{A}\right)\varphi'' + q\frac{I_0}{A}\varphi' - qz(x_0 v'' - y_0 u'') - q(x_0 v' - y_0 u') = 0. \tag{2.29}$$

Since the axes x and y have been fixed on the upper free cross section of the bar, Eq. (2.29) has the following boundary conditions.

a) At the top of the bar torsional deflection equals zero:

$$\varphi(O) = 0\,. \tag{2.30}$$

b) At the bottom of the bar there is a built-in end, so there is no warping there:

$$\varphi'(H) = 0\,. \tag{2.31}$$

c) At the top of thear on the upper cross section no compressive stresses are acting, so warping stresses cannot be reacted, either:

$$\varphi''(O) = 0\,. \tag{2.32}$$

d) At the top of the bar torsional moments equal zero:

$$GI_t\varphi'(O) - EI_\omega\varphi'''(O) = 0\,. \tag{2.33}$$

Eqs (2.9), (2.10) and (2.29) are the three simultaneous differential equations for buckling by bending and torsion and can be used to calculate the critical load. The simultaneous differential equations (2.9), (2.10), (2.29) are identical to those derived by V. Z. VLASOV [9]. It is seen that the angle of rotation appears in all three equations, so that, in the general case, torsional buckling and bending of the axis occurs simultaneously.

The order of the differential equation (2.29) can be reduced. Carrying out the integration of Eq. (2.29) we obtain

$$\varphi''' - \frac{GI_t}{EI_\omega}\varphi' + \frac{q\,I_0}{EI_\omega\,A}\int z\varphi''\,dz + q\frac{I_0}{A}\varphi - \frac{q\,x_0}{EI_\omega}\int z\,v''\,dz +$$

$$+ \frac{q\,y_0}{EI_\omega}\int zu''\,dz - \frac{q\,x_0}{EI_\omega}v + \frac{q\,y_0}{EI_\omega}u + c_1 = 0$$

wherein

$$\int z\,\varphi''\,dz = z\,\varphi' - \varphi + c_2\,,$$

$$\int z\,v''\,dz = z\,v' - v + c_3\,,$$

$$\int z\,u''\,dz = z\,u' - u + c_4\,,$$

so the reduced equation is as follows:

$$\varphi''' - \left(\frac{GI_t}{EI_\omega} - \frac{q\,I_0}{EI_\omega\,A}z\right)\varphi' - \frac{q}{EI_\omega}z(x_0v' - y_0u') + c_5 = 0\,. \tag{2.34}$$

The boundary condition (2.33) gives zero for the value of the constant of integration c_5. Thus — since the order of Eqs (2.9), (2.10) can be reduced in a very similar way — the simultaneous differential equations for buckling by torsion and flexure may be written in the following simple form:

$$u''' + \frac{qz}{EI_y}(u' + y_0\varphi') = 0\,, \tag{2.35}$$

$$v''' + \frac{qz}{EI_x}(v' - x_0\varphi') = 0\,, \tag{2.36}$$

$$\varphi''' - \left(\frac{GI_t}{EI_\omega} - \frac{q\,I_0\,z}{EI_\omega\,A}\right)\varphi' - \frac{q\,z}{EI_\omega}(x_0v' - y_0u') = 0\,. \tag{2.37}$$

The boundary conditions for the differential equations are:

$$u(O) = v(O) = \varphi(O) = 0\,, \tag{2.38}$$

$$u'(H) = v'(H) = \varphi'(H) = 0\,, \tag{2.39}$$

$$u''(O) = v''(O) = \varphi''(O) = 0\,. \tag{2.40}$$

3. Bars with double-symmetric cross section

In the particular case when the shear center coincides with the centroid we have

$$x_0 = y_0 = 0\,,$$

and therefore Eqs (2.35) (2.36), (2.37) become

$$u''' + \frac{qz}{EI_y} u' = 0 , \tag{3.1}$$

$$v''' + \frac{qz}{EI_x} v' = 0 , \tag{3.2}$$

$$\varphi''' - \left(\frac{GI_t}{EI_\omega} - \frac{q\, I_0}{EI_\omega\, A} z \right) \varphi' = 0 . \tag{3.3}$$

Each of these equations contains only one unknown quantity and can be treated separately, so that torsional buckling is independent of flexural buckling in both planes.

Accordingly, through the stability analysis of bars having two axes of symmetry the following possibilities have to be taken into account.

a) The centrally compressed bar may buckle in the plane of the axis of x or y.

In these cases the critical load can be obtained by solving the eigenvalue problems given by Eqs (3.1) and (3.2). The parts of Eqs (2.38) to (2.40) concerning with deflections only give the boundary conditions. These kinds of equations are traceable to Bessel-equations [8]. Having used Bessel-functions the critical loads may be given by the following formulae:

$$N_{cr,x} = q\, H = \frac{7{,}84\, EI_y}{H^2} , \tag{3.4}$$

or

$$N_{cr,y} = q\, H = \frac{7{,}84\, EI_x}{H^2} . \tag{3.5}$$

b) Under compression a pure torsional buckling may occur. The axis of the bar remains straight, while the cross sections rotate about the shear-center axis. The approximate and accurate calculation of the critical load of the pure torsional buckling $N_{cr,\varphi}$ will be presented in the following sections.

Only the lowest of the three values of the critical load — $N_{cr,x}$, $N_{cr,y}$, $N_{cr,\varphi}$ — is of interest in practical applications.

3.1 The approximate calculation of the critical load for pure torsional buckling

Considering the physical aspect of the problem a lower bound can be given for the approximate calculation of the critical load. It is the theorem of R. V. SOUTHWELL that provides a possibility for us to do so. This lower bound will also be used for determining the accurate critical load.

In general the Southwell-theorem holds for eigenvalue-problems [1] stating that if the

$$S(v) = \chi^2 N(v) \tag{3.6}$$

eigenvalue-problem can be written in the form

$$S_1(v) + S_2(v) + \ldots = \chi^2 N(v)\,, \tag{3.7}$$

that is, it can be composed of "parts of problems", then — under certain conditions concerning the S and N functions — the lowest eigenvalue χ^2 is never smaller than the sum of the eigenvalues of the "parts of problems":

$$\chi^2 \geq \chi_1^2 + \chi_2^2 + \ldots = \sum_{i=1}^{n} \chi_i^2\,. \tag{3.8}$$

The theorem applying to stability problems is as follows. If the stiffness of an elastic beam may be considered the sum of the stiffnesses of several superposed parts, then the critical load is approximately or maybe exactly equal to the sum of the critical loads of the component parts but never smaller than that. Accordingly, the sum considered an approximate value is always smaller than the actual critical load or may be equal to it [7].

Thus, if the rigidity of a bar is considered the sum of torsional and warping rigidities, then the sum of the critical loads associated with this parts of rigidities gives a lower bound of the actual critical load. As a next step these "part" critical loads will be presented.

a) The critical load associated with the warping rigidity

In this case Eq. (3.3) becomes

$$\varphi''' + \frac{q\, i_p^2}{EI_\omega} z\varphi' = 0 \tag{3.9}$$

where the notation

$$i_p^2 = \frac{I_0}{A}$$

has been introduced.

The boundary conditions for Eq. (3.9) are those of (2.38). When solving the differential equation Timoshenko's deduction may be applied again. Thus the critical load associated with the warping rigidity may be calculated from the formula

$$N_{cr,\varphi}^{\omega} = \frac{1}{i_p^2} \frac{7{,}83\, EI_\omega}{H^2}\,. \tag{3.10}$$

The deflection-curve is shown in Fig. 5a.

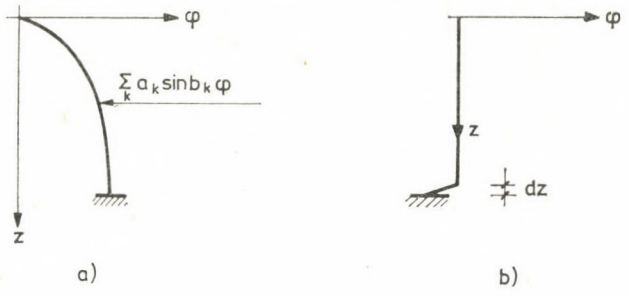


Fig. 5

b) The critical load associated with the torsional rigidity

In this case Eq. (3.3) yields the following differential equation:

$$\varphi'(q\, i_p^2 z - GI_t) = 0\,, \tag{3.11}$$

and the boundary condition for the equation is

$$\varphi(O) = 0\,.$$

On the section $O \leqq z < H$

$$\varphi' = 0$$

meets the requirements, and the value of the expression

$$q\, i_p^2 H - GI_t$$

is different from zero.

On the other hand, at the point $z=H$ the value of φ' is different from zero, thence

$$q\, i_p^2 H - GI_t = 0\,. \tag{3.12}$$

Since — according to the boundary condition — at the top of the bar the value of the twisting equals zero, there is no twisting on the whole section of

$$O \leqq z < H\,,$$

and at the point $z = H$ the bar undergoes a "sudden" twisting (Fig. 5b).

The critical load associated with the torsional rigidity can be obtained from Eq. (3.12):

$$N^t_{cr,\varphi} = q_{cr}\, H = \frac{1}{i_p^2} GI_t\,. \tag{3.13}$$

Applying the Southwell theorem, now we can give a lower bound for the approximate calculation of a bar developing pure torsional deformation. It is Eqs (3.10) and (3.13) that have to be summarised:

$$N_{cr,\varphi} = q_{cr} H \geq \frac{1}{i_p^2} \left(\frac{7{,}84 EI_\omega}{H^2} + GI_t \right) . \tag{3.14}$$

The above expression is identical to that obtained by L. KOLLÁR by an elementary deduction [6].

3.2. The exact analysis of the critical load of pure torsional buckling

Let us consider again the differential equation of pure torsional buckling (3.3) and introduce the following notations in order that non-dimensional quantities can be worked with:

$$\alpha = \frac{q H^3 i_p^2}{7{,}83 EI_\omega} = \frac{N_{cr}}{N_{cr,\varphi}^{\omega}} , \quad [-] \tag{3.15}$$

$$\beta = \frac{GI_t H^2}{7{,}83 EI_\omega} = \frac{N_{cr,}^{t}}{N_{cr,\varphi}^{\omega}} , \quad [-] \tag{3.16}$$

With these notations Eq. (3.3) may be rewritten as follows:

$$\varphi''' + \left(\frac{7{,}83\alpha}{H^3} z - \frac{7{,}84\beta}{H^2} \right) \varphi' = 0 . \tag{3.17}$$

On the base of (2.38) to (2.40) this third-order linear, homogeneous differential equation of variable coefficients has the following boundary conditions:

$$\varphi(O) = 0 , \tag{3.18}$$

$$\varphi'(H) = 0 , \tag{3.19}$$

$$\varphi''(O) = 0 . \tag{3.20}$$

The differential equation has no closed form solution. Let us try to find the solution in the form of the power series:

$$\varphi = c_0 + \sum_{k=1}^{\infty} c_k z^k \tag{3.21}$$

where c_k denotes the coefficients yet unknown. Having substituted the necessary derivatives

$$\varphi' = \sum_{k=1}^{\infty} k c_k z^{k-1} , \tag{3.22}$$

$$\varphi'' = \sum_{k=2}^{\infty} k\,(k-1)\,c_k\,z^{k-2}\,, \tag{3.23}$$

$$\varphi''' = \sum_{k=3}^{\infty} k(k-1)\,(k-2)\,c_k\,z^{k-3} \tag{3.24}$$

into Eq. (3.17) it follows that the co-efficients of each power z have to be equal to zero. Thence a formula can be generated for the sucessive determination of the unknown co-efficients c_k:

$$c_k = \frac{7{,}83\,\beta\,c_{k-2}}{(k-1)k\,H^2} - \frac{(k-3)\,7{,}83\,\alpha\,c_{k-3}}{(k-2)(k-1)\,k\,H^3}\,, \quad k \geqq 3\,. \tag{3.25}$$

The initial values needed for the calculation are obtained by making use of the boundary conditions (3.18), (3.20) and of the power series (3.21) and its derivative (3.23). The coefficient c_1 appears in the expressions of all other co-efficients, so its value has no influence on the determination of the eigenvalue. Thus

$$c_0 = 0\,, \tag{3.26}$$

$$c_2 = 0\,, \tag{3.27}$$

and

$$c_1 = c_1\,. \tag{3.28}$$

By making use of the last boundary condition (3.19) — which has not been used up yet — and of the expression (3.22) a "conditional" equation may be obtained:

$$F(\alpha) = c_1 + \sum_{k=3}^{\infty} k\,c_k\,H^{k-1} = 0\,. \tag{3.29}$$

The coefficients c_k in the equation can be calculated from (3.25).

The function $F(\alpha)$ vanishes only if the eigenvalue of the problem, that is the critical load — or more precisely its proportional part (see Eq. 3.15) — is substituted into the expressions of c_k. In this way the value of the critical load can be calculated with the required accuracy by trial and error. For this end the "conditional" equation has been solved for different ratios of rigidity β; the error limit was $\Delta = 10^{-8}$. The eigenvalues needed for the calculation of the critical load are shown in Fig. 6, Fig. 7 and Table 1.

The values of the critical load have been calculated by the aid of Southwell's theorem (3.14) as well:

$$N_{cr,\varphi} = \frac{1}{i_p^2}\left(\frac{7{,}83\,EI_\omega}{H^2} + GI_t\right)\,. \tag{3.30}$$

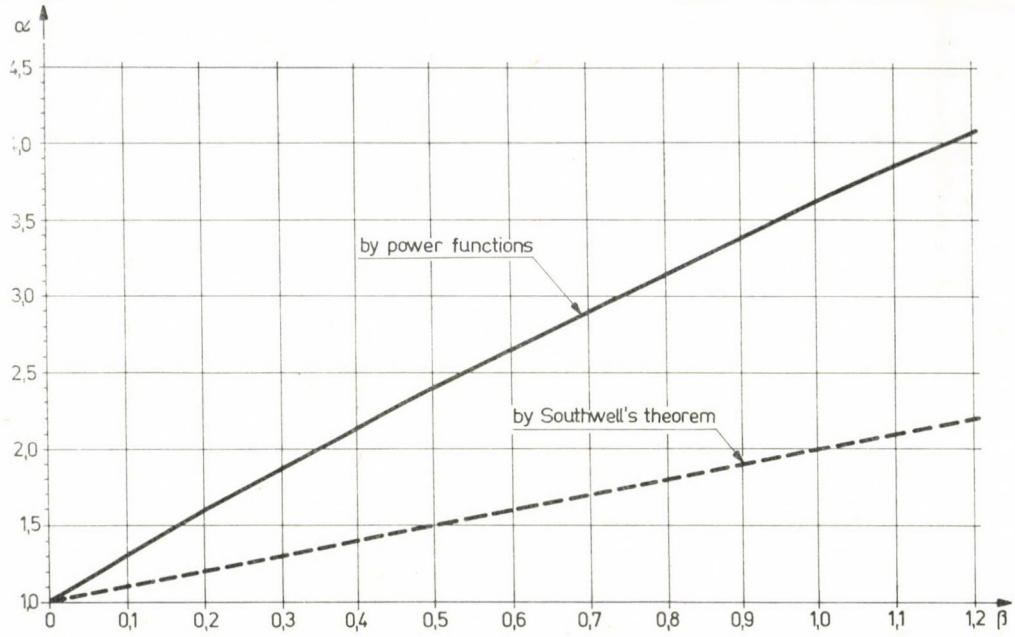

α
4,5
4,0
3,5
3,0
2,5
2,0
1,5
1,0
by power functions
by Southwell's theorem
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1,0
1,1
1,2
β

Fig. 6

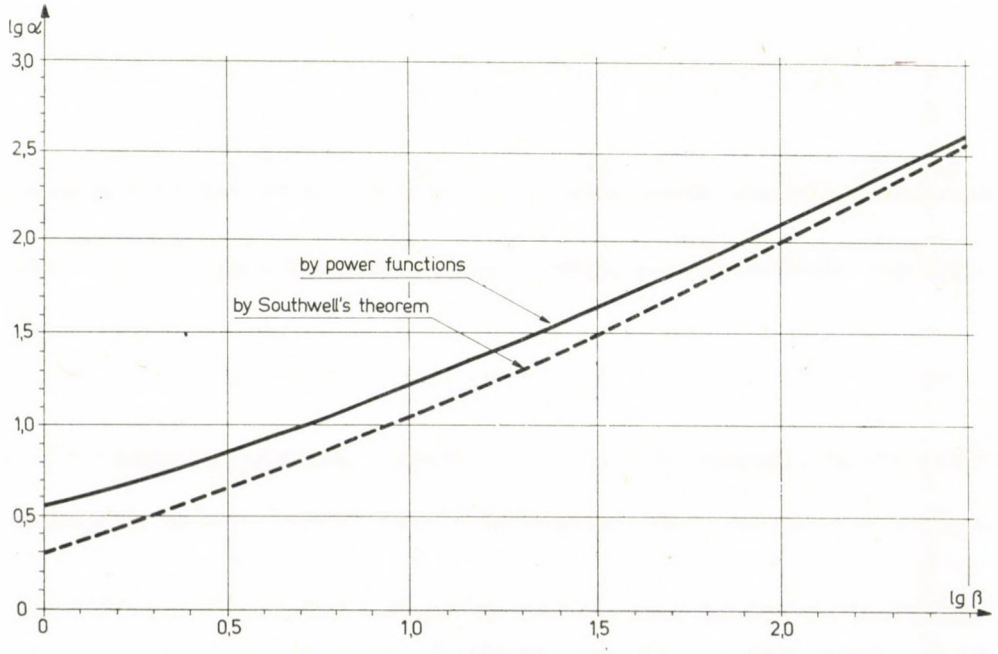

lg α
3,0
2,5
2,0
1,5
1,0
0,5
0
by power functions
by Southwell's theorem
0
0,5
1,0
1,5
2,0
lg β

Fig. 7

Table 1

β Formula (3.16)	α Formula (3.15)	β Formula (3.16)	α Formula (3.15)	β Formula (3.16)	α Formula (3.15)
0,0	1,000	2	5,62	60	82,26
0,001	1,003	3	7,43	70	94,40
0,01	1,03	4	9,10	80	106,4
0,1	1,29	5	10,70	90	118,4
0,2	1,58	6	12,24	100	130,2
0,3	1,86	7	13,75	120	153,8
0,4	2,12	8	15,23	140	177,2
0,5	2,38	9	16,68	160	200,3
0,6	2,63	10	18,12	180	223,3
0,7	2,88	20	31,82	200	246,2
0,8	3,12	30	44,86	300	359,4
0,9	3,35	40	57,54	400	471,4
1,0	3,58	50	69,99	500	580,8

These values, represented by broken lines in Figs 6 and 7, have been considered suitable initial values for the successive approximation. As it can be seen from Fig. 6 and 7, formula (3.30) gives the value of the critical load with an approximation always on the safe side. The maximum deviation from the exact value is 47 per cent. The reason of the deviation is that Southwell theorem gives exact result only if the basic problems have identical buckling forms in nature. If the buckling curves are different — in our case even the style of the curves are entirely different (Fig. 5) — the theorem gives the value of the critical load to the benefit of safety corresponding with the "degree" of the difference.

4. Bars with mono-symmetric cross section

Let us assume that the axis y is an axis of symmetry. In this case we have

$$x_0 = 0 ,$$

and the differential equations of the problem become

$$v''' + \frac{qz}{EI_x} v' = 0 , \tag{4.1}$$

$$u''' + \frac{qz}{EI_y} (u' + y_0 \varphi') = 0 , \tag{4.2}$$

$$\varphi''' - \left(\frac{GI_t}{EI_\omega} - \frac{q\, z\, I_0}{EI_\omega A} \right) \varphi' + \frac{q\, z\, y_0}{EI_\omega} u' = 0 . \tag{4.3}$$

The boundary conditions are still Eqs (2.38) to (2.40). In this case pure torsional buckling interacts with flexural buckling only in the plane xz. Accordingly, the stability analysis of a bar with mono-symmetric cross section may be carried out bearing in mind the following possibilities.

a) The centrally compressed bar may buckle in the plane of symmetry. Having solved Eq. (4.1) the value of the critical load is given by the formula

$$N_{cr,y} = \frac{7{,}83\ EI_x}{H^2}\,. \tag{4.4}$$

b) Buckling perpendicular to the plane of symmetry is combined with torsion resulting in buckling by torsion and flexure.

The exact value of the critical load for torsional-flexural buckling would be obtained by solving the simultaneous differential equations (4.2) and (4.3), but in this paper only an approximate solution will be given instead.

The critical load can be easily calculated using the critical loads of the basic problems by the aid of Föppl's theorem [10]:

$$\frac{1}{N_{cr,\,comb}} = \frac{1}{N_{cr,x}} + \frac{1}{N_{cr,\varphi}}\,, \tag{4.5}$$

wherein

$N_{cr,\varphi}$ critical load for pure torsional buckling. Its value can be obtained by using Fig. 6. and Fig. 7. according to instructions in Section 3.2.,

$N_{cr,x} = \dfrac{7{,}83\ EI_y}{H^2}$ critical load for flexural buckling in direction x,

Formula (4.5) yields the critical load with an approximation always on the safe side.

Only the lowest of the critical loads given by (4.4), (4.5) is of practical importance.

5. Bars with unsymmetric cross section

In the general case of a bar of thin-walled open cross section with no axes of symmetry, buckling failure occurs by a combination of bending in both xz and yz planes and torsion. The solution of the symultaneous differential equations (2.35) to (2.37) would furnish an exact result for the critical load, but instead of this we again give an approximate formula based on Föppl's theorem:

$$\frac{1}{N_{cr,\,comb}} = \frac{1}{N_{cr,x}} + \frac{1}{N_{cr,y}} + \frac{1}{N_{cr,\varphi}}\,. \tag{5.1}$$

The notation in (5.1) are:

$N_{cr,y} = \dfrac{7{,}83\ EI_x}{H^2}$ critical load for flexural buckling in direction y,

$N_{cr,x} = \dfrac{7{,}83\ EI_y}{H^2}$ critical load for flexural buckling in direction x,

$N_{cr,\varphi}$ critical load for pure torsional buckling. See Section 3.2.

REFERENCES

1. COLLATZ, L.: Eigenwertaufgaben mit technischen Anwendungen. Leipzig 1949, 388
2. CSELLÁR, Ö.: Spatial Buckling of Compressed Thin-walled Bars.* *Magyar Építőipar* (1971), 338—343
3. CSELLÁR, Ö.—HALÁSZ, O.: Die Stabilitätsuntersuchung gedrückter dünnwandiger Stäbe. *Acta Techn. Hung.* **36** (1961), 545—554
4. CSELLÁR, Ö.—HALÁSZ, O.—RÉTI, V.: Thin-walled Steel Structures.* Műszaki Könyvkiadó, Budapest 1965.
5. HALÁSZ, O.: Steel Structures III/1. Stability Analysis.* Tankönyvkiadó, Budapest 1970
6. KOLLÁR, L.: Stiffening of Buildings against Torsional Buckling. Regional Colloquium on Stability of Steel Structures. Budapest—Balatonfüred, 1977.
7. KORÁNYI, I.: Stability Problems in Engineering Practice. Plane-buckling.* Akadémia Kiadó, Budapest 1965, 111—114
8. TIMOSHENKO, S.—GERE, J.: Theory of Elastic Stability. McGraw-Hill, New York 1961
9. VLASOV, V. Z.: Thin-walled Elastic Bars. (In Russian), Moscow 1940
10. FÖPPL, L.: Über das Ausknicken von Gittermasten, insbesondere von hohen Funktürmen. *Zeitschrift für Angewandte Mathematik und Mechanik* **13** (1933), 1—10.

* in Hungarian

Torsionsknickung eines einseitig eingspannten Stabes unter verteilten Längsbelastungen.— Die Arbeit bringt die räumliche Stabilitätsanalyse eines Balkens mit eingespanntem unteren und freiem oberen Ende unter entlang seiner Achse verteilten Längsbelastungen. Die simultanen Differentialgleichungen für kombinierte Torsions- und Biegeknickung werden aufgestellt. Die exakte und die angenäherte Berechnung der kritischen Last eines Balkens mit zweifach symmetrischem Querschnitt werden eingehend behandelt. Für die angenäherte Analyse der kritischen Belastung eines Balkens mit einfach symmetrischem und mit asymmetrischem Querschnitt werden geschlossene Formeln angegeben.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (1—2), pp. 109—132 (1980)

STABILITY INVESTIGATION OF RETARDED DIFFERENTIAL EQUATIONS

G. STÉPÁN*

[Manuscript received May 4, 1979]

This paper presents and proves a proposition on the basis of which an algorithm may be established for the stability investigation of homogeneous, linear, retarded differential equations of constant coefficients. This algorithm may be used for the retarded differential equations in the same way as the Routh—Hurwitz criterion for the ordinary differential equations. In the very same way as the theory of the retarded differential equations involves as a special case the theory of the ordinary differential equations, also the algorithm may be utilized in lieu of the Routh—Hurwitz criterion for the ordinary differential equations. Finally, it showes the practical applicability of the results for the case of the reproductive machine tool vibrations.

1. Introduction

The mathematical research for retarded differential equations began to develop in recent years. Numerous economical and biological models as well as some models of population development, of technical problems and of time-retarded feed-back control systems are retarded differential equations. In investigating the problems mentioned above, it is necessary that the stability investigation of the linearized, homogeneous retarded differential equations of constant coefficients can be carried out as simply as possible as a function of the coefficients as parameters.

The general form of ordinary, linear, homogeneous differential equations of constant coefficients is as follows:

$$\dot{x} = Ax\,, \tag{1.1}$$

and the associated characteristic equation is

$$\det(\lambda I - A) = 0\,. \tag{1.2}$$

Herein $\lambda \in C$ where C is the set of the complex numbers, A the coefficient matrix of nxn dimension of the differential equation, $x{:}R \to R^n$ a function, I unity

* STÉPÁN G. Bánki Donát u 14/A, H-1148 Budapest, Hungary

matrix of nxn dimension and R the set of the real numbers. (1.2) is a polynome of nth order of λ. The solution $x \equiv 0$ to the differential equation (1.1) is asymptotically stable in the case when the real part of each root of (1.2) is negative. The calculation of these complex roots is difficult, wherefore, numerous attempts have been made to develop an algorithm in order to make the stability investigation easier.

Such an algorithm was originated by ROUTH in 1875, and in 1895 HURWITZ developed a similar algorithm independently of the former. In 1916 these were simplified to some extent by LIÉNARD and CHIPART. The Routh—Hurwitz criterion [2] has ever since been widely used. The verification of the criterion is to be found, for example, in [3].

The general shape of the retarded, linear, homogeneous differential equation of constant coefficients is the following

$$\dot{x}(t) = Ax(t) + Bx(t - r) , \tag{1.3}$$

wherein $r \in R$; B_{nxn}, A_{nxn} are real matrices, $r > 0$. The characteristic equation of (1.3) is obtained by the substitution of the solution $x(t)$ for ce^{zt}, $c \in R^n$, $z \in C$; in this case the role of λ has been taken over by z:

$$\det (zI - A - Be^{-rz}) = 0 . \tag{1.4}$$

If the real part of each root of (1.4) is negative, the differential equation (1.3) is asymptotically stable [1]. The problem is now far more complicated than that described above because the transcendent equation (1.4) has, in general, infinitely many complex roots. The stability investigation seems to be hopeless in this way although HALE presented a theorem which characterizes the positions of the roots on the complex plane (see [1], Lemma 4.1., page 18). Theoretically the stability investigation may be carried out by establishing Ljapunov functions. The method presented by HALE for the unidimensional case may also be applied; with its help the determination of the boundaries of the stability regions in the field of coefficients may be reduced to the solution of a boundary-value problem related to an ordinary differential equation. HALE's comment on this method is that in case of higher dimensions, the problem is so complex that it becomes insoluble.

The one and only procedure which deserves attention from the viewpoint of practical calculation originates from PONTRJAGIN [4] (1942). He investigated the positions of the zeroes of the polynomes of two variables $P(z,w)$ in the case of $w = e^z$. The characteristic equation (1.4) may be reduced to the polynome of the form $P(z,w)$ so that Pontrjagin's theorems may be applied (see [1], Th.A. 1—3., p. 338). However, their application is rather difficult because prior to it a time-transformation should be carried out and a transcendent equation should be examined from the point of view whether it has a complex root at all. This

problem is intricate even in case one applies another theorem of PONTRJAGIN which makes this investigation easier. Although it has no practical significance, it should be noted that in this way only comparable delays are permitted. HAYES, BELLMANN and COOKE (see [1], Th.A.5—6, p. 339) obtained their results in the investigation of the stability of somewhat simpler retarded differential equations by referring to PONTRJAGIN.

PONTRJAGIN achieved his results by extending and generalizing the methods presented in the verification of the Routh—Hurwitz criterion.

In the following another algorithm will be presented for the stability investigation. It is different from the one discussed above and can easily and generally be applied. The algorithm will be presented for the sake of simplification only in connection with the retarded differential equation of nth order

$$\sum_{k=0}^{n} a_k \frac{d^k}{dt^k} x(t) = bx(t-r)\,, \quad a_k, r \in R, a_n > 0\,, \quad r > 0, n > 1 \tag{1.5}$$

and the results obtained will be generalized later.

2. The stability investigation

A theorem will be stated and proved as a basis of the algorithm. The integer part of k will be symbolized by $[k]$.

Theorem:

Consider the retarded differential equation

$$\sum_{k=0}^{n} a_k \frac{d^k}{dt^k} x(t) = bx(t-r); \quad a_k, b, r \in R, k = 1, \ldots, n\,; \quad a_n > 0;\ r > 0;\ n > 1 \tag{2.1}$$

the characteristic equation of which is

$$D(z) = \sum_{k=0}^{n} a_k z^k - be^{-rz}\,. \tag{2.2}$$

Be

$$M(y) = Re(D(iy)) = \sum_{j=0}^{\left[\frac{n}{2}\right]} a_{2j}(-1)^j y^{2j} - b\cos(ry)\,, \tag{2.3}$$

$$S(y) = Im(D(iy)) = \sum_{j=0}^{\left[\frac{n-1}{2}\right]} a_{2j+1}(-1)^j y^{2j+1} + b\sin(ry)\,, \tag{2.4}$$

wherein $y \in R$. Let $y_1 \geq y_2 \geq \ldots \geq y_m \geq 0$ designate the positive real zeroes of $M(y)$ and y_1^* the maximum zero-locus of $S(y)$. q is an integer to which $y_q > y_1^* > y_{q+1}$. The solution $x \equiv 0$ to the Eq. (2.1) is asymptotically stable, if and only if

$$\left(\sum_{k=\left(n-2\left[\frac{n}{2}\right]\right)q+1}^{m} (-1)^{k+1} \operatorname{sign}(S(y_k)) + \right.$$

$$\left. + \frac{1}{2}\left(n - 2\left[\frac{n}{2}\right]\right)(-1)^{\left[\frac{n-1}{2}\right]+q+1} \right) \operatorname{sign}(M(y_1+1)) + \frac{n}{2} = 0 , \quad (2.5)$$

and

$$S(y_k) \neq 0 , \quad k = 1, \ldots, m . \quad (2.6)$$

Proof:

The solution $x \equiv 0$ is asymptotically stable if and only if the real parts of the roots of the characteristic equation associated with the retarded differential Eq. (2.1) are negative [1]. In the course of verification it will be realized that under the conditions of the *Theorem* the characteristic equation has no roots either in the positive half of the complex plane, or on the imaginary axis. The characteristic polynome (2.2) in the *Theorem* is a specific case of the Eq. (1.4).

First, one should realize that due to the condition (2.6) $D(z)$ will have no root on the imaginary axis. To prove this statement the real and imaginary parts of $D(z)$ are needed. These are, in case of $z = x + iy$, as follows:

$$\operatorname{Re}(D(z)) = \sum_{k=0}^{n} a_k \sum_{l=0}^{\left[\frac{k}{2}\right]} \binom{k}{2l} (-1)^l x^{k-2l} y^{2l} - be^{-rx} \cos(ry) , \quad (2.7)$$

$$\operatorname{Im}(D(z)) = \sum_{k=0}^{n} a_k \sum_{l=0}^{\left[\frac{k-1}{2}\right]} \binom{k}{2l+1} (-1)^l x^{k-2l-1} y^{2l+1} + be^{-rx} \sin(ry) . \quad (2.8)$$

On the imaginary axis $z = iy_0$, here the expressions (2.7) and (2.8) will be simplified according to (2.3) and (2.4).

Let us suppose that $z_0 = iy_0$ is a root of $D(z)$. In this case $M(y_0) = 0$ and $S(y_0) = 0$. However, by virtue of (2.6) $S(y_k) \neq 0$, wherein y_k is the root of $M(y)$, wherefore, $D(z)$ has no root whose real part is equal to zero.

Now let us realize that $D(z)$ has no root in the positive half plane either. For this purpose consider the following curves on the complex plane of numbers:
(K_R): $z = \operatorname{Re}^{i\varphi}$, the value of φ ranges from $-\pi/2$ to $\pi/2$;
(e_1): $z = iy$, the value of y ranges from R to 0;
(e_2): $z = iy$, the value of y ranges from 0 to $-R$.

$$(g^*) = (K_R) \cup (e_1) \cup (e_2) .$$

(g^*) is a closed curve. Be $(g) = \lim_{R\to\infty}(g^*)$ which encloses the whole positive half plane. $D(z)$ has no pole and by virtue of [1] (Lemma 4.1., p. 18) it has at most roots of finite number within (g). In this case:

$$\lim_{R\to\infty} \oint_{(g^*)} \frac{\frac{dD(z)}{dz}}{D(z)}\, dz = \lim_{R\to\infty} \ln\,(D(z))\,|_{(g^*)} =$$

$$= \lim_{R\to\infty} (\ln \mid D(z) \mid + i\varphi(z))\,|_{(g^*)} = i\Delta_{(g)}\varphi(z) = N \cdot 2\pi i\,, \tag{2.9}$$

wherein

$$\varphi(z) = \arctan \frac{\mathrm{Im}(D(z))}{\mathrm{Re}(D(z))}\,, \tag{2.10}$$

$\Delta_{(g)}$ is the change along the curve (g) while N is the number of the $D(z) = 0$ roots of positive real parts, that is, of those which fall into the region bordered by the curve (g). Thus, the condition of the stability may be formulated as follows: the value of N must be equal to zero

$$\oint_{(g)} \frac{\frac{dD(z)}{dz}}{D(z)}\, dz = 0\,, \tag{2.11}$$

that is,

$$\Delta_{(g)}\varphi(z) = 0\,. \tag{2.12}$$

Let us first calculate the value of the integral along (K_R):

$$\lim_{R\to\infty} \int_{(K_R)} \frac{\frac{dD(z)}{dz}}{D(z)}\, dz = \lim_{R\to\infty} \int_{(K_R)} \frac{\sum_{k=1}^{n} k a_k z^{k-1} + rbe^{-rz}}{\sum_{k=0}^{n} a_k z^k - be^{-rz}}\, dz =$$

$$= \lim_{R\to\infty} \int_{\varphi=-\pi/2}^{\pi/2} \frac{\sum_{k=1}^{n} k a_k R^{k-1} e^{i\varphi(k-1)} + rbe^{-rR(\cos\varphi + i\sin\varphi)}}{\sum_{k=0}^{n} a_k R^k e^{i\varphi k} - be^{-rR(\cos\varphi + i\sin\varphi)}}\, Re^{i\varphi}\, id\varphi =$$

$$= \int_{\varphi=-\pi/2}^{\pi/2} \frac{na_n e^{i\varphi n} + \lim_{R\to\infty} \left(\sum_{k=1}^{n-1} k a_k R^{k-n} e^{i\varphi k} + rbR^{1-n} e^{-rR\cos\varphi}\, e^{i(\varphi - rR\sin\varphi)} \right)}{a_n e^{i\varphi n} + \lim_{R\to\infty} \left(\sum_{k=0}^{n-1} a_k R^{k-n} e^{i\varphi k} - bR^{-n} e^{-rR\cos\varphi}\, e^{-rRi\sin\varphi} \right)} i\, d\varphi\,,$$

$$\lim_{R\to\infty} \int_{(K_R)} \frac{\frac{dD(z)}{dz}}{D(z)}\, dz = n\pi i\,. \tag{2.13}$$

In the calculation, the advantage that the integrand is continuous, has been utilized. The condition of the stability may, on the basis of Eq. (2.11) be changed as follows:

$$\lim_{R\to\infty}\int_{(e_1)\cup(e_2)}\frac{\frac{dD(z)}{dz}}{D(z)}\,dz=\lim_{R\to\infty} i\Delta_{(e_1)\cup(e_2)}\varphi(z)=-n\pi i\,. \tag{2.14}$$

$\varphi(z)$ is made with the help of Eq. (2.10). $S(y)$ is an uneven, $M(y)$ is an even function, therefore the equality

$$\Delta_{(e_1)}\varphi=\operatorname{arc\,tan}\frac{S(y)}{M(y)}\Big|_{y=R}^{0}=\operatorname{arc\,tan}\frac{-S(y)}{M(y)}\Big|_{y=-R}^{0}=\operatorname{arc\,tan}\frac{S(y)}{M(y)}\Big|_{y=0}^{-R}=\Delta_{(e_2)}\varphi \tag{2.15}$$

is true.

Thus the condition of the stability may be reformulated as follows:

$$\lim_{R\to\infty}\Delta_{(e_1)}\varphi=\operatorname{arc\,tan}\frac{\sum_{k=0}^{[(n-1)/2]}a_{2k+1}(-1)^k y^{2k+1}+b\sin(ry)}{\sum_{k=1}^{[n/2]}a_{2k}(-1)^k y^{2k}+a_0-b\cos(ry)}\Bigg|_{y=\infty}^{0}=-\frac{n}{2}\pi\,. \tag{2.16}$$

The real positive roots of $M(y)$ are $y_1\geq y_2\geq\ldots\geq y_m$, wherein the root of k multiplicity is marked k times with subscripts. (For example, if 1,7 is a triple root, then $y_1=3$; $y_2=1{,}7$; $y_3=1{,}7$; $y_4=1{,}7$; $y_5=1{,}1$.)

Be the real, non-negative roots of $S(y)$ $y_1^*, y_2^*, \ldots, y_p^*$, where $y_1^*\geq y_2^*\geq\ldots,\geq y_p^*\geq 0$.

Let us now examine the value of $\lim_{R\to\infty}\Delta_{(e_1)}\varphi$ with the aid of (2.16); this will be carried out in several steps.

$$a\rangle\ \text{Be } M(y_1+1)>0\,,\quad \text{that is,}\quad \lim_{y\to\infty}M(y)=+\infty\,.$$

$\alpha\rangle$ Be in $I_i=(y_{i+1}^*, y_i^*)$ one single root of $M(y)$ and this should have an odd subscript: y_{2k+1}.

1) Be in I_i $S(y)>0$. It can easily be seen that

$$M(y)>0,\quad \text{if}\quad y=y_{2k+1}+0\,; \tag{2.17}$$

$$M(y)<0\,,\quad \text{if}\quad y=y_{2k+1}-0\,. \tag{2.18}$$

Thus, the following statements are true:

$$\frac{S(y_i^*)}{M(y_i^*)}=0\,,\quad \tan\varphi=0\,,\quad \varphi=0\,;$$

$$\frac{S(y_{2k+1}+0)}{M(y_{2k+1}+0)} \to +\infty\,, \quad \tan\varphi \to +\infty\,, \quad \varphi = \frac{\pi}{2} - 0\,;$$

$$\frac{S(y_{2k+1}-0)}{M(y_{2k+1}-0)} \to -\infty\,, \quad \tan\varphi \to -\infty\,, \quad \varphi = \frac{\pi}{2} + 0\,;$$

$$\frac{S(y^*_{i+1})}{M(y^*_{i+1})} = 0\,, \qquad \tan\varphi = 0\,, \qquad \varphi = \pi\,.$$

Therefore:

$$\Delta_{I_i}\varphi = \pi$$

wherein Δ_{I_i} means the change while y decreases from y^*_i to y^*_{i+1}.

2) In case, if in I_i $S(y) < 0$, it can be easily seen that

$$\Delta_{I_i}\varphi = -\pi\,.$$

3) From the cases 1) and 2)

$$\Delta_{I_i}\varphi = \operatorname{sign} S(y_{2k+1}) \cdot \pi\,. \tag{2.19}$$

$\beta\rangle$ Be in I_i one single root of $M(y)$ and it should have an even subscript: y_{2k}. In this case $M(y)$ changes its sign in an inverse way in comparison with the case in (2.17) and (2.18). Therefore, the result found in $\alpha\rangle$ will be valid with an opposite sign:

$$\Delta_{I_i}\varphi = -\operatorname{sign}(S(y_{2k}))\pi\,. \tag{2.20}$$

$\gamma\rangle$ In case where in I_i there are two roots of $M(y)$ one of them is of even and the other of odd subscript. Here $\Delta_{I_i}\varphi = 0$. If there are three roots, the change in the angle will be either $-\pi$ or $+\pi$, depending on the majority of odd or even subscripts out of the three roots, and so on. If, for example, the roots y_{2k}, y_{2k+1}, y_{2k+2} and y_{2k+3} of the $M(y)$ fall into I_i, then the value of

$$\Delta_{I_i}\varphi = \sum_{j=k}^{k+1} \left(\operatorname{sign}(S(y_{2j+1})) - \operatorname{sign}(S(y_{2j}))\right)\pi \tag{2.21}$$

is zero in the given case.

$\delta\rangle$ The root of the highest value of $S(y)$ is y^*_1. In the interval $[0, y^*_1]$ $S(y)$ has p roots which determine $p-1$ intervals similar to I_i. Since $y^*_p = 0$, so

$$I_{p-1} = (0, y^*_{p-1})\,;$$
$$I_i = (y^*_{i+1}, y^*_i)\,, \qquad i = 2, \ldots, p-2\,;$$
$$I_1 = (y^*_2, y^*_1)\,.$$

For these intervals one should summarize the changes of φ. If one assumes that the function $M(y)$ has q roots whose values are higher than that of y_1^*, that is,

$$y_1 \geq \cdots \geq y_q > y_1^* > y_{q+1} \geq \cdots \geq y_m ,$$

then, considering (2.19), (2.20) and (2.21), one obtains

$$\Delta_{[0,y_1^*]}\varphi = \sum_{k=q+1}^{m} (-1)^{k+1} \operatorname{sign}\left(S(y_k)\right)\pi . \tag{2.22}$$

In Eq. (2.22), the expression in the exponent of (-1) determines if the sign should be negative in case y is of even subscript or positive when y is of odd subscript.

$\varepsilon\rangle$ Let us now examine the interval (y_1^*, ∞), too.

1) Consider the case where n is an even number. Since $a_n > 0$,

$$\tan\varphi = \frac{S(y)}{M(y)} \to 0 , \quad \text{if} \quad y \to \infty ,$$

because Eqs (2.3) and (2.4) show that the order of $M(y)$ is higher than that of $S(y)$. Therefore, from the viewpoint of the investigation of $\tan\varphi$, the interval (y_1, ∞) is just the same as the I_i mentioned in the paragraph $\alpha\rangle$ because

$$\begin{aligned} &\text{if} \quad y \to +\infty \qquad \text{so} \qquad \tan\varphi = 0 ; \\ &\text{if} \quad y = y_1^* \qquad\quad \text{so} \qquad \tan\varphi = 0 . \end{aligned}$$

Thus Eq. (2.18) is true also in the interval (y_1, ∞):

$$\lim \Delta_{(e_1)}\varphi = \sum_{k=1}^{m} (-1)^{k+1} \operatorname{sign}\left(S(y_k)\right)\pi . \tag{2.23}$$

2) Be now n odd. Then $a_n > 0$, and by virtue of (2.3) and (2.4)

$$\text{if} \quad y \to +\infty \quad \text{so} \quad \tan\varphi = \frac{S(y)}{M(y)} \to +\infty \quad \text{or} \quad -\infty .$$

First let us assume that $S(y_1^* + 1) > 0$, that is, $\lim\limits_{y\to\infty} S(y) = +\infty$. Since, in $a\rangle$ it has been specified that $M(y_1 + 1) > 0$, in case of $q = 0$ (that is, if $y_1 < y_1^*$):

$$\begin{aligned} &y \to +\infty , \quad \tan\varphi \to +\infty , \quad \varphi = \frac{\pi}{2} - 0 ; \\ &y = y_1^* , \qquad \tan\varphi = 0 , \qquad \varphi = 0 . \end{aligned}$$

Thus:

$$\Delta_{(y_1^*,\infty)}\,\varphi = -\frac{\pi}{2}\,.$$

In case of $q = 1$ (that is, if $y_2 < y_1^* < y_1$):

$$y \to +\infty\,,\quad \tan\varphi \to +\infty\,,\quad \varphi = \frac{\pi}{2} - 0\,;$$

$$y \to y_1 + 0\,,\ \tan\varphi \to +\infty\,,\quad \varphi = \frac{\pi}{2} - 0\,;$$

$$y \to y_1 - 0\,,\ \tan\varphi \to -\infty\,,\quad \varphi = \frac{\pi}{2} + 0\,;$$

$$y = y_1^*\,,\qquad \tan\varphi = 0\,,\qquad \varphi = \pi\,.$$

Thus:

$$\Delta_{(y_1^*,\infty)}\,\varphi = \frac{\pi}{2}\,.$$

$\Delta_{(y_1^*,\infty)}\,\varphi$ is equal to $+\pi/2$ or $-\pi/2$ depending on whether the number of the roots of $M(y)$ in the interval (y_1^*, ∞) is even or odd.
Thus

$$\Delta_{(y_1^*,\infty)}\,\varphi = (-1)^{q+1}\frac{\pi}{2}\,. \tag{2.24}$$

Be now $S(y^*{}_1 + 1) < 0$, then
in case of $y \to \infty$, $\tan\varphi \to -\infty$, $\varphi = -\pi/2 + 0$, and, as compared with the above-said, all of the changes of signs occur in the opposite way, i.e.,

$$\Delta_{(y_1^*,\infty)}\,\varphi = -(-1)^{q+1}\frac{\pi}{2}\,. \tag{2.25}$$

By reducing the expressions (2.24) and (2.25) we get

$$\Delta_{(y^*,\infty)}\varphi = (-1)^{q+1}\frac{\pi}{2}\operatorname{sign}\left(S(y_1^* + 1)\right). \tag{2.26}$$

Ensuing from Eq. (2.4) and due to the fact that n is an odd number and $a_n > 0$:

$$\operatorname{sign}\left(S(y_1^* + 1)\right) = \operatorname{sign}\left((-1)^{\left[\frac{n-1}{2}\right]} a_n\right) = (-1)^{\left[\frac{n-1}{2}\right]}. \tag{2.27}$$

Hereafter, considering (2.22), (2.26) and (2.27), in case where n is an odd number

$$\lim_{R\to\infty} \Delta_{(e_1)}\,\varphi = \Delta_{[0,y_1^*]}\,\varphi + \Delta_{(y_1^*,\infty)}\,\varphi =$$

$$= \sum_{k=q+1}^{m} (-1)^{k+1} \operatorname{sign}\left(S(y_k)\right)\pi + (-1)^{q+1}(-1)^{\left[\frac{n-1}{2}\right]}\frac{\pi}{2}\,. \tag{2.28}$$

3) By making use of the fact that

$$n - 2\left[\frac{n}{2}\right] = \begin{cases} 0 \text{ in case if } n \text{ is even} \\ 1 \text{ in case if } n \text{ is uneven,} \end{cases}$$

the results (2.23) and (2.28) formulated in the preceding paragraphs 1) and 2) may be expressed in case of an arbitrary $n > 1$ as follows:

$$\lim_{R\to\infty} \Delta_{(e_1)}\varphi = \sum_{k=\left(n-2\left[\frac{n}{2}\right]\right)q+1}^{m} (-1)^{k+1} \operatorname{sign}\left(S(y_k)\right)\pi + \\ + \left(n - 2\left[\frac{n}{2}\right]\right)(-1)^{\left[\frac{n-1}{2}\right]+q+1}\frac{\pi}{2}. \tag{2.29}$$

b⟩ Be now $M(y_1 + 1) < 0$ (i.e., $\lim_{y\to\infty} M(y) = -\infty$). Then, it is clear that the order of ideas shown in paragraph *a*⟩ remains true, only the signs are changed into opposite ones. The change of φ will then be the (-1)-fold of the expression (2.29).

c⟩ Considering paragraph *b*⟩, from (2.16) and (2.29) one obtains the condition of the stability:

$$\left(\sum_{k=\left(n-2\left[\frac{n}{2}\right]\right)q+1}^{m} (-1)^{k+1} \operatorname{sign}\left(S(y_k)\right)\pi + \right.$$

$$\left. + \left(n - 2\left[\frac{n}{2}\right]\right)(-1)^{\left[\frac{n-1}{2}\right]+q+1}\frac{\pi}{2}\right) \operatorname{sign}\left(M(y_1 + 1)\right) = -\frac{n}{2}\pi$$

which agrees with the condition (2.5) of the *Theorem*.

Thus, the *Theorem* has been proved. The *Theorem* seems to be intricate, however, its use is rather simple considering the following:

The roots of the equation $M(y) = 0$ may easily be found with the aid of a computer program; the root of highest value of the function $S(y)$ may be found in the very same way, however, this is only needed in case of an odd n. The checking of the fulfilment of the condition (2.5) does not make any difficulty, the examination of (2.6) may be carried out at the same time.

With the aid of the algorithm in question a stability investigation may be quickly performed on a minicomputer with the aid of a simple programme. In this way there are possibilities for a great number of investigations, that is, the influence of the change of a_k, b, r upon the stability can be charted.

With the aid of the *Theorem* several new statements may be formulated. From among these, one will be presented and applied in examining the examples to be discussed later. The statement is as follows:

Let the order (n) of the retarded differential equation (2.1) be even, and then it is unstable for $a_0 < b$.

The prove of this statement is as follows:

Be $n = 4j + 2$, $j = 0, 1, \ldots$, so $n/2 = 2j + 1$, i.e. uneven; and since $a_n > 0$, sign $(M(y_1 + 1)) = -1$, in accordance with Eq. (2.3). The conditions (2.5) and (2.6) will be reduced as follows

$$\sum_{k=1}^{m} (-1)^{k+1} \operatorname{sign}\left(S(y_k)\right) = \frac{n}{2}, \tag{2.30}$$

$$\operatorname{sign}\left(S(y_k)\right) \neq 0\,, \qquad k = 1, \ldots, m\,.$$

At the left side of Eq. (2.30) the sum is even if m is even. $M(y) = 0$ has positive real roots of an even number in case where

$$M(0) = a_0 - b < 0\,.$$

Considering that at the right side of Eq. (2.30) the number is odd, (2.30) is not fulfilled if $a_0 < b$, and (2.1) is unstable. Be $n = 4j$, $j = 1, 2, \ldots$, so $n/2 = 2j$ is even and according to (2.3) sign $(M(y_1 + 1)) = 1$. The constraint equations are

$$\sum_{k=1}^{m} (-1)^{k+1} \operatorname{sign}\left(S(y_k)\right) = -\frac{n}{2}, \tag{2.31}$$

$$\operatorname{sign}\left(S(y_k)\right) \neq 0\,, \qquad k = 1, \ldots, m\,.$$

In Eq. (2.31) the sum at the left side is odd if the number of the positive real roots m is odd. This is also the case if

$$M(0) = a_0 - b < 0\,.$$

Thus, in case of $a_0 < b$ (2.31) is not fulfilled, (2.1) is unstable.

The formulation of a similar statement for an uneven number n is more intricate. In turn, for an even n-number it was clear that in one of the half-spaces of the parameters $a_n, \ldots, a_1, a_0, b, r$ the retarded differential equation (2.1) is unstable (for $a_0 < b$).

3. Generalization of the validity of the Theorem

The stability investigation of a great number of differential equations containing, for example, several retarded terms is more intricate than that of (1.5):

$$\sum_{k=0}^{n} a_k \frac{d^k}{dt^k} x(t) = \sum_{l=1}^{s} b_l x(t - r_l) \; ; \quad b_l, a_k, r_l \in R; \; r_l > 0; \; a_n > 0;$$

$$k = 1, \ldots, n; \; l = 1, \ldots, s \, .$$

It may be seen that the proposition remains valid, only the formulae (2.3) and (2.4) will change:

$$M(y) = \sum_{k=0}^{\left[\frac{n}{2}\right]} a_{2k}(-1)^k y^{2k} - \sum_{l=1}^{s} b_l \cos(r_l y) \, , \tag{3.1}$$

$$S(y) = \sum_{k=0}^{\left[\frac{n-1}{2}\right]} a_{2k+1}(-1)^k y^{2k+1} + \sum_{l=1}^{s} b_l \sin(r_l y) \, . \tag{3.2}$$

Also the investigated Eq. (1.5) may be reduced to the retarded differential equation (1.3) of general shape and n-dimensions. If n is even, the stability investigation of (1.3) is comparable to that of (1.5), in the *Theorem*, with the exception here the (2.3) and (2.4) will change. The expressions are more complicated because the equation is:

$$\dot{x}(t) = Ax(t) + Bx(t - r)$$

and its characteristic polynome may be found on the basis of the equation

$$D(z) = \det(zI - A - Be^{-rz})$$

by the development of the determinant. After some mathematical transformations it is obvious that the new shapes of (2.3) and (2.4) will be the following (3.3) and (3.4), but it has to be stated that the meaning of

$$\sum_{(l,h,k,j)}^{(0)} \text{ is } \sum_{\substack{l+2h+k=n \\ j=2h \\ l,h,k \geq 0}}$$

and the meaning of

$$\sum_{(l,h,k,j)}^{(-1)} \text{ is } \sum_{\substack{l+2h+k=n-1 \\ j=2h \\ l,h,k \geq 0}}$$

$$M(y) = \mathrm{Re}\big(D(iy)\big) = \sum_{(l,h,k,j)}^{(0)} \sum_{f=1}^{\binom{n}{l}\binom{n-l}{j}} \det\left(F^f_{ljk}\right)(-1)^h y^j \cos(rky) + \\ + \sum_{(l,h,k,j)}^{(-1)} \sum_{f=1}^{\binom{n}{l}\binom{n-l}{j}} \det\left(F^f_{l,j+1,k}\right)(-1)^h y^{j+1} \sin(rky)\,, \tag{3.3}$$

$$S(y) = \mathrm{Im}\big(D(iy)\big) = \sum_{(l,h,k,j)}^{(0)} \sum_{f=1}^{\binom{n}{l}\binom{n-l}{j}} \det\left(F^f_{ljk}\right)(-1)^{h+1} y^j \sin(rky) + \\ + \sum_{(l,h,k,j)}^{(-1)} \sum_{f=1}^{\binom{n}{l}\binom{n-l}{j}} \det\left(F^f_{l,j+1,k}\right)(-1)^h y^{j+1} \cos(rky)\,. \tag{3.4}$$

F^f_{ljk} in (3.3) and (3.4) is a matrix of nxn dimension and may be developed from the matrices A, B and $\Delta = -I$ of nxn dimension in the following way: the $u_1, \ldots, u_l$-th rows of matrix A, the $u_{l+1}, \ldots, u_{l+j}$-th rows of the matrix Δ, the $u_{l+j+1}, \ldots, u_{l+j+k}$-th rows of the matrix B should be selected where $l + j + k = n$ and $u_i \neq u_m$, if $i \neq m$, $i,m = 1, \ldots, n$. In this way n rows have been selected, which written under each other, in the order of the increasing series of the u-values, yield the matrix F. Thus, $\binom{n}{l}\binom{n-l}{j}$ different F-matrices may be established. The upper limit of the summation relating to f refers to this fact. Thus, for example:

$$F^1_{ljk} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{l1} & \cdots & a_{ln} \\ \delta_{l+1,1} & \cdots & \delta_{l+1,n} \\ \vdots & & \vdots \\ \delta_{l+j,1} & \cdots & \delta_{l+j,n} \\ b_{l+j+1,1} & \cdots & b_{l+j+1,n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix}$$

The coefficients of the expressions $M(y)$ and $S(y)$ should also be conveniently calculated in case of $n > 3$ with the use of a computer. However, to avoid an excessive increase in the calculation-time resulting from the development of the great number of determinants, such numerical methods should be applied which make use of the fact that among the elements of the matrices F many zeros are to be found. To find the zeros of the positive real values of $M(y)$ represents an increased problem in comparison with (2.3), however, also this may

be facilitated by using certain proper methods. As a matter of course, if possible, in lieu of the general formulae (3.3) and (3.4) it is worthwhile using rather (3.1), (3.2) or (2.3), (2.4) because they afford calculations significantly simpler than in the former case.

4. Specific cases of the stability investigation

By making use of the *Theorem*, also the stability of the ordinary differential equations may be investigated in case the coefficient of the retarded term is equal to zero in the retarded differential equation (1.5), that is, $b = 0$. In that case $M(y)$ and $S(y)$ will be a polynome each; the determination of the roots of $M(y)$ may be made even by making use of the resolving formula of the equation of second order up to the value 5 of n. For the case of $n > 5$ it is very easy to establish a numerical method for the calculation of the roots. Therefore, the method of stability investigation ensuing from the *Theorem* may be more rapid than, for example, by developing Hurwitz determinants, mainly in cases of high values of n.

The *Theorem* is not valid for one dimensional differential equations of first order. In fact, in case of $n = 1$, $M(y)$ has infinitely many zeros of real value, thus the values m and q entering in the *Theorem* will have no meaning at all. Such an amplification of the *Theorem* which also contains the case of $n = 1$ would unnecessarily complicate the formulae; anyway, this case has already been solved by HALE (see [1], page 108).

5. Examples

One of the advantages of the *Theorem* is that it offers an algorithm to the stability investigation, however, it is also suitable for deriving relations of a closed form. In the following some examples are presented with the help of which we can learn how to use the *Theorem*.

Firstly, it should be investigated when the real parts of the zeros of the polynome

$$D(z) = a_3 z^3 + a_2 z^2 + a_1 z + a_0 , \qquad a_3 > 0 \tag{5.1}$$

will be negative. In this case, according to (2.3) and (2.4):

$$M(y) = - a_2 y^2 + a_0 ,$$
$$S(y) = - a_3 y^3 + a_1 y .$$

Let us assume that $a_0 a_2 < 0$, where $M(y) = 0$ has no real root, $m = 0$, $q = 0$ and thus according to (2.5):

$$\left(0 + \frac{1}{2}(-1)^2\right) \operatorname{sign}(-a_2) + \frac{3}{2} = \frac{3}{2} \pm \frac{1}{2} \neq 0 .$$

So there will be a root of positive real part of (5.1). Therefore, $a_0 a_2 > 0$, thus $M(y) = 0$ has a real root

$$y_1 = \sqrt{\frac{a_0}{a_2}}.$$

Let us assume that $y_1^* > y_1^*$ (y_1^* positively exists), then $m = 1$, $q = 1$ therefore, according to (2.5):

$$\left(0 + \frac{1}{2}(-1)^3\right) \operatorname{sign}(-a_2) + \frac{3}{2} = \frac{3}{2} \pm \frac{1}{2} \neq 0.$$

Thus, (5.1) will again have a root of positive real part. Consequently, $y_1 < y_1^*$, that is, the root of the highest value of $S(y)$

$$y_1^* = \sqrt{\frac{a_1}{a_3}}, \quad a_1 > 0$$

and

$$\sqrt{\frac{a_0}{a_2}} < \sqrt{\frac{a_1}{a_3}}$$

i.e.

$$\begin{vmatrix} a_1 & a_0 \\ a_3 & a_2 \end{vmatrix} > 0.$$

In this case $m = 1$, $q = 0$ and, according to (2.5):

$$\left((-1)^2 \operatorname{sign}\left(\left(-a_3 \frac{a_0}{a_2} + a_1\right)\sqrt{\frac{a_0}{a_2}}\right) + \frac{1}{2}(-1)^2\right) \operatorname{sign}(-a_2) + \frac{3}{2} +$$

$$= \frac{3}{2} \operatorname{sign}(-a_2) + \frac{3}{2} = 0.$$

This will be fulfilled if $a_2 > 0$. Thus we have obtained the well known condition ensuing from the Routh–Hurwitz criterion: the coefficients and the determinant built up from them must be positive in case the real parts of the roots of (5.1) are negative.

Let us now consider a theorem established by HALE (see [1], Th.A.6., p. 340). In proving the theorem he used the Pontrjagin's statements also mentioned in the introduction of this paper. We shall see that Hale's theorem may be said to be trivial on the basis of the *Theorem*. The statement is as follows:

All roots of the equation $(z^2 + az)e^z + 1 = 0$ are of negative real part if and only if

$$a > \frac{\sin \zeta}{\zeta}$$

wherein ζ is the root of the equation

$$\zeta^2 = \cos \zeta \tag{5.2}$$

in the interval $[0, \pi/2]$.

The retarded differential equation, whose characteristic equation is mentioned in the statement is

$$\frac{d^2}{dt^2} x(t) + a \frac{d}{dt} x(t) + x(t-1) = 0 .$$

The characteristic equation of this, according to the definition, is:

$$D(z) = z^2 + az + e^{-z} .$$

On the basis of (2.1) and (2.2)

$$M(y) = -y^2 + \cos y ,$$
$$S(y) = ay - \sin y .$$

$M(y)$ has only one single positive root: y_1, thus $m = 1$. Since $n = 2$, it is unnecessary to determine the value of q. By substituting the above values into (2.5) one obtains:

$$\left(\sum_{k=1}^{1} (-1)^{k+1} \operatorname{sign} (S(y_k)) \right) \operatorname{sign} \left(M(y_1 + 1) \right) + \frac{2}{2} = 0 ,$$

i.e.,

$$(-1)^2 \operatorname{sign} (S(y_1))(-1) + 1 = 0 ,$$
$$\operatorname{sign} (S(y_1)) = 1 ,$$
$$ay_1 - \sin y_1 > 0 .$$

Accordingly

$$a > \frac{\sin y_1}{y_1} ,$$

wherein y_1 is the root of $M(y) = 0$, i.e., of Eq. (5.2).

As it can be seen, Hale's theorem is a very specific case of the *Theorem* made in this paper.

By using the *Theorem* other results may easily be obtained, for example, for the case of retarded differential equations of second order. Let us take the following example:

$$\frac{d^2}{dt^2} x(t) + a_0 x(t) = bx(t-1) \tag{5.3}$$

the stability investigation of which should be performed in the functions of the parameters a_0 and b. The characteristic equation of (5.3) is

$$D(z) = z^2 + a_0 - be^{-z} .$$

Let us apply the *Theorem*. According to (2.3) and (2.4)

$$M(y) = -y^2 + a_0 - b \cos y , \tag{5.4}$$

$$S(y) = b \sin y . \tag{5.5}$$

The positive real roots of $M(y) = 0$ are required. According to the statement proved at the end of the part 2, (5.3) is unstable, if $a_0 < b$. In the following, let us consider the case $a_0 > b$. (2.5) will be reduced as follows:

$$\sum_{k=1}^{m} (-1)^{k+1} \operatorname{sign} (b \sin y_k) - 1 = 0 . \tag{5.6}$$

Be first $b > 0$. If the solution to (5.3) is asymptotically stable, then there is at least one $k(k = 0,1,\ldots)$, where the value of $M(y)$ is positive at $y = 2k\pi$, and negative at $y = (2k + 1)\pi$, i.e.,

$$\begin{aligned} &-(2k\pi)^2 + a_0 - b > 0 , \\ &a_0 > b + 4k^2\pi^2 ; \end{aligned} \tag{5.7}$$

$$\begin{aligned} &-((2k + 1)\pi)^2 + a_0 + b_0 < 0 , \\ &a_0 < -b + (2k + 1)^2\pi^2 . \end{aligned} \tag{5.8}$$

If, in turn (5.7) and (5.8) are fulfilled, so

$$\begin{aligned} &\sqrt{a_0 - b} > 2k\pi , \\ &\sqrt{a_0 + b} < (2k + 1)\,\pi \end{aligned}$$

is true, which means that all positive real roots of $M(y)$ fall into the interval $(2k\pi,(2k + 1)\pi)$, wherein $S(y)$ is positive; the condition (5.6) is fulfilled and the solution to (5.3) is asymptotically stable. Also to $b < 0$, a calculation may be made. On the whole, it can be concluded that the solution $x \equiv 0$ to (5.3) is asymptotically stable if and only if

$$b > 0,\ a_0 > b + 4k^2\pi^2 \text{ and } a_0 < -b + (2k + 1)^2\pi^2,\ k = 0, 1,\ldots$$

or

$$b < 0,\ a_0 > -b + (2j - 1)^2\pi^2 \text{ and } a_0 < b + 4j^2\pi^2,\ j = 1, 2,\ldots$$

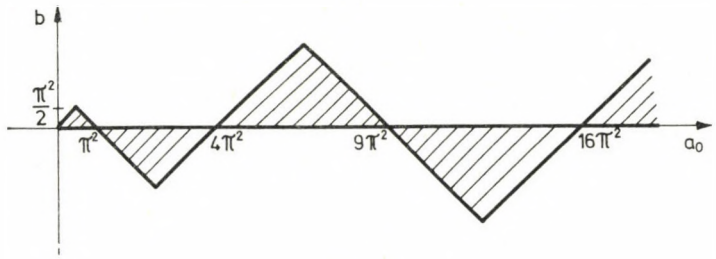


Fig. 1

According to Fig. 1 (5.3) will be asymptotically stable in case the parameters a_0 and b are selected from the regions represented by a dashed line.

Consider the last example, the general retarded differential equation of second order:

$$\frac{d^2}{dt^2}x(t) + a_1\frac{d}{dt}x(t) + a_0 x = bx(t-1)\,. \tag{5.9}$$

Its characteristic equation is

$$D(z) = z^2 + a_1 z + a_0 - be^{-z} \tag{5.10}$$

whereby

$$M(y) = -y^2 + a_0 - b\cos y\,,$$
$$S(y) = a_1 y + b\sin y\,,$$

and the constraint equation (2.5) is:

$$\sum_{k=1}^{m}(-1)^{k+1}\,\mathrm{sign}\,(a_1 y_k + b\sin y_k) - 1 = 0\,. \tag{5.11}$$

In connection with the stability of (5.9) one can prove several statements relating to the asymptotical behaviour of the boundary curves of the stability regions or the statements, that with the increase of a_1 the regions of stability will widen, that in case of $a_0 > b$, $b > 0$ and $a_1 > b$ or $a_0 > b$, $b < 0$ and $a_1 > -b$ (5.9) is surely asymptotically stable, and that in case of the values $a_1 < -2$ (5.9) is positively unstable. This time only the latter statement will be proved.

Since

$$\frac{dM}{dy} = -2y + \sin y\,,$$

the loci of the extreme values of $M(y)$ are given by the points of intersections of the curves $2y$ and $b\sin y$ (Fig. 2). By taking into account that the zero of

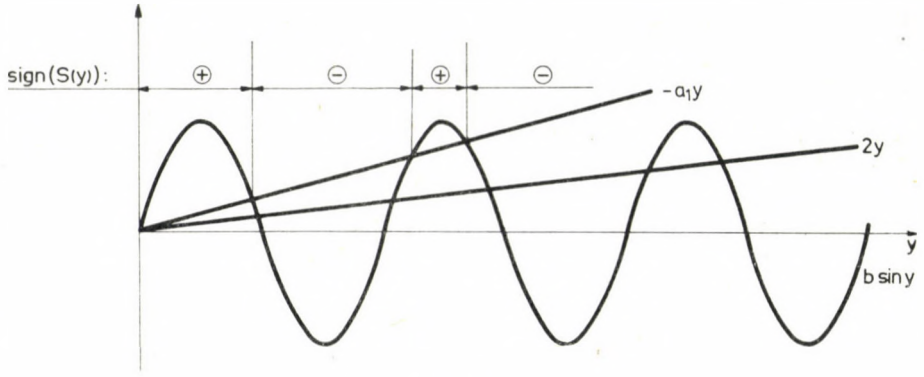


Fig. 2

uneven subscript of $M(y)$ may only be between a maximum locus and a subsequent minimum locus, Fig. 2 clearly show that in case of $a_1 < -2$, $S(y) = a_1y + b \sin y$ is always negative, in these intervals wherefore, (5.11) will not be fulfilled.

Also from Fig. 2 it may be seen that with the increase of a_1 the region of stability will be widened.

By programming the algorithm ensuing from the *Theorem* one obtains the chart of stability in the field of the parameters a_0, a_1 and b. In Fig. 3, the approximate curves plotted on the basis of the results obtained show the constant contour lines a_1.

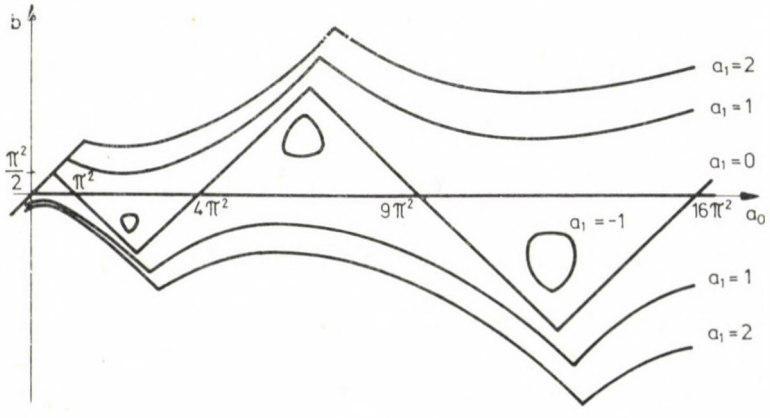


Fig. 3

6. Reproductive machine tool vibration

A very significant field of application of the mathematical issues is the investigation of the reproductive machine tool vibrations. Simply explained, the essential part of the phenomenon is that the vibration of the tool is generated by the periodically changing cutting force which is due to the surface, by some reason wavily machined at the preceding work process. The deduction of the mathematical model is to be found in [5, 6, 7, 8], its result is always a retarded differential equation of second order. According to TOBIAS and FISCHWICK [9] the model of the so-called A-type vibration is shown in Fig. 4.

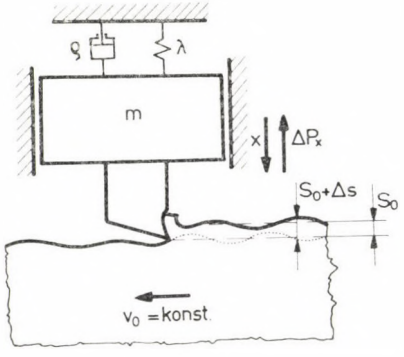


Fig. 4

According to TOBIAS [5], the differential equation is as follows:

$$m\frac{d^2x}{dt^2} + \varrho\frac{dx}{dt} + \lambda x = -\Delta P_x\,, \tag{6.1}$$

$$\Delta P_x = z_c k_1 \left(x(t) - \mu x\left(t - \frac{T}{z_f}\right)\right) + \frac{z_c}{z_f}(k_s - k_1)\frac{2\pi}{\Omega}\Delta r\,. \tag{6.2}$$

In the above and following equations these symbols are used:

ΔP_x — change of component of cutting force of x-direction;
s_0 — theoretical chip thickness;
z_f — number of cutting edges;
z_c — average number of cutting edges in contact;
μ — overlap factor;
Ω — angular velocity of workpiece (or of tool in case of rotary tool);
T — period of revolution: $T = \dfrac{2\pi}{\Omega}$;
N — r.p.m.: $N = \dfrac{\Omega}{2\pi}60$;
k_s — static chip thickness coefficient; $k_s = \left.\dfrac{\partial P}{\partial s_0}\right|_{\Delta\Omega=0}$;

k_1 — dynamic chip thickness coefficient, $k_1 = \left.\frac{\partial P}{\partial s}\right|_{\Delta\Omega=\Delta r=0}$;

Δr — change in feed rate, $\Delta r = \frac{dx}{dt}$;

ω_0 — frequency, $\omega^2 = \frac{\lambda}{m}$;

D — relative damping coefficient, $D = \frac{\varrho\omega_0}{2\lambda}$.

The retarded differential equation obtained from (6.1) and (6.2) is:

$$\frac{d^2x}{dt^2} + \left(2D\omega_0 + \frac{\omega_0^2}{\lambda}\frac{z_f}{z_c}(k_s - k_1)\,T\right)\frac{dx}{dt} + \\ + \omega_0^2\left(1 + z_c\frac{k_1}{\lambda}\right)x = \omega_0^2\, z_c\frac{k_1}{\lambda}\mu x\left(t - \frac{T}{z_f}\right). \tag{6.3}$$

The stability chart may be interesting in the plane of any two parameters. TOBIAS [5] shows a chart as the function of $Q = 1/(2D)$ and $z_f N/(2\pi\omega_0)$, i.e., in fact, as the function of D and N, however, he does not deal with the calculation in detail, owing to its complexity. The stability investigation of (6.3) will be carried out on the basis of what has been said in parts *2* and *5*. The characteristic equation will be:

$$D^*(\zeta) = \zeta^2 + \left(2D\omega_0 + \frac{\omega_0^2}{\lambda}\frac{z_c}{z_f}(k_s - k_1)\,T\right)\zeta + \\ + \omega_0^2\left(1 + z_c\frac{k_1}{\lambda}\right) - \omega_0^2\, z_c\frac{k_1}{\lambda}\mu e^{-\frac{T}{z_f}\zeta}$$

(ζ being a complex number). Carrying out the transformation $z = T\zeta/z_f$, one gets:

$$D(z) = z^2 + \frac{T}{z_f}\left(2D\omega_0 + \frac{\omega_0^2}{\lambda}\frac{z_c}{z_f}(k_s - k_1)\,T\right)z + \\ + \frac{T^2}{z_f^2}\omega_0^2\left(1 + z_c\frac{k_1}{\lambda}\right) - \frac{T^2}{z_f^2}z_c\frac{k_1}{\lambda}\mu e^{-z}. \tag{6.4}$$

Since (6.4) agrees with the characteristic equation (5.10), the stability chart presented in [5] is a section of the chart presented in Fig. 3. By assuming that $k_s = k_1$, this section is shown by the part of Fig. 5 vertically shaded, wherein

$$a_1 = \frac{2\pi\omega_0}{z_f N}2D, \qquad a_0 = \left(\frac{2\pi\omega_0}{z_f N}\right)^2\left(1 + z_c\frac{k_1}{\lambda}\right), \\ b = \left(\frac{2\pi\omega_0}{z_f N}\right)^2\frac{z_c}{\omega_0^2}\frac{k_1}{\lambda}\mu.$$

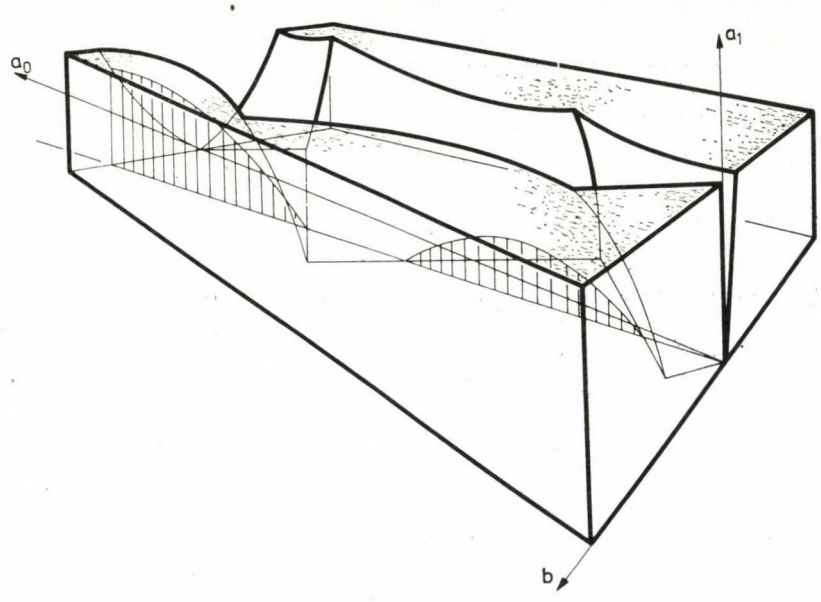


Fig. 5

Thus, the chart Q, $z_f N/(2\pi\omega_0)$ may be constructed according to Fig. 6. In the shaded areas of Figs 5 and 6 the vibration of the tool is unstable. In Fig. 6

$$\tan\gamma = \frac{\dfrac{z_c}{\omega_0^2}\dfrac{k_1}{\lambda}\mu}{1+z_c\dfrac{k_1}{\lambda}}.$$

TOBIAS presents the calculation (Fig. 6) of the constant Q_0 (see [5], p. 177).

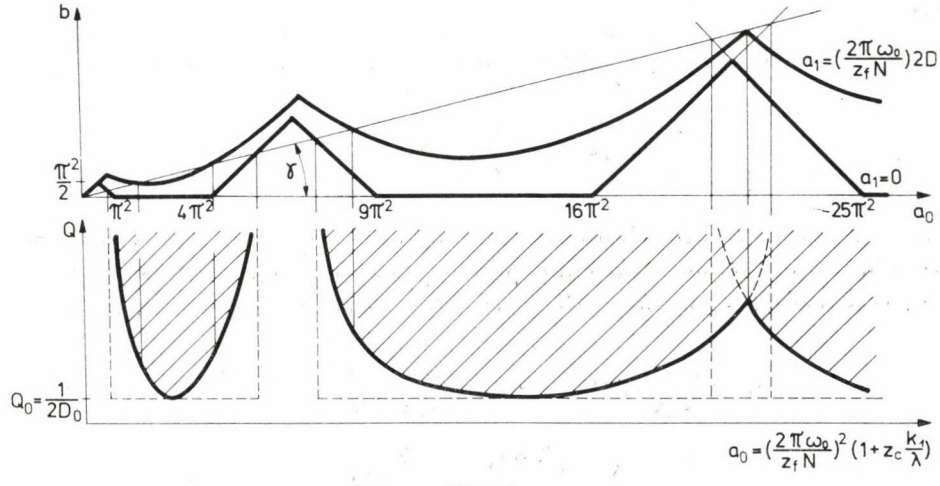


Fig. 6

It is unnecessary to construct the stability chart because by changing the parameters N and Q, a computer program will produce it with the aid of the algorithm ensuing from the *Theorem*.

In carrying out the technological design it is of great significance to know the regions of the field of feed rate, depth of cut and revolutions per minute, where an increasing tool vibration is produced. In this respect the stability chart k_1, N is of interest to which one can find an example in [10]. Also this may be produced as a section of the chart shown in Fig. 3.

There has been made a computer program to investigate the stability of (6.3) according to the algorithm proved in part *2*. In case of given parameters, it makes certain if (6.3) is stable or not in 1 to 1,5 seconds, so we have the opportunity to carry out numerous investigations. For two parameters arbitrarily selected, the program produces the stability chart.

A stability chart k_1, N produced by the computer program is presented. The data are as follows

$$m = 50 \text{ kg}; \qquad k_1 = k_s;$$

$$\lambda = 30\frac{MN}{m}; \qquad \mu = 1;$$

$$\varrho = 3{,}87\frac{kNs}{m}; \quad z_f = z_c = 1.$$

The chart is represented in Fig. 7 (the shaded area designates unstableness).

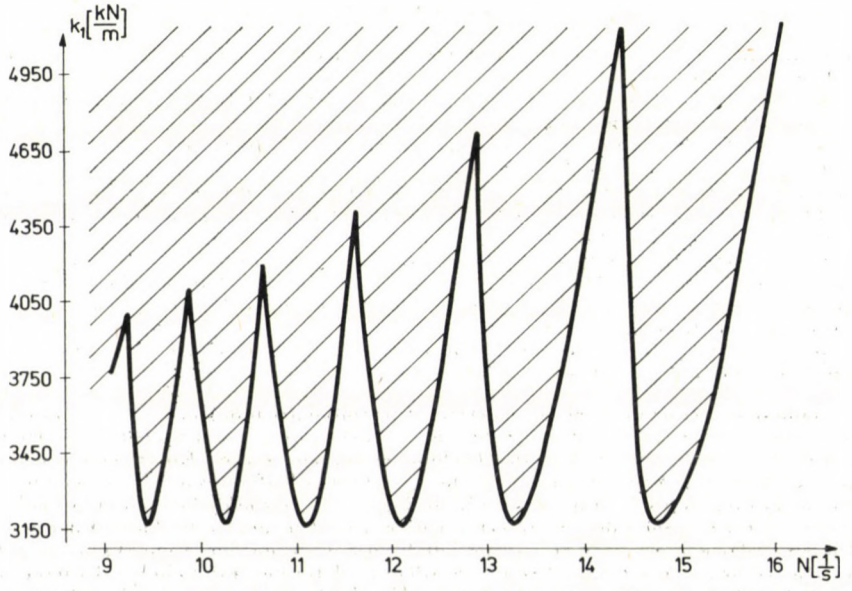


Fig. 7

The value of $k_{1,0}$ (Fig. 7) has been determined in an analytic way by TLUSTY [6]. With the data presented above, this yields:

$$D = \frac{\varrho}{2\lambda}\sqrt{\frac{\lambda}{m}} = 0{,}05\,,$$

$$k_{1,0} = 2D\lambda(1 + D) = 3{,}15\,\frac{MN}{m}\,.$$

This agrees with the value of map $k_{1,0}$ produced by the computer. (The limit of error of the values k_1 is 0,075 MN/m on the chart.)

There has also been constructed a stability chart which concerns the parameters k_1,α; in this case, the tool to be seen in Fig. 4 does not vibrate in the direction x but in that inclined at angle α to the direction x. Its character agrees with that on the chart presented in [6].

The computer program produced with the aid of the algorithm resulting from the *Theorem* may investigate any desired parameter. The algorithm established by the author is standard, its operation is rapid and simple, it may readily be built into the computerised technology-planning procedure.

However, in this paper only one of the fields of applications of the algorithm has been presented in connection with the vibration of cutting tools. Presumably, the method may conveniently be applied in the control technique where no analytic procedure has so far been at disposal for the stability investigation of retarded feed-back systems.

REFERENCES

1. HALE, J. K.: Theory of Functional Differential Equations, Springer Verlag, New York—Heidelberg—Berlin 1977
2. GANTMACHER, F.: Lectures in Analytical Mechanics, Mir Publishers, Moscow 1970
3. БАБАКОВ, И. М.: Теория коиданий, Издательство Наука, Москва, 1965.
4. PONTRJAGIN, L. S.: On the Zeros of Some Elementary Transcendental Functions, *Amer. Math. Soc. Transl.*, 1955
5. TOBIAS, S. A.: Schwingungen an Werkzeugmaschinen, Carl Hauser Verlag, München 1961
6. TLUSTY—POLACEK—DANEK—SPACEK: Selbsterregte Schwingungen an Werkzeugmaschinen, VEB Verlag Technik, Berlin 1962
7. АРИОРЕГО—ТОРАУН: Обработка металлоо резанием Машиностролние, Москва, 1977.
8. LUDVIG, GY.: Dynamic of Machines, Műszaki Könyvkiadó, Budapest 1973 (in Hungarian)
9. TOBIAS-FISCHWICK: The Cutter of Lathe Tools under Orthogonal Cutting Conditions, *Transaction A.S.M.E.* **80**, (1959)
10. SLAVICEK, J.: The Effect of Irregular Tooth Pitch on Stability of Milling, 6. International M.T.D.R. Conference, 1965

Stabilitätsuntersuchung der retardierten Differentialgleichungen. — Ein Lehrsatz wird vorgeführt und nachgewiesen mit Hilfe dessen ein Algorithmus zur Untersuchung der homogenen, linearen, retardierten Differentialgleichungen von konstanten Koeffizienten hergestellt werden kann. Dieser Algorithmus kann für die retardierten Differentialgleichungen ebenso benutzt werden wie das Routh—Hurwitzsche Kriterium für die gewöhnlichen Differentialgleichungen. Ebenso, wie die Theorie der retardierten Differentialgleichungen die Theorie der gewöhnlichen Differentialgleichungen als einen Spezialfall enthält, kann der Algorithmus anstatt des Routh—Hurwitzschen Kriteriums benutzt werden, auch für die gewöhnlichen Differensialgleichungen. Schließlich wird die praktische Nutzbarkeit der Ergebnisse für den Fall der reproduktiven Meisselvibrationen nachgewiesen.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (1—2), pp. 133—150 (1980)

GENERALIZATION AND SOME APPLICATIONS OF THE CONCEPT OF THE IDEAL CONSTRAINT

J. HERING*

[Manuscript recieved May 4, 1979]

Author describes the classic explanation of ideal constraint, and points out that this explanation can only be applied to geometric constraints. Therefore, for the definition of the ideal constraint a new explanation is also given including the classic explanation, however, that it may also be applied to kinematic constraints. Based on this definition of the ideal constraint, minimum principles may be established leading to the most general principle, i.e., to Gauss's principle defining the motion of mechanical systems. It is reported how Appell's equation of motion for anholonomic systems, Langrange's equation of motion for holonomic systems and Euler's equation of motion for rigid bodies may be determined on the base of the principles exposed above by the author.

1. Classic explanation of the ideal constraint

Let us consider a material particle of mass m (material point) subjected at a given moment of time t to an active force F and to a constraint defined by the geometrical constraint equation

$$\mathrm{f}(r, \mathrm{t}) = 0 \,, \tag{1.1}$$

or by the kinematic constraint equation

$$\mathrm{f}(r, \dot{r}, \mathrm{t}) = 0 \,, \tag{1.2}$$

where:

r — the position vector of mass point in a certain system of inertia
$\dot{r} = v$ — velocity of mass point.

Under the effect of the active force F and reactive force R induced by the constraint, the mass point carries out a motion with an acceleration $\ddot{r} = a$ and thus, it may be written

$$F + R = \mathrm{m} \cdot a \,. \tag{1.3}$$

The constraint is named in the traditional way as ideal, when the virtual performance

$$\tilde{P}^* = R \cdot \tilde{v} \tag{1.4}$$

* Dr. J. HERING, Szikla u. 8, H-1221 Budapest, Hungary

of the reactive force R is equal to zero,

where: $\tilde{v}$ — virtual velocity of mass point m, i.e., the difference of two possible velocities v' and v'':

$$\tilde{v} = v' - v'' \tag{1.5}$$

In the case of geometric constraint in the form (1.1), the condition (1.4 fully determines the direction of the ideal constraint force R. In fact, derivation of the function (1.1) with respect to time yields:

$$\frac{\partial \mathrm{f}}{\partial r} \cdot \dot{r} + \frac{\partial \mathrm{f}}{\partial \mathrm{t}} = 0 \,. \tag{1.6}$$

Be velocities $\dot{r}'$ and $\dot{r}''$ compatible with the constraint equation (1.6), i.e.,

$$\frac{\partial \mathrm{f}}{\partial r} \cdot \dot{r}' + \frac{\partial \mathrm{f}}{\partial \mathrm{t}} = 0 \tag{1.7a}$$

and

$$\frac{\partial \mathrm{f}}{\partial r} \cdot \dot{r}'' + \frac{\partial \mathrm{f}}{\partial \mathrm{t}} = 0 \,. \tag{1.7b}$$

From the difference of the two equations, by taking into consideration (1.5), we have

$$\frac{\partial \mathrm{f}}{\partial r} \cdot \tilde{v} = 0 \tag{1.8}$$

which compared with (1.4) evidently results in

$$R \,||\, \frac{\partial \mathrm{f}}{\partial r}. \tag{1.9}$$

However, the kinematic constraint of the form (1.2) depends, not only on the time t and the position vector (r), also the velocity vector $\dot{r} = v$, wherefore, the possible velocities may, in a general case, be of arbritrary direction, whence ensures that the difference of the two possible velocities, i.e., the virtual velocities $\tilde{\imath}$ may also follow any optional direction. Thus, the equation defining the ideal constraint

$$R \cdot \tilde{v} = 0$$

might, in general, be fulfilled in case where $R = 0$, which means that no ideal constraint force exists in that sense.

For the explanation of what has been said above in connection with the kinematic constraint, let us consider the constraint defined by the equation

$$\dot{r}^2 - v^2(t) = 0 \tag{1.10}$$

where $v^2(t)$ is a prescribed function. The condition (1.10) means that at a given moment the value of the velocity of the mass point is prescribed, but its direction may be arbitrary. The geometric loci of the end points of the possible velocity vectors define a sphere surface with the radius v(t). Thus, the virtual velocity may be of any direction, consequently, the condition (1.4) cannot be applied for the definition of the concept of the ideal kinematic constraint.

In the following, the generalization of the concept of the ideal constraint will be presented which may be applied both to the geometric and to the kinematic constraints and leads to the remarkable principles of the mechanics from which the equations of motion of mechanical systems may be established.

2. General interpretation of the ideal constraint

In case, where to a material particle of mass m, submitted to a constraint and moving at the moment t with a velocity v, is acted on by a constraint force R, so the performance of the reactive force is

$$\mathrm{P}^* = R \cdot v \,. \tag{2.1}$$

The constraint is named in the common meaning as ideal if, in case of given moment (t), position (r) and velocity (v) the possible changes of the constraint performance P^* (in a sufficiently short time Δt) are of the same magnitude, which means that the constraint performance-velocity P^* is stationary, i.e., its variation is equal to zero:

$$\delta \mathrm{P}^* = 0 \,. \tag{2.2}$$

Let in the moment $t + \Delta t$ be:

— $R + \Delta R$ — constraint force,
— $v + \Delta v_1$ and $v + \Delta v_2$ — two possible velocities of the mass point.

Thus, the two possible performances of the constraint force will be at the moment $t + \Delta t$:

$$\begin{aligned} \mathrm{P}_1^* &= \mathrm{P}^* + \Delta \mathrm{P}_1^* = (R + \Delta R) \cdot (v + \Delta v_1) \,, \\ \mathrm{P}_2^* &= \mathrm{P}^* + \Delta \mathrm{P}_2^* = (R + \Delta R) \cdot (v + \Delta v_2) \,. \end{aligned} \tag{2.3}$$

The difference of the possible performances at the moment $t + \Delta t$ is, by virtue of (2.2), equal to zero:

$$\Delta \mathrm{P}_2^* - \Delta \mathrm{P}_1^* = (R + \Delta R) \cdot (\Delta v_2 - \Delta v_1) = 0 \,. \tag{2.4}$$

By using the possible accelerations a_1 and a_2 at the moment t, permitted by the constraint, it may be written as:

$$\Delta v_2 = a_2 \cdot \Delta t \quad \text{and} \quad \Delta v_1 = a_1 \cdot \Delta t \tag{2.5}$$

which substituted into (2.4) gives:

$$\Delta P_2^* - \Delta P_1^* = (R + \Delta R) \cdot (a_2 - a_1) \cdot \Delta t = 0 \,. \tag{2.6}$$

By dividing (2.6) with Δt and carrying out the transition $\Delta t \to 0$ so, for the difference of the velocities of the possible constraint performances at the moment t one obtains:

$$\dot{P}_2^* - \dot{P}_1^* = R \cdot (a_2 - a_1) = 0 \,. \tag{2.7}$$

In lieu of the possible acceleration a_1 using the real acceleration a, for the general definition of the ideal constraint may be written

$$\delta \dot{P}^* = R \cdot \delta a = 0 \,, \tag{2.8}$$

i.e., based on (2.8):

— The constraint is named ideal if the scalar product of the constraint force and variation of the acceleration is equal to zero or,

— in case of ideal constraint the velocity of the constraint performance is stationary, in the possible field of acceleration permitted by the constraint.

Let us investigate what does the condition (2.8), in the case of the geometric constraint (1.1) and kinematic constraint (1.2) mean:

— In the case of the geometric constraint, by deriving two times Eq. (1.1) with respect to time,

$$\frac{\partial \dot{f}}{\partial r} \cdot \dot{r} + \frac{\partial f}{\partial r} \cdot \ddot{r} + \frac{\partial^2 f}{\partial t^2} = 0 \tag{2.9}$$

is obtained, whence to the variation of acceleration, compatible with the constraint:

$$\frac{\partial f}{\partial r} \cdot \delta r = 0 \tag{2.10}$$

compared with (2.8) gives

$$R \,||\, \frac{\partial f}{\partial r} \,. \tag{2.11}$$

(2.11) agrees with the relationship (1.9) obtained by the traditional interpretation of the ideal constraint.

— In the case of a kinematic constraint deriving from the constraint equation (1.2) with respect to time results in

$$\frac{\partial \mathrm{f}}{\partial r} \cdot \dot{r} + \frac{\partial \mathrm{f}}{\partial \dot{r}} \cdot \ddot{r} + \frac{\partial \mathrm{f}}{\partial \mathrm{t}} = 0 \,, \tag{2.12}$$

whence

$$\frac{\partial \mathrm{f}}{\partial \dot{r}} \cdot \delta \ddot{r} = 0 \,. \tag{2.13}$$

Combination of (2.13) and (2.18) gives

$$R \,||\, \frac{\partial \mathrm{f}}{\partial \dot{r}} \tag{2.14}$$

which could not be pointed out by the classical interpretation.

In the example given by the constraint equation (1.10)

$$R \,||\, \frac{\partial \mathrm{f}}{\partial \dot{r}} = 2 \cdot \dot{r} \,, \tag{2.14a}$$

i.e.,

$$R \,||\, \dot{r} = v \,. \tag{2.15}$$

Substituting into the relationship (2.8) the value

$$R = \mathrm{m} \cdot a - F \tag{2.16}$$

of the constraint force R defined by (1.3), Gauss's principle concerning the material particle is obtained

$$(\mathrm{m} \cdot a - F) \cdot \delta a = 0 \,. \tag{2.17}$$

In lieu of (2.17)

$$\delta \left(\frac{1}{2} \mathrm{m} \cdot a^2 - F \cdot a \right) = 0 \tag{2.18}$$

may be written.

According to APPELL, introducing the expression of the energy of acceleration

$$\mathrm{S} = \frac{1}{2} \mathrm{m} \cdot a^2 \tag{2.19}$$

the principle of the velocity of the constraint performance expressed by (2.8) may be written in the form

$$\delta \mathrm{S} = F \cdot \delta a \,. \tag{2.20}$$

3. Extreme properties of the ideal constraint

Let us consider a material particle of mass m moving with an acceleration a which is submitted at the moment t to an active force F and to a reactive force R. By dividing in thought the effects of the forces F and R one can say on the one hand that at the free motion (without constraint) the force F causes an acceleration a°:

$$F = \mathrm{m} \cdot a^\circ \tag{3.1}$$

on the other hand, the reactive force modifies the acceleration a° by a value α:

$$R = \mathrm{m} \cdot \alpha \,. \tag{3.2}$$

Therefore, the resultant acceleration of the constraint motion is

$$a = a^\circ + \alpha \tag{3.3}$$

and thus

$$F + R = \mathrm{m}(a^\circ + \alpha) = \mathrm{m} \cdot a \,. \tag{3.4}$$

A possible value of the acceleration defined by the relationship (3.3) compatible with the constraint is

$$a_1 = a^\circ + \alpha_1 \,, \tag{3.5}$$

wherefrom the variation of the acceleration will be

$$\delta a = a_1 - a = \alpha_1 - \alpha = \delta\alpha \,. \tag{3.6}$$

By substituting the expressions (3.2) and (3.6) into (2.8) we have

$$\mathrm{m} \cdot \alpha \cdot \delta\alpha = 0 \,, \tag{3.7}$$

i.e.,

$$\delta \left(\frac{1}{2} \mathrm{m}\, \alpha^2 \right) = 0 \,, \tag{3.8}$$

or

$$\delta \left(\frac{1}{2} \frac{R^2}{\mathrm{m}} \right) = 0 \,, \tag{3.9}$$

or, with the value of the energy of the constraint acceleration

$$\mathrm{S}^* = \frac{1}{2} \mathrm{m}\, \alpha^2 \,, \tag{3.10}$$

$$\delta \mathrm{S}^* = 0 \,. \tag{3.11}$$

Summing up the results obtained, the following minimum principles may be established in connection with the ideal constraints:

a) The ideal constraint changes the free motion carried out without constraint on the effect of the active force F by the kinematically possible lowest value of acceleration α:

$$\delta\left(\frac{1}{2}\,\mathrm{m}\,\alpha^2\right) = 0\,. \tag{3.11a}$$

b) The ideal constraint causes a minimum reactive force:

$$\delta\left(\frac{1}{2}\,\frac{R^2}{\mathrm{m}}\right) = 0\,. \tag{3.11b}$$

c) The energy of the acceleration induced by the ideal constraint is minimal

$$\delta S^* = 0\,.$$

d) The value of the velocity of the ideal constraint performance, in the field of possible accelerations permitted by the constraint is minimal:

$$\delta \dot{P}^* = \delta S^* = 0\,.$$

Based on the above statements, it may be stated that the ideal constraint disturbs the motion (acceleration) carried out without any constraint of a point of mass to the least possible degree.

4. Determination of the force of the ideal constraint

Let us consider the motion of the mass point submitted to the geometric constraint (1.1):

$$\mathrm{f}(r, \mathrm{t}) = 0\,, \tag{4.1}$$

whence

$$\frac{\partial \mathrm{f}}{\partial r} \cdot \dot{r} + \frac{\partial \mathrm{f}}{\partial \mathrm{t}} = 0 \tag{4.2}$$

and

$$\dot{\frac{\partial \mathrm{f}}{\partial r}} \cdot \dot{r} + \frac{\partial \mathrm{f}}{\partial r} \cdot \ddot{r} + \frac{\partial^2 \mathrm{f}}{\partial \mathrm{t}^2} = 0\,. \tag{4.3}$$

The direction of the ideal constraint force may be defined with the aid of (2.11)

$$R = \mathrm{R} \cdot n \,||\, \frac{\partial \mathrm{f}}{\partial r}\,, \tag{4.4}$$

where: R — absolute value of constraint force R,
n — unit vector of R-direction

Multiplication of Eq. (4.3) with m and replacement of (1.3) into it, yields

$$\mathrm{m}\frac{\partial \dot{\mathrm{f}}}{\partial r} \cdot \dot{r} + \frac{\partial \dot{\mathrm{f}}}{\partial r} \cdot (F + R) + \mathrm{m}\frac{\partial^2 \mathrm{f}}{\partial \mathrm{t}^2} = 0\,, \tag{4.5}$$

whence, considering (4.4), we have

$$R = -\left[\frac{\mathrm{m}}{\left|\dfrac{\partial \mathrm{f}}{\partial r}\right|}\left(\frac{\partial \dot{\mathrm{f}}}{\partial r} \cdot \dot{r} + \frac{\partial^2 \mathrm{f}}{\partial \mathrm{t}^2}\right) + F \cdot n\right] \cdot n. \tag{4.6}$$

Let for example an instationary geometric constraint be given

$$\mathrm{f}(r, \mathrm{t}) = r^2 - \mathrm{h}^2(\mathrm{t}) = 0\,, \tag{4.7}$$

wherein h(t) is a given function of time. In that case:

$$\begin{aligned} &\frac{\partial \mathrm{f}}{\partial r} = 2 \cdot r, \quad n = \frac{r}{\mathrm{h}}, \quad \frac{\partial \mathrm{f}}{\partial \mathrm{t}} = -2 \cdot \mathrm{h} \cdot \dot{\mathrm{h}}\,, \\ &\frac{\partial^2 \mathrm{f}}{\partial \mathrm{t}^2} = -2 \cdot \mathrm{h} \cdot \ddot{\mathrm{h}} - 2 \cdot \dot{\mathrm{h}}^2, \quad \frac{\partial \dot{\mathrm{f}}}{\partial r} = 2 \cdot \dot{r}\,, \end{aligned} \tag{4.8}$$

thus, the constraint force, from (4.6) and (4.8) is

$$R = \left[\mathrm{m}\left(\frac{\dot{\mathrm{h}}^2}{\mathrm{h}} + \ddot{\mathrm{h}} - \frac{\dot{r}^2}{\mathrm{h}}\right) - \frac{F \cdot r}{\mathrm{h}}\right] \cdot \frac{r}{\mathrm{h}}\,. \tag{4.9}$$

In case where the motion of the mass point is hindered by a kinematic constraint of the form (1.2), so

$$\mathrm{f}(r, \dot{r}, \mathrm{t}) = 0 \tag{4.10}$$

and

$$\frac{\partial \mathrm{f}}{\partial r} \cdot \dot{r} + \frac{\partial \mathrm{f}}{\partial \dot{r}} \cdot \ddot{r} + \frac{\partial \mathrm{f}}{\partial \mathrm{t}} = 0\,. \tag{4.11}$$

The constraint force, from (2.14) is

$$R = \mathrm{R} \cdot n \,||\, \frac{\partial \mathrm{f}}{\partial \dot{r}} . \tag{4.12}$$

By multiplying (4.11) with m and replacing (1.3) into it, one obtains

$$\mathrm{m} \frac{\partial \mathrm{f}}{\partial r} \cdot \dot{r} + \frac{\partial \mathrm{f}}{\partial \dot{r}} (F + R) + \mathrm{m} \frac{\partial \mathrm{f}}{\partial \mathrm{t}} = 0 , \tag{4.13}$$

wherefrom, by considering (4.12),

$$R = -\left[\frac{\mathrm{m}}{\left| \frac{\partial \mathrm{f}}{\partial \dot{r}} \right|} \left(\frac{\partial \mathrm{f}}{\partial r} \cdot \dot{r} + \frac{\partial \mathrm{f}}{\partial \mathrm{t}} \right) + F \cdot n \right] \cdot n \tag{4.14}$$

is obtained.

Let for example, the mass point subjected to the kinematic constraint (1.10) be given

$$\mathrm{f}(r, \dot{r}, \mathrm{t}) = r^2 - \mathrm{v}^2(\mathrm{t}) = 0 . \tag{4.15}$$

Then

$$\frac{\partial \mathrm{f}}{\partial r} = 0, \quad \frac{\partial \mathrm{f}}{\partial \dot{r}} = 2 \cdot \dot{r} , \tag{4.16}$$

$$\frac{\partial \mathrm{f}}{\partial \mathrm{t}} = -2 \cdot \mathrm{v} \cdot \dot{\mathrm{v}}, \quad n = \frac{\dot{r}}{\mathrm{v}} .$$

Replacing (4.16) into (4.14), gives the ideal constraint force

$$R = \left(\mathrm{m} \cdot \dot{\mathrm{v}} - \frac{F \cdot \dot{r}}{\mathrm{v}} \right) \frac{\dot{r}}{\mathrm{v}} . \tag{4.17}$$

5. Mechanical systems subjected to ideal constraints

The ith element of a mechanical system consisting of N material particles subjected to constraints is acted on by the active force F_i and by the reactive force R_i. Then

$$F_i + R_i = \mathrm{m}_i \cdot a_i \tag{5.1}$$

wherein: m_i — mass of i-th particle
a_i — acceleration of i-th particle.

In case, where all constraints are ideal in the system, i.e., the mechanical system in question is ideal, so, by virtue of (2.8), to every constraint force R_i it is true that

$$R_i \delta a_i = 0 \tag{5.2}$$

which summing up to the whole system results in

$$\sum_{i=1}^{N} R_i \cdot \delta a_i = 0 \,. \tag{5.3}$$

Expressing R_i from (5.1) and replacing in (5.3) gives Gauss's principle:

$$\sum_{i=1}^{N} (\mathrm{m}_i a_i - F_i) \cdot \delta a_i = 0 \,. \tag{5.4}$$

Gauss's principle expressed by Eq. (5.4) is the most general principle determining the motion of the mechanical system which may also be formulated as a minimum principle. Let us replace into (5.3) the relationships (3.2) and (3.6):

$$\delta \sum_{i=1}^{N} \left(\frac{!R_i}{\sqrt{m_i}} \right)^2 = 0 \,, \tag{5.5}$$

i.e., in the ideal mechanical systems the sum of squares of the specific constraint forces related to the square roots of the masses m_i is of minimal value.

Transformation of the Eq. (5.4) yields

$$\delta \sum_{i=1}^{N} \left(\frac{1}{2} \mathrm{m}_i \cdot a_i^2 - F_i \cdot a_i \right) = 0 \tag{5.6}$$

which expressed by the acceleration energy of the system

$$\mathrm{S} = \frac{1}{2} \sum_{i=1}^{N} \mathrm{m}_i \, a_i^2 \tag{5.7}$$

may be written in the form:

$$\delta \mathrm{S} = \sum_{i=1}^{N} F_i \cdot \delta a_i \,. \tag{5.8}$$

Let us consider, as an example for the ideal mechanical system a rigid body. The rigid body is defined by the stationary geometric constraint equations taking the form

$$(r_i - r_j)^2 = \text{constant}, \tag{5.9}$$

where r_i and r_j are the position vectors of the particles i-th and j-th, respectively, in a certain inertia system.

Let us try to find out to what internal set of forces is the constraint (5.9) equivalent.

In order to be able to determine the effects of the mass points i-th and j-th exerted on each other, let us divide the forces acting on the mass points m_i and m_j into three parts

$$F_i + B_i + R_{ij} = \mathrm{m}_i \cdot a_i \,, \tag{5.10}$$
$$F_j + B_j + R_{ji} = \mathrm{m}_j \cdot a_j \,,$$

wherein:

F_i and F_j are active (external) forces acting on mass points considered,

$B_i = \sum_{\substack{k \neq i \\ k \neq j}}' B_{ik}$ and $B_j = \sum_{\substack{k \neq j \\ k \neq i}}' B_{jk}$ are the effects of the particles other than m_i and m_j of the rigid body exerted upon the mass points m_i and m_j, which from the viewpoint of the system, consisting of mass points m_i and m_j considered, should be assumed to be external forces,

R_{ij} and R_{ji} effect on each other defined by constraint equation of mass points considered:

$$R_{ij} + R_{ji} = 0 \,. \tag{5.11}$$

Deriving twice the constraint equation (5.9) with respect to time results

$$(v_i - v_j)^2 + (r_i - r_j) \cdot (a_i - a_j) = 0 \,, \tag{5.12}$$

whence the variation of the accelerations is

$$(r_i - r_j) \cdot (\delta a_i - \delta a_j) = 0 \,. \tag{5.13}$$

Let us assume that the constraint (5.9) connecting the mass points m_i and m_j is an ideal one. In this case by virtue of (2.8)

$$R_{ij} \cdot \delta a_i = 0 \,, \tag{5.14}$$
$$R_{ji} \cdot \delta a_j = 0$$

which considering (5.11) may be written in the form

$$R_{ij} \cdot (\delta a_i - \delta a_j) = 0 \,. \tag{5.15}$$

Based on (5.11), (5.13) and (5.15)

$$R_{ij} \,||\, (r_i - r_j) \,||\, R_{ji} \,, \tag{5.16}$$

which completes Newton's third axiom defined by (5.11) with the establishments that the constraint forces defined by (5.9) are not only of the same magnitude, but also of common action line.

6. Equations of motion of anholonomic mechanical systems

In the mechanical system consisting of N material particles be

$$f_\alpha(r_1, \ldots r_N, t) = 0 \qquad (\alpha = 1 \ldots d) \tag{6.1}$$

an ideal geometric and

$$h_\beta(r_1, \ldots r_N, \dot{r}_1, \ldots \dot{r}_N, t) = 0 \qquad (\beta = 1 \ldots g) \tag{6.2}$$

an ideal kinematic constraint. Thus, the number of degrees of freedom is

$$n = 3 \cdot N - d - g\,. \tag{6.3}$$

Taking into account the geometric constraints,

$$m = 3N - d = n + g \tag{6.4}$$

by choosing the general coordinate q_k, the position vectors r_i may be written as the functions of the general coordinates. Introducing the general coordinate vector

$$\boldsymbol{q} = \begin{bmatrix} q_1 \\ \vdots \\ \vdots \\ q_m \end{bmatrix} \tag{6.5}$$

the positional vectors r_i are:

$$r_i = r_i(\boldsymbol{q}, t) \qquad (i = 1 \ldots N) \tag{6.6}$$

which put into kinematic constraint equation (6.2) yields

$$h_\beta(\boldsymbol{q}, \dot{q}, t) = 0\,, \qquad (\beta = 1 \ldots g) \tag{6.7}$$

In the following only kinematic constraints are dealt with which are the linear functions of the velocities $\dot{r}_i$ and general velocities $\dot{q}$, i.e.,

$$l_\beta(\boldsymbol{q}, t) = \boldsymbol{h}_\beta \cdot \dot{q}\,, \qquad (\beta = 1 \ldots g) \tag{6.8}$$

wherein

$$\boldsymbol{h}_\beta = h_\beta(\boldsymbol{q}, t) = \begin{bmatrix} h_{\beta 1} \\ \vdots \\ \vdots \\ h_{\beta m} \end{bmatrix}\,. \tag{6.9}$$

Let us now set up the convenient linear combinations $\mathrm{n} = \mathrm{m}-\mathrm{g}$ of the general velocities $\dot{q}_k$, the pseudovelocities

$$\dot{\mathrm{s}}_j = b_j \cdot \dot{q}\,, \qquad (\mathrm{j} = 1 \ldots \mathrm{n}) \tag{6.10}$$

where

$$b_j = b_j(q, \mathrm{t}) = \begin{bmatrix} \mathrm{b}_{j1} \\ \vdots \\ \mathrm{b}_{jm} \end{bmatrix}. \tag{6.11}$$

By interpretation with the column vector or matrix

$$\dot{s} = \begin{bmatrix} \dot{\mathrm{s}}_1 \\ \cdot \\ \cdot \\ \cdot \\ \dot{\mathrm{s}}_n \\ l_1 \\ \cdot \\ \cdot \\ \cdot \\ l_g \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} \mathrm{b}_{11} \ldots\ldots \mathrm{b}_{1m} \\ \cdot \quad \ldots\ldots \quad \cdot \\ \cdot \quad \ldots\ldots \quad \cdot \\ \cdot \quad \ldots\ldots \quad \cdot \\ \mathrm{b}_{n1} \ldots\ldots \mathrm{b}_{nm} \\ \mathrm{h}_{11} \ldots\ldots \mathrm{h}_{1m} \\ \cdot \quad \ldots\ldots \quad \cdot \\ \cdot \quad \ldots\ldots \quad \cdot \\ \cdot \quad \ldots\ldots \quad \cdot \\ \mathrm{h}_{g1} \ldots\ldots \mathrm{h}_{gm} \end{bmatrix} \tag{6.12}$$

one obtains

$$\dot{s} = \mathbf{B} \cdot \dot{q}\,. \tag{6.13}$$

By a convenient selection of the functions $\mathrm{b}_{ji}(q,\mathrm{t})$ it may always be achieved that

$$\det \mathbf{B} \neq 0\,, \tag{6.14}$$

whence, by using the notation $\mathbf{B}^{-1} = \mathbf{C}$

$$\dot{q} = \mathbf{C} \cdot \dot{s}\,. \tag{6.15}$$

From (6.6) the velocity of the mass point m_i is

$$\dot{r}_i = \frac{\partial r_i}{\partial q} \cdot \dot{q} + \frac{\partial r_i}{\partial \mathrm{t}} = \mathbf{D}_i \cdot \dot{q} + d_i\,, \tag{6.16}$$

or, by replacing (6.15)

$$\dot{r}_i = \mathbf{D}_i \cdot \mathbf{C} \cdot \dot{s} + d_i = \mathbf{H}_i \cdot \dot{s} + d_i\,, \tag{6.17}$$

wherein the notation $\mathbf{D}_i \cdot \mathbf{C} = \mathbf{H}_i$ was introduced.

The acceleration of the mass point m_i, calculated from (6.17) is

$$\ddot{r}_i = \dot{\mathbf{H}}_i \cdot \dot{s} + \mathbf{H}_i \cdot \ddot{s} + d_i \,. \tag{6.18}$$

Thus, the acceleration energy of the system is

$$S = \frac{1}{2} \sum_{i=1}^{N} m_i \ddot{r}_i^2 = S(q, \dot{s}, \ddot{s}, t) \,, \tag{6.19}$$

and its variation according to the pseudoacceleration $\ddot{s}$

$$\delta S = \frac{\partial S}{\partial \ddot{s}} \cdot \delta \ddot{s} \,, \tag{6.20}$$

wherein

$$\delta \ddot{s} = \begin{bmatrix} d\ddot{s}_1 \\ \vdots \\ \delta \ddot{s}_n \\ 0 \\ \vdots \\ 0 \end{bmatrix} , \tag{6.21}$$

because

$$\delta l_\beta = 0 \,, \qquad (\beta = 1 \ldots g) \tag{6.22}$$

The variation of the acceleration $\ddot{r}_i = a_i$, according to (6.18) is:

$$\delta a_i = \delta \ddot{r}_i = \mathbf{H}_i \cdot \delta \ddot{s} \,. \tag{6.23}$$

Replacement of the Eqs (6.20) and (6.23) into (5.8) yields

$$\frac{\partial S}{\partial \ddot{s}} \cdot \delta \ddot{s} = \sum_{i=1}^{N} F_i^T \cdot \mathbf{H}_i \cdot \delta \ddot{s} = Q \cdot \delta \ddot{s} \,, \tag{6.24}$$

wherein F_i^T is the transposition of the matrix F_i and

$$\sum_{i=1}^{N} F_i^T \cdot \mathbf{H}_i = Q = [Q_1, \ldots Q_n] \tag{6.25}$$

is the general force which may be defined by making use of the identity

$$\delta S = F_1 \cdot \delta a_1 + \ldots + F_N \cdot \delta a_N = Q_1 \delta \ddot{s}_1 + \ldots + Q_n \delta \ddot{s}_n \tag{6.26}$$

Since all the $\delta\ddot{s}_j$-s are independent of each other, from (6.24) it is evident that

$$\frac{\partial S}{\partial \ddot{s}} = Q\,,$$

or

$$\frac{\partial S}{\partial \ddot{s}_j} = Q_j \quad (j = 1, \ldots n) \tag{6.27}$$

are the equations of motion according to APPELL, of the anholonomic systems.

7. Equations of motion of holonomic systems

In the holonomic systems there are only geometric constraints, i.e.,

$$n = 3 \cdot N - d = m\,, \tag{7.1}$$

accordingly, in lieu of the pseudovelocities one can calculate by using the general velocities $\dot{q}_k$

$$\dot{s} = \dot{q}\,, \tag{7.2}$$

whence, by virtue of (6.13) and (7.2),

$$\mathbf{B} = \mathbf{C} = \mathbf{I} \tag{7.3}$$

which is the unit matrix and, according to (6.25), the general force is

$$Q = \sum_{i=1}^{N} F_i^T \cdot \mathbf{H}_i = \sum_{i=1}^{N} F_i^T \cdot \mathbf{D}_i \cdot \mathbf{C} = \sum_{i=1}^{N} F_i^T \cdot \mathbf{D}_i = \sum_{i=1}^{N} F_i^T \cdot \frac{\partial r_i}{\partial q}\,. \tag{7.4}$$

Thus, Appell's equation is

$$\frac{\partial S}{\partial \ddot{q}} = Q\,, \tag{7.5}$$

wherein:

$$\begin{aligned}
\frac{\partial S}{\partial \ddot{q}} &= \frac{\partial}{\partial \ddot{q}}\left[\frac{1}{2}\sum_{i=1}^{N} m_i \cdot \ddot{r}_i^2\right] = \sum_{i=1}^{N} m_i \cdot \ddot{r}_i \cdot \frac{\partial \ddot{r}}{\partial \ddot{q}} = \\
&= \frac{d}{dt}\left[\sum_{i=1}^{N} m_i \cdot \dot{r}_i \cdot \frac{\partial \dot{r}_i}{\partial q}\right] - \sum_{i=1}^{N} m_i \cdot \dot{r}_i \cdot \frac{\partial \ddot{r}_i}{\partial \dot{q}} = \\
&= \frac{d}{dt}\frac{\partial}{\partial \dot{q}}\left(\frac{1}{2}\sum_{i=1}^{N} m_i \cdot \dot{r}_i^2\right) - \frac{\partial}{\partial q}\left(\frac{1}{2}\sum_{i=1}^{N} m_i \cdot \dot{r}_i^2\right) = \\
&= \frac{d}{dt}\frac{\partial T}{\partial \dot{q}} - \frac{\partial T}{\partial q}\,.
\end{aligned} \tag{7.6}$$

Here, it has been taken into account that by considering (6.16), (6.17), (6.18) and (7.3):

$$\frac{\partial \ddot{r}_i}{\partial \ddot{q}} = \frac{\partial \dot{r}_i}{\partial \dot{q}} = \frac{\partial r_i}{\partial q} = \mathbf{D}_i \quad \text{and} \quad \frac{\partial \ddot{r}_i}{\partial \dot{q}} = \frac{\partial \dot{r}_i}{\partial q}, \tag{7.7}$$

and the value of the kinetic energy

$$\mathrm{T} = \frac{1}{2} \sum_{i=1}^{N} \mathrm{m}_i \cdot \dot{r}_i^2 \tag{7.8}$$

has been replaced into (7.6). Finally, by taking into account (7.5) and (7.6) one obtains Lagrange's equation of the second kind valid for holonomic systems:

$$\frac{\mathrm{d}}{\mathrm{dt}} \frac{\partial \mathrm{T}}{\partial \dot{q}} - \frac{\partial \mathrm{T}}{\partial q} = Q ,$$

i.e.,

$$\frac{\mathrm{d}}{\mathrm{dt}} \frac{\partial \mathrm{T}}{\partial \ddot{q}_\mathrm{k}} - \frac{\partial \mathrm{T}}{\partial q_\mathrm{k}} = \mathrm{Q}_k, \quad (k = 1 \ \ldots \ n) \tag{7.9}$$

8. Equations of motion of the rigid body

The acceleration a_s, angular velocity ω and angular acceleration ε of the centre of gravity S of the rigid body of mass m the acceleration energy is expressed by the formula:

$$\mathrm{S} = \frac{1}{2} \int_m [a_s + \varepsilon \times r + \omega \times (\omega \times r)]^2 \, \mathrm{dm}, \tag{8.1}$$

wherefrom, by taking into account the equation we obtain

$$\int_m r \cdot \mathrm{dm} = 0 , \tag{8.2}$$

$$\mathrm{S} = \frac{1}{2} \mathrm{m} \cdot a_s^2 + \frac{1}{2} \int_m [\varepsilon \times r + \omega \times (\omega \times r)]^2 \, \mathrm{dm} . \tag{8.3}$$

Defining the value of the integral at the right-hand side of the equation it may be written:

$$\mathrm{S} = \frac{1}{2} \mathrm{m} a_s^2 + \frac{1}{2} \varepsilon^T \cdot \mathbf{\Theta}_s \cdot \varepsilon + (\varepsilon \omega \pi_s) + [\omega \circ \omega - \mathbf{I} \cdot \omega^2]^2 \, .. \, (\mathbf{I} \cdot \Theta_s - \mathbf{\Theta}_s), \tag{8.4}$$

wherein

$$\mathbf{\Theta}_s = \int_m [\mathbf{I} \cdot r^2 - r \circ r] \cdot \mathrm{dm} \tag{8.5}$$

is the tensor of inertia calculated with respect to the centre of gravity of the rigid body,

$$\pi_s = \mathbf{\Theta}_s \cdot \omega \tag{8.6}$$

the vector of the moment of momentum calculated with respect to the centre of gravity of the rigid body,

$$\Theta_s = \int_m r^2 \cdot \mathrm{dm} \tag{8.7}$$

the moment of inertia calculated with respect to the centre of gravity of the rigid body.

Considering the velocity and angular-velocity coordinates of the pseudo-velocity:

$$\dot{s}_1 = v_s\,, \quad \text{and} \quad \dot{s}_2 = \omega\,. \tag{8.8}$$

Since, from the point of view of the derivation of Appell's equations the last term in Eq. (8.4) is indifferent, it is sufficient to determine the part

$$\mathrm{S}^0 = \frac{1}{2}\,\mathrm{m}a_s^2 + \frac{1}{2}\,\varepsilon^T \cdot \mathbf{\Theta}_s \cdot \varepsilon + (\varepsilon\omega\pi_s)\,. \tag{8.9}$$

For the determination of the general forces the motor to the center of gravity S of the external set of forces acting on the rigid body

$$[F, M_S]_S$$

should be taken as a basis. The variation of the acceleration-energy may be written according to Eq. (6.26) as follows:

$$F \cdot \delta a_\mathrm{S} + M_\mathrm{S} \cdot \delta\varepsilon = Q_1 \cdot \delta\ddot{\mathrm{s}}_1 + Q_2 \cdot \delta\ddot{\mathrm{s}}_2\,, \tag{8.10}$$

whence, by making use of (8.8),

$$Q_1 = F \qquad \text{and} \qquad Q_2 = M_\mathrm{S} \tag{8.11}$$

is obtained.

Accordingly, Appell's equations give, on the one hand the expression

$$\frac{\partial \mathrm{S}}{\partial a_\mathrm{s}} = Q_1\,,$$

i.e., by using (8.9) and (8.11), the theorem of momentum:

$$\mathbf{m} \cdot \boldsymbol{a}_S = F \tag{8.12}$$

and, on the other hand,

$$\frac{\partial S}{\partial \varepsilon} = Q_2 ,$$

i.e.,

$$\boldsymbol{\Theta}_S \cdot \varepsilon + \omega \times \pi_S = M_S \tag{8.13}$$

give the Euler's equation of motion of the rigid body.

REFERENCES

1. Appell, P.: Traité de mécanique rationnelle. Tome 2. Gauthier-Villars, Paris 1953, pp. 337—395, 492—499
2. Pérès, J.: Mécanique générale. Masson et Cie. Paris 1962, pp. 77—99, 192—226
3. ПАРС. Л. А. Аналитическая динамика (A treatise on analytical dynamics). Наука, Москва 1971. стр. 15—18, 220—243.
4. Gantmacher, F.: Lectures in analytical mechanics. Mir. Moscow 1975, pp. 9—65. [Лекции по аналитической механике].

Verallgemeinerung und einige Anwendungen des idealen Zwangsbegriffes. — Beschrieben wird die klassische Deutung des idealen Zwanges, die sich nur auf geometrische Zwänge anwenden läßt. Im weiteren gibt der Autor zur Definition des idealen Zwanges eine Formulierung an, die auch die klassische Deutung enthält, jedoch auch auf kinematische Zwänge angewandt werden kann. Aufgrund dieser Definition des idealen Zwanges können Minimumprinzipien festgestellt werden, die zum allgemeinsten, die Bewegung der mechanischen Systeme definierenden Grundsatz, zum Gaußschen Grundsatz führen. Es wird nachgewiesen, daß aufgrund der obenerwähnten Prinzipien die Appellschen Bewegungsgleichungen der Anholonomsysteme, die Lagrangeschen Bewegungsgleichungen der Holonomsysteme und die Eulerschen Bewegungsgleichungen des steifen Körpers ermittelt werden können.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (1—2), pp. 151—162 (1980)

MATRIX ITERATION ANALYSIS OF RECTANGULAR PLATES WITH SIMPLY SUPPORTED AND FREE EDGES

I. HEGEDŰS*

[Manuscript received November 24, 1977]

An efficient computer method is presented for the analysis of elastic deflections of rectangular plates with free and simply supported edges. It is based on the direct solution of the matrix equation of nodal deflections, hence it is a development of the method for all around simply supported rectangular plates where, rather than assuming fictitious load complexes, internal iteration is applied to yield results of arbitrary accuracy for a boundary condition seemingly preventing direct solution. Convergency of iteration is demonstrated, and on the basis of the constructive demonstration, a condition sufficient for the convergency of similar types of iteration is given.

1. Introduction

Difference equations deduced from the Lagrange differential equation for small deflections of elastic plates are usually of a special built-up, therefore, further analysis can be simplified from several aspects. Among the simplified methods, the direct solution of the matrix equation of nodal deflections by spectral decomposition of coefficient matrices, due to EGERVÁRY [1, 2], is made outstanding by its economy and elegance of mathematical treatment. The economy of this method established for analyzing simply supported rectangular plates appears from its application in problems of analyzing even plates of rather general form and boundary conditions. In these cases the rectangular plate with four Navier edges acts as a primary beam of the tested structure where the system of deflections to be determined of the tested plate domain inside the "primary plate" consists of a linear combination of deflections produced by properly chosen special load systems (determinable by solving a linear equation system).

Now, a development of the method will be applied for the analysis of rectangular plates with one or two free edges replacing the quoted method of "primary plates" by internal iteration, to be refined at will.

A typical case — matrix equation of nodal deflections of slabs with two parallel free edges — will be involved to illustrate deviations from the basic

* Dr. HEGEDŰS I., Váci Mihály u. 10, H-2083 Solymár, Hungary

problem — analysis of a rectangular plate with four Navier edges, — presenting the solution method and demonstrating the applied iteration to converge. Finally, possibilities of generalizing the method will be discussed.

2. Matrix equation for nodal displacements of rectangular plates

The method of finite differences is based on approximating the required function by a finite set of values, its derivatives by quotients of differences, writing difference equations of the same purport as the differential equation of the problem — irrespective of approximation errors — and determining the elements of the substituting set of values by solving the difference equation system. In case of a linear differential equation, the system of difference equations is a vectorial equation taking the form:

$$\mathbf{A}\,\mathbf{w} = \frac{1}{K}\,\mathbf{p} \tag{1}$$

where $\mathbf{w}$ and $\mathbf{p}$ are deflection and load vectors, resp., K is the flexural rigidity of the plate, and $\mathbf{A}$ the coefficient matrix of the equation system.

Applying the method of finite differences with a rectangular network for determining deflections of rectangular plates with boundary conditions constant along each edge, $\mathbf{A}$ obtains a special built-up: it can be produced as sum of three simple direct products [4]. For a plate with four Navier edges (Fig. 1a), applying the usual difference operators:

$$\mathbf{A} = \underset{m}{\mathbf{C}^2} \otimes \underset{n}{\mathbf{E}} + 2\,\underset{m}{\mathbf{C}} \otimes \underset{n}{\mathbf{C}} + \underset{m}{\mathbf{E}} \otimes \underset{n}{\mathbf{C}^2} = (\underset{m}{\mathbf{C}} \otimes \underset{n}{\mathbf{E}} + \underset{m}{\mathbf{E}} \otimes \underset{n}{\mathbf{C}})^2 \tag{1.a}$$

subscripts referring to matrix sizes and $\otimes$ to the direct multiplication

$$\underset{m}{\mathbf{C}} = \frac{1}{h_y^2}\begin{bmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & \ddots & \ddots & \ddots \\ & & 1 & -2 \end{bmatrix}, \qquad \underset{n}{\mathbf{C}} = \frac{1}{h_x^2}\begin{bmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & \ddots & \ddots & \ddots \\ & & 1 & -2 \end{bmatrix}$$

Be $\mathbf{W}$ and $\mathbf{P}$ matrices of size $m \times n$ including elements of $\mathbf{w}$ and $\mathbf{p}$ in conformity with the arrangement of the nodal system. The correlation between linear equation systems of direct product-polynomial coefficient matrices and matrix equations (e.g. [3]) permits the writing of matrix equation having a content perfectly identical to that of (1):

$$\underset{m}{\mathbf{C}^2}\mathbf{W} + 2\,\underset{m}{\mathbf{C}}\,\mathbf{W}\,\underset{n}{\mathbf{C}^*} + \mathbf{W}\,\underset{n}{\mathbf{C}^{2*}} = \frac{1}{K}\,\mathbf{P} \tag{2}$$

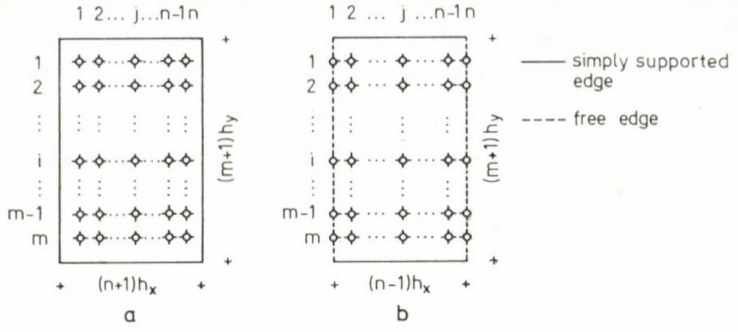


Fig. 1

where ( )* indicates transposition that can be even omitted due to symmetry in $\mathbf{C}$.

Considering the spectral decomposition of coefficient matrices to be known:

$$\underset{m}{\mathbf{C}} = \underset{m}{\mathbf{U}}\langle\lambda_i\rangle\underset{m}{\mathbf{V}}, \quad \underset{m}{\mathbf{C}^2} = \underset{m}{\mathbf{U}}\langle\lambda_i^2\rangle\underset{m}{\mathbf{V}}, \quad i = 1, 2, \ldots, m$$

$$\underset{n}{\mathbf{C}} = \underset{n}{\mathbf{U}}\langle\lambda_j\rangle\underset{n}{\mathbf{V}}, \quad \underset{n}{\mathbf{C}^2} = \underset{n}{\mathbf{U}}\langle\lambda_j^2\rangle\underset{n}{\mathbf{V}}, \quad j = 1, 2, \ldots, n.$$

Using variable transformation $\mathbf{X} = \underset{m}{\mathbf{V}}\,\mathbf{W}\,\underset{n}{\mathbf{U}}$, $\mathbf{Q} = \dfrac{1}{K}\,\underset{m}{\mathbf{V}}\,\mathbf{P}\,\underset{n}{\mathbf{U}}$ Eq. (2) can be transformed into a matrix equation of diagonal coefficients:

$$\langle\lambda_i^2\rangle\mathbf{X} + 2\langle\lambda_i\rangle\mathbf{X}\langle\lambda_j\rangle + \mathbf{X}\langle\lambda_j^2\rangle = \mathbf{Q}\,. \tag{2.a}$$

Introducing the logical multiplication interpreted for matrices of equal size:

$$[a_{ij}] \curlywedge [b_{ij}] = [a_{ij} \cdot b_{ij}],$$

(2.a) may be written in the form:

$$[(\lambda_i^2 + 2\,\lambda_i\lambda_j + \lambda_j^2)] \curlywedge \mathbf{X} = [(\lambda_i + \lambda_j)^2] \curlywedge \mathbf{X} = \mathbf{Q}, \tag{2.b}$$

hence (since $\lambda_i + \lambda_j \neq 0$):

$$\mathbf{X} = [(\lambda_i + \lambda_j)^{-2}] \curlywedge \mathbf{Q}, \quad \mathbf{W} = \mathbf{U}_m\{[(\lambda_i + \lambda_j)^{-2}] \curlywedge \mathbf{Q}\}\mathbf{V}_n\,. \tag{3.a,b}$$

Trigonometric equations easy to program are available (in e.g. [3]) for the spectral decomposition of $\underset{m}{\mathbf{C}}$, $\underset{n}{\mathbf{C}}$ (and of subsequently factorized coefficient matrices) so the solution is simple to write, so to say, in a closed formula. Radical simplification — as seen by deduction steps — is provided partly by the possibility to

"contract" differential equation system (1) into matrix equation (2), and partly by the identify of modals in the decomposed right and left-hand side coefficient matrices (2) so that the equation can be transformed into a simpler one having purely diagonal coefficient matrices.

For analysing the plate with each two parallel Navier types and free edges in Fig. 1b, in case of $\nu = 0$, of important computational advantage, difference equations of an analogous simplicity can be written [4]:

$$\mathbf{A}\,\mathbf{w} = [\underset{m}{\mathbf{C}^2} \otimes \underset{n}{\tilde{\tilde{\mathbf{E}}}} + 2\,\underset{m}{\mathbf{C}} \otimes \underset{n}{\tilde{\tilde{\mathbf{C}}}} + \underset{m}{\mathbf{E}} \otimes \underset{n}{\tilde{\tilde{\mathbf{B}}}}]\,\mathbf{w} = \frac{1}{K}\,(\underset{m}{\mathbf{E}} \otimes \underset{n}{\tilde{\tilde{\mathbf{E}}}})\,\mathbf{p} \tag{4}$$

where

$$\underset{n}{\tilde{\tilde{\mathbf{E}}}} = \underset{n}{\mathbf{E}} - \underset{n}{\tilde{\tilde{\mathbf{F}}}}, \quad \underset{n}{\tilde{\tilde{\mathbf{F}}}} = \langle 0{,}5\ 0\ \ldots\ 0\ 0{,}5\rangle;$$

$$\underset{n}{\tilde{\tilde{\mathbf{C}}}} = \underset{n}{\mathbf{C}} + \frac{2}{h_x^2}\,\underset{n}{\tilde{\tilde{\mathbf{F}}}}; \tag{5}$$

$$\underset{n}{\tilde{\tilde{\mathbf{B}}}} = \underset{n}{\tilde{\tilde{\mathbf{C}}}^2} - \underset{n}{\tilde{\tilde{\mathbf{D}}}}, \quad \underset{n}{\tilde{\tilde{\mathbf{D}}}} = \begin{bmatrix} 1 & -1 & & & & \\ -1 & 1 & & & & \\ & & 0 & & & \\ & & & 0 & & \\ & & & & 1 & -1 \\ & & & & -1 & 1 \end{bmatrix} \frac{1}{h_x^4}.$$

Set of Eqs (4) contracted into a matrix equation:

$$\underset{m}{\mathbf{C}^2}\,\mathbf{W}\,\underset{n}{\tilde{\tilde{\mathbf{E}}}} + 2\,\underset{m}{\mathbf{C}}\,\mathbf{W}\,\underset{n}{\tilde{\tilde{\mathbf{C}}}} + \mathbf{W}\,\underset{n}{\tilde{\tilde{\mathbf{B}}}} = \frac{1}{K}\,\mathbf{P}\,\underset{n}{\tilde{\tilde{\mathbf{E}}}}. \tag{6}$$

A great difference against (2) is that coefficients of (6) cannot be simultaneously transformed into diagonals similar to (2.a), preventing the direct application of that method of solution. Rearranging (6) by means of Eqs (5), to leave only terms with simultaneously diagonizable coefficient matrices:

$$\underset{m}{\mathbf{C}^2}\,\mathbf{W} + 2\,\underset{m}{\mathbf{C}}\,\mathbf{W}\,\underset{n}{\tilde{\tilde{\mathbf{C}}}} + \mathbf{W}\,\underset{n}{\tilde{\tilde{\mathbf{C}}}^2} = \frac{1}{K}\,\mathbf{P}\,\underset{n}{\tilde{\tilde{\mathbf{E}}}} + \mathbf{W}\,\underset{n}{\tilde{\tilde{\mathbf{D}}}} + \underset{m}{\mathbf{C}^2}\,\mathbf{W}\,\underset{n}{\tilde{\tilde{\mathbf{F}}}}. \tag{6.a}$$

Again, considering spectral decomposition $\underset{m}{\mathbf{C}} = \underset{m}{\mathbf{U}} \langle \lambda_i \rangle \underset{m}{\mathbf{V}}$ and $\underset{n}{\tilde{\tilde{\mathbf{C}}}} = \underset{n}{\tilde{\tilde{\mathbf{U}}}} \langle \tilde{\lambda} \rangle \underset{n}{\tilde{\tilde{\mathbf{V}}}}$ as being known, variable transformation similar to the former one will be applied:

$$\mathbf{X} = \underset{m}{\mathbf{V}}\,\mathbf{W}\,\underset{n}{\tilde{\tilde{\mathbf{U}}}}. \qquad \mathbf{Q} = \frac{1}{\mathbf{K}}\,\underset{m}{\mathbf{V}}\,\mathbf{P}\,\underset{n}{\tilde{\tilde{\mathbf{E}}}}\,\underset{n}{\tilde{\tilde{\mathbf{U}}}}.$$

Introducing matrices

$$\underset{n}{\tilde{\tilde{\mathbf{M}}}} = \underset{n}{\tilde{\tilde{\mathbf{V}}}}\,\underset{n}{\tilde{\tilde{\mathbf{D}}}}\,\underset{n}{\tilde{\tilde{\mathbf{U}}}} \quad \text{and} \quad \underset{n}{\tilde{\tilde{\mathbf{N}}}} = \underset{n}{\tilde{\tilde{\mathbf{V}}}}\,\underset{n}{\tilde{\tilde{\mathbf{F}}}}\,\underset{n}{\tilde{\tilde{\mathbf{U}}}},$$

Eq. (6.a) can be rewritten as:

$$[(\lambda_i + \tilde{\tilde{\lambda}}_j)^2] \curlywedge \mathbf{X} = \mathbf{Q} + \mathbf{X}\underset{n}{\tilde{\tilde{\mathbf{M}}}} + \langle \lambda_i^2 \rangle \mathbf{X}\underset{n}{\tilde{\tilde{\mathbf{N}}}}. \tag{6.b}$$

Although no direct solution formula similar to (3.a,b) can be given for both sides of (6.b) containing the unknown matrix, $\lambda_i + \tilde{\lambda}_j \neq 0$ permits the following iteration:

$$\mathbf{X}_0 = 0\,.$$

$$\mathbf{X}_{k+1} = [(\lambda_i + \tilde{\tilde{\lambda}}_j)^{-2}] \curlywedge \{\mathbf{Q} + \mathbf{X}_k \tilde{\tilde{\mathbf{M}}} + \langle \lambda_i^2 \rangle \mathbf{X}_k \tilde{\tilde{\mathbf{N}}}\}\,. \tag{7}$$

Provided the iteration is convergent, the limiting value of $\mathbf{X}_k$ meets Eq. (6.b) hence $\lim_{k\to\infty} \mathbf{X}_k = \mathbf{X}$.

The solution of Eq. (6) is to produce by "retransformation" using modals of $\mathbf{C}_m$ and $\tilde{\tilde{\mathbf{C}}}_n$:

$$\mathbf{W} = \underset{m}{\mathbf{U}}\,\mathbf{X}\,\underset{n}{\tilde{\tilde{\mathbf{V}}}}\,.$$

A non-zero Poisson's ratio involves a further modification $\tilde{\tilde{\mathbf{C}}}$ and $\tilde{\tilde{\mathbf{B}}}$, thus, trigonometric formulae in (3) do not suit spectral decomposition, but the solution principle is the same.

3. Proof of the iteration convergence

Direct observation of the element by element convergence of the series of matrices $\mathbf{X}_k$ is hampered by different type matrix operations in the iteration formula, therefore, a different way will be taken to prove the convergame of matrix iteration. First, convergence of the matrix iteration in case of convergence of an adequately selected vector iteration will be recognized, then this vector iteration will be proved to converge.*

The first, essential statement will refer to the possibility of achieving the transformed matrix of nodal deflection by the used iteration. Namely, operations involved in iteration formula (7) permit any element of $\mathbf{X}$ to assume an arbitrary real value. Now then, also elements of $\mathbf{W}_k$ can assume any value of

* The author found a direct proof of the matrix iteration convergence after handing in the manuscript of the paper. It is based on the invariance of trace of matrix products against simultaneous orthogonal transformation of the matrices.

concrete mechanical purport, $\mathbf{X}$ and $\mathbf{W}_k$ being in a mutually unambiguous relation. Again, convergence of series of matrices $\mathbf{X}_k$ is equivalent to the convergence of the series of matrices $\mathbf{W}_k$ interpreted by "retransformation" and of the series of vectors $\mathbf{w}_k$ composed of columns of $\mathbf{W}_k$ and vice versa. Thus, demonstrating the convergence of $\mathbf{w}_k$ means the granting convergence of iteration for $\mathbf{X}_k$.

Convergence of the series of vectors will be proved by using the principle of minimum potential energy (Castigliano theorem). The Castigliano functional being in finitized form:

$$\Pi_c(\mathbf{x}) = \frac{K}{2}\mathbf{x}^* \mathbf{A}\,\mathbf{x} - \mathbf{x}^* (\underset{m}{\mathbf{E}} \otimes \underset{n}{\tilde{\tilde{\mathbf{E}}}})\,\mathbf{p} = \text{minimum!} \tag{8}$$

The first term in the expression multiplied by $h_x \cdot h_y$ is strain energy belonging to deformation $\mathbf{x}$, and the second one the potential loss of external forces, of course, with neglects due to approximations in the method of finite differences. Deflection vector $\mathbf{x} = \mathbf{w}$ providing the minimum of Π_c contains nodal displacements yielding a solution of the problem, and is equal exactly to the solution of (4), obvious from the identity between the set of conditional equations requiring the first partial derivatives of Π_c with respect to all elements of vector $\mathbf{x}$ to vanish, and (4). (Let us point out here the importance of symmetry of the coefficient matrix of the difference equation system: positiveness of the first term is proportional to the strain energy of the plate in (8) for any $\mathbf{x} \neq 0$ is granted mathematically only if $\mathbf{A}$ is symmetric.).

Modifying function in (8) by a term containing a vector parameter $\mathbf{t}$:

$$\Pi(\mathbf{x}; \mathbf{t}) = \frac{K}{2}\{\mathbf{x}^* \mathbf{A}\,\mathbf{x} + (\mathbf{x} - \mathbf{t})^* (\hat{\mathbf{A}} - \mathbf{A})(\mathbf{x} - \mathbf{t}) - \\ - \mathbf{x}^* (\underset{m}{\mathbf{E}} \otimes \underset{n}{\tilde{\tilde{\mathbf{E}}}})\,\mathbf{p} = \text{minimum!} \tag{9}$$

where

$$\hat{\mathbf{A}} = \underset{m}{\mathbf{C}^2} \otimes \underset{n}{\mathbf{E}} + 2\,\underset{m}{\mathbf{C}} \otimes \underset{n}{\tilde{\tilde{\mathbf{C}}}} + \underset{m}{\mathbf{E}} \otimes \underset{n}{\tilde{\tilde{\mathbf{C}}}^2} = [\underset{m}{\mathbf{C}} \otimes \underset{n}{\mathbf{E}} + \underset{m}{\mathbf{E}} \otimes \underset{n}{\tilde{\tilde{\mathbf{C}}}}]^2$$

$\hat{\mathbf{A}}$ is the matrix of an operation of the same effect as the combination in the left-hand side of matrix equation (6.a) on the deflection vector $\mathbf{w}$ formed by columns of $\mathbf{W}$. Therefore, iterated matrices of iteration (7) can be matched exactly with the following iterated vectors:

$$\mathbf{w}_0 = 0$$

$$\begin{aligned} &\Pi(\mathbf{x};\ \mathbf{w}_0) = \text{minimum!} && \mathbf{x} = \mathbf{w}_1\,; \\ &\Pi(\mathbf{x};\ \mathbf{w}_1) = \text{minimum!} && \mathbf{x} = \mathbf{w}_2\,; \\ &\quad\vdots \\ &\Pi(\mathbf{x};\ \mathbf{w}_k) = \text{minimum!} && \mathbf{x} = \mathbf{w}_{k+1}\,. \end{aligned} \tag{10}$$

Applying the Egerváry theorem [1] on hypermatrix eigenvalues permits to directly prove that for $\mathbf{x} \neq 0$:

$$\mathbf{x}^*\hat{\mathbf{A}}\mathbf{x} \geq \mathbf{x}^*\mathbf{A}\mathbf{x} > 0\,, \qquad (11)$$
$$\mathbf{x}^*(\hat{\mathbf{A}} - \mathbf{A})\mathbf{x} \geq 0\,.$$

Taking the uniqueness of the solution into account, condition (9) and inequality (11) yield: For $\mathbf{x} \neq \mathbf{t}$

$$\Pi(\mathbf{x}\,;\,\mathbf{t}) \geq \Pi(\mathbf{x}\,;\,\mathbf{x}) = \Pi_c(\mathbf{x}) \qquad (12)$$

Applying inequality (12) on vectorial iteration (10):

$$\begin{aligned}
&\Pi(\mathbf{w}_1\,;\,\mathbf{w}_0) \geq \Pi(\mathbf{w}_1\,;\,\mathbf{w}_1))\,, && \text{due to (12)} \\
&\Pi_c(\mathbf{w}_1) = \Pi(\mathbf{w}_1\,;\,\mathbf{w}_1) > \Pi(\mathbf{w}_2\,;\,\mathbf{w}_1)\,, && \text{due to (10)} \\
&\Pi(\mathbf{w}_2\,;\,\mathbf{w}_1) \geq \Pi(\mathbf{w}_2\,;\,\mathbf{w}_2)\,, && \text{due to (12)} \qquad (13) \\
&\Pi_c(\mathbf{w}_2) = \Pi(\mathbf{w}_2\,;\,\mathbf{w}_2) > \Pi(\mathbf{w}_3\,;\,\mathbf{w}_2)\,, && \text{due to (10) etc.}
\end{aligned}$$

Definite inequalities inside the sequence (13) of inequalities permit to write the series of definite inequalities:

$$\Pi_c(\mathbf{w}_0) > \Pi_c(\mathbf{w}_1) > \ldots > \Pi_c(\mathbf{w}_k) > \Pi_c(\mathbf{w}_{k+1}) > \ldots \qquad (14)$$

Thus, vectorial iteration yields a monotonous decreasing series of $\Pi_c(\mathbf{x})$. $\Pi_c(\mathbf{x})$ having a lower bound the series tends to $\Pi_c(\mathbf{w})$ and the series of iterated vectors $\mathbf{w}_k$ tends to $\mathbf{w}$, solution of (14). Iterated vectors are of finite size, therefore, convergence is met for all vector elements.

By proving the convergence of vectorial iteration, also that of matrix iteration has been proved.

4. Further applications

This method, developed for plates simply supported along two opposite edges, can be applied — with some alterations unaffecting its essential — on rectangular plates simply supported along three edges and the fourth edge free (Fig. 2.a) or simply supported along two adjacent edges and unsupported along the other two (Fig. 2.b).

In the former case, transformation and arrangement of matrix equation

$$\underset{m}{\mathbf{C}^2}\,\underset{n}{\mathbf{W}\,\tilde{\mathbf{E}}} + 2\,\underset{m}{\mathbf{C}}\,\underset{n}{\mathbf{W}\,\tilde{\mathbf{C}}} + \underset{n}{\mathbf{W}\,\tilde{\mathbf{B}}} = \frac{1}{K}\,\underset{n}{\mathbf{P}\,\tilde{\mathbf{E}}}$$

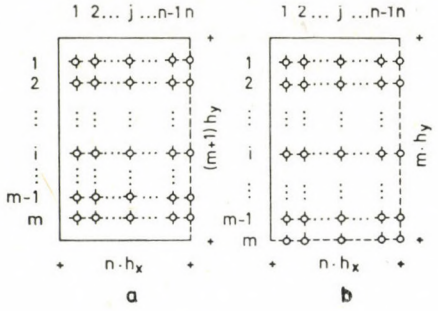


Fig. 2

into iteration formula

$$\mathbf{x}_{k+1} = [(\lambda_i + \tilde{\lambda}_j)^{-2}] \curlywedge \{\mathbf{Q} + \mathbf{X}\tilde{\mathbf{M}} + \langle\lambda_i^2\rangle \mathbf{X}\tilde{\mathbf{N}}\},$$

in the second case, a similar transformation of

$$\underset{m}{\tilde{\mathbf{B}}}\,\mathbf{W}\,\underset{n}{\tilde{\mathbf{E}}} + 2\,\underset{m}{\tilde{\mathbf{C}}}\,\mathbf{W}\,\underset{n}{\tilde{\mathbf{C}}} + \underset{m}{\tilde{\mathbf{E}}}\,\mathbf{W}\,\underset{n}{\tilde{\mathbf{B}}} = \frac{1}{K}\,\underset{m}{\tilde{\mathbf{E}}}\,\mathbf{P}\,\underset{n}{\tilde{\mathbf{E}}}$$

yielding the iteration formula

$$\mathbf{X}_{k+1} = [(\tilde{\lambda}_i + \tilde{\lambda}_j)^{-2}] \curlywedge \{\mathbf{Q} + \underset{m}{\tilde{\mathbf{M}}}\mathbf{X}_k(\mathbf{E} - \underset{n}{\tilde{\mathbf{N}}}) + \langle\tilde{\lambda}_i^2\rangle\, \mathbf{X}_k\underset{n}{\tilde{\mathbf{N}}} +$$

$$+ (\underset{m}{E} - \underset{m}{\tilde{\mathbf{N}}})\,\mathbf{X}_k\,\underset{n}{\tilde{\mathbf{M}}} + \underset{m}{\tilde{\mathbf{N}}}\,\mathbf{X}_k\langle\tilde{\lambda}_j^2\rangle\}$$

leads to the required transformation of the nodal deflection matrix. Elements of the upper left and lower right corner of matrices $\tilde{\mathbf{B}}$, $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{E}}$ are identical to those of matrices $\mathbf{C}^2$, $\mathbf{C}$, $\mathbf{E}$ and $\tilde{\tilde{\mathbf{B}}}$, $\tilde{\tilde{\mathbf{C}}}$, $\tilde{\tilde{\mathbf{E}}}$, respectively; $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{N}}$ are derived analogous to the presented matrices $\tilde{\tilde{\mathbf{M}}}$ and $\tilde{\tilde{\mathbf{N}}}$ as products of difference matrices $\tilde{\mathbf{D}}$ and $\tilde{\mathbf{F}}$ by modals of $\tilde{\mathbf{C}}$.

The increased number of the matrix products in this latter iteration formula require little excess of computation if the fact that $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{N}}$ are simple diads is made use of.

In computer application, convergence speed of iteration appears to depend both on the network and the load system, but computation of adequate accuracy required much less time, even under adverse convergence conditions, than the direct solution of a system of difference equations of the same purport. An essential advantage of the method is to require small storage capacity, improving applicability of small-capacity computers. Typically, in the computer

ODRA-1024, applied to test this method, storage capacity shortage made establishment and solution of the difference equation system of the network with 50 to 60 nodes difficult. This method poses no difficulty in determining deflections of a network with 350 to 400 nodes without using peripheral storage.

5. Generalization of applicability conditions of iteration

To prove the convergence of iteration for solving the set problem, is a constructive proof by permitting to define the applicability criterion with a rather general validity.

To this aim, two basic concepts will be defined: those of insignificantly deviating systems, and of a uniformly stiffer system. Let us consider elastic systems A and $\hat{A}$ as insignificantly deviating ones, if the generalized vector $\{x\}$ describing the equilibrium form of A under an arbitrary load can be interpreted as a system of values describing the equilibrium form of $\hat{A}$ under an adequately chosen load and vice versa.

Let the strain energy belonging to the identical equilibrium form $\{x\}$ of A and $\hat{A}$ be given by the non-negative and positive quadratic terms $1/2Q(\{x\}\ ;\ \{x\}) \geq 0$ and $1/2\ \hat{Q}(\{x\}\ ;\ \{x\}) > 0$ respectively.

Assuming both internal works to be related by the inequality

$$\frac{1}{2} Q(\{x\};\ \{x\}) \leq \frac{1}{2} \hat{Q}(\{x\};\ \{x\}) \tag{15}$$

for an arbitrary $\{x\} \neq 0$ then system $\hat{A}$ can be spoken of as uniformly stiffer than A.

Let the load on both systems be described by generalized vectors $\{p\}$ so as to give equal losses of potentials of the load system $\{p\}$ on both systems in case of equal $\{x\}$. Let the potential loss be determined by the bilinear term $P(\{x\}\ ;\ \{p\})$.

This mode of solution, in a more general wording: determination of the equilibrium form $\{w\}$ belonging to $\{p\}$ of an elastic system A by the convergent series of directly computable equilibrium forms of a substituting system $\hat{A}$ insignificantly deviating from, and uniformly stiffer than, the original system A.

According to the Castigliano theorem, criteria of the equilibrium of the two systems are:

$$\text{System } A\text{:}\quad \Pi_c(\{x\}) = \frac{1}{2} Q(\{x\}\ ;\ \{x\}) - P(\{x\}\ ;\ \{p\}) = \text{min!} \tag{16}$$

$$\text{System } \hat{A}\text{:}\quad \hat{\Pi}_c(\{x\}) = \frac{1}{2} \hat{Q}(\{x\}\ ;\ \{x\}) - P(\{x\}\ ;\ \{p\}) = \text{min!} \tag{17}$$

Uniting both functionals by using a parameter vector $\{t\}$ of the same basis as $\{x\}$:

$$\Pi(\{x\}\ ;\ \{t\}) = \frac{1}{2}\left[Q(\{x\}\ ;\ \{x\}) + \hat{Q}(\{x-t\}\ ;\ \{x-t\}) - \right.$$
$$\left. -Q\left(\{x-t\}\ ;\ \{x-t\}\right)\right] - P(\{x\}\ ;\ \{p\})\ . \tag{18}$$

Different specifications of parameter vector $\{t\}$ yields functionals $\hat{\Pi}_c$ and Π_c as follows:

$$\Pi(\{x\}\ ;\ \{0\}) = \hat{\Pi}_c(\{x\}) \tag{19}$$

$$\Pi(\{x\}\ ;\ \{x\}) = \Pi_c(\{x\}) \tag{20}$$

Inequality (15) shows that, for identical $\{p\}$:

$$\Pi_c(\{x\}) \leq \hat{\Pi}_c(\{x\})\ . \tag{21}$$

Following from the structure of Π and from the non-negativity of quadratic terms Q and $\hat{Q}$ reflecting the physical purport, for arbitrary $\{x\}$ and $\{t\}$:

$$\Pi(\{x\}\ ;\ \{t\}) \geq \Pi(\{x\}\ ;\ \{x\})\ . \tag{22}$$

Functional Π can be written in the form:

$$\Pi(\{x\}\ ;\ \{t\}) = \frac{1}{2}\hat{Q}(\{x\}\ ;\ \{x\}) - P(\{x\}\ ;\ \{p\}) - \frac{1}{2}\left[\hat{Q}(\{x;\ \{t\}) - Q(\{x\};\ \{t\})\right] +$$
$$+ \frac{1}{2}\left[\hat{Q}(\{t\}\ ;\ \{t\}) - Q(\{t\}\ ;\ \{t\})\right]\ . \tag{23}$$

In case of a fixed $\{t\}$, the fourth term in the righth-and side of (23) is a constant, irrelevant to $\{\mathbf{x}\}$ yielding an extremal value, the second and third term, united into one bilinear form, can be considered as loss of potential of a fictitious load depending on $\{x\}$.

Assuming, — in conformity with the original problem setting — the equilibrium form of A for an arbitrary load to be directly determinable, so also is $\{x\}$ yielding the minimum of (18) belonging to an arbitrary parameter vector $\{t\}$, in conformity with (23). Therefore, single steps of generalized vector iteration

$$\Pi(\{x\}\ ;\ \{0\}) = \text{minimum !} \quad \rightarrow \{w_1\} \tag{24}$$
$$\Pi(\{x\}\ ;\ \{w_k\} = \text{minimum !} \quad \rightarrow \{w_{k+1}\}$$

can directly be performed. Inequalities (12), (24) permit to state iteration to result in a monotonous decreasing series of Π, having as lower bound the mini-

mum of Π_c because of (20) and (22). Since limit equality

$$\lim_{k \to \infty} [\Pi(\{w_{k+1}\} ; \{w_k\}) - \Pi(\{w_k\} ; \{w_k\})] = 0$$

is only true if $\{w_k\}$ differs from $\{w_{k+1}\}$ at most by elements of a zero-measure set, iteration results in the deformation vector $\{w\}$ sought for, — exactly, in the case of a deformation vector of finite degrees of freedom, and "with a deviation at most in the elements of a zero-measure set" in case of infinite degrees of freedom.

Thus, a necessary condition of the iteration to be applied is that the set of elements (functions, values) describing or approximating the equilibrium form of the examined system is to describe the equilibrium form of a substituting elastic system insignificantly different from the former exact or approximate one and adequate condition of convergence is to have this substituting system uniformly stiffer than the substituted one. Of course, in case of finite degrees of freedom, the original and the substituting system should have equal degrees of freedom. If the examined system A is a substituting system itself (as e.g. in our case, because of finitization), then significant or insignificant deviation between A and $\hat{A}$ should be decided by confronting them. A clear illustration is the paradoxical application below of the computation principle:

Rectangular plates with four Navier edges or with three Navier edges and one free edge are usually treated as significantly deviating elastic systems. Applying, however, a network on the plate with simply supported edges and another on the plate with one edge free, one mesh narrower in the direction perpendicular to the free edge as seen in Figs 3a and 3b, from the aspect of matrix equations of the nodal deflections, both systems are of equal degrees of freedom, hence insignificantly deviating. Since the plate supported all around is uniformly stiffer than its counterpart with one edge free, matrix iteration formulae similar to those above can be deduced for determining nodal deflections of the free-edged plate, applying solutions for the plate with four Navier edges.

In fact, although by much more iteration steps than in the method presented before, the solution is achievable.

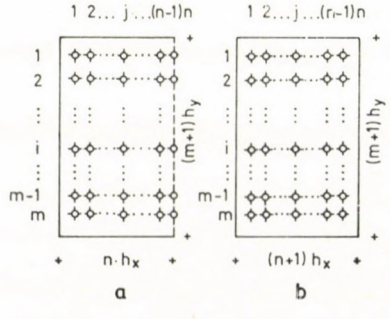


Fig. 3

REFERENCES

1. EGERVÁRY, J.: On Hypermatrices of Pairwise Exchangeable Blocks and their Applications in Grid Dynamics.* *MTA Alk. Mat. Intézetének Közleménye* **3** (1954), Budapest
2. SZABÓ, J.: Application of the Matrix Theory for the Approximate Solution of Partial Differential Equations Describing Certain Structural Problems.* *MTA Mat. Kut. Intézetének Közleményei*, **1.** No. 4 (1956), Budapest
3. RÓZSA, P.: Linear Algebra and Applications.* Műszaki Könyvkiadó, Budapest 1974
4. HEGEDŰS, I.: Matrix Algorithm Analysis of Specially Supported Reinforced Concrete Slabs.* *KAMM Füzetek* **70**, 45. (1970) Budapest

* In Hungarian

Matrix-Iteration für die Berechnung der Durchbiegung von Rechteckplatten mit freien und frei aufliegenden Rändern. — Das beschriebene maschinelle Verfahren eignet sich zur Untersuchung von Rechteckplatten mit freien und frei aufliegenden Rändern. Es beruht auf der direkten Lösung der Matrizengleichung der Knotenpunktdurchbiegungen, demgemäß ist es eine Weiterentwicklung des Verfahrens für die Berechnung ringsherum frei aufliegender Rechteckplatten, das anstatt der Annahme fiktiver Lastgruppen, durch innere Iteration beliebig genaue Ergebnisse liefert im Falle von Randbedingungen, die eine direkte Lösung scheinbar verhindern. Nach Beweis der Konvergenz der Iteration, wird aufgrund des konstruktiven Beweises eine hinreichende Bedingung für die Konvergenz ähnlicher Iterationstypen gegeben.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (1–2), pp. 163–185 (1980)

GENERAL ANALYSIS OF MONOCHROMATIC SIGNALS PROPAGATING ALONG INHOMOGENEOUS TRANSMISSION LINES. PART I

I. ÁRKOS, Mrs. FERENCZ*

[Manuscript received: September 8, 1978]

The *first part* of this paper gives a short summary of the fundamental relations and common methods of investigation of inhomogeneous transmission lines. Then using the method of inhomogeneous basic modes shows a process for determining the wave pattern along the line which seems generally to be sufficiently useful. Based on this, one may have different categories of transmission lines. Finally, by using a further method too, the way of determining the wave pattern and the different categories will be shown.

The *second part* of the paper gives the manner of determining the group velocity by using a perturbation process based on the method of inhomogeneous basic modes. In this case homogeneous, inhomogeneous and within this strictly linear and dispersive lines are analysed. Finally, an example is shown for determining the wave pattern and group velocity by using the method of inhomogeneous basic modes.

1. Introduction

Nowadays the exact analysis of electromagnetic wave propagation (signal propagation) is becoming more and more important because of the rather wide spreading of information, energy transmission and the general use of electromagnetic measuring methods. One essential, actual and not at all closed topic among these is the analysis of inhomogeneous transmission lines.

A lot of problems arise, from finding the parameters for inhomogeneous transmission lines [1] to studying the possibilities of solving the transmission lines equations [2]. No general way is known for analytically treating dispersive lines, exceeding the computerized methods for some special cases. However, the theoretical investigation of the inhomogeneous case may be very important, e.g. in case of protecting systems of railways [3, 4, 5], signal propagation [22], etc.

The intention of this paper is to give an analysis of the electromagnetic wave pattern and group velocity along the inhomogeneous linear transmission lines, adapting the methods of examining electromagnetic wave propagation in free space [1, 6, 7, 8]. The discussion will be based on the analysis of $u(x)$ and $i(x)$ that is, space and time variations of voltage and current along the line, respectively.

* Ferenczné, Árkos Ilona, H-1088, Budapest, Puskin u. 24. Hungary

The waveguides will be excluded from the discussion now, that is, only the TEM configuration will be considered, without giving a description of the electromagnetic field in detail. The discussion will be restricted to stationary, linear, motionless transmission lines, and strictly monochromatic solutions will be sought in the form of $e^{j\omega_0 t}$, where ω_0 is the angular frequency and t is the time. This choice may be reasonable where considering that the analysis of propagation of signals having general $S(\omega)$ spectra will be based just on the investigations of propagation of monochromatic signals [1, 2, 9].

The general conditions of existence of the solutions will not be dealt with in this paper. The criteria of existence must be considered separately in each practical cases when the parameters and boundary conditions are known. In the followings only the determination of the solution will be discussed, automatically assuming its existence.

It also extends beyond the limits of this paper to give a detailed analysis for the parameters of transmission lines, whether by using general [1] or special [2] methods. The well known results will be used, and when the necessity arises, the way to develop the investigations of parameters will be shown.

As was mentioned above, only linear transmission lines will be dealt with, that is, transmission lines the characters of which are independent of signal parameters (amplitude, etc.), apart from the case of dispersion, when the parameters of the line depend on frequency. In this case the linearity itself will also be examined. It is, however, obvious, that when $\omega_0 =$ constant, the dispersivity causes no problem in the determination of the wave pattern. The only difference caused by this "nonlinearity in time" under the conditions defined above appears in the determining of the group velocity.

2. The basic relations

The basic equations of the transmission lines may be obtained by using the model of Fig. 1 in the known manner.

The case, when differentiating is not defineable along the line (e.g. because of its being divided into different layers) will not be dealt with here.

As is well-known, with the commonly used assumptions for Stoke's and Gauss's theorems, the following relations hold for $\bar{E}$ (electric field strength), $\bar{B}$ (magnetic induction), and $\bar{J}$ (current density) and ϱ (space charge density):

$$\oint_l \bar{E}\, d\bar{l} = -\frac{\partial}{\partial t}\int_{A_l} \bar{B}\, d\bar{A}_l$$

and

$$\oint_A \bar{J} d\bar{A} = -\frac{\partial}{\partial t}\int_V \varrho dV\,, \tag{2.1}$$

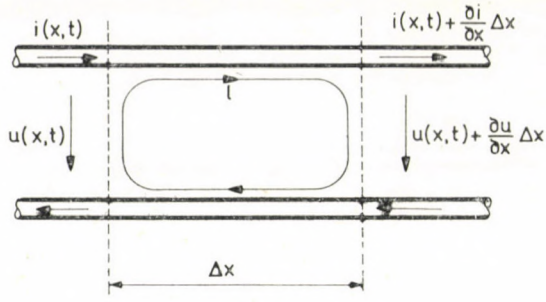


Fig. 1

respectively, where $d\bar{l}$ is the length in metres of a conductor element along the line l, dA_l is an elementary surface area in m^2 bounded by the curve l, $d\bar{A}$ is an elementary surface area of a closed surface A around the volume V.

Using these relations a possibility arises for introducing the "integrated" R, G, L and C mediumparameters. E.g., the R resistance per unit length will be definiable in that case, when using the relation $\bar{J} = \sigma . \bar{E}$ of isotropic case (σ is specific conductivity), from the integral

$$u = \int_l \bar{E} d\bar{l} = \int_l \frac{\bar{J}}{\sigma} d\bar{l} \tag{2.2}$$

one may derive

$$\Delta u = Ri\Delta x ,$$

that is, if along the cross section of the conductor the $\bar{J}/\sigma$ is integrable. If this condition cannot be satisfied, the cable cannot be investigated as a transmission line [22].

Let us assume in the followings, that R, L, C and G are defineable.

Considering the section of the line of Fig. 1., from Eqs (2.1) it follows, that

$$\begin{aligned} iR\Delta x + \frac{\partial u}{\partial x}\Delta x &= -\frac{\partial \Delta\Phi}{\partial t}, \\ uG\Delta x + \frac{\partial i}{\partial x}\Delta x &= -\frac{\partial \Delta Q}{\partial t} \end{aligned} \tag{2.3}$$

where Φ stands for the magnetic flux and Q for the charge.

As for the defineability of L and C, and the linearity of the line will be assumed, it follows, that

$$\Delta\Phi \cong Li\Delta x \quad \text{and} \quad \Delta Q \cong Cu\Delta x ,$$

so

$$-\frac{\partial u}{\partial x} = Ri + \frac{\partial}{\partial t}(Li) \tag{2.4}$$

and

$$-\frac{\partial i}{\partial x} = Gu + \frac{\partial}{\partial t}(Cu)$$

will be obtained.

If the line parameters are independent on the time t, it follows, that

$$\begin{aligned} -\frac{\partial u}{\partial x} &= Ri + L\frac{\partial i}{\partial t}, \\ -\frac{\partial i}{\partial x} &= Gu + C\frac{\partial u}{\partial t}. \end{aligned} \tag{2.5}$$

Let now "effective" parameters L_T and C_T be introduced here, using the following definitions:

$$\frac{\partial(L_T i)}{\partial t} = Ri + \frac{\partial}{\partial t}(Li) \tag{2.6}$$

and

$$\frac{\partial(C_T u)}{\partial t} = Gu + \frac{\partial}{\partial t}(Cu)\,. \tag{2.7}$$

These definitions are analogous to the way of rewriting Maxwell's equations e.g. for plasmas [21]. The conditions of this "rewriting" are given more generally in [10].

Now the transmission line equations are:

$$\begin{aligned} -\frac{\partial u}{\partial x} &= \frac{\partial(L_T i)}{\partial t}, \\ -\frac{\partial i}{\partial x} &= \frac{\partial(C_T u)}{\partial t}, \end{aligned} \tag{2.8}$$

or written even more shortly

$$\frac{\partial \bar{F}}{\partial x} = -\frac{\partial(\mathbf{Z}_T \bar{F})}{\partial t} \tag{2.9}$$

where

$$\bar{F} = \begin{pmatrix} u \\ i \end{pmatrix}, \quad \mathbf{Z}_T = \begin{pmatrix} 0 & L_T \\ C_T & 0 \end{pmatrix}.$$

3. General survey of the commonly used discussion

The discussion of transmission lines is usually based on Eqs (2.5), assuming that the solutions are of the following form:

$$u(x, t) = u(x)\, e^{j\omega_0 t} \tag{3.1}$$

and

$$i(x, t) = i(x) \cdot e^{j\omega_0 t}\,,$$

respectively.

For investigating the existence of the solutions, generally it will do to justify the existence of the dispersion equation [10, 13] and the solution of the Floquet-differential equation [11].

By using Eqs (3.1), from (Eqs 2.5) it follows, that

$$\frac{du(x)}{dx} = -(R + j\omega_0 L) \cdot i(x) = -r(x)\, i(x)$$

and (3.2)

$$\frac{di(x)}{dx} = -(G + j\omega_0 C) \cdot u(x) = -y(x) \cdot u(x)\,,$$

respectively. From these the following linear, second ordered differential equations can be obtained:

$$\frac{d^2u(x)}{dx^2} - \frac{1}{r(x)} \frac{dr(x)}{dx} \frac{du(x)}{dx} - r(x)\, y(x)\, u(x) = 0 \tag{3.3}$$

and

$$\frac{d^2(x)}{dx^2} - \frac{1}{y(x)} \frac{dy(x)}{dx} \frac{di(x)}{dx} - r(x)\, y(x)\, i(x) = 0\,, \tag{3.4}$$

respectively, the coefficients of which are constant. The solutions of these cannot generally be given in a closed form, but in each practical case, when already determined, the corresponding $u(x)$ and $i(x)$ solutions may be connected together on the base of Eqs (3.3)—(3.4).

It is also possible to determine $r(x)$ and $y(x)$ for a known wave pattern [14].

However, this manner of solving will not be suitable for drawing general conclusions, considering the properties of the propagating signal, the complete sphere of phenomena, to be expected, and the interpretation of occasional "irregular" results of the measurements.

3.1. Exponential transmission line

The transmission line equations can be easily solved and the wave pattern is known [1, 2] if $r(x)$ and $y(x)$ are exponential functions of position.

Let an ideal line be considered, that is, $R = 0$, $G = 0$, and let it be

$$L(x) = L_0 e^{-kx} \quad \text{and} \quad C(x) = C_0 e^{kx} \tag{3.5}$$

where the constants: k, L_0 and C_0 are known.

By substituting, Eqs (3.3) and (3.4), yield:

$$\frac{d^2u(x)}{dx^2} - \left[\frac{d}{dx} \ln \left(\gamma(x) \cdot Z_0(x)\right)\right] \frac{du(x)}{dx} - \gamma^2(x)\, u(x) = 0 \tag{3.6}$$

and

$$\frac{d^2i(x)}{dx^2} + \left[\frac{d}{dx} \ln \frac{Z_0(x)}{\gamma(x)}\right] \frac{di(x)}{dx} - \gamma^2(x)\, i(x) = 0\,, \tag{3.7}$$

respectively, where

$$\gamma(x) = \sqrt{r(x) \cdot y(x)}$$

and

$$Z_0(x) = \sqrt{\frac{r(x)}{y(x)}} = e^{-kx} \sqrt{\frac{L_0}{C_0}} = Z_{00}\, e^{-kx} \tag{3.8}$$

As for these differential equations being linear and of constant coefficients, the solutions [11] will be sought in the form of

$$u(x) = A_1 e^{\Gamma_1 x} + A_2 e^{\Gamma_2 x}\,.$$

Then

$$\Gamma_{1,2} = -\frac{k}{2} \pm j\omega_0 \sqrt{L_0 C_0} \sqrt{1 - \frac{k^2}{4\omega_0^2 L_0 C_0}} = -\frac{k}{2} \pm j\beta, \tag{3.9}$$

and now follows, that

$$u(x,t) = U_0^+ e^{-\frac{k}{2}x} e^{j(\omega_0 t - \beta x)} + U_0^- e^{-\frac{k}{2}x} e^{j(\omega_0 t + \beta x)} \tag{3.10}$$

and

$$i(x,t) = \frac{\frac{k}{2} + j\beta}{j\,\omega_0 L_0} U_0^+ e^{\frac{k}{2}x} e^{j(\omega_0 t - \beta x)} - \frac{-\frac{k}{2} + j\beta}{j\,\omega_0 L_0} U_0^- e^{\frac{k}{2}x} e^{j(\omega_0 t + \beta x)}, \tag{3.11}$$

respectively.

Now, the $Z_1(x) = u(x)/i(x)$ input impedance can be determined, the boundary conditions (e.g., the terminating impedance) can be taken into consideration, reflection coefficient can be defined as $u^-(x)/u^+(x)$, etc.

*

So the well-known way of solving and a known case was shown.

In the following for the sake of a more general analysis we will try to adapt a new method which was elaborated for investigating the wave propagation in inhomogeneous "free space".

4. Determination of the wave pattern along inhomogeneous transmission lines by using the method of "inhomogeneous basic modes"

4.1. A short survey of the method

The method presents a possibility of investigating electromagnetic wave propagation in inhomogeneous media [7]. It may be applied in all the (not strongly) inhomogeneous cases, when the

$$\frac{\partial^2 f}{\partial x_1 \, \partial x_2} = \frac{\partial^2 f}{\partial x_2 \, \partial x_1} \tag{4.1}$$

condition holds for all the functions playing any role here [7, 10].

The solutions of Maxwell's equations will be sought in the form of

$$\bar{F} = \bar{F}_0 e^{j(\omega_0 t - \varphi)} , \tag{4.2}$$

for monochromatic signals, for bianisotropic media, assuming the existence of the solutions [6, 7]. Further, for the sake of generality, let

$$\begin{aligned} \bar{D} &= \varepsilon \cdot \bar{E} + \varkappa \cdot \bar{H} , \\ \bar{B} &= \nu \cdot \bar{E} + \mu \cdot \bar{H} , \end{aligned} \tag{4.3}$$

where $\bar{F}_0$ stands for the complex amplitude, φ for the phase, $\bar{H}$ for the magnetic field, $\bar{D}$ for the electric displacement, the ε, $\varkappa$, ν and μ tensors for the characteristic parameters of the medium.

It can be justified [7], that the solution to be sought for has always the form of

$$\begin{aligned} \bar{F}(x) &= \sum_{n=1}^{N} [A_n(x) - jB_n(x)] \cdot [\bar{F}_{0nA}(x) + j\bar{F}_{0nB}(x)] \cdot e^{j[\omega_0 t - \varphi_n(x)]} = \\ &= \sum_{n=1}^{N} C_n(x) \, e^{-j\varphi_{cn}(x)} \, \bar{F}_{0n}(x) \, e^{j[\omega_0 t - \varphi_n(x)]} = \sum_{n=1}^{N} \bar{F}_n(x) , \end{aligned} \tag{4.4}$$

where for different values of n the different inhomogeneous basic modes

$$\bar{F}_{0n}(x) \, e^{j[\omega_0 t - \varphi_n(x)]}$$

can be obtained. It is obvious, that

$$C_n(x) = \sqrt{A_n^2(x) + B_n^2(x)} ,$$

$$\varphi_{cn}(x) = \arctan \frac{B_n(x)}{A_n(x)} ,$$

and

$$\bar{F}_{0n} = \bar{F}_{0nA} + j\bar{F}_{0nB} .$$

Introducing the notation

$$\operatorname{grad} \varphi_n = \bar{K}_n ,$$

the different inhomogeneous basic modes may be obtained from the following equations:

$$\begin{aligned} \bar{K}_n \times \bar{H}_{0n} &= -\omega_0 \varepsilon_0 (\boldsymbol{\varepsilon} \bar{E}_{0n} + \boldsymbol{\varkappa} \bar{H}_{0n}) , \\ \bar{K}_n \times \bar{E}_{0n} &= \omega_0 \mu_0 (\boldsymbol{\nu} \bar{E}_{0n} + \boldsymbol{\mu} \bar{H}_{0n}) , \end{aligned} \tag{4.5}$$

and

$$| \boldsymbol{\varepsilon}^{-1} (\mathbf{K}_n + \omega_0 \varepsilon_0 \mathbf{K}) \, \boldsymbol{\mu}^{-1} (\mathbf{K}_n - \omega_0 \mu_0 \boldsymbol{\nu}) + k_0^2 \mathbf{1} | = 0 ,$$

where

$$k_0^2 = \omega_0^2 \varepsilon_0 \mu_0 ,$$

ε_0 and μ_0 stand for the permittivity and permeability of vacuum, respectively, and

$$\mathbf{K}_n = \begin{bmatrix} 0 & -K_{n3} & K_{n2} \\ K_{n3} & 0 & -K_{n1} \\ -K_{n2} & K_{n1} & 0 \end{bmatrix} .$$

The functions in solution (4.4) which are not determined yet, — after solving Eqs (4.5) — fulfil the equations

$$\sum_{n=1}^{N} \{ \operatorname{grad} (\ln C_n - j\varphi_{cn}) \times \bar{H}_n + \nabla_{H0n} \bar{H}_n \} = 0 \tag{4.6}$$

and

$$\sum_{n=1}^{N} \{ \operatorname{grad} (\ln C_n - j\varphi_{cn}) \times \bar{E}_n + \nabla_{EOn} \bar{E}_n \} = 0 ,$$

respectively, where

$$\nabla_{F0n} = \begin{bmatrix} 0 & -\dfrac{\partial \ln F_{20n}}{\partial x_3} & \dfrac{\partial \ln F_{30n}}{\partial x_2} \\ \dfrac{\partial \ln F_{10n}}{\partial x_3} & 0 & -\dfrac{\partial \ln F_{30n}}{\partial x_1} \\ -\dfrac{\partial \ln F_{10n}}{\partial x_2} & \dfrac{\partial \ln F_{20n}}{\partial x_1} & 0 \end{bmatrix} .$$

This is the method to be applied at transmission lines, along which propagation is one-dimensional and the polarization need not be taken into consideration.

4.2. *Applying the method to transmission lines*

First of all it must be underlined that only that kind of inhomogeneous lines will be discussed, along which all the functions satisfy condition (4.1) that is

$$\frac{\partial^2 f}{\partial x_1 \, \partial x_2} = \frac{\partial^2 f}{\partial x_2 \, \partial x_1} \, .$$

If the condition for several position cannot be satisfied, the line must be divided into sections which already do for (4.1). The method probably may be generalized in this case in analogous manner to the case of wave propagation in free space [19, 20].

When the above mentioned conditions hold, the problem will be described by Eq (2.9), that is

$$\frac{\partial \bar{F}}{\partial x} = - \frac{\partial (\mathbf{Z}_T \, \bar{F})}{\partial t} \, .$$

The solutions take the form of (4.2):

$$u(x, t) = u_0(x) \; e^{j[\omega_0 t - \varphi_u(x)]}$$

and

$$i(x, t) = i_0(x) \; e^{j[\omega_0 t - \varphi_i(x)]} \, ,$$

respectively, where $\varphi_u(x)$ and $\varphi_i(x)$ are generally complex. The solution then is — see Eq. (4.4) —:

$$\bar{F} = \begin{bmatrix} u(x) \\ i(x) \end{bmatrix} = \sum_{n=1}^{N} \bar{F}_n = \sum_{n=1}^{N} [C_n(x) \; \bar{F}_{0n} \, e^{-j\varphi_{cn}(x)}] \cdot e^{j[\omega_0 t - \varphi_n(x)]} \, , \qquad (4.8)$$

and this monochromatic solution is only an approximation from a practical point of view, however, in many practical cases and in a group of theoretical investigations it may be considered as being completely sufficient.

Substituting this into Eq (2.9) and by differentiating:

$$\sum_{n=1}^{N} \left[\nabla_{F0n} - j \, \frac{d\varphi_n(x)}{dx} \cdot \mathbf{1} + j\omega_0 \, \mathbf{Z}_T \right] \cdot \bar{F}_n = 0$$

or using the notation,

$$j\omega_0 \mathbf{Z}_T = \mathbf{Z} ,$$

$$\sum_{n=1}^{N} \left[\nabla_{F0n} - j \frac{d\varphi_n(x)}{dx} \cdot \mathbf{1} + \mathbf{Z} \right] \cdot \bar{F}_n = 0 . \tag{4.9}$$

By using Eqs (4.5) let inhomogeneous basic modes be chosen in such a way, that they obey a formal variant of the (4.9) condition for homogeneous or quasi-homogeneous case, that is to

$$\left(j \frac{d\varphi_n}{dx} \cdot \mathbf{1} - \mathbf{Z} \right) \cdot \bar{F}_n = 0 . \tag{4.10}$$

The solution differs from the trivial one if the dispersion equation

$$\left| j \frac{d\varphi_n(x)}{dx} \cdot \mathbf{1} - \mathbf{Z} \right| = 0 \tag{4.11}$$

holds. After determining the different inhomogeneous basic modes the still unknown functions in the solution can be obtained by solving the following equation

$$\sum_{n=1}^{N} \nabla_{F0n} \bar{F}_n = 0 \tag{4.12}$$

where

$$\nabla_{F0n} = \begin{vmatrix} \dfrac{d \ln u_{0n}(x)}{dx} & 0 \\ 0 & \dfrac{d \ln i_{0n}(x)}{dx} \end{vmatrix} . \tag{4.13}$$

4.3. *Evolution of the method. Calculation of the homogeneous and quasi-homogeneous cases*

In the homogeneous case $\mathbf{Z}$ or $\mathbf{Z}_T$ is constant along the line. A case will be considered as quasi-homogeneous if the solution obtained in a strictly analogous way to the homogeneous case will be satisfactorily precise. In this case the application of a small correction for the amplitude may be necessary [15].

Now from Eq. (4.11) it can be obtained, that

$$\begin{vmatrix} \underline{v)} & -\omega_0 L_T \\ C_T & \dfrac{d\varphi_n(x)}{dx} \end{vmatrix} = 0$$

that is

$$\left(\frac{d\varphi_n(x)}{dx}\right)^2 - \omega_0^2 L_T C_T = 0 \,. \tag{4.14}$$

It follows, that

$$\frac{d\varphi_n(x)}{dx} = \pm\, \omega_0 \sqrt{L_T C_T} = \pm\, \Gamma_0 = \pm\, j\gamma_0 \,, \tag{4.15}$$

the number of modes: $N = 2$,

$$\varphi_0(x) = \pm\, \Gamma_0 x = \pm\, j\gamma_0 x \,,$$
$$\gamma_0 = \sqrt{(R + j\omega_0 L)\,(G + j\omega_0 C)} \,.$$

Further, $\bar{F}_n$ must fulfil (4.10) or, as the equations are not linearly independent, at least one of them

$$\pm\, j\gamma_0 u_{0n} - \omega_0 L_T\, i_{0n} = 0 \,,$$

from which follows that

$$\frac{u_{0n}(x)}{i_{0n}(x)} = \pm \sqrt{\frac{R + j\omega_0 L}{G + j\,\omega_0 C}} = \pm\, Z_0 \,. \tag{4.16}$$

By using this, Eq. (4.12) gives

$$\nabla_{F0n} = \mathbf{0} \,.$$

The solution is:

$$\bar{F} = \bar{F}_{01}\, e^{j\omega_0 t + j\gamma_0 x} + \bar{F}_{02}\, e^{j\omega_0 t - j\gamma_0 x} \,, \tag{4.17}$$

and the boundary conditions may be taken into consideration in the usual manner.

Consequently, the method in this case gives the known results.

Determining the inhomogeneous basic modes

These modes will be obtained as solution of Eqs (4.10) and (4.11). Similarly as above

$$\frac{d\varphi_n(x)}{dx} = \pm\, \omega_0 \sqrt{L_T C_T} = \pm\, j \sqrt{r(x)\, y(x)} = \pm\, j\Gamma(x) \,. \tag{4.18}$$

Integrating this we get:

$$\varphi_n(x) = \pm\, j \int \Gamma(x) dx + \varphi_{0n} \,. \tag{4.19}$$

The number of modes is $N = 2$, namely

$$\begin{aligned} n &= 1: \qquad +j\int \Gamma(x)dx\,, \\ n &= 2: \qquad -j\int \Gamma(x)dx\,. \end{aligned} \tag{4.20}$$

The inhomogeneous basic modes are:

$$\bar{F}_n = \bar{F}_{0n}\, e^{j\omega_0 t \pm \int \Gamma(x)dx - j\varphi_{0n}}\,, \qquad n = 1, 2, \tag{4.21}$$

when substituting these into Eq. (4.10) we find, that

$$u_{0n}(x) = (-1)^n \cdot Z_0(x) \cdot i_{0n}(x)\,, \tag{4.22}$$

where

$$Z_0(x) = \sqrt{\frac{r(x)}{y(x)}} = \sqrt{\frac{L_T}{C_T}}\,.$$

So the final form of (4.21) is:

$$\bar{F}_n = i_{0n}(x)\begin{bmatrix} (-1)^n \cdot Z_0(x) \\ 1 \end{bmatrix} e^{j\omega_0 t + (-1)^{n-1}\int \Gamma(x)dx - j\varphi_{0n}}\,. \tag{4.23}$$

Determining the total wave pattern

Let the (4.23) form of $\bar{F}_n$ be substituted into Eq. (4.12) and by using (4.22) t follows, that

$$\begin{aligned} &-\frac{d[Z_0(x)\, i_{01}(x)]}{dx}\, e^{j\omega_0 t + \int \Gamma(x)dx - j\varphi_{01}(x)} + \\ &+\frac{d[Z_0(x)\, i_{02}(x)]}{dx}\, e^{j\omega_0 t - \int \Gamma(x)dx - j\varphi_{02}(x)} = 0\,, \end{aligned}$$

and

$$\frac{di_{01}(x)}{dx}\, e^{j\omega_0 t + \int \Gamma(x)dx - j\varphi_{01}(x)} + \frac{di_{02}(x)}{dx}\, e^{j\omega_0 t - \int \Gamma(x)dx - j\varphi_{02}(x)} = 0\,. \tag{4.24}$$

The complete wave pattern can be obtained by solving Eqs (4.24). The further steps of the solving, however, depend on the character of $Z_0(x)$.

a) If $Z_0(x) = 0$, it follows from Eq. (4.22), that u_{0n} and so $u_n(x) = 0$ $(n = 1, 2)$, therefore

$$\frac{di_{01}}{dx}\, e^{\int \Gamma dx - j\varphi_{01}} + \frac{di_{02}}{dx}\, e^{-\int \Gamma dx - j\varphi_{02}} = 0\,,$$

that is, one current of the two may be chosen arbitrarily. Physically, that means a short-circuit.

b) $$\frac{dZ_0(x)}{dx} = 0 .$$

Now it follows from Eq. (4.24) that

$$-\frac{di_{01}(x)}{dx} e^{\int \Gamma dx - j\varphi_{01}} + \frac{di_{02}(x)}{dx} e^{-\int \Gamma dx - j\varphi_{02}} = 0$$

and (4.25)

$$\frac{di_{01}(x)}{dx} e^{\int \Gamma dx - j\varphi_{01}} + \frac{di_{02}}{dx} e^{-\int \Gamma dx - j\varphi_{02}} = 0 ,$$

and the solution of this system of equations are:

$$\begin{aligned} i_{01}(x) &= \text{constant} , \\ i_{02}(x) &= \text{constant} . \end{aligned} \tag{4.26}$$

u_{01} and u_{02} can be determined in the known manner together by taking the boundary conditions into consideration.

However, it can be seen, that not only homogeneous transmission lines but all those can be drawn into this category for which

$$Z_0(x) = k = \text{constant}$$

that is

$$\begin{aligned} L_T(x) &= k^2 C_T(x) , \\ r(x) &= k^2 y(x) , \end{aligned}$$

and making use of Eq. (4.18):

$$\Gamma(x) = k y(x) .$$

Let this category be named "*homogeneous-class*" of transmission lines. It is to be noted, that the exponential line does not belong to this group.

4.4. *Calculation of the current wave in general*

In the following let it be assumed that

$$Z_0(x) \not\equiv 0 \tag{4.27}$$

and

$$\frac{dZ_0(x)}{dx} \not\equiv 0 .$$

It follows from Eq. (4.24), that

$$\frac{dZ_0(x)}{dx}\left(-i_{01}\,e^{\int\Gamma dx-j\varphi_{01}}+i_{02}\,e^{-\int\Gamma dx-j\varphi_{02}}\right)+$$

$$+Z_0\left(-\frac{di_{01}}{dx}\,e^{\int\Gamma dx-j\varphi_{01}}+\frac{di_{02}}{dx}\,e^{-\int\Gamma dx-j\varphi_{02}}\right)=0 \tag{4.28}$$

and

$$\frac{di_{01}}{dx}\,e^{\int\Gamma dx-j\varphi_{01}}=-\frac{di_{02}}{dx}\,e^{-\int\Gamma dx-j\varphi_{02}}.$$

Now a further profit of this method may be seen: it is sufficient to investigate one of $i_{01}(x)$ or $i_{02}(x)$ and using what was told above, generalized conclusions can be drawn. E.g.:

$$\frac{di_{02}}{dx}=-\frac{di_{01}}{dx}\,e^{2\int\Gamma dx}\,e^{j(\varphi_{02}-\varphi_{01})}\,,$$

$$i_{02}(x)=-e^{j(\varphi_{02}-\varphi_{01})}\int\frac{di_{01}(x)}{dx}\,e^{2\int\Gamma dx}\,dx+C_2\,. \tag{4.29}$$

By changing the indices, a similar result can be obtained for $i_{01}(x)$. Using $i_{02}(x)$ Eq. (4.28) can be rearranged. Meanwhile, φ_{01} and φ_{02} will be lost as they are arbitrary constants determined by the boundary conditions. We obtain as a result:

$$2Z_0i'_{01}+Z'_0i_{01}+Z'_0\,e^{-2\int\Gamma dx}\left[\int i'_{01}\,e^{2\int\Gamma dx}\,dx-C_2\right]=0 \tag{4.30}$$

and

$$2Z_0\,i'_{02}+Z'_0\,i_{02}+Z'_0\,e^{2\int\Gamma dx}\left[\int i'_{02}\,e^{-2\int\Gamma dx}\,dx-C_1\right]=0\,,$$

respectively.

By differentiating this it can be written as:

$$\frac{Z'_0}{Z_0}\,i''_{01}+\left[2\left(\frac{Z'_0}{Z_0}\right)^2-\frac{Z''_0}{Z_0}+2\Gamma\frac{Z'_0}{Z_0}\right]i'_{01}+\Gamma\left(\frac{Z'_0}{Z_0}\right)^2i_{01}=0$$

and (4.31)

$$\frac{Z'_0}{Z_0}\,i''_{02}+\left[2\left(\frac{Z'_0}{Z_0}\right)^2-\frac{Z''_0}{Z_0}-2\Gamma\frac{Z'_0}{Z_0}\right]i'_{02}-\Gamma\left(\frac{Z'_0}{Z_0}\right)^2i_{02}=0\,,$$

respectively. (The solutions of these differential equation should be substituted in the individual cases into Eq. (4.28) for controlling the results made by the boundary conditions.)

Let be introduced

$$f(x) = \frac{Z_0'}{Z_0} = \frac{d}{dx} \ln Z_0 , \qquad n = 1, 2$$

$$g(x) = (-1)^{n-1} \Gamma(x) \tag{4.32}$$

and

$$\eta(x) = \begin{cases} i_{01}(x) \\ i_{02}(x) \end{cases}$$

into Eq. (4.31), so we arrive at:

$$\eta'' + \left[-\frac{f'}{f} + f + 2g \right] \eta' + (f \cdot g)\eta = 0 . \tag{4.33}$$

Choosing here [11, 12]

$$\eta(x) = \zeta(x) \cdot u(x) = \zeta(x) \cdot e^{-\int \left(g + \frac{1}{2} f - \frac{1}{2} \frac{f'}{f}\right) dx} \tag{4.34}$$

it follows that

$$\zeta'' + \left[\left(-g + \frac{f'}{2f} \right)' - \left(-g + \frac{f'}{2f} \right)^2 - \frac{f^2}{4} \right] \zeta = 0 . \tag{4.35}$$

Let further be

$$\zeta(x) = C_z \cdot e^{\int \left(g - \frac{1}{2} \frac{f'}{f} - \nu(x)\right) dx} \tag{4.36}$$

then

$$\eta(x) = C_z \cdot e^{-\int \frac{f}{2} dx} \; e^{-\int \nu(x) dx} , \tag{4.37}$$

and

$$\nu' = \nu^2 + \left(\frac{f'}{f} - 2g \right) \nu - \left(\frac{f}{2} \right)^2 . \tag{4.38}$$

Now let

$$G(x) = \frac{f'}{f} - 2g$$

$$H(x) = -\frac{f^2}{4} ,$$

one has for $\nu(x)$ the following differential equation of the Riccati-type

$$\nu' = \nu^2 + G(x) \cdot \nu + H(x) . \tag{4.39}$$

The solution of this cannot generally be given in a closed form, but for the individual problems, — perhaps by using approximative methods — it can still be determined.

In the followings several type of transmission lines will be shown for which the solution will also be given.

4.5. "*Polynomial Class*" *of transmission lines*

It is known [11] that if $G(x)$ and $H(x)$ are polynoms, Eq. (4.39) has no solution in polynomial form if the grade of the polynom

$$\Delta = g^2 - G' - 4H \tag{4.40}$$

is odd, and the solutions are the following functions:

$$v(x) = -\frac{1}{2}\left(G \pm \langle\sqrt{\Delta}\rangle\right) = -\frac{1}{2}(G + D) \tag{4.40a}$$

if Δ is even. Here $\langle\sqrt{\Delta}\rangle$ means the rational integer part of $\sqrt{\Delta}$, e.g $\sqrt{x^4 - 2x^3 + x - 6} = x^2 - x - 1/2$. Solutions can be found to both signs of $\langle\sqrt{\Delta}\rangle$ only in case of Δ being constant. In case of polynomial class, however, $\Delta \neq$ $\neq$ constant according to the definition, that is, in this case solution can be found belonging only to one of the two signs. Let for that one of $\langle\sqrt{\Delta}\rangle$'-s the symbol D be used.

Now it follows that

$$\Delta = 3\left(\frac{f'}{f}\right)^2 - 4\frac{f'}{f}g - 2\frac{f''}{f} + 4g^2 + 4g' - f^2\,. \tag{4.41}$$

The expression (4.41) will be polynomial in case of $f'/f = 0$, that is $f = K$, and

$$\frac{Z_0'}{Z_0} = K = \frac{d}{dx}(\ln Z_0)\,, \tag{4.42}$$

$$\ln Z_0 = Kx + C\,, \tag{4.43}$$

$$Z_0 = Z_{00}\, e^{Kx} = \sqrt{\frac{r(x)}{y(x)}}\,. \tag{4.44}$$

In this case g may be of polynomial form:

$$g = \pm\, \Gamma(x) = \pm\sqrt{r(x)\, y(x)} = \pm\, P(x) \tag{4.45}$$

and from this

$$G = \mp\, P(x)\,, \tag{4.46}$$

$$H = -\frac{k^2}{4}\,. \tag{4.47}$$

It is obvious that in this case Δ is always *even*.

The solution can be obtained in the following form:

$$\begin{aligned} i_{01} &= \zeta(x) \cdot u(x) = C_z\, e^{-k|x|}\; e^{-\int \nu(x)dx} = \\ &= C_z\, e^{-k|x|/2}\, e^{\frac{1}{2}\int(\pm P(x)+D)dx} = \\ &= C_z\, e^{-k|x|/2}\, e^{\pm\int P(x)dx}\, e^{+\frac{1}{2}\int Ddx}\,. \end{aligned} \tag{4.48}$$

From Eq. (4.32) it follows that

$$i_{01}(x) = e^{-k|x|/2}\left[C_{z1}\, e^{-\int P(x)dx} + C_{z2}\, e^{\int P(x)dx}\right] e^{\frac{1}{2}\int Ddx}\,. \tag{4.49}$$

The constants C_{z1} and C_{z2} are to be determined by the boundary conditions.

Further i_{02} and the inhomogeneous basic modes may be determined by Eq. (4.29) and (4.23), respectively.

In those cases when

$$\left|\frac{f'}{f}\right| \ll g \quad \text{or} \quad |Z_0''| \ll |Z_0'| \ll |Z_0|\,,$$

it seems to be a possible and expedient way of approximation to treat the transmission line as a polynomial one. It depends on the necessary accuracy of the investigations, that whether or not this approximation will be acceptable.

4.6. "*Exponential class*" *of transmission lines*

A special group of the "polynomial class" of lines can be obtained when assuming that

$$\Delta = G^2 - 2G' - 4H = \text{constant}, \tag{4.50}$$

that is

$$-2\left(\frac{f''}{f}\right) + 3\left(\frac{f'}{f}\right)^2 - 4g\left(\frac{f'}{f}\right) + 4g^2 + 4g' + f^2 = \text{const.} \tag{4.51}$$

In this case both the variant of

$$\nu(x) = -\frac{1}{2}\left[G(x) \pm \langle\sqrt{\Delta}\,\rangle\right] \tag{4.52}$$

belonging to the two signs will exist. As for the exponential line mentioned in Part 3.1 belong to this group, let this category be named "*exponential-class*".

Let the exponential transmission line itself be analyzed as an example, Chapter 7.

4.7 "*Periodical class*" *of transmission lines*

For the sake of adapting ourselves to known mathematical methods [11] let us return to Eq. (4.35), which takes the form:

$$\zeta'' + \mathcal{F}(x) \cdot \zeta = 0 \,.$$

When $\mathcal{F}(x)$ has the form:

$$\mathcal{F}(x) = \Phi(x) + \lambda \tag{4.53}$$

where λ is a constant, $\Phi(x)$ is periodical, let the line be named "periodical".

In this case we can choose among several mathematical methods [11]. Let the Floquet-solution be chosen which uses the Fourier-series of $\Phi(x)$. We shall determine a periodical or nearly periodical solution, the

$$\zeta(x + 2\pi) = e^{2\pi\mu}\zeta(x) \,, \tag{4.54}$$

where μ is complex. Let

$$\Phi(x) = \sum_{n=-\infty}^{\infty} a_n \, e^{jnx} \tag{4.55}$$

and

$$\zeta(x) = e^{\mu x} \sum_{n=-\infty}^{\infty} b_n \, e^{jnx} \,, \tag{4.56}$$

where not all of the b_n coefficients are 0; μ can be determined from the determinant of proper values, the b_n coefficients are provided by a linear system of equations.

Other methods of solving provide other groups of solutions — e.g. Hill, Strutt, Markovic, Whittaker-Watson [11] etc. — if they exist.

The periodical inhomogeneous transmission lines were more generally discussed, [1, etc.], therefore in this paper it will not be further dealt with.

The structure of the complete solution was shown under Eq. (4.34).

4.8 *Further possible groups. Summary*

When a practical problem is given, if necessary, further categories may be defined.

This categorizing may be based on Eq. (4.35), and e.g. there may be a possibility for determining "Bessel-type" groups of lines, or

$$-\mathcal{F}(x) = m^2(x) + m'(x) \tag{4.57}$$

— type lines, etc.

The categorizing may also be executed based on Eq. (4.39). Such groups may be defined, as

$$\begin{aligned} &1.\quad 1 + G(x) + H(x) \equiv 0\,, \\ &2.\quad H(x) = C_0^2 e^{\int G(x)dx}\,, \qquad H > 0 \\ &3.\quad H(x) = \psi^2 - \Phi\psi + \psi', \quad \text{where} \quad \nu(x) = \psi = \frac{\Phi - G}{2}\,, \end{aligned} \tag{4.58}$$

and the solution exists if the given

$$\begin{aligned} &\Phi \text{ exists, e.g.} \\ &\Phi = 0 \quad \text{if} \quad 4H = G^2 - 2G'\,, \\ &\Phi = G - 2\sqrt{H} \quad \text{if} \quad 2G = 4\sqrt{H} + \frac{H'}{H}\,. \end{aligned} \tag{4.59}$$

etc.

The above mentioned method makes the further defining of certain groups possible, which need not by all means belong into one of the groups above, nevertheless we can in these cases by this method also try to make analytical deductions.

The reversing of the method may give an important chance of determining the transmission line parameters to be realized, using the known solutions for a given wave pattern. However, in this case one has to take the possibility into account, that because of the lack of general solution, after realizing, besides the wanted wave pattern "parasite" wave pattern may also occur.

NB.: It is obvious [7] that the inhomogeneous basic modes themselves are not the actual propagating modes but at a certain position momentarily originating basic signals the interference of which results in a propagating signal. Their sum results in an actual propagating wave pattern, as it was shown.

5. The direct analysis of resulting wave pattern

There are methods which are suitable for the direct determination of the complete wave pattern [6, 10 15, 16,] for wave propagation in inhomogeneous media and are of very good use in many cases. As in this case polarization cannot be defined and the investigated phenomenon is one-dimensional in space, the possible sphere of use of our method may be wide spread and it seems to be expedient to show it.

5.1 *The short review of the method*

It is assumed that the electromagnetic plane wave, when propagating in space, may be sought for in the form

$$\bar{F} = \bar{F}_0\, e^{j(\omega_0 t - \varphi)}$$

where $\bar{F}_0$ may be considered to be *constant*, and $\varphi = \varphi(\bar{r})$. It will also be assumed that the time variations are excluded in accordance with the earliers.

Let the medium parameters be, for instance, as follows:

$$\epsilon,\ \mu = 1,\ \varkappa = \nu = 0\ .$$

(The discussion may be evolved for general, linear media also [6, 10].) The equations which describe the phenomena can be obtained by elementary steps:

$$\bar{\nabla} \times \bar{H} = j\omega_0 \varepsilon_0 \epsilon \bar{E}\ ,$$
$$\bar{\nabla} \times \bar{E} = -j\omega_0 \mu_0 \bar{H}\ ,$$

from which by eliminating $\bar{H}$ it follows, that:

$$\nabla \times \nabla \times \bar{E} - k_0^2 \epsilon \bar{E} = 0 \tag{5.1}$$

After developing $\bar{\nabla} \times \bar{\nabla} \times \bar{E}$, it can be seen that a nontrivial solution of (5.1) can be obtained in case of the fulfilment of

$$| \mathbf{KK} + k_0^2 \epsilon - j(\overline{\mathrm{Grad}}\ \bar{K} + \nabla \bar{K} \mathbf{1}) | = 0\ , \tag{5.2}$$

the dispersion equation. It can easily be seen [16] that in one-dimensional cases the use of the method is especially advantageous. If

$$\epsilon = \varepsilon(x) \cdot \mathbf{1}\ ,$$

the choice of

$$\bar{E}(x) = \bar{E}_0\, e^{-\int_0^x \gamma(x)dx}$$

is expedient. $\bar{H}(x)$ than can be obtained from Maxwell's equations. From Eq. (5.2) it now follows, that

$$-\frac{d\gamma}{dx} + \gamma^2 + k_0^2\, \varepsilon(x) = 0$$

where

$$\gamma = jK.$$

In the followings this procedure will be applied to transmission lines. It is new as concerning the method that the medium by transmission lines is somewhat complicated as for L_T and C_T, which are analogous to μ and ε respectively, cannot be taken for 1.

5.2 *Applying the method to transmission lines*

The solution will be sought for in a form of Eq. (3.1) or Eq. (4.7), however, is not decomposed according to (4.8), but keeping the form mentioned before:

$$u(x, t) = U_0 \, e^{-\int_0^x \gamma_u(x)dx} \, e^{j\omega_0 t} = u(x) \, e^{j\omega_0 t} \, ,$$

$$i(x, t) = J_0 \, e^{-\int_0^x \gamma_i(x)dx} e^{j\omega_0 t} = i(x) \, e^{j\omega_0 t} \, , \tag{5.3}$$

where U_0 and I_0 are constant.

The method presents equations of more useful though somewhat different form than Eqs (3.3) and (3.4) [14, 16]:

Let us start from the following equations:

$$\frac{\partial u(x, t)}{\partial x} = - L_T \frac{\partial i(x, t)}{\partial t}$$

and

$$\frac{\partial i(x, t)}{\partial x} = - C_T \frac{\partial u(x, t)}{\partial t}$$

where L_T and C_T are independent of time. From this follows that

$$\frac{\partial^2 i}{\partial x^2} = - \frac{\partial}{\partial x} \left(C_T \frac{\partial u}{\partial t} \right)$$

and (5.4)

$$\frac{\partial^2 u}{\partial x^2} = - \frac{\partial}{\partial x} \left(L_T \frac{\partial i}{\partial t} \right) ,$$

than

$$\frac{\partial^2 i}{\partial x^2} + \frac{dC_T}{dx} \frac{\partial u}{\partial t} + C_T \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial x} \right) = 0$$

and (5.5)

$$\frac{\partial^2 u}{\partial x^2} + \frac{dL_T}{dx} \frac{\partial i}{\partial t} + L_T \frac{\partial}{\partial t} \left(\frac{\partial i}{\partial x} \right) = 0 ,$$

respectively. By using Eqs (5.3) and (5.4) and by differentiating, it can be obtained that the nontrivial $i(x)$ and $u(x)$ exist if

$$- \frac{\partial \gamma_i}{\partial x} + \gamma_i \left(\gamma_t + \frac{C_T'}{C_T} \right) + \Gamma^2(x) = 0$$

and (5.6)

$$- \frac{\partial \gamma_u}{\partial x} + \gamma_u \left(\gamma_u + \frac{L_T'}{L_T} \right) + \Gamma^2(x) = 0 ,$$

respectively, where according to Eq. (4.18)

$$jT(x) = \omega_0 \sqrt{L_T C_T} .$$

Now, if

$$\frac{C_T'}{C_T} = \frac{L_T'}{L_T},$$

follows that

$$\gamma_u = \gamma_i . \tag{5.7}$$

To this group belongs the "homogeneous-class" of transmission lines. The corresponding u and i will be outlined by Eq. (5.4).

The Eqs (5.6) are Riccati-differential equations. It can be seen, that now an analysis based on the complete picture, is more difficult than the one based on the "inhomogeneous basic modes". Further, this method is usable also in case of function F fulfils condition

$$\frac{\partial^2 F(x_1, x_2)}{\partial x_1 \, \partial x_2} = \frac{\partial^2 F(x_1, x_2)}{\partial x_2 \, \partial x_1} .$$

It can be seen, that using the denotes of

$$\begin{aligned} y &= \gamma_u \quad \text{or} \quad \gamma_i , \\ f &= L_T \quad \text{or} \quad C_T , \\ g &= C_T \quad \text{or} \quad L_T , \end{aligned}$$

respectively, Eq. (5.6) may be rewritten in the form:

$$y' = y^2 + \frac{f'}{f} y + gf . \tag{5.8}$$

Formally this equation does not differ from Eq. (4.39), so the considerations on the possibility of solving, made above, hold for Eq. (5.8), too, and the categorizing of transmission lines when approximating the question in this way, may also be accepted.

It is to be noted, that e.g. by using [11], further groups — characteristic for the resultant pattern — can be outlined. So:

$$\text{a}^\circ. \quad (fg)^{1/2} = e^{\int \frac{f'}{f} dx} \left[\beta + \alpha \int fg \, e^{-\int \frac{f'}{f} dx} \, dx \right]$$

where α and β are constant. The characteristic of the group is

$$\sqrt{fg} = \beta f + \alpha \int g dx \tag{5.9}$$

and the solution is known [11]

$$\text{b}^\circ. \quad a^2 \frac{g}{f} = 1 ,$$

where a is constant and the solution is known in this case also [11].

Further, it can be seen from Eq. (5.8), that if ϵ, μ, $\varkappa$ and ν are partially or approximately totally independent, it can occur that the geometry of the line produces exact connection between f and g.

REFERENCES

1. Collin, R. E.: Grundlagen der Mikrowellentechnik; VEB Verlag Technik, Berlin 1973
2. Simonyi, K.: Theoretische Elektrotechnik; VEB Deutsches Verlag der Wissenschaften, Berlin 1968
3. Ebel, I. L.—List, F. D.: Izmerityel szaprotyivljenyija ballaszta tipa ISzB-1; Avtomatyika i Telemechanika ("Szovjaz"), No. **5.**, 3 (1967)
4. Erdős L.: Ballasztellenállás vizsgálata rádióizotópos mérési módszerrel, tekintettel az ágyazat szennyezettségére és az aljak minőségi állapotára; Vasúti Tudományos Kutató Intézet, 421/1972
5. Erdős L.: A ballasztellenállás vizsgálati módszereinek továbbfejlesztése, automatizálása, Vasúti Tudományos Kutató Intézet, OP 12/1972
6. Ferencz Cs.: Elektromágneses hullámterjedés inhomogén, lineáris közegekben; kandidátusi értekezés, MTA Könyvtár, Budapest 1970
7. Ferencz, Cs.: Electromagnetic Wave Propagation in Inhomogeneous Media; Method of Inhomogeneous Basic Modes; *Acta Techn. Hung.* **86,** (1978), 79
8. Ferencz, Cs.: Electromagnetic Wave Propagation: The Analysis of the Group Velocity; *Acta Techn. Hung.* **86** (1978), 169
9. Fodor Gy.—Simonyi K.—Vágó I.: Elméleti villamosságtan példatár; Tankönyvkiadó, Budapest 1967
10. Ferencz, Cs.: Electromagnetic Wave Propagation in Inhomogeneous Media: Strong and Weak Inhomogeneities; *Acta Techn. Hung.* **85** (1977), 433
11. Kamke, E.: Differentialgleichungen, Lösungsmethoden und Lösungen; Akademische Verlagsgesellschaft Geest & Portig K.-G., Leipzig 1951
12. Baule, B.: Die Mathematik des Naturforschers und Ingenieurs, Band IV. Gewöhnliche Differentialgleichungen; S. Hirzel Verlag, Leipzig 1953
13. Arnaud, J. A.—Saleh, A. A. M.: Theorems for Bianisotropic Media; *Proc. IEEE*, **60,** (1972) 639
14. Ralston, A.: Bevezetés a numerikus analízisbe; Műszaki Könyvkiadó, Budapest 1969
15. Budden, K. G.: Radio Waves in the Ionosphere; Cambridge at the University Press, 1966
16. Heald, M. A.—Wahrton, C. B.: Plasma Diagnostics with Microwaves; John Wiley & Sons, New York 1965
17. Arkos I.: Elektronfizika, Jegyzet, Budapest 1967
18. Allis, W. P.—Buchsbaum, S. J.—Bers, A.: Waves in Anisotropic Plasma; M. I. T. Press, Cambridge, Mass. 1963
19. Idemen, M.: The Maxwell's Equation in the Sense of Distributions; *IEEE Trans. Ant. and Prop.*, **AP-21,** (1973), 736
20. Ferencz, Cs.: Electromagnetic Wave Propagation in Inhomogeneous Media: The Analysis of the Rotation of Polarization and the Application of the Principle of Modified Ray Tracing; *Acta Techn. Hung.* **86** (1978), 363 and **87** (1978), 263
21. Simonyi, K.: Foundations of Electrical Engineering; Pergamon Press, 1963
22. Reiter, Gy.: Generalized Telegraphist's Equations for Waveguide of Varying Cross-sections. *Proc. of the Inst. of the Electrical Engineers*, Part B. Supplement, Convention on long distance transmission by waveguide. (1959) January, pp 54—57

Allgemeine Analyse von monochromatischen Signalen bei Fortpflanzung entlang inhomogener Übertragungsleitungen, Teil I. — Der erste Teil der Arbeit bringt eine kurze Übersicht der grundlegenden Beziehungen und der üblichen Untersuchungsmethoden von inhomogenen Übertragungsleitungen. Dann wird unter Verwendung der Methode der inhomogenen grundlegenden Moden ein Verfahren für Bestimmung der Wellenform entlang der Leitung gezeigt, welches allgemein als genügend nützlich erscheint. Hierauf basiert eine Kategorisierung der Übertragungsleitungen. Schließlich wird mit Hilfe einer weiteren Methode der Weg zur Bestimmung der Wellenformen und die verschiedenen Kategorien gezeigt.

INDEX

Printed in Hungary

A kiadásért felel az Akadémiai Kiadó igazgatója — Műszaki szerkesztő: Zacsik Annamária

A kézirat nyomdába érkezett: 1980. III. 6. — Terjedelem: 16,45 (A/5) ív, 63 ábra

81.8023 Akadémiai Nyomda, Budapest — Felelős vezető: Bernát György

Acta Techn. Hung. **90** (1980), pp. 3—14

BARTA, J.: *Theorems on the Minimum of Strain Energy in Elastostatics*

External forces are acting on an elastic body. Certain conditions are prescribed for their points of application, magnitudes, directions or for the displacements caused by these. Under the prescribed conditions. one raises the question; what are the characteristics of the active external forces, if the strain energy should be minimum. The theorems which are formulated and proved in this paper, yield contributions to answering the raised question. A numerical example explains the application of the theorems.

Acta Techn. Hung. **90** (1980), pp. 15—27

S. PRAKASH-HORVÁTH, Z.: *Effect of Titanium Oxide Content of the Bauxite Change on the Settling Properties of the Redmud*

The settling characteristic of the redmud is effected by the type of impurity minerals in the bauxite and their particle size, crystal structure and mode of distribution. Titanium oxide content of the bauxite also efects the settling property. The settling ability of the redmud is decreased with increase in titanium oxide content of the bauxite charge. Addition of settling agent (flour) does not significantly improve the settling ability of the redmuds in laboratory scale investigation.

Acta Techn. Hung. **90** (1980), pp. 29—50

FERENCZ, CS.: *Electromagnetic Wave Propagation in Moving Media with Special Regard to Frequency-Shifts in Astronomy („Anomalous" Frequency-Shifts in Astronomy)*

In the first part after a comprehensive review of problems which appear to be fundamental in the study of electromagnetic wave propagation in moving media it is shown how to eliminate the oserved discrepancy and the possible ways of analyzing the electromagnetic wave propagation in moving media are discussed. One of the discussed methods, developed under the name of relativistic ray tracing, is described in detail in the second part. It is shown how to determine the signal spectrum and the so-called basic effect for which expressions formulated in cases thought to be most important. The weighting functions and the proofs of their form are given.

Finally, in the third part results of earlier similar analyses are utilized for explaining „anomalous" frequency shifts and other associated wave propagation phenomena (line broadening, polarisation rotation, etc.) observed in the solar system. It is suggested that similar, apparently anomalous frequency shifts as well as other associated phenomena observed in galactic and extragalactic wave propagation could be interpreted in the same way.

Acta Techn. Hung. **90** (1980), pp. 59—85

ÖLLŐS, G.: *Hydraulic Bases of the Ground-Water Lowering with Vacuum Wells*

The earlier methods of treating the vacuum-well hydraulics were based mainly on the aspect of the gravitational well hydraulics. The lowering of the ground-water level as a physical process itself and the boundary conditions have not yet been cleared up. The author brought the gravitational and vacuum-well hydraulics to a common theoretical basis, and continued to develop the interpretation of the boundary conditions and physical processes. He established a new general formula for the yielding capacity of the filter galery which also accounts for the water conveyance of the capillary fringe, and from which also the yielding capacity of the gravitational filter galery, the single vacuum and gravitational well may be derived. Also the formula for the approximate calculation of the curve separating the zones of the air and water streams has been established.

Acta Techn. Hung. **90** (1980), pp. 87—90

ECSEDI, I.: *A Note Connection with the Deformation Energy*

This paper deals with bodies made of linearly elastic material. It verifies an niequality relation in connection with the deformation energy, by applying the usual assumption of elastostatics.

Acta Techn. Hung. **90** (1980), pp. 91—108

ZALKA, K.: *Torsional Buckling of a Cantilever Subjected to Distributed Normal Loads*

The paper deals with the spatial stability analysis of a bar with a built-in lower and a free upper end, subjected to distributed normal loads along its axis. The simultaneous differential equations of combined torsional and flexural buckling are set up. The exact and approximate computation of the critical load of a bar with double-symmetric cross section is dealt with in detail. Closed formulae are given for the approximate analysis of the critical load of a bar with mono-symmetric cross section.

Acta Techn. Hung. **90** (1980), pp. 109—132

STÉPÁN, G.: *Stability Investigation of Retarded Differential Equations*

This paper presents and proves a proposition on the basis of which an algorithm may be established for the stability investigation of homogeneous, linear, retarded differential equations of constant coefficiens. This algorithm may be used for the retarded differential equations in the same way as the Routh-Hurwitz criterion for the ordinary differential equations. In the very same way as the theory of the retarded differential equations involves as a special case the theory of the ordinary differential equations, also the algorithm may be utilized in lieu of the Routh-Hurwitz criterion for the ordinary differential equations. Finally, it shows thee practical applicability of the results for the case of the reproductive machine tool vibrations.

Acta Techn. Hung. **90** (1980), pp. 133—150

HERING, J.: *Generalization and Some Applications of the Concept of the Ideal Constant*

Author describes the elassic explanation of ideal constraint, and points out that this explanation can only be applied to geometric constrains. Therefore, for the definition of the ideal constraint a new explanation;is also given including the elasic explanation, however, that it may also be applie to kinemaic constraints. Based on this definition of the ideal constraint, minimum principles may be established leading to the most general principle, i. e., to Gauss's principle defining the motion of mechanical systems. It is reported how Appell's equation of motion for anholonomic systems. Language's equation of motion for holonomic systems and Euler's equation of motion for rigid bodies may be determined on the base of the principles;exposed above by the author.

Acta Techn. Hung. **90** (1980), pp. 151—162

HEGEDŰS, I.: *Matrix Iteration Analysis of Rectangular Plates with Simply Supported and Free Edges*

An efficient computer method is presented for the analysis of elastic deflections of rectangular plates with free and simply supported edges. It is based on the direct solution of the matrix equation of nodal deflections, hence it is a development of the method for all around simply supported rectangular plates where, rather than assuming fictitious load complexes, internal iteration is applied to yield results of arbitrary accuracy for a boundary condition seemingly preventing direct solution. Convergency of iteration is demonstrated , and on the basis of the constructive demonstration, a;condition sufficient for the convergency of similar types of iteration is given.

Acta Techn. Hung. **90** (1980), pp 163—185

Mrs. Ákos, I.: *General Analysis of Monochromatic Signals Propagating along Inhomogeneous Transmission Lines, Part I.*

The *first part* of this paper gives a short summary of the fundamental relations and common methods of investigation of inhomogeneous transmission lines. Then using the method of inhomogeneous basic modes shows a process for determining the wave pattern along the line which seems generally to be sufficiently useful. Based on this, one may have different categories of transmission lines. Finally, by using a further method too, the way of determining the wave pattern and the different categories will he shown.

The *second part* of the paper gives the manner of determining the group velocity by using a perturbation process based on the method of inhomogeneous basic modes. In this case homogeneous and within this strictly linear and dispersive lines are analysed. Finally, an example is shown for determining the wave pattern and group velocity by using the method of inhomogeneous basic modes.

The *Acta Technica* publish papers on technical subjects in English, French, German and Russian.

The *Acta Technica* appear in parts of varying size, making up one volume.

Manuscripts should be addressed to

Acta Technica
H-1051 Budapest
Münnich Ferenc u. 7.
Hungary

Correspondence with the editors and publishers should be sent to the same address

Orders may be placed with "Kultura" Foreign Trading Company (H-1389 Budapest 62, P.O.B. 149. Account No. 218-10990) or its representatives abroad.

Les *Acta Technica* paraissent en français, allemand, anglais et russe et publient des travaux du domaine des sciences techniques.

Les *Acta Technica* sont publiés sous forme de fascicules qui seront réunis en volumes.

On est prié d'envoyer les manuscrits destinés à la rédaction à l'adresse suivante:

Acta Technica
H-1051 Budapest
Münnich Ferenc u. 7.
Hongrie

Toute correspondance doit être envoyée à cette même adresse.

On peut s'abonner à l'Entreprise du Commerce Extérieur «Kultura» (H-1389 Budapest 62, P.O.B. 149. Compte courant No. 218-10990) ou chez représentants à l'étranger.

«*Acta Technica*» публикуют трактаты из области технических наук на русском, немецком, английском и французском языках.

«*Acta Technica*» выходят отдельными выпусками разного объема. Несколько выпусков составляют один том.

Предназначенные для публикации рускописи следует направлять по адресу:

Acta Technica
H-1051 Budapest
Münnich Ferenc u. 7.
Венгрия

По этому же адресу направлять всякую корреспонденцию для редакции и администрации.

Заказы принимает предприятие по внешней торговле «Kultura» (H-1389 Budapest 62, P.O.B. 149. Текущий счет № 218-10990) или его заграничные представительства и уполномоченные.

Reviews of the Hungarian Academy of Sciences are obtainable at the following addresses:

AUSTRALIA
C.B.D. LIBRARY AND SUBSCRIPTION SERVICE, Box 4886, G.P.O., *Sydney N.S.W. 2001*
COSMOS BOOKSHOP, 145 Ackland Street, *St. Kilda (Melbourne), Victoria 3182*

AUSTRIA
GLOBUS, Höchstädtplatz 3, *1200 Wien XX*

BELGIUM
OFFICE INTERNATIONAL DE LIBRAIRIE, 30 Avenue Marnix, *1050 Bruxelles*
LIBRAIRIE DU MONDE ENTIER, 162 Rue du Midi, *1000 Bruxelles*

BULGARIA
HEMUS, Bulvar Ruszki 6, *Sofia*

CANADA
PANNONIA BOOKS, P.O. Box 1017, Postal Station "B", *Toronto, Ontario M5T 2T8*

CHINA
CNPICOR, Periodical Department, P.O. Box 50, *Peking*

CZECHOSLOVAKIA
MAD'ARSKÁ KULTURA, Národní třída 22, *115 66 Praha*
PNS DOVOZ TISKU, Vinohradská 46, *Praha 2*
PNS DOVOZ TLAČE, *Bratislava 2*

DENMARK
EJNAR MUNKSGAARD, Norregade 6, *1165 Copenhagen*

FINLAND
AKATEEMINEN KIRJAKAUPPA, P.O. Box 128, *SF-00101 Helsinki 10*

FRANCE
EUROPÉRIODIQUES S.- A., 31 Avenue de Versailles, *78170 La Celle St.-Cloud*
LIBRAIRIE LAVOISIER, 11 rue Lavoisier, *75008 Paris*
OFFICE INTERNATIONAL DE DOCUMENTATION ET LIBRAIRIE, 38 rue Gay Lussac, *75240 Paris Cedex 05*

GERMAN DEMOCRATIC REPUBLIC
HAUS DER UNGARISCHEN KULTUR, Karl-Liebknecht-Strasse 9, *DDR-102 Berlin*
DEUTSCHE POST ZEITUNGSVERTRIEBSAMT, Strasse der Pariser Kommüne 3—4, *DDR-104 Berlin*

GERMAN FEDERAL REPUBLIC
KUNST UND WISSEN ERICH BIEBER, Postfach 46, *7000 Stuttgart 1*

GREAT BRITAIN
BLACKWELL'S PERIODICALS DIVISION, Hythe Bridge Street, *Oxford OX1 2ET*
BUMPUS, HALDANE AND MAXWELL LTD., Cowper Works, *Olney, Bucks MK46 4BN*
COLLET'S HOLDINGS LTD., Denington Estate, *Wellingborough, Northants NN8 2QT*
W.M. DAWSON AND SONS LTD., Cannon House, *Folkestone, Kent CT19 5EE*
H. K. LEWIS AND CO., 136 Gower Street, *London WC1E 6BS*

GREECE
KOSTARAKIS BROTHERS, International Booksellers, 2 Hippokratous Street, *Athens-143*

HOLLAND
MEULENHOFF-BRUNA B.V., Beulingstraat 2, *Amsterdam*
MARTINUS NIJHOFF B.V., Lange Voorhout 9—11, *Den Haag*
SWETS SUBSCRIPTION SERVICE, 347b Heereweg, *Lisse*

INDIA
ALLIED PUBLISHING PRIVATE LTD, 13/14 Asaf Ali Road, *New Delhi 110001*
150 B-6 Mount Road, *Madras 600002*
INTERNATIONAL BOOK HOUSE PVT. LTD., Madame Cama Road, *Bombay 400039*
THE STATE TRADING CORPORATION OF INDIA LTD., Books Import Division, Chandralok, 36 Janpath, *New Delhi 110001*

ITALY
EUGENIO CARLUCCI, P.O. Box 252. *70100 Bari*
INTERSCIENTIA, Via Mazzé 28, *10149 Torino*
LIBRERIA COMMISSIONARIA SANSONI, Via Lamarmora 45, *50121 Firenze*
SANTO VANASIA, Via M. Macchi 58, *20124 Milano*
D. E. A., Via Lima 28, *00198 Roma*

JAPAN
KINOKUNIYA BOOK-STORE CO. LTD., 17-7 Shinjuku-ku 3 chome, Shinjuku-ku, *Tokyo 160-91*
MARUZEN COMPANY LTD., Book Dpartment, P.O. Box 5056 Tokyo International, *Tokyo 100-31*
NAUKA LTD. IMPORT DEPARTMENT, 2-30-11 Minami Ikebukuro, Toshima-ku, *Tokyo 171*

KOREA
CHULPANMUL, *Phenjan*

NORWAY
TANUM-CAMMERMEYER, Karl Johansgatan 41—43, *1000 Oslo*

POLAND
WĘGIERSKI INSTYTUT KULTURY, Marszałkowska 80, *Warszawa*
CKP I W ul. Towarowa 28 *00-958 Warszawa*

ROMANIA
D. E. P., *Bucureşti*
ROMLIBRI, Str. Biserica Amzei 7, *Bucureşti*

SOVIET UNION
SOJUZPETCHATJ — IMPORT, *Moscow*
and the post offices in each town
MEZHDUNARODNAYA KNIGA, *Moscow G-200*

SPAIN
DIAZ DE SANTOS, Lagasca 95, *Madrid 6*

SWEDEN
ALMQVIST AND WIKSELL, Gamla Brogatan 26, *101 20 Stockholm*
GUMPERTS UNIVERSITETSBOKHANDEL AB, Box 346, *401 25 Göteborg 1*

SWITZERLAND
KARGER LIBRI AG, Petersgraben 31, *4011 Basel*

USA
EBSCO SUBSCRIPTION SERVICES, P.O. Box 1943, *Birmingham, Alabama 35201*
F. W. FAXON COMPANY, INC., 15 Southwest Park, *Westwood, Mass. 02090*
THE MOORE-COTTRELL SUBSCRIPTION AGENCIES, *North Cohocton, N. Y. 14868*
READ-MORE PUBLICATIONS, INC., 140 Cedar Street, *New York, N. Y. 10006*
STECHERT-MACMILLAN, INC., 7250 Westfield Avenue, *Pennsauken, N. J. 080110*

VIETNAM
XUNHASABA, 32, Hai 3a Trung, *Hanoi*

YUGOSLAVIA
JUGOSLAVENSKA KNJIGA, Terazije 27, *Beograd*
FORUM, Vojvode Mišiča 1, *21000 Novi Sad*

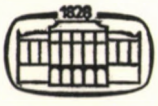
20. III. 1981. HU ISSN 0001—5407 **Index: 26.007**

ACTA TECHNICA

ACADEMIAE SCIENTIARUM HUNGARICAE

REDIGIT: M. MAJOR

TOMUS 90
FASCICULI 3—4



AKADÉMIAI KIADÓ, BUDAPEST 1980

ACTA TECHNICA

Az *Acta Technica* angol, francia, német és orosz nyelven közöl értekezéseket a műszaki tudományok köréből.

Az *Acta Technica* változó terjedelmű füzetekben jelenik meg, több füzet alkot egy kötetet.

A közlésre szánt kéziratok a következő címre küldendők:

Acta Technica
1051 Budapest, Münnich Ferenc u. 7.

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Megrendelhető a belföld számára az „Akadémiai Kiadó"-nál (1363 Budapest Pf. 24. Bankszámla 215-11448), a külföld számára pedig a „Kultura" Külkereskedelmi Vállalatnál (1389 Budapest 62, P.O.B. 149 Bankszámla: 218-10990) vagy annak külföldi képviseleteinél és bizományosainál.

Die *Acta Technica* veröffentlichen Abhandlungen aus dem Bereiche der technischen Wissenschaften in deutscher, englischer, französischer und russischer Sprache.

Die *Acta Technica* erscheinen in Heften wechselnden Umfanges. Vier Hefte bilden einen Band.

Die zur Veröffentlichung bestimmten Manuskripte sind an folgende Adresse zu senden:

Acta Technica
H-1051 Budapest
Münnich Ferenc u. 7.
Ungarn

An die gleiche Anschrift ist auch jede für die Schriftleitung und den Verlag bestimmte Korrespondenz zu richten.

Bestellbar bei »Kultura« Außenhandelsunternehmen (H-1389 Budapest 62, P.O.B. 149, Bankkonto Nr. 218-10990) oder seinen Auslandsvertretungen.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (3—4), pp. 187—199 (1980)

STATIC PENETRATION TEST RESULTS WITH SOILS HAVING SLIGHT OR MEDIUM COHESION

Á. KÉZDI*
MEMBER OF THE HUNG. AC. OF SCI.
and
Zb. MLYNAREK**

[Manuscript received January 10, 1980]

The paper analyses the influence of the phase composition: (volume percentages of solid, air and water) on the cone resistance during static penetration test. The analysis is based on data obtained in investigations using the "boulder loam" in Great Poland Lowland ("Konin clay" and "Rudnicze clay") and loess from Törökbálint, Hungary. The static souding test was carried out using the laboratory device. The function which furnishes the relationship between cone resistance and phase composition was determined numerically. The analysis of the cohesion and angle on internal friction effect has been done using statistical methods.

Introduction

Static penetration is a well known method of soil investigation. Test data collected in several countries helped to find some relationships. Based on these it is possible to estimate approximately the condition and type of soil and its mechanical properties.

On the other hand, — as the European Symposium on Penetration Testing in Stockholm in 1974 has shown — the effect of several factors on cone resistance are still unknown.

Relationships between cone resistance and strength parameters are usually expressed by empirical formulae for two different soils: cohesionless sands and clays.

Relationships between cone resistance and consistency index (for cohesive soils) — *viz.* relative density (for cohesionless soils) are also in general use.

However, it can be very misleading in soils having slight or medium cohesion to determine cone resistance on the basic of soil consistency. We often observed that the value of cone resistance remained constant in spite of the different moisture content of the soil. KÉZDI (1974) drew attention to the fact that knowledge of one phase in the soil volume may be insufficient for interpretation of soil behaviour and that the whole phase composition was important. This paper intends to investigate the effect of phase composition on static cone resistance.

* Prof. Dr. KÉZDI, Á, H-1012 Budapest, Logodi u. 9. Hungary
** Dr. hab. MLYNAREK, ZB., Institute of Water Engineering, Poznań, Poland

1. Theoretical assumptions

BANACH's equation (1950) describes the resistance of a material point penetrating in to a frictional medium. Based on this equation MLYNAREK (1970, 1979) gave a general relationship determining cone resistance in cohesive soils in the following form:

$$F_1(q_c, d, h, H, w, \gamma_d, \sigma_0, v_s, c, \Phi, \delta, \varkappa_0, E_q) = 0\,. \tag{1}$$

The structure of the soil, the mineralogical composition of the soil and the grain size distribution of the soil have also an influence.

Here: q_c — cone resistance per unit area (kN/m²)
d — cone diameter (m)
h — cone height (m)
H — depth of penetration (m)
w — moisture content (l)
γ_d — effective dry unit weight of soil (kN/m³)
σ_0 — overburden pressure
v_s — speed of penetration (m/min)
c — cohesion (kN/m²)
Φ — angle of internal friction
δ — angle of friction between cone material and the soil
$\varkappa_0$ — coefficient depending on the degree of consolidation
E_q — modulus of deformation of the grains of the soil (kN/m²)
F — area of the base of the cone (m²)
U — circumference of the base of the cone (m).

Using a dimensional matrix and substituting $R = F/U$ one can obtain the following structural formula:

$$\frac{q_c}{\gamma_d \cdot d} = F_2\left(\frac{d}{h}, \frac{H}{d}, \frac{\sigma_0}{c}\frac{\delta}{\Phi}, \frac{\gamma_d R}{E_q}, w, v_s, \varkappa_0\right) \tag{2}$$

Eq. (2) is analogous to DE BEER's equation (1974) for cohesionless soils. It is evident from Eq. (2) that the function describing cone resistance cannot be written in a dimensionally homogeneous form. Dimensional products d/h, and H/d determine respectively: geometrical similarity and dimensionless depth (DE BEER, 1964; LANGHAAR, 1964, LUNDGREEN, 1957). In order to examine the influence of the phase composition on cone resistance the relative numbers proposed by KÉZDI (1969) will be used.

The volume percentage of solids is represented by a relationship:

$$s = \frac{\gamma_d}{\gamma_s} 100\ (\%)\,, \tag{3}$$

the relative volume percentage of water:

$$v = \frac{w \cdot \gamma_d}{\gamma_w} 100\ (\%) \tag{4}$$

where: γ_s — specific gravity of solid particles.

The relative volume percentage of gas is given by

$$l = 100 - (s\% + v\%) \ (\%). \tag{5}$$

To be able to find out how the changes in the soil consistency affect cone resistance, it is necessary to examine the following partial functions:

$$\frac{q_c}{\gamma_d \cdot d} = f_1(s, v) \text{ at } \frac{d}{h} = \text{const}, \frac{H}{d} = \text{const}, v_s = \text{const}, \tag{6}$$

and

$$\frac{q_c}{\gamma_d \cdot d} = f_2\left(\frac{H}{d}\right) \text{ at } s = \text{const}, \frac{d}{h} = \text{const}, v_s = \text{const}, v = \text{const}, \frac{\delta}{\Phi} = \text{const}, \frac{\sigma_0}{c} = \text{const}. \tag{7}$$

The easiest way to determine the above relationships is a model test on known mineralogical type of soil having an identical structure and varied graining.

From the above analysis may therefore be concluded that — as in the case of cohesionless soils — it is not possible to obtain an absolute relationship between cone resistance and the index describing the state of the soil: consistency index or relative density.

It can be seen from Eq. (2) that the bearing capacity equation used to calculate cone resistance (MITCHELL—DURGUNOGLU, 1975; SANGLERAT, 1969), i.e.

$$\frac{q_c}{\gamma_d \cdot d} = \frac{c}{\gamma_d \cdot d} N_c \cdot \xi_c + N_\gamma q \, \xi q_\gamma \tag{8}$$

is a special case of Eq. (2).

2. Proposed interpretation of the static penetration diagram

To explain the influence of the phase composition on cone resistance, model tests were made with three different soils, representing characteristic deposits from the territories of Poland and Hungary, namely: sandy clay (loam) from *Konin* (2) loam from *Rudnicze* (3) — Poland and loess from *Törökbálint* (1) — Hungary.

Loesses genetically belong to aeolian soils. They belong usually to the group of so-called transition soils (see e.g. KÉZDI, 1979), since their Atterberg limits are as follows: $w_p \simeq 20$ percent, $w_L \simeq 25 \div 28$ percent. Investigation of the mineralogical content helped to establish that loess usually contains 50 to 70 percent quartz, $10 \div 25$ percent feldspar, $10 \div 30$ percent calcite. Typical clay minerals (illite, kaolinite) occur in small quantities.

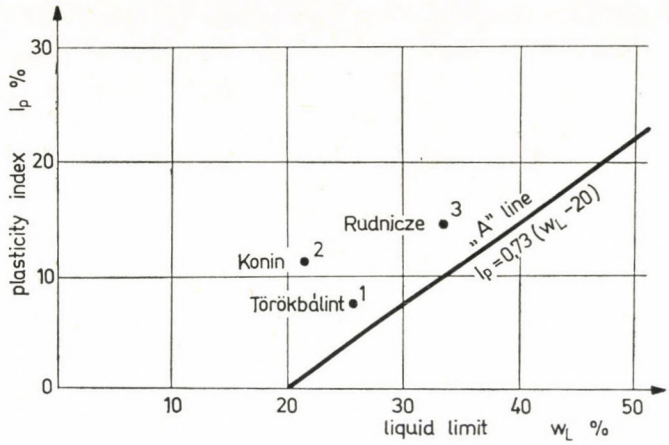


Fig. 1. Plasticity chart

Clay from *Konin* is a boulder loam type of Central Polish glaciation (Riss). ($I_p = 20$ percent, $w_L = 34$ percent.) Loam from *Rudnicze* represents a similar mineralogical type of clay fraction, i.e. illite with kaolinite admixture. In Fig. 1 the points representing the examined soils are shown on the plasticity chart.

Cone resistance was measured in the laboratory by an apparatus (made by *Leonard Farnell Ltd.*) shown in Fig. 2. The cone penetration and the cone resistance were measured with a Kyowa tensometric bridge and a Videoton registrator X-Y or Sefram registrator. A cone with an apex angle of 60° and base surface $F = 2{,}0$ cm^2 penetrated at a speed of 1,0 mm/min. To obtain a required phase composition, samples were compacted dynamically with 45 to 60 mkN/m^3 energy.

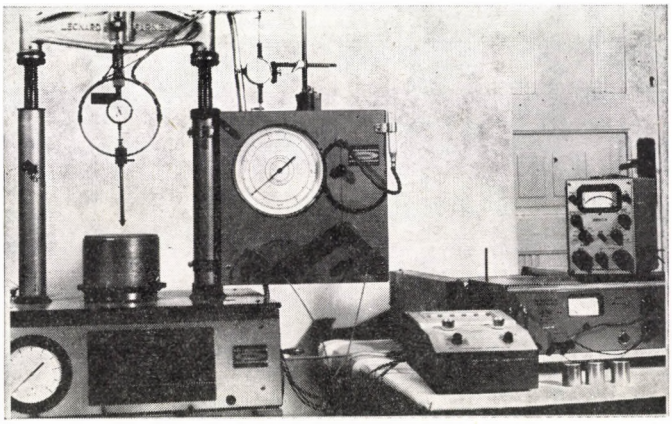


Fig. 2. Sounding device

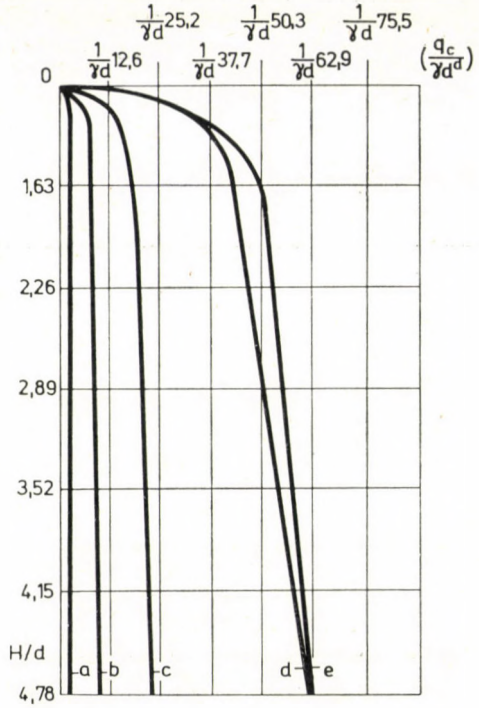


Fig. 3. Static penetration curves (loess Törökbálint) for samples of different phase composition: *a)* s = 63%, v = 32%; *b)* s = 66%, v = 29%; *c)* s = 67%, v = 26%; *d)* s = 64%, v = = 19%; *e)* s = 64%, v = 25%

Figure 3 represents a relationship between cone resistance and related depth (H/d) for *Törökbálint* loess. Curves illustrate changes in cone resistance for the samples for which values of s, v, l phase composition parameters were close to constants. The shape of the curves indicates that the resistance of the cone increases with depth of penetration. The effect of the depth on cone resistance is illustrated by a histogram from the X-Y registrator (Fig. 4).

The histograms reveal three interesting facts:

(a) initial depth, which conforms with the generally accepted definition of settlement for a shallow foundation, depends in case of cone load on the phase composition and consequently on the value of cone resistance;

(b) histograms reveal a sudden and pronounced decrease of cone resistance when the above mentioned initial depth has been reached (point A in Fig. 4). This has been particularly striking at cone resistance values above 1960 kN/m^2. The sudden drop of cone resistance is an evidence of dilatation which occurs after the first limiting state has been reached. This phenomenon was also observed in model tests made by DURGUNOGLU and MITCHELL (1974);

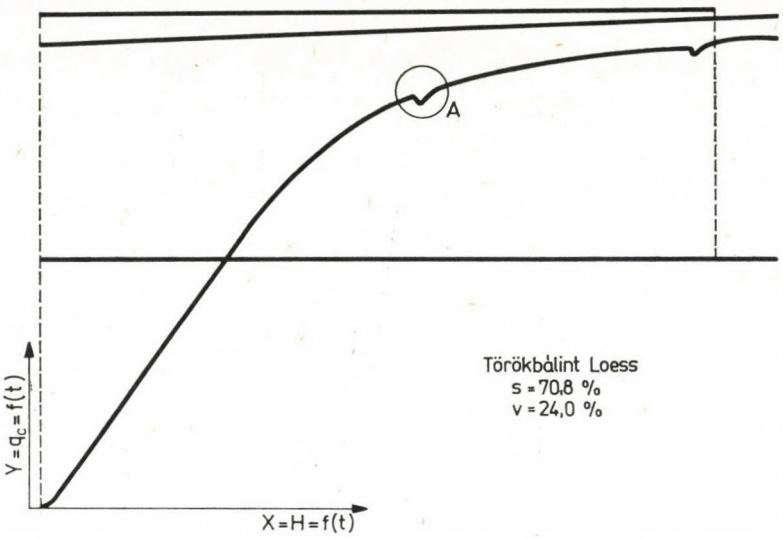


Fig. 4. Cone resistance versus depth (histogram X—Y registrator, loess Törökbálint)

(c) the shape of penetration curves clearly shows the importance of the phase composition of the soil sample.

Figure 5 shows the cone resistance as a function of the phase composition. The diagrams for the dimensionless depth $H/d = 2{,}89$ very well illustrate that for an estimation of the cone resistance it is indispensable to know the phase composition of the soils. For lines $s = \text{const}$ and $v = \text{const}$ we obtain different values of cone resistance. On the contrary, the plasticity index is of limited importance in this range. This phenomenon was earlier observed by JUMIKIS (1965), however, he did not give an explanation.

It is also evident from Fig. 5 that the effect of s and v on cone resistance depends on the granulometric and mineralogical composition of the soil. For *Törökbálint* loess the lines for equal cone resistance values run close to each other, while for *Konin* loam they are wider apart. An explanation of this phenomenon is the different grain size distribution of the respective soils.

Clay content of *Konin* loam was 16 percent, clay and silt fraction together 42 percent and for *Törökbálint* loess as much as 73 percent. It may therefore be assumed that for soils having a clay content of 20 percent or more, the changes in cone resistance depend not only on the clay fraction but more importantly on the joint content of clay and silt. KISIEL (1975) observes that an increased amount of silt and sand fraction brings about an additional increase of dry friction. It can therefore be expected — in the light of those observations — that in the case of loess the cone resistance is higher than for loams, for an analogous range of s and v values. A detailed explanation of the

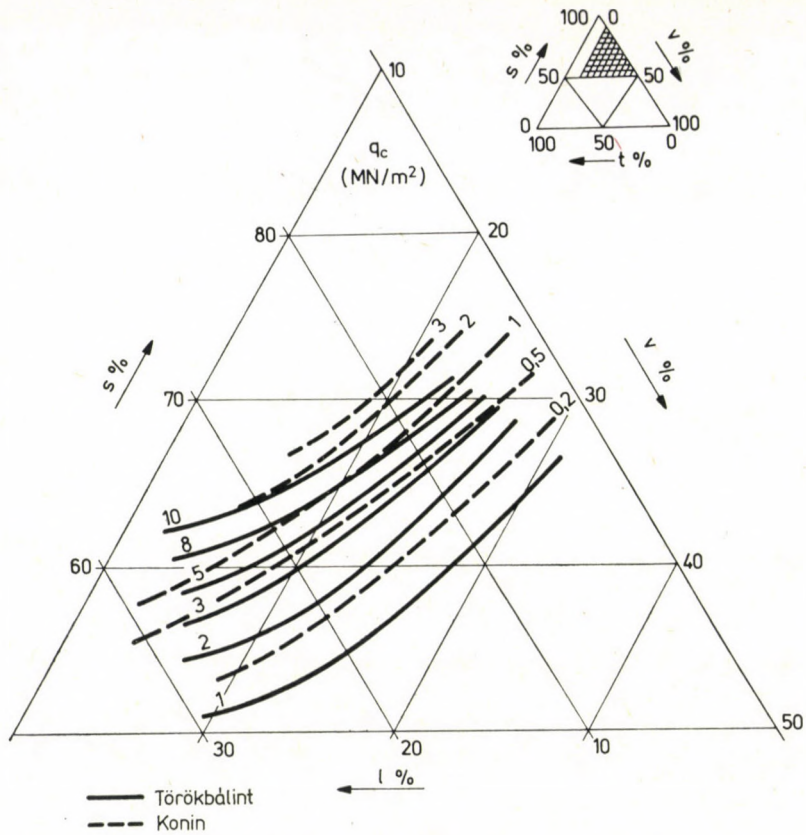


Fig. 5. Cone resistance as a function of phase composition ($H/d = 2{,}89$; Törökbálint)

phenomenon of cementation and friction forces will be presented below on the basis of the shear parameters of the examined soils.

An additional reason for the differences in cone resistance in particular soils was their mineralogical composition. In loess samples minerals from the quartz group prevail, having characteristically higher values of friction coefficients than loams.

Of practical significance is the knowledge of the function which describes a relationship between the cone resistance and the s, v, l values. The above relationship geometrically represents a surface (Fig. 6). Its equation for particular soils was determined empirically. The following function was assumed:

$$Y = \exp\left(Ax^2 + Bxy + Cy^2 + Dx + Ey + F\right) + \alpha \tag{9}$$

where : Y — cone resistance at liquid limit of soil
A, B, C, D, E, F — coefficients
α — asymptote of the function

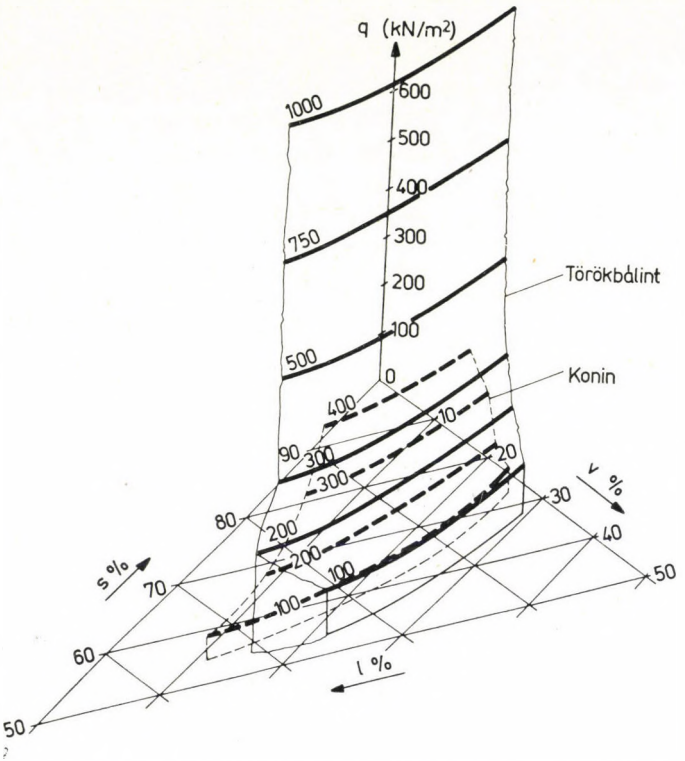


Fig. 6. Cone resistance as a function of phase composition

The choice of surface from the family of functions given by Eq. (9) is based on the criterion of minimisation of the sum:

$$\sum_{i=1}^{n} [f(x_i, y_i) - Y_i]^2, \tag{10}$$

where:

$$Y_i = q_{ci},\ x_i = s_i,\ y_i = v_i\ .$$

An analysis shows that the accepted type of function approximates the results for particular soils with high precision. The coefficients of Eq. (9) are presented in Table 1. Calculations of the coefficients were carried out on Odra 1304 computer.

The equation of the relationship between the cone resistance and s, v variables shows a certain analogy to JÁKY's (1944) empirical law which determines the relationship between the unconfined compression strength (σ_d) and moisture content (w) in cohesive soil:

$$\sigma_d = a \exp\left(-\frac{w}{k}\right). \tag{11}$$

Table 1

The coefficients of Eq. (9)

Locality	H/d	$A \cdot 10^{-3}$	$B \cdot 10^{-3}$	$C \cdot 10^{-3}$	$D \cdot 10^{-2}$	$E \cdot 10^{-2}$	F	α
1. Törökbálint	2,89	0,5	— 9,5	—3,8	26,3	57,8	—12,7	1,0
2. Konin	2,89	2,2	—14,0	—0,4	11,2	75,4	—10,7	0,2
3. Rudnicze	2,89	—10,8	2,9	3,4	141,8	—50,3	—37,6	0,2

JÁKY's law was later generalized by KÉZDI (1971); for transition soils having three phases, the following relationship has been established:

$$\ln \sigma_d = as - bv - cl. \tag{12}$$

3. Statistical analysis of the effect of cohesion and internal friction angle on cone resistance

As already mentioned, the static penetration process is used in many countries to determine the values of cohesion and angle of internal friction of soils. Analysing materials from the *Stockholm Symposium* (1974) it may be concluded that opinions in this matter are different. On the one hand, it is accepted that a close interdependence exists between the cone resistance and the cohesion; on the other hand, however, it is emphasized that it is not possible to establish a close correlation between the cone resistance and the cohesion or the angle of internal friction (e.g. AMAR, 1974). It seems therefore worthwhile to analyze the problem for soils, where shear strength depends on both components of shear strength, i.e. cohesion and angle of internal friction.

A convenient method is the statistical method of multiple linear regression (VOLK, 1965). Details necessary for the analysis: cohesion and internal friction angle determined in the triaxial apparatus (unconsolidated — undrained test; BISHOP, 1962).

On Figs 7 and 8 the effect of the phase composition (s, v, l) on the shear parameters is presented. It may be observed that they are very similar to the changes in the cone resistance. For that reason an approximation of the relationship between the cohesion and s, v variables was carried out with Eqs (9) and (13):

$$Y = \exp(Ax + By + c) + \alpha \tag{13}$$

where

$$x = s;\ y = v.$$

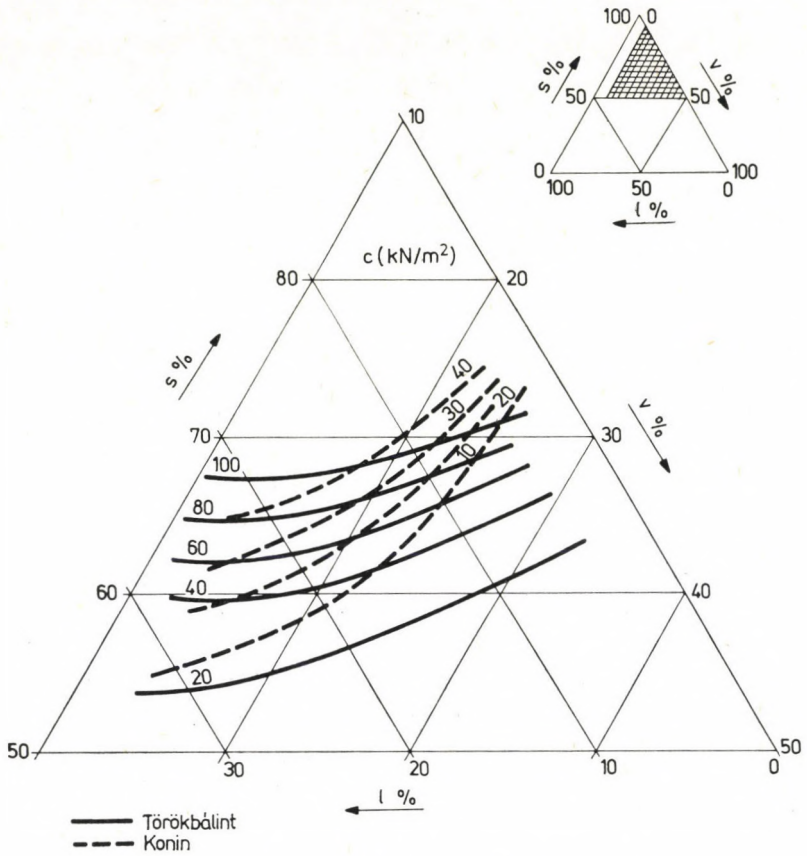


Fig. 7. Cohesion as a function of phase composition

Tab e 2

The coefficients of Eq. (13)

Locality	Dependent variable	$A \cdot 10^{-2}$	$B \cdot 10^{-2}$	$C \cdot 10^{-2}$	α
Törökbálint	"c"	12,3	— 6,4	—731	0,02
Konin	"c"	42,4	— 2,5	—408	0,02
Rudnicze	"c"	12,2	—11,9	—651	0,02
Törökbálint	tan Φ	4,5	— 7,1	—190	0,0
Konin	tan Φ	0,6	0,9	—153	0,0
Rudnicze	tan Φ	3,3	— 2,6	—205	0,0

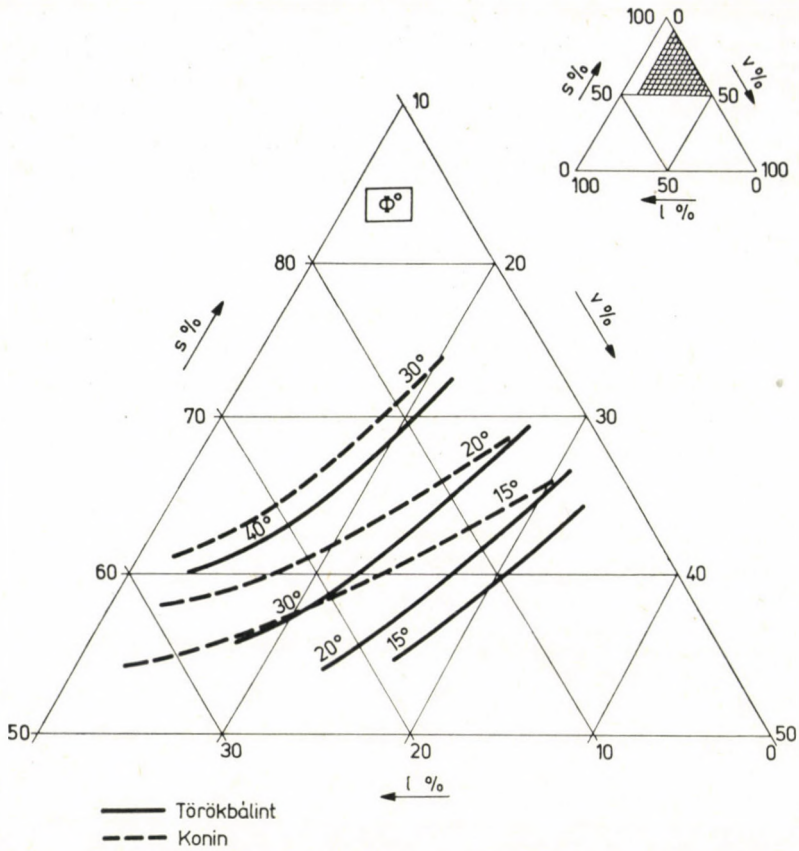


Fig. 8. Angle of internal friction as a function of phase composition

Calculations revealed that for loess and clay samples from *Rudnicze* a better approximation of the results is obtained from Eq. (9), while for *Konin* loam Eq. (13) is acceptable. Numerical values of the coefficients are given in Table 2.

Approximation of the relationship between cohesion and s, v variables by means of function with Eq. (13) confirms the function proposed by JÁKY. The existence of this relationship was confirmed by numerous tests with clayey soils (see for example HORN, 1964; KÉZDI, 1974).

4. Conclusions

Investigations of cone resistance in soils with up to 20 percent clay content revealed a decisive significance of several factors. One of them is the phase composition of the soil. In soils for which the effect of porosity on

strength properties cannot be ignored, estimation of strength properties on the basis of the relationship between the cone resistance and the plasticity index is not reliable.

The influence of the phase composition depends quite naturally on the granulometry and on the mineralogical composition of the soil. A definite relationship between cone resistance and (s, v) values is despite its empirical character, significant. The derived equation proved to be also useful for the interpretation of the penetration process in all examined soils. It may therefore be assumed that for other soils the relationship between cone resistance and (s, v) values may be approximated with Eq. (9).

Cohesion and internal friction angle have a significant influence on cone resistance. However, the varied influence of these parameters indicates that the recommended formulae for evaluation of cohesion on the basis of cone resistance may be applied only for soils for which the internal friction effect can be disregarded.

Acknowledgements

The authors gratefully appreciate the help given by Dr. Gy. HORVÁTH, research engineer in performing the tests and in bringing the manuscript in its final shape.

REFERENCES

AMAR, S.: The Use of the Static Penetrometer in the Laboratories des Ponts et Chaussées *Proc. E.S.O.P.T.*, Stockholm, vol. 2—2 (1974)

BANACH, S.: Mechanika, Czytelnik, 1954

DE BEER, E.: The Scale Effect in the Transposition of the Results of Deep Sounding Tests in the Ultimate Bearing Capacity of Piles and Caisson Foundation. *Geotechnique*, **12** (1963)

DE BEER, E.: Dimensional Analysis of the Problem of the Use of the Results of the Static Sounding Tests. *Proc. E.S.O.P.T.* (1974)

BISHOP, A. W.—HENKEL, D. J.: The Measurement of Soil Properties in the Triaxial Test. Arnold, London 1962

GILL, W.: Soil Dynamics in Tillage and Traction. *Research Service. VS Dep. of Agriculture* 1967

HORN, A.: Die Scherfestigkeit von Schluff. Westdeutscher Verlag, Köln 1964

KÉZDI, Á.: Handbuch der Bodenmechanik. Akadémiai Kiadó, Budapest 1964.

KÉZDI, Á.—VARGA, L.—TIMÁR, A.: Strength of transition Soils. *Proc. 4th Budapest Conference on Soil Mech. and Found. Engg.* (1971)

KÉZDI, Á.: Closing Address. *Proc. E.S.O.P.T., Stockholm*, vol. **1—2**, (1974)

KÉZDI, Á.: Fragen der Bodenphysik. Akadémiai Kiadó, Budapest, 1976

KISIEL, I.—SOBOCINSKI, J.—STACHON, M.: Reologiczne wlasciwosci gruntów ilowych. *Archiwum Hydrotechniki*, **23** (1976)

JÁKY, J.: Talajmechanika. Egyetemi Nyomda, Budapest 1944

JUMIKIS, A.: Soil Mechanics. van Nostrand and Co. Princeton, N.J. 1965

MITCHELL, J.—DURGUNOGLU, T.: Static Penetration Resistance of Soils. University of California, Berkeley, California 1975

MLYNAREK, Zb.: Metoda sondowania statycznego w zastosownai do charakterystyki konsystencji i cech wytrzymalosciowych gliny piaszczystej. *Archiwum Hydrotechniki*, **20** (1973)

MLYNAREK, Zb.: Sondowania statyczne i dynamiczne w badanich geotechnicznych. *Geotechnika w Wielkopolsce* (1979)

LANGHAAR, H.: Dimensional Analysis and Theory of Models, J. Wiley, London, 1964

LUNDGREEN, H.: Dimensional Analysis in Soil Mechanics. *Acta Polytechnica* **237,** Copenhagen (1957)
SANGLERAT, C.: The Penetrometer and Soil Exploration. Elsevier Pub. Comp. Amsterdam 1972
VOLK, M.: Statystyka stosowana dla inzynierow. Warszawa, 1965
WOLSKI, W.: Interpretacja wynikow sondowania gruntów. Typescript, 1977

Ergebnisse der statischen Penetrometerversuche in schwach- oder mittelbindigen Böden. — Die theoretischen Prinzipien der Benutzung von statischen Sonden und die Penetrationstiefe beeinflussenden Parameter werden untersucht. Die Autoren haben nachgewiesen, daß zwischen dem Eindringungswiderstand des Kegels und den Zustandskennwerten des Bodens keine eindeutige Abhängigkeit besteht. Mit drei Bodenarten von schwacher oder mittelhoher Kohäsion haben Versuche angestellt um die Abhängigkeit zwischen der Penetrationstiefe und der Druckkraft mit modernen Meßapparaten zu messen und zu registrieren. Die infolge der Bodendilatation auftretenden Widerstandsabnahme wurde beobachtet. Die Auswirkung der Phasenzusammensetzung der Bodenprobe wurde durch ausführliche Versuchsserien untersucht und festgestellt, daß der Eindringungswiderstand von der Schlamm- und Tongehalt des Erdreichs abhängt. Die Auswirkung der Phasenzusammensetzung ist teils in dreieckdiagrammatischer, teils durch räumlicher Abbildung dargestellt. Die Versuchsergebnisse wurden durch Anwendung des Verfahrens der kleinsten Quadrate ausgeglichen. Schließlich wurde die Einwirkung der Kohäsion und des Winkels der inneren Reibung auf den Eindringungswiderstand mit statistischen Verfahren untersucht.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (3—4), pp. 201—220 (1980)

EIN VORSCHLAG ZUR KODIERUNG DER THEORIEN DER BAUMECHANIK

M. V. SOARE*

[Eingegangen am 18. Dezember 1979]

Ausgehend von der Feststellung, daß er bisher keine systematische Einteilung der Theorien der Baumechanik gibt, wird eine Kodierung vorgeschlagen, die sich auf vier Hauptfaktoren stützt, nämlich die Art des Ausdrucks des Gleichgewichts der Struktur, das Verhalten des Werkstoffs, die Größe der Verformungen und die Betrachtungsweise des Ausdrucks der Krümmung. Den Abschluß bilden Beispiele der Kodierung anhand Aufgaben der Baumechanik.

1. Einführung

Die verschiedenen Kapitel der Baumechanik als Ingenieurfach, das im weitesten Sinne das Studium des Spannungs- und Verformungzustandes in Bauelementen und Strukturen zum Gegenstand hat, stützen sich auf eine Reihe von vereinfachenden Hypothesen hinsichtlich des Stoffgesetzes des Materials, der Größe der Verformungen im Verhältnis zu den allgemeinen Abmessungen des Körpers u. a.

Betrachtet man das vor dreihundert Jahren (1678) von Hooke aufgestellte Gesetz als »Geburtsurkunde« der Baumechanik, so muß man anerkennen, daß die bedeutenden Beiträge verschiedener Forscher im Laufe der Zeit zum Aufbau einer gut organisierten Fachwissenschaft geführt haben, mit Unterschieden in der Aufeinanderfolge der einzelnen Kapitel, in der Beweisführung oder im Behandlungsumfang, je nach Veranlagung des Verfassers oder der Fachsparte.

Gleichzeitig ist aber zu bemerken, daß die Fortschritte der Mathematik, der Rechentechnik, die mittels Modellen im verkleinerten oder natürlichen Maßstab erhaltenen Versuchsdaten oftmals eine Verfeinerung der grundlegenden Hypothesen nach sich gezogen haben. So führten etwa im Fall des Hookeschen Gesetzes, das das linear-elastische Verhalten verformbarer Festkörper beschreibt, die modernen Tendenzen einer Erforschung der Sicherheitsreserven zur Entwicklung der Berechnung im plastischen Bereich und zur Aufstellung von Gesetzen, die den ganzen Verhaltensbereich des Werkstoffs decken.

* Prof. Dr.-Ing. habil. MIRCE V. SOARE, R-76327, Str. Pictor Băncilă 22, Bucureşti 5, sector 6, Romania.

An zahlreichen Beispielen könnte gezeigt werden, wie manche Grundhypothese (mit vereinfachendem Charakter) bei der Lösung bestimmter Aufgaben berücksichtigt wird oder nicht.

Im Bestreben, die verschiedenen Hypothesen und Theorien der Baumechanik zusammenzufassen, stieß der Verfasser nur auf schematische Tendenzen einer Einteilung ohne allgemeinen Charakter, mit unterschiedlichen Formulierungen bei den verschiedenen Autoren.

Zu diesen Formulierungen gehören: mechanische, geometrische und physikalische Linearität bzw. Nichtlinearität, Fragen der Festigkeit und der (elastischen) Stabilität, Theorien erster, zweiter und — seltener — dritter Ordnung.

All diese unsystematischen und manchmal auch widersprüchlichen Aspekte veranlagten den Vorschlag einer Kodierung der verschiedenen Theorien in der Baumechanik, die eine genaue Kenntnis der bei der Lösung bestimmter Aufgaben zugrundegelegten Hypothesen und gleichzeitig auch des Entwicklungsgrades der betreffenden Theorie ermöglicht.

In dem hier folgenden 2. Abschnitt soll kurz auf die bestehenden Einteilungen eingegangen werden, im 3. und 4, Abschnitt sind die Hauptfaktoren zusammengestellt, auf die sich der Ansatz einer Aufgabe der Baumechanik stützt und die Kodierungsvorschläge formuliert.

Der 5. Abschnitt enthält Beispiele für einige Fälle und der letzte Abschnitt eine Diskussion und Schlußfolgerungen.

2. Bestehende Einteilungen

Eine erste Einteilung der Theorien der Baumechanik bezieht sich — gewöhnlich stillschweigend — auf die Linearität bzw. Nichtlinearität der Beziehungen, die ein bestimmtes Problem beherrschen. Sie kann sich auf die Ausdrucksweise der Gleichgewichtsbedingungen, das Verhalten des Werkstoffs bei Belastung und die Größenordnung der Verformungen beziehen.

So sind zu unterscheiden:

— die mechanische (statische) Linearität, wenn die Gleichgewichtsbedingungen für den unverformten Zustand des betreffenden Körpers ausgedrückt sind;

— die mechanische (statische) Nichtlinearität, wenn die Gleichgewichtsbedingungen für den verformten Zustand des Systems ausgedrückt sind;

— die physikalische Linearität, wenn die eindeutige Beziehung zwischen den Spannungen und den Verformungen mathematisch durch eine lineare Gleichung von der Form $\sigma = E\varepsilon$ oder $n = G\gamma$ ausgedrückt wird;

— die physikalische Nichtlinearität, wenn die Beziehungen zwischen den Spannungen und den Verformungen einen mathematischen Ausdruck von

der Form $\varepsilon = f(\sigma)$ oder $\sigma = g(\varepsilon)$ erhalten (vgl. einige Beispiele hierfür im 3. Abschnitt);

— die geometrische Linearität, wenn beim Ausdruck der Dehnungen in Potenzreihenentwicklungen der Verschiebungskomponenten nur die linearen Terme mitgenommen werden, etwa $\varepsilon_x = \partial u/\partial x$;

— die geometrische Nichtlinearität, wenn beim Ausdruck der Dehnungen in Potenzreihenentwicklungen der Verschiebungskomponenten auch die Terme höherer Ordnung berücksichtigt werden, etwa

$$\varepsilon_x = \frac{\partial u}{\partial x} + \frac{1}{2}\left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial x}\right)^2\right]. \tag{2.1}$$

Die Frage der Theorien verschiedener Ordnung ist strittiger.

Nach E. CHWALLA [9] »liegt es also nahe, den Einfluß, den die elastischen Verformungen auf die Gleichgewichtsbedingungen nehmen, zu vernachlässigen und die Gleichgewichtsbedingungen unmittelbar für das unverformte Tragwerk aufzustellen; diese vereinfachte Berechnungsweise wird als Theorie erster Ordnung (also als 1. Stufe der Annäherung an die vollkommen exakte Lösung!) bezeichnet«.

Dem gleichen Autor zufolge wäre die genauere Berechnung mit Berücksichtigung des Einflusses des elastischen Tragwerksverformung auf die Gleichgewichtsbedingungen die Theorie zweiter Ordnung.

»Erwähnt sei, daß auch die Theorie 2. Ordnung keine exakte Theorie darstellt, da sie zwar den Einfluß der Formänderungen berücksichtigt, aber an die Voraussetzungen gebunden ist, daß diese Formänderungen in Vergleich zu den Tragwerkabmessungen klein bleiben ... Würden wir beliebig große Verformungen zulassen, dann müßte die Rechnung noch weiter zugeschärft und damit noch erheblich verwickelter gestaltet werden (Theorie 3. Ordnung!).«

Eine andere Formulierung ist bei V. GIONCU [15] anzutreffen, nämlich:

— die Theorie erster Ordnung (oder die lineare Theorie) ist dadurch gekennzeichnet, daß sie den Einfluß der Verformungen auf den Spannungszustand vernachlässigt;

— in der Theorie 2. Ordnung (oder der linearisierten Stabilitätstheorie) werden die Verformungen als unendlich klein betrachtet und die Produkte der zweiten Ableitungen der Durchbiegungen sind daher vernachlässigbar;

— die Theorie 3. Ordnung ist eine nichtlineare Stabilitätstheorie (vgl. [15] S. 331—332).

Wieder eine andere Formulierung fand S. FALK [11], der die drei Theorien folgendermaßen charakterisiert:

— Theorie erster Ordnung: Gleichgewicht und Reduktion der Spannungen am unverschobenen Element;

— Theorie zweiter Ordnung: Gleichgewicht und Reduktion der Spannungen am verschobenen aber nicht verdrehten Element;

— Theorie dritter Ordnung: Gleichgewicht und Reduktion der Spannungen am verschobenen und verdrehten Element.

Der gleiche Autor stellt fest, daß bei seiner Formulierung alle drei Theorien von Haus aus nichtlinear sind. Durch bestimmte Vereinfachungen lassen sie sich aber linearisieren (vgl. S. FALK [11] und J. LENSING [22]).

In der Tat gilt S. FALKS Behauptung daß in dieser Hinsicht in der Literatur ein heilloses Durcheinander herrscht [12].

Wir erinnern noch an das Werk von R. ROSMAN [28], welches die Stockwerkrahmen behandelt, in dem es auch den Einfluß der Deformationen über die Größe der Schnittkräften in Betracht zieht.

Durch ein schrittweises Näherungsverfahren wird die unverformte Lage, die in Wirklichkeit die Theorie I. Ordnung darstellt, nach der die nächste Stufe der Iteration als Theorie 2. Ordnung und die letzte Stufe (wegen der schnellen Konvergenz der Korrektionen) als Theorie 3. Ordnung angesehen wird, untersucht.

Nach E. CHWALLA [9] »gibt es noch eine große Gruppe von Problemen, bei denen die Anwendung der Theorie 2. Ordnung eine innere Notwendigkeit darstellt und daher grundsätzlich immer erforderlich ist. Es sind dies die Stabilitätsprobleme.«

Bei der Definition der Stabilität stößt man auf folgendes: »Die Stabilitätstheorie sucht nun für ein gegebenes Tragwerk und eine gegebene Belastungsweise jene kritische Laststufe auf, unter der die Stabilitätsgrenze erreicht wird und daher das Gleichgewicht zwischen der inneren und äußeren Kräften die Eigenschaft verliert, stabil zu sein. Wir sprechen von Biegeknickung eines Stabes oder Stabwerkes, wenn die einzelnen Stabquerschnitte an der Stabilitätsgrenze — beim erwähnten zwanglosen Übergang von der ersten zur zweiten, benachbarten Gleichgewichtsfigur — nur parallel verschoben, aber nicht verdreht werden, wenn also die einzelnen Stabelemente nur Verbiegungen, jedoch keine Verdrillungen erfahren« (a. s. O. [9], S. 49).

Anschließend werden die Drillknickung, die Biegedrillknickung und als Sonderfall der letzteren die Kippung definiert, bei der sich der Querschnitt um den Schubmittelpunkt verdreht bzw. um diesen Punkt verschiebt und verdreht. Hier ist eine Annäherung an die Definitionen von S. FALK für die Theorien 2. bzw. 3. Ordnung festzustellen.

Eine nuancierte Formulierung findet man bei Al. GHEORGHIU [14]; er betont »den wesentlichen Unterschied zwischen der Stabilitätsberechnung und der Rechnung 2. Ordnung:

(a) Bei der Stabilitätsrechnung ist die Größe der Lasten unbekannt und ihr Wert P_{kr} zu bestimmen, bei dem der Stabilitätsverlust im Sinne des Knikkens eintritt.

(b) Bei der Rechnung 2. Ordnung sind die Quer- und Längslasten bekannt und die Größe der Schnittkräften und Verschiebungen sollen bestimmt

werden, die auftreten, wenn die Außenkräfte auf die verformte Lage bezogen werden.

Der Stabilitätsverlust des Tragwerks erscheint als Grenzzustand bei der Berechnung 2. Ordnung« (a. a. O. [14], S. 513).

Zu einer ähnlichen Äußerung gelangen BRUSH und ALMROTH: »Die Knickberechnung ist im wesentlichen eher ein Teil der nichtlinearen als der linearen Mechanik« [7].

Hinzugefügt sei die Tatsache, daß die Rechnung 2. Ordnung auch den Fall einschließt, in dem die Längskraft eine Zugkraft ist.

Als Abschluß dieser summarischen Inkursion in die Fachliteratur sei nur noch darauf hingewiesen, daß bei den bestehenden Definitionen nicht alle Hauptfaktoren berücksichtigt sind; dieser Tatbestand soll im folgenden Abschnitt untersucht werden.

3. Die Hauptfaktoren im Ansatz der Baumechanikaufgaben

Eine aufmerksame Prüfung beweist, daß der Formulierung der Baumechanikaufgaben vier Hauptfaktoren zugrundeliegen:

A) die Art des Ausdrucks für die Gleichgewichtsbedingungen;

B) das Werkstoffverhalten;

C) die Größe der Verformungen;

D) der Ausdruck der Krümmung der verformten Mittelfaser bzw. der Krümmungen der verformten Mittelfläche.

Die verschiedenen Aspekte dieser Faktoren sollen nun untersucht werden.

A)*Die Bedingungen des statischen Gleichgewichts* lassen sich mit Bezug auf

1° die ursprüngliche (unverformte) Lage des Tragwerks oder

2° die (verformte) Endlage des Tragwerks ausdrücken.

Die beiden Varianten entsprechen der im 2. Abschnitt beschriebenen mechanischen Linearität bzw. Nichtlinearität.

Wie E. CHWALLA betont, müssen bei Stabilitätsaufgaben die statischen Gleichgewichtsbedingungen unbedingt in bezug auf den verformten Zustand der Struktur ausgedrückt werden.

Mathematisch führen die beiden Ausdrucksweisen für die statischen Gleichgewichtsbedingungen zu verschiedenen Aspekten; so entspricht nähmlich die erstere den Aufgaben mit Randbedingungen, während sich die zweite auf Eigenwertprobleme bezieht. Dabei ist aber auch die Hypothese des linearen Werkstoffgesetzes und die der infinitesimalen Verformungen impliziert.

B) *Das Verhalten des Werkstoffs* kann

1° linear oder

2° nichtlinear sein.

Lineares Verhalten entspricht der im 2. Abschnitt erwähnten physikalischen Linearität und wird durch das Hookesche Gesetz ausgedrückt.

Nichtlineares Verhalten entsprechend physikalischer Nichtlinearität gibt je nach dem für den Werkstoff geltenden Stoffgesetz zu mannigfaltigen Lösungen Anlaß.

Im Fachschrifttum wurden verschiedene Vorschläge zur Schematisierung der das Werkstoffgesetz wiedergebenden Kurve gemacht, und zwar für

— ideal-elastoplastisches Verhalten (Bild 1a), auch Prandtlsche Kurve genannt;

— elastoplastisches Verhalten mit Verfestigungszone (Bild 1b);

— elastoplastisches Verhalten mit Fließstufe und Verfestigungszone (Bild 1c);

— ideal-plastisches Verhalten (Bild 1d);

— ideal-plastisches Verhalten mit Verfestigungszone (Bild 1e).

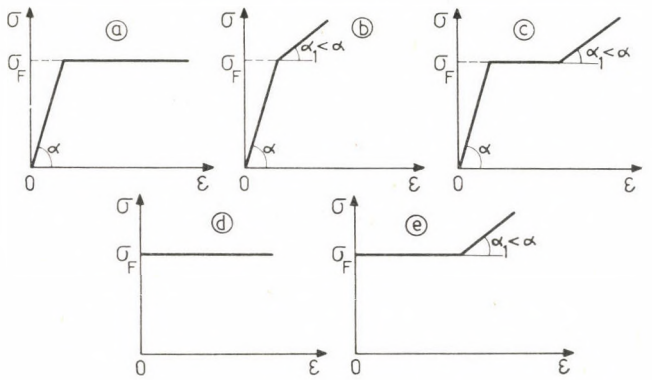


Bild 1

Unter den Vorschlägen zum Ansatz eines einzigen analytischen Ausdrucks für den gesamten Variationsbereich seien erwähnt:

$$\varepsilon = \frac{\sigma^n}{E} \qquad (\text{mit } n \gtrless 1)\,;$$

— die in den alten deutschen Vorschriften für Stahlbauten angegebene Kurve

$$\sigma = \sigma_p + (\sigma_F - \sigma_p) \tanh \frac{E\varepsilon - \sigma_p}{\sigma_F - \sigma_p} \tag{3.1}$$

mit σ_F als Fließgrenze, σ_p als Proportionalitätsgrenze;

— die Ramberg—Osgoodsche Kurve (1943), geschrieben in einer der Formen

$$\frac{\varepsilon}{\varepsilon_0} = \frac{\sigma}{\sigma_0} + k\left(\frac{\sigma}{\sigma_0}\right)^n, \tag{3.2a}$$

$$\varepsilon = \frac{\sigma}{R} + \left(\frac{\sigma}{B}\right)^n, \tag{3.2b}$$

auch Dreiparameterkurve genannt [24], [27];

— die Formel von YLINEN (1956) [33]

$$E\varepsilon = c\sigma - (1-c)\sigma_F \ln\left(1 - \frac{\sigma}{\sigma_F}\right),$$

in der $c \neq 1$ eine Stoffkonstante bedeutet; für $c = 1$ kommt man auf das Hookesche Gesetz;

— die Formel von J. E. GOLDBERG und R. M. RICHARD (1963) [16]

$$\frac{\sigma}{\sigma_0} = \frac{\dfrac{\varepsilon}{\varepsilon_0}}{\left(1 + \left|\dfrac{\varepsilon}{\varepsilon_0}\right|^n\right)^{1/n}}$$

mit σ_0 als Maximalspannung, ε_0 als der dieser entsprechenden Dehnung und n als dem Parameter, der die nichtlineare Beziehung zwischen σ und ε wiedergibt.

C) *Die Verformungen* können

1° infinitesimal oder

2° endlich sein.

Bei infinitesimalen (oder kleinen) Verformungen können die Verschiebungen der verschiedenen Querschnitte als unendlich kleine Größen im Vergleich zu den geometrischen Maßen der Struktur betrachtet werden. Die Hypothese ermöglicht die Vernachlässigung der Dehnungen im Verhältnis zur Einheit (etwa $(1 + \varepsilon \approx 1)$, so wie man bei der Ermittlung des Ausdrucks für die Gleitung in der ebenen Elastizitätstheorie vorgeht

$$\gamma_{xy} = \frac{\dfrac{\partial v}{\partial x}}{1 + \varepsilon_x} + \frac{\dfrac{\partial u}{\partial y}}{1 + \varepsilon_y} \approx \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y};$$

diese Vernachlässigung gestattet die Linearisierung des geometrischen Aspektes und entspricht der im 2. Abschnitt erwähnten geometrischen Linearität.

Endliche Verformungen können ihrerseits wiederum klein oder beliebig groß im Verhältnis zu den allgemeinen Maßen der Struktur sein. Gewöhnlich werden sie als klein angesehen, so wie etwa Beziehung (2.1) ausdrückt, und entsprechen der geometrischen Nichtlinearität.

D) *Der Ausdruck der Krümmung* ist in erster Linie von der Art der Spannungen in den Stabquerschnitten (bzw. über die Plattendicke) und ferner von der Ausdrucksweise der Krümmung abhängig.

Die Krümmung kann hierin
0) überhaupt nicht,
1) angenähert,
2) genau
erscheinen.

Die Fälle, in denen bei der Untersuchung eines Bauteils oder einer Struktur die Krümmung der verformten Stabachse (bzw. die Krümmungen der verformten Mittelfläche) nicht eingreifen, sind selten; sie kommen etwa bei Hängesystemen oder bei kritischen Systemen wie dem im 5. Abschnitt gezeigten vor.

Der exakte Ausdruck der Krümmung einer ebenen Kurve mit der Gleichung $w = w(x)$ ist wohlbekannt

$$\frac{1}{\varrho} = \frac{\dfrac{d^2 w}{dx^2}}{\left[1 + \left(\dfrac{dw}{dx}\right)^2\right]^{3/2}} .$$

Vernachlässigt man das Quadrat von dw/dx gegenüber der Einheit, betrachtet also

$$1 + \left(\frac{dw}{dx}\right)^2 \approx 1 ,$$

so erhält man den approximativen Ausdruck für die Krümmung

$$\frac{1}{\varrho} \approx \frac{d^2 w}{dx^2} ,$$

der in der Festigkeitslehre und in der Platten- und Schalentheorie weit verbreitet ist.

Bemerkenswert ist, daß die Größenordnung des Fehlers von der im Fall C abweicht, wenn infinitesimale Verformungen angenommen werden.

Der approximative Ausdruck der Krümmung ermöglicht die Linearisierung der Aufgabe, reiht sich also in die geometrische Linearität ein; gleichzeitig trifft er sowohl in der mechanischen (statischen) Linearität als auch in der Nichtlinearität zu.

Akzeptiert man stillschweigend die Hypothese ebener Querschnitte nach J. BERNOULLI für Stabsysteme bzw. die Love— Kirchhoffsche Hypothese der geraden Normalen für Platten und Schalen, so erkennt man, daß der letztere Fall die Untersuchungen einschränkt, in dem er die Aufgaben der Elastizitätstheorie (etwa diejenigen des lokalen Kontaktes elastischer Körper — nach HERTZ) ausschließt.

4. Ein Kodierungsvorschlag

Drei der weiter oben untersuchten vier Faktoren lassen also je drei Varianten zu, der erstgenannte nur zwei. Bei der ersten Variante tritt der betreffende Faktor nicht in Erscheinung, die zweite zeigt eine schwächere, die dritte eine stärkere Approximation oder sogar den exakten mathematischen Ansatz.

Diese Aufzählung ließe sich gewiß noch durch Anfügung neuer Fälle verfeinern, etwa durch die Berücksichtigung der Axialverformungen gleichzeitig mit den Biegeverformungen; einer einfachen Formulierung halber beschränken wir uns hier jedoch auf die Faktoren, A, B, C, D. Bezeichnet man mit 1 und 2 bzw. mit 0, 1, 2 die möglichen Fälle, so läßt sich jede Aufgabe der Baumechanik mit einem ABCD-Kode verschlüsseln, wobei der Term A die Werte $i = 1$ und 2 und jeder der Terme B, C, D die Werte $i = 0, 1, 2$ erhalten kann. Diese Kodierung ist in Tabelle 1 zusammengestellt.

Tabelle 1

Faktor / i =	Ausdruck des Gleichgewichts	Werkstoffgesetz	Größe der Verformungen	Ausdruck der Krümmung
	A	B	C	D
0	—	kommt nicht vor	kommt nicht vor	kommt nicht vor
1	bei unverformter Struktur	linear	infinitesimal	approximativ
2	bei verformter Struktur	nichtlinear	endlich	exakt

Aus der theoretischen Kombination aller in Tabelle 1 enthaltener Möglichkeiten ergeben sich 54 Fälle:

1000	1100	1200	2000	2100	2200
1001	1101	1201	2001	2101	2201
1002	1102	1202	2002	2102	2202
1010	1110	1210	2010	2110	2210
1011	1111	1211	2011	2111	2211
1012	1112	1212	2012	2112	2212
1020	1120	1220	2020	2120	2220
1021	1121	1221	2021	2121	2221
1022	1122	1222	2022	2122	2222

Offenbar kann A nur die Werte 1 und 2 haben, da die Gleichgewichtsbedingungen immer ausgedrückt werden müssen.

$B = 0$ kommt in Einzelfällen vor, etwa bei statisch bestimmten Systemen, wenn nur der Spannungszustand interessiert.

$C = 0$ trifft man bei hauptsächlich biegebelasteten Strukturen an, wenn die Dehn- und Gleitverformungen vernachläßigbar sind, wie es bei den statisch bestimmten Systemen vorkommt, für die nur der Spannungszustand interessiert.

$D = 0$ kennzeichnet die nur auf Zug beanspruchten Systeme.

Die übrigen Aufgabenkategorien bieten die interessantesten Aspekte der Diskussion.

Ohne die Theorien nach den Werten ordnen zu können, die der $ABCD$-Kode annehmen kann, ist die erste Theorie (1000) offenbar die einfachste, die letzte (2222) die am weitesten entwickelte.

Ferner ist zu bemerken, daß in der Baumechanik nicht für alle Kodefälle Lösungen bekannt sind; es gibt also noch leere Stellen, für die entweder bisher keine Lösungen gefunden wurden oder aber einfach kein theoretisches oder praktisches Interesse besteht.

Für $B = 2$ (nichtlineares Verhalten der Werkstoffe) bestehen zahlreiche Möglichkeiten je nach dem für das betreffende Material geltenden Stoffgesetz.

Im folgenden Abschnitt sollen die wichtigsten Kodes kurz exemplifiziert werden. In manchen Fällen erhalten hierbei die Ergebnisse eine bisher in der Fachliteratur nicht veröffentlichte Form.

Die Diskussion der übrigen Fälle und die Zusammenhänge mit den bisherigen Einteilungen folgen im letzten Abschnitt.

5. Aufgabenbeispiele in der vorgeschlagenen Kodierung

Wir beginnen mit den Fällen $ABC0$ ($D = 0$).

Betrachtet sei das Gelenkstabsystem aus Bild 2a, ein klassisches Beispiel eines zugbeanspruchten, statisch unbestimmten Stabsystems. Solange sich der Werkstoff völlig elastisch verhält, entsprechen die Spannungen in den Stäben den Beziehungen [19], [29]:

$$X_1 = X_3 = \frac{P \cos^2 \alpha}{1 + 2 \cos^3 \alpha}, \quad X_2 = \frac{P}{1 + 2 \cos^3 \alpha}.$$

Offenbar handelt es sich hier um den Kodefall 1110.

Wird jedoch eine Beanspruchung über die Elastizitätsgrenze des Werkstoffs hinaus zugelassen und die Prandtlsche Kurve (Bild 1a) berücksichtigt, so geht der Verlauf der Kräfte $X_1 = X_3$ und X_2 in Abhängigkeit von der Kraft P aus Bild 2b hervor. Dem entspricht Fall 1210.

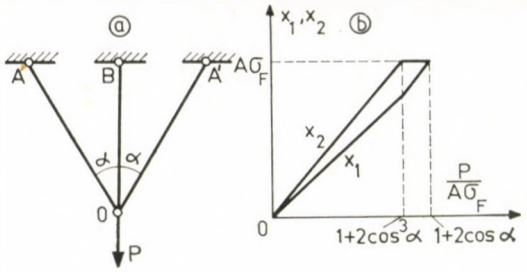


Bild 2

Ferner sei das im Bild 3a wiedergegebene und von TIMOSHENKO und YOUNG [32] untersuchte, kritische System als ein Beispiel herangezogen, in dem das Superpositionsprinzip keine Gültigkeit mehr hat. Bei Annahme infinitesimaler Verformungen entspricht dieser Fall dem Kode 2110.

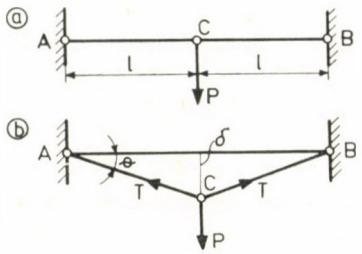


Bild 3

Betrachten wir nun die Aufgabe bei endlichen Verformungen (Kode 2120). Sei Θ der Neigungswinkel der beiden verformten Stäbe gegenüber ihrer ursprünglichen Lage und T seien die Schnittkräfte in den Stäben.

Der Verschiebung des Mittelgelenks um die Grösse δ entspricht eine Verdrehung Θ der beiden Stäbe, so daß

$$\tan \Theta = \frac{\delta}{l}$$

gilt. Dann werden die Schnittkräfte in den Stäben zu

$$T = \frac{P}{2 \sin \Theta} = \frac{P}{2} \frac{\sqrt{1 + \operatorname{tg}^2 \Theta}}{\operatorname{tg} \Theta} = \frac{P}{2} \frac{\sqrt{l^2 + \delta^2}}{\delta} . \tag{5.1}$$

Da die Länge jedes Stabes infolge Dehnung zu $l' = \sqrt{l^2 + \delta^2}$ wird, lautet die Dehnung

$$\varepsilon = \frac{\sqrt{l^2 + \delta^2} - l}{l} = \frac{T}{EA} . \tag{5.2}$$

Die Elimination von T aus den Beziehungen (5.1) und (5.2) ergibt

$$\frac{\sqrt{l^2+\delta^2}-l}{l}=\frac{P}{2EA}\,\frac{\sqrt{l^2+\delta^2}}{\delta},$$

oder in expliziter Form

$$P=2EA\,\frac{\delta}{l}\left(1-\frac{1}{\sqrt{1+\dfrac{\delta^2}{l^2}}}\right). \tag{5.3a}$$

Entwickelt man das Radikal in eine Potenzreihe und berücksichtigt nur die ersten beiden Terme, so findet man die früher erwähnte Lösung

$$P=EA\left(\frac{\delta}{l}\right)^3 \quad \text{bzw.} \quad \delta=l\sqrt[3]{\frac{P}{EA}}. \tag{5.3b}$$

Die Aufgabe läßt sich unschwer auf den Fall eines anderen Stoffgesetzes, etwa mittels der Formel von RAMBERG—OSGOOD (3.2b) übertragen. Mit den dimensionslosen Bezeichnungen

$$\xi=\frac{\delta}{l},\ \beta=\frac{P}{AE}$$

läßt sich die Beziehung $f(P, \delta)=0$

$$\sqrt{1+\xi^2}-1=\frac{\beta}{2}\,\frac{\sqrt{1+\xi^2}}{\xi}+\left(\frac{E}{B}\,\frac{\beta}{2}\,\frac{\sqrt{1+\xi^2}}{\xi}\right)^n$$

schreiben.

Für $n\to\infty$ kommt man auf die der Prandtlschen Kurve entsprechende Lösung zurück. In beiden Fällen gilt der Kode 2220.

*

Kode 1111. Bekannt ist die angenäherte Differentialgleichung für die verformte Stabachse bei einfacher Biegung

$$\frac{1}{\varrho}\approx\frac{d^2w}{dx^2}=-\frac{M}{EI}. \tag{5.4}$$

In dem in Bild 4a wiedergegebenen Fall ist der Stab auf reine Biegung beansprucht und die Durchbiegung erhält den einfachen Ausdruck

$$w=\frac{M_0}{2EI}x(l-x);$$

die verformte Stabachse ist eine Parabel und die maximale Durchbiegung hat die Größe

$$w_0 = \frac{M_0 l^2}{8EI}. \tag{5.5}$$

Im Belastungsfall des Bildes 4b ist der Stab einfacher Biegung unterworfen ($M = -P(l - x)$) und die Durchbiegungsgleichung lautet:

$$w = \frac{Px^2}{2EI}\left(l - \frac{x}{3}\right).$$

mit dem Maximalwert (am freien Ende $x = l$):

$$w_0 = \frac{Pl^3}{3EI}. \tag{5.6}$$

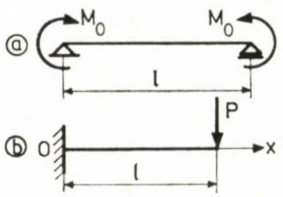


Bild 4

Kode 1112. Führt man bei den vorstehenden Beispielen den genauen mathematischen Ausdruck der Krümmung ein, so lautet die exakte Differentialgleichung der verformten Stabachse bei einfacher Biegung

$$\frac{\dfrac{d^2w}{dx^2}}{\left[1 + \left(\dfrac{dw}{dx}\right)^2\right]^{3/2}} = -\frac{M}{EI}. \tag{5.7}$$

Für den Fall des Bildes 4a ist eine einfache geometrische Lösung möglich; die verformte Stabachse ist ein Kreisbogen mit dem Radius $\varrho = EI/M_0$ und die maximale Durchbiegung erhält den Ausdruck [29]:

$$w_{\max} = \varrho\left[1 - \sqrt{1 - \left(\frac{l}{2\varrho}\right)^2}\right].$$

Setzt man anstelle von ϱ, wieder den vorher angegebenen Wert und entwickelt das Radikal in eine Potenzreihe (da $l/2\varrho \ll 1$ gilt), so erhält man

$$w_{\max} = w_0\left(1 + 4\frac{w_0^2}{l\varrho^2} + 32\frac{w_0^4}{\varrho^4} + \ldots\right), \tag{5.8}$$

wobei w_0 aus (5.5) hervorgeht.

Aus (5.8) folgt, daß der mittels der klassischen Lösung erhaltene Wert dem ersten Term der Reienentwicklung der exakten Lösung entspricht. Angesichts der Größenordnung des Verhältnisses w_0/l ist die Approximation durchaus gerechtfertigt.

Für den im Bild 4b gezeigten Fall wird die Differentialgleichung zu

$$\frac{\dfrac{d^2w}{dx^2}}{\left[1 + \left(\dfrac{dw}{dx}\right)^2\right]^{3/2}} = \frac{P}{EI}(l - x)\ . \tag{5.9}$$

Mit den Bezeichnungen $p = \dfrac{dw}{dx}$, $\dfrac{dp}{dx} = \dfrac{d^2w}{dx^2}$, läßt sich die Gleichung einmal integrieren

$$P = \frac{\dfrac{P}{EI}\left(lx - \dfrac{x^2}{2}\right) + C_1}{\sqrt{1 - \left[\dfrac{P}{EI}\left(lx - \dfrac{x^2}{2}\right) + C_1\right]^2}}\ ;$$

mit Rücksicht auf die Randbedingung $p = dw/dx = 0$ für $x = 0$, woraus $C_1 = 0$ folgt, erhält man

$$\frac{dw}{dx} = \frac{\dfrac{P}{EI}\left(lx - \dfrac{x^2}{2}\right)}{\sqrt{1 - \left(\dfrac{P}{EI}\right)^2\left(lx - \dfrac{x^2}{2}\right)^2}}\ .$$

Direkte Integration ergibt

$$w = \int \frac{\dfrac{P}{EI}\left(lx - \dfrac{x^2}{2}\right)}{\sqrt{1 - \left(\dfrac{P}{EI}\right)^2\left(lx - \dfrac{x^2}{2}\right)^2}}\,dx + C_2\ . \tag{5.10}$$

Die Substitution

$$\xi = \frac{P}{EI}\left(lx - \frac{x^2}{2}\right)$$

reduziert das Integral in (5.10) auf das elliptische Integral [18]

$$w = C_2 - \frac{EI}{Pl}\int \frac{\xi d\xi}{\sqrt{1-\xi^2}\cdot\sqrt{1-\frac{2EI}{Pl^2}\xi}}\,.$$

Selbst in den einfachsten Fällen ist also die Berechnung der (hinsichtlich der Krümmung) exakte Lösung recht umständlich. Eine angenäherte Lösung erhält man, indem man das Radikal in (5.10) in eine Potenzreihe entwickelt und dann gliedweise integriert. Für die maximale Durchbiegung erhält man

$$w_{\max} = w_0\left(1 + \frac{27}{35}\,\frac{w_0^2}{l^2} + \ldots\right),$$

wobei w_0 aus (5.6) hervorgeht.

Die durch die Berücksichtigung des genauen Ausdrucks für die Krümmung der verformten Stabachse eingeführte Korrektur ist auch hier ohne weiteres vernachlässigbar.

Kode 2111. Ein an einem Ende eingespannter und am anderen Ende freistehender Stab sei einer Druckkraft P ausgesetzt. Erreicht P einen bestimmten kritischen Wert, so bleibt das Gleichgewicht bekanntlich nicht mehr stabil, und es werden in der Umgebung der ursprünglichen geradlinigen Gleichgewichtsform gekrümmte Formen möglich.

Die Knickdifferentialgleichung lautet in der vereinfachten Form von EULER:

$$\frac{d^2w}{dx^2} + \frac{P}{EI}w = 0\,. \tag{5.11}$$

Da sie linear und homogen ist, werden von Null verschiedenen Lösungen nur für bestimmte Werte des Parameters $\beta = \sqrt{P/EI}$, sogenannte Eigenwerte, vorkommen. Den Mindestwert bildet die kritische Belastung nach EULER

$$P_E = \frac{\pi^2 EI}{4l^2}\,.$$

Zu bemerken ist, daß die verformte Stabachse bis auf einen Proportionalitätsfaktor (in den Grenzen der infinitesimalen Verformungen) genau bestimmt ist

$$w = w_0 \sin\frac{\pi x}{2l}\,,$$

da die Auflösung der Gleichung (5.11) ein Eigenwertproblem darstellt.

Der mathematische Aspekt ändert sich, wenn eine Quereinwirkung, etwa der Angriff der Kräfte P mit einer Exzentrizität e eintritt.

Dann erhält die maximale Durchbiegung die Form

$$\frac{w_{\max}}{e} = \sec\left(\frac{\pi}{2}\sqrt{\frac{P}{P_E}}\right) - 1 ,$$

ein als Sekantenformel bekannter Ausdruck [33].

Kode 2112. Wird anstelle des approximativen Ausdrucks der Krümmung der exakte Ausdruck gewählt, so wird Gleichung (5.11) zu

$$\frac{\dfrac{d^2w}{dx^2}}{\left[1+\left(\dfrac{dw}{dx}\right)^2\right]^{3/2}} + \frac{P}{EI}w = 0 .$$

Dann wird als unabhängige Variabel der Bogen s und als abhängige die Tangentenneigung Θ gewählt; die substitution $dw/ds = \sin\Theta$ und die Vernachlässigung der Stabverkürzung infolge des Druckes [33] führen zu der Differentialgleichung

$$EI\frac{d^2\Theta}{ds^2} + P\sin\Theta = 0 .$$

Die kritische Belastung ermittelt man mittels des vollständigen elliptischen Integrals erster Art [33]. Eine analoge Lösung erhält man wenn man auch die Axialverformungen in Betracht zieht [26].

Kode 2121. Das überkritische Verhalten ebener Platten und das Biegen mit großen Verformungen lassen sich nach den Kármánschen Gleichungen beschreiben:

$$K\left(\frac{\partial^4 w}{\partial x^4} + 2\frac{\partial^4 w}{\partial x^2\partial y^2} + \frac{\partial^4 w}{\partial y^4}\right) = Z + \\ + \frac{\partial^2 F}{\partial x^2}\frac{\partial^2 w}{\partial v^2} - 2\frac{\partial^2 F}{\partial x\partial y}\frac{\partial^2 w}{\partial x\partial y} + \frac{\partial^2 F}{\partial v^2}\frac{\partial^2 w}{\partial x^2}, \tag{5.12a}$$

$$\frac{\partial^4 F}{\partial x^4} + 2\frac{\partial^4 F}{\partial x^2\partial y^2} + \frac{\partial^4 F}{\partial y^4} = E\delta\left[\left(\frac{\partial^2 w}{\partial x\partial y}\right)^2 - \frac{\partial^2 w}{\partial x^2}\frac{\partial^2 w}{\partial y^2}\right]. \tag{5.12b}$$

Die Nichtlinearität wird bei dieser Aufgabe durch die geometrischen Beziehungen vom Typ (2.1), etwa

$$\varepsilon_x = \frac{\partial u}{\partial x} + \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2, \quad \varepsilon_y = \frac{\partial v}{\partial y} + \frac{1}{2}\left(\frac{\partial w}{\partial y}\right)^2.$$

$$\gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} + \frac{\partial w}{\partial x}\frac{\partial w}{\partial y}$$

und durch den Ausdruck des Gleichgewichts für den verformten Zustand der Platte berücksichtigt.

Die Lösungen erhält man auf variationsanalytischem Wege (vgl. etwa [25], [34]) oder numerisch [30].

Ist die Komponente der normalen Plattenbelastung nicht Null ($Z \neq 0$), so hat man es mit einem Biegeproblem mit großen Verformungen zu tun. Gilt $Z = 0$ und ist die Platte in ihrer Mittelebene beansprucht, so liegt Beulen mit großen Verformungen vor und das überkritische Verhalten läßt sich quantitativ bestimmen.

Die klassische (linearisierte) Beulaufgabe ergibt sich, wenn man nur Gleichung (5.12a) berücksichtigt, $Z = 0$ setzt und konstante Werte (die Randwerte) für

$$N_x = \frac{\partial^2 F}{\partial y^2}, \; N_y = \frac{\partial^2 F}{\partial x^2}, \; N_{xy} = N_{yx} = -\frac{\partial^2 F}{\partial x \, \partial y}$$

wählt.

Kode 1211. Die Berechnung der auf reine Biegung beanspruchten Stäbe kann für jedes Werkstoffgesetz (im allgemeinen graphisch, versuchsmäßig [5]) vor sich gehen; sie vereinfacht sich wesentlich, wenn man die Prandtlsche Kurve (Bild 1a) zuläßt.

Kode 2211. Das Knicken von Stahlstäben im elastisch-plastischen Bereich wurde in den alten österreichischen Vorschriften (ÖNORM B 4300-1953) und den deutschen Normen (DIN 4114 Blatt 1 u. 2) ausgehend von Beziehung (3.1) untersucht und daraus der Tangentenmodul abgeleitet

$$E_t = \frac{d\sigma}{d\varepsilon} = E\left[1 - \left(\frac{\sigma - \sigma_p}{\sigma_F - \sigma_p}\right)^2\right].$$

6. Diskussion der Kodierung. Schlußfolgerungen

Im vorangehenden Abschnitt wurden zur Exemplifizierung des hier beschriebenen Kodierungsvorschlags sieben der im Rahmen desselben möglichen Fällen — unter Weglassung von $D = 0$ — behandelt. Es ergibt sich die natürliche Frage, was mit den übrigen Fällen geschieht.

Jeder derselben impliziert offenbar zumindest eine Nichtlinearität, d. h. eine mathematisch schwer zu lösende Aufgabe.

Die Nichtlinearität des Werkstoffs ($B = 2$) wird immer mehr für Stahlbetonstrukturen (Duktilität des Bauwerks) im Hinblick auf die bessere Nutzung der seismischen Sicherheitsreserven und für verschiedene Legierungen untersucht.

Hinsichtlich der geometrischen und statischen Nichtlinearität scheint infolge der Komplexität der mathematischen Aspekte die Berücksichtigung

der einen und die Vernachlässigung der anderen nicht gerechtfertigt. Daher sind wohl die Kodes 1122, 1212, 1221, 1222 nicht interessant, ihre zukünftige Anwendung ist aber nicht ausgeschlossen.

Sicher wird die weite Verbreitung der Rechenautomaten die Theorien der Baumechanik auf eine neue Entwicklungsstufe heben.

Auf die bestehenden Einteilungen zurückkommend, sei der Zusammenhang zwischen denselben und der hier vorgeschlagenen Kodierung hervorgehoben:

— mechanische Linearität	1BCD	(A = 1)
— mechanische Nichtlinearität	2BCD	(A = 2)
— physikalische Linearität	A1CD	(B = 1)
— physikalische Nichtlinearität	A2CD	(B = 2)
— geometrische Linearität	AB1D	(C = 1)
— geometrische Nichtlinearität	AB2D	(C = 2)
— Theorie 1. Ordnung	1111	
— Theorie 2. Ordnung	2111	
— Stabilitätsprobleme	2111	
— Theorie 3. Ordnung	2121 u. 2122	

Daraus lassen sich einige Schlußfolgerungen ziehen.

1° Diese Kodierung deckt alle bestehenden Einteilungen, ist allgemeiner und gibt direkte Hinweise für die Grundfaktoren, auf die sich eine Theorie stützt.

2° Sie zeigt die bestehenden Lücken und die Wege, die zu einer verfeinerten Theorie führen.

3° Sie deckt die innere Verknüpfung zwischen dem Stabilitätsproblem und der Theorie 2. Ordnung auf, wobei das erstere als ein Sonderfall der letzteren erscheint.

4° Die vorgeschlagene Kodierung läßt die Fragen der Stabilität in ihrem wahren Licht erscheinen; sie können durch den mathematischen Aspekt von Eigenwertproblemen erklärt werden, die eine Vermeidung von kritischen Werten aber kein Studium des überkritischen Verhaltens ermöglichen.

Zu bemerken wäre noch, daß ein Kode nicht einer bestimmten Aufgabe entspricht, sondern eine Aufgabenkategorie kennzeichnet, die auf den gleichen ABCD-Hypothesen beruhen.

Danksagung

Abschließend möchte der Verfasser den Kollegen Dr.-Ing. J. APPELTAUER, Dozent Dr.-Ing. V. BĂNUT, Prof. Dr.-Ing. I. BELEŞ, Dozent Dr.-Ing. I. Gh. BOCĂ, Prof. Dipl.-Ing. A. CARACOSTEA, Prof. Dr.-Ing. V. ILLE, Dozent Dr.-Ing. V. PETCU, Prof. Dr.-Ing. F. GOBESZ, und Prof. Dr.-Ing. habil. P. P. TEODORESCU für Vorschläge und Beurteilungen im Zusammenhang mit diesen Ausführungen herzlichst danken.

SCHRIFTTUM

1. APPELTAUER, J.: Lehrgang der Festigkeitslehre für Bauwesenstudenten. I. Grundlagen. Inst. Politechnic »Traian Vuia« Timisoara, 1973
2. BĂNUȚ, V.: Berechnung von elastischnichtlinearen Strukturen. *Studii si cercetări de mecanică aplicata* (rumänisch), **36** (1977), 635—646
3. BĂRSAN, G. M.: Baudynamik und Baustabilität (in rumänischer Sprache). Ed. didactic și pedagogică, București 1980
4. BELEȘ, A. A.—SOARE, V. M.: Fragen der elastischen Stabilität in Stahlbauten. Betrachtungen über die Festlegung der Knickbeiwerte (in rumänischer Sprache). *Standardizarea*, **10** (1958), 7—20, 61—74
5. BELEȘ, A. A.—VOINEA, R. P.: Festigkeitslehre für Bauingenieure. II. Ed. Technică, București, 1958, S. 307
6. BERNARD, A.—FREY, F.—JANSS, J.—MASSONNET, CH.: Recherches sur le comportement au flambemenent de barres en aluminium. *Mémoires de l'Association Internationale de Ponts et Charpentes*, **33**-I, Zürich 1973, S. 1—3
7. BRUSH, D. O.— ALMROTH, B. O.: Buckling of Bars, Plates and Shells. McGraw-Hill, Inc., New York, 1975
8. BUBENHEIM, H. J.: Beitrag zur Ermittlung des Schnittkraftverlaufs in räumlichen Stabwerken unter Berücksichtigung nichtlinearelastischer Werkstoffgesetze. *Beton- und Stahlbetonbau* **69** (1974), 109—117
9. CHWALLA, E.: Einführung in die Baustatik. Neudruck der zweiten, verbesserten und ergänzten Auflage, Stahlbau-Verlag-GMBH, Köln (Rhein), 1954
10. CORNELIUS, W.: Stabilitätsproblem-Festigkeitsproblem. *Die Bautechnik* **26** (1949), 257—261
11. FALKS, S.: Die linearisierte Theorie dritter Ordnung des geraden elastischen Balkens. *Z.A.M.M.*, **55** (1975), 79—81
12. FALKS, S.: Private Mitteilung
13. FRANZ, G.: Konstruktionslehre des Stahlbetons. 2ter Bd. Tragwerke. Springer-Verlag Berlin, Heidelberg, New York, 1969, 279—287.
14. GHEORGHIU, AL.: Statik, Stabilität und Dynamik der Bauten (in rumänischer Sprache). 2te Aufl., Ed. didactică și pedagogică, București 1974
15. GIONCU, V.: Thin Reinforced Concrete Shells. Special Analysis Problems. John Wiley, London, 1979
16. GOLDBERG, J. E.—RICHARD, R. M.: Analysis of Nonlinear Structures. *Journal of the Structural Division, Proceedings of the ASCE*, **89**, No. St 4, 333—351 (1963)
17. HAWRANEK, A.—STEINHARDT, O.: Theorie und Berechnung der Stahlbrücken. Springer-Verlag, Berlin (Göttingen) Heidelberg 1958
18. HORT, W.—THOMA, A.: Die Differentialgleichungen der Technik und Physik (7te Aufl.), Johann Ambrosius Barth Verlag, Leipzig 1956
19. ILLE, V.: Festigkeitslehre. Bd. II. Inst. Politehnic Cluj-Naposa, 1977
20. KEINZEL, E.: Vereinfachte Berechnung von Stockwerkrahmen nach der Spannungstheorie II und höher Ordnung. *Die Bautechnik* (1969), H. 9
21. KRÖPLIN, B.-H.: Beulen ausgesteifter Blechfelder mit geometrischer und stofflicher Nichtlinearität. Bericht Nr. 77—22 aus dem Institut für Statik der Technischen Universität Braunschweig 1977
22. LENSING, J.: Die verallgemeinerte linearisierte Theorie dritter Ordnung des geraden elastischen Balkens. Dissertation, Technische Universität Carolo-Wilhelmina zu Braunschweig, Juli 1976
23. MASSONNET, GH.: Résistance des Matériaux. 2ème édition. Sciences et Lettres-Liège, **1** (1967), **2** (1974)
24. MASSONNET, CH.: Comportement à la ruine des plaques raidies comprimées. Théorie et expérience. Comptes rendus du quatrième congrès canadien de mécanique appliqué, Montréal, 28 mai—1er juin 1973, p. G 51—G 74
25. MÜLLER, H.: Differentialgleichungen orthotroper Platten im geometrisch und beschränkt physikalisch nichtlinearen Bereich. *Z.A.M.M.* **44** (1967), 539—557
26. PFLÜGER, A.: Stabilitätsprobleme der Elastostatik. Dritte neubearbeitete Auflage. Springer-Verlag Berlin/Heidelberg/New York 1975
27. POROV, E. P.: Introduction to Mechanics of Solids. Prentice-Hall, Inc. Englewood Cliffs, N. J., 1968
28. ROSMAN, R.: Beitrag zur Berechnung von Stockwerkrahmen nach der Spannungstheorie II. und III. Ordnung mit Hilfe des stellvertretenden Kragträgers. *Die Bautechnik* **39** (1962), 275—279

29. SOARE, M. V.: Festigkeitslehre. Lehrgang und Anwendungen, Bd. 1 u. 2, Institut für Bauwesen, Bukarest, 1978.
30. SOARE, M. V.: Die Beulung ebener Platten als Problem großer Verformungen. *Studii şi cercetări de mecanică aplicată* (in rumänischer Sprache), **37** (1978), 547—562
31. STEINBACH, W.: Die Theorie 2. Ordnung für den räumlich belasteten Stab mit dünnwandigem, offenem Querschnitt. Stahlbau und Baustatik. Aktuelle Probleme. Springer-Verlag, Wien, 1965, S. 292—335
32. TIMOSHENKO, ST. P.—YOUNG, D. H.: Theory of Structures. 2nd Ed., McGraw-Hill Book Company, New York 1965, 221—3
33. TIMOSHENKO, ST. P.—GERE, J. M.: Theory of Elastic Stability. 2nd Ed. McGraw-Hill Book Company, New York 1961
34. YLINEN, A.: A Method of Determining the Buckling Stress and the Required Cross-Sectional Area for Centrally Loaded Straight Columns in Elastic and Inelastic Range. Publications of the *International Association for Bridge and Structural Enginering*, **16** (1956) 529—550
35. WOLMIR, A. S.: Biegsame Platten und Schalen (Deutsche Bearbeitung A. Duda). VEB Verlag für Bauwesen, Berlin 1962.

A Proposal of Codification of Theories in the Theory of Structures. — Starting from the statement that there is no systematic classification of theories in the theory of structures, a codification is proposed based on four main factors and namely: the manner of expressing the equilibrium of the structure, the constitutive law of the material, the magnitude of deformations and the curvature expressions. The paper is concluded with examples of codification of some problems in the theory of structures.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (3–4), pp. 221–245 (1980)

NOTES ON THE FIELD EQUATIONS WITH STRESSES AND ON THE BOUNDARY CONDITIONS IN THE LINEARIZED THEORY OF ELASTOSTATICS

I. KOZÁK*
CANDIDATE OF TECHN. SCI.

[Manuscript received March 14, 1979]

The necessary and sufficient conditions of the single-valued displacements can be determined from the strain field, that is the six Saint-Venant's compatibility equations are not independent of each other. To eliminate the interdependence, K. WASHIZU in 1957 has given a solution by allowing for Bianchi's identity. In this paper, on the basis of the condition of the single-valued displacements concerning the boundary (i. e. on the base of the compatibility boundary condition) similarly with the aid of Bianchi's identity, the solution given by K. WASHIZU is extended to all feasible cases and, at the same time, the results are generalized to any curvilinear coordinate systems. In this way, the set of the governing equations of the linearized theory of the elastostatics, in connection with the six stress coordinates, is composed of the three equilibrium equations and of three suitably selected compatibility equations, however, the customary boundary conditions should be complemented by three compatibility boundary conditions.

1. Introduction

1.1 The set of the governing equations of the linearized theory of elastostatics is, in the system of the curvilinear coordinates x^1, x^2, x^3, if no couple stresses are applied, and $\mathbf{T} = t^{kl}\,\mathbf{g}_k\,\mathbf{g}_l$ is the stress field, $\mathbf{A} = a_{kl}\mathbf{g}^k\,\mathbf{g}^l$ is the strain field, $\mathbf{u} = u_k\mathbf{g}^k$ is the displacement field, further, $\mathbf{q} = q^k\,\mathbf{g}_k$ is the system of the volume forces, as follows:

$$t^{kl}_{\cdot\cdot;\,l} + q^k = 0\,, \tag{1.1}_1$$

$$\varepsilon_{klm}\,t^{kl} = 0\,, \tag{1.1}_2$$

$$a_{kl} = c_{klpq}\,t^{pq}, \tag{1.2}$$

$$2a_{kt} = u_{k;\,l} + u_{l;\,k}\,. \tag{1.3}$$

The subscript following the semicolon is the sign of the covariant derivative, and summing should be carried out according to the same subscripts and superscripts, furthermore;

ε_{klm} = covariant permutation tensor;
c_{klpq} = coefficient tensor of constitutive equations (which involves at most 21 independent material behaviour constants).

* KOZÁK, I. Dózsa Gy. u. 14, H-3525 Miskolc, Hungary

Let us denote the region of the space occupied by the solid with V, its boundary with S (Fig. 1), then the boundary conditions should be as follows:

$$t^{kl} n_l = p^k \qquad x \in S_t \tag{1.4}_1$$

$$u_k = \tilde{u}_k \qquad x \in S_u \tag{1.4}_2$$

in which
n_k = normal vector directed outwards of the solid;
p_k = load applied at boundary section S_t;
$\tilde{u}_k$ = displacement prescribed at boundary section S_u;
x = denotes all three coordinates of the points and
$S_t \cup S_u = S$.

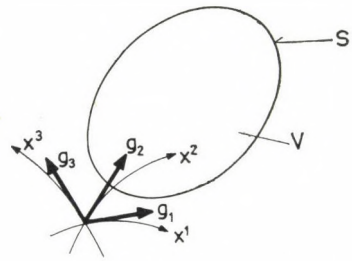


Fig. 1

In the following, it is assumed that the field of displacements satisfies the following conditions of normalization:

$$\int_V \mathbf{u}\, dV = 0\,, \tag{1.5}_1$$

$$\int_V \mathbf{r} x \mathbf{u}\, dV = 0 \tag{1.5}_2$$

wherein $\mathbf{r}$ is the position vector.

The conditions (1.5) eliminate the displacements which differ from each other only by the character of a rigid-body motion of the solid.

It is known that the boundary-value problems (1.1)—(1.4) of the linearized theory of elastostatics completed by the conditions (1,5) yield a single-valued solution to the fields $u_k(x)$, $t^{kl}(x)$ and $a_{kl}(x)$.

The solution of the actual problems may be found in two different ways.

According to the first such field equations should be established from the set of the governing equations (1.1)—(1.3) in which only the coordinates of the displacement should be included. Navier's equation for the field $u_k(x)$ is to be found in this way from the unique solution to which, in case of given boundary conditions, according to Eqs (1.3) and (1.2) also single-valued strain and stress fields may be produced.

According to the second method of solution such field equations are to be established wherein only stress coordinates enter, and the solution, to

which by using appropriate boundary conditions, is single-valued to the stress field and also to the field of displacements to be deduced from the former.

The constitutive Eq. (1.2) furnishes an unambiguous interconnection between the fields $a_{kl}(x)$ and $t^{kl}(x)$. In this paper, from those mentioned, always the stress field is considered which is to be investigated whereas the strain field as well as the incompatibility tensor field are always considered as being expressed by the stress field, making use of the formula (1.2).

Thus, if the strain field and by that the stress field are known, the displacements may be determined by making use of Cesaro's formula [1], [2: pp. 63—64]. In the solid, bounded by one single connected surface, the necessary and sufficient condition of the single-valued displacements, which is determined in the previously written way and satisfies the conditions (1.5), is the fulfilment of six Saint-Venant's compatibility equations [1], [3: pp. 25—29], [2: p. 64], [4: pp. 188—190].

The shapes of the compatibility equations expressed with the aid of the stress coordinates in the case of the homogeneous isotropic solid, by considering the equilibrium equations, are the so-called Beltrami—Michell's compatibility equations [5], [6].

In connection with the second way for finding the solutions, three problems arise which are correlated to each other.

I. For the six coordinates of the symmetrical stress tensor three equilibrium and six compatibility equations are available. Thus, the question is, which are the three Saint-Venant's compatibility equations, written down with the stresses, which should be used together with the three equilibrium equations if only six field equations are wanted for treating the problem.

II. What are the appropriate boundary conditions?

III. The theorem of minimum complementary energy declares that from among the statically admissible stress fields of a boundary-value problem, in case of the actual (i.e., also kinematically possible) solution, the complementary energy has a minimum value. Thus, it may be anticipated that from the theorem of minimum complementary energy Saint-Venant's six compatibility equations might be derived as the necessary and sufficient condition of the single-valued displacements determined from the stress field. On the other hand, R. V. Southwell [7: pp. 211—217] in investigating the solid bounded by one single closed surface, found that based either upon Maxwell's [8] or Morera's [9] stress functions, the theorem of minimum complementary energy resulted only three compatibility equations. Therefore, the question arises what is the solution to the problem described in the foregoing, known as "Southwell's paradox".

1.2 K. Washizu points out in the Cartesian system of coordinates on the basis of Bianchi's identically fulfilled formula of Riemann's geometry [10] that if the three compatibility equations, which are in connection with

Maxwell's stress functions (with Morera's stress functions), followed by the theorem of minimum complementary energy as was written by SOUTHWELL, will be satisfied inside the solid bounded by a single, simply connected surface, and the other three at the boundary of the solid, then, also the other three equations will be fulfilled within the solid. By that, WASHIZU furnishes solutions to all three problems.

1.3 This paper considers the field equations of the linearized theory of elastostatics established by exclusively using the stresses, as well as the associated boundary conditions for the case of a solid bounded by a simply connected surface. The necessary and sufficient field equations and boundary conditions associated with the stress field are determined assuring the uniqueness of the stress field as well as that of the displacement field to be produced from the stress field by Cesaro's formula, according to the relationships (1.5). By that, the paper complements and generalizes Washizu's solution described under point 1.2.

Under point 2 the linearized equations related to the displacement, rotation and strain fields of the continua, as well as all the possible shapes of the stress functions, are explained and summarized.

Under point 3 the conditions of the single-valued displacements to be produced on the surface from the stress field are then determined, under point 4 Washizu's solution reported in [10] is completed, whereafter, all the possible cases are pointed out, the compatibility boundary conditions are determined and the results generalized to an arbitrary system of curvilinear coordinates.

Under point 5, on the basis of the foregoing, the necessary and sufficient field equations and the boundary conditions of the boundary-value problem formulated to the stress field of the linearized theory of elastostatics, according to the concept of the first paragraph of point 1.3 are presented.

Point 6 summarizes the issues yielded by the points 3 to 5.

Point 7 presents, as an example, the application of the results in the system of Cartesian coordinates.

1.4 It is assumed that the considered solid is homogeneous, and occupies, in the way already mentioned, the region V of space and its boundary S is a simply connected closed surface. The system of volume forces may be arbitrary.

In the paper it is pointed out that the necessary three compatibility field equations may be selected depending on the applied system of coordinates, for reasons of generalization, in the following, the system of coordinates will be assumed to be arbitrary and curvilinear (the sequence of the covariant derivations might be interchanged). The notations introduced having been retained, new symbols will be explained when they are first applied.

2. Further knowledge on linearized elastostatics used in this paper

2.1 Besides the notation used under point 1.1, the following symbols are applied:

$\mathbf{U} = u_{kl}\,\mathbf{g}^k\,\mathbf{g}^l$ — gradient of displacement field;
$\Psi = \psi_{kl}\,\mathbf{g}^k\,\mathbf{g}^l$ — rotation tensor;
∇ — Hamilton's differential operator.

The comma before the subscript designates the partial derivation, the point between the values, the scalar, and the italicized cross-sign, the vectorial product.

In case of small displacements and strains, the following linearized field equations are valid:

$$\mathbf{U} = \mathbf{u}\nabla = \frac{\partial \mathbf{u}}{\partial x^l}\,\mathbf{g}^l = \mathbf{u}_{,l}\,\mathbf{g}^l \qquad u_{kl} = u_{k;l} \tag{2.1}$$

$$\mathbf{U} = \mathbf{A} + \Psi \qquad u_{kl} = a_{kl} + \psi_{kl} \tag{2.2}$$

$$\mathbf{A} = \frac{1}{2}(\mathbf{u}\nabla + \nabla\mathbf{u}) \qquad a_{kl} = \frac{1}{2}(u_{k;l} + u_{l;k}) \tag{2.3}$$

$$\Psi = \frac{1}{2}(\mathbf{u}\nabla - \nabla\mathbf{u}) \qquad \psi_{kl} = \frac{1}{2}(u_{k;l} - u_{l;k}) \tag{2.4}$$

further, from Eqs (2.3) and (2.4) by the interchangeability of the sequence of the covariant derivations

$$\psi_{kl;m} = a_{km;l} - a_{lm;k} \tag{2.5}_1$$

may be obtained, wherein

$$\Psi\nabla = \Psi_{,m}\,\mathbf{g}^m = \psi_{kl;l}\,\mathbf{g}^k\,\mathbf{g}^l\,\mathbf{g}^m = d_{klm}\,\mathbf{g}^k\,\mathbf{g}^l\,\mathbf{g}^m \tag{2.5}_2$$

is the gradient of the rotation field.

Eq. (2.5) is, as a matter of fact, the condition that the gradient of the rotation field is a total derivative.

2.2 Be the arbitrarily selectable points O and P of the solid connected by an arbitrary curve h defined by the equation $x^k = x^k(s)$, (s being a parameter measured along the curve). Along the curve (Fig. 2.) where

$$\mathbf{t} = t^k\,\mathbf{g}_k = \frac{d\mathbf{r}}{ds}, \quad t^k = \frac{dx^k}{ds} \tag{2.6}$$

is the tangent vector and $d\mathbf{r} = \mathbf{t}\,ds$, the derivative of the displacement field:

$$d\mathbf{u} = \mathbf{U}\cdot d\mathbf{r} = \mathbf{U}\cdot\mathbf{t}\,ds. \tag{2.7}$$

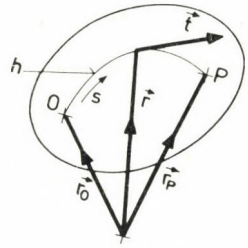


Fig. 2

By taking account of Eqs (2.2) and (2.6) and carring out identic transformations, the following may be obtained:

$$d\mathbf{u} = \mathbf{A} \cdot \mathbf{t}\, ds + \psi \cdot \frac{d\mathbf{r}}{ds} ds = \\ = \mathbf{A} \cdot \mathbf{t}\, ds + \left\{ - \frac{d}{ds} [\psi \cdot (\mathbf{r}_P - \mathbf{r})] + \frac{d\psi}{ds} \cdot (\mathbf{r}_P - \mathbf{r}) ds \right\} \tag{2.8}_1$$

in which $\mathbf{r}_P$ is the position vector of the fixed point P.

The derivative of the rotation field may be produced by making use of formula (2.5):

$$\frac{d\psi}{ds} = (\psi \nabla) \cdot \mathbf{t} = (a_{km;l} - a_{lm;k})\, \mathbf{g}^k\, \mathbf{g}^l\, \mathbf{g}^m \cdot \mathbf{t}\,. \tag{2.8}_2$$

The integral of Eq. (2.7) along the curve h from O to P, also by taking into account the transformations made in Eq. (2.8), yields for the calculation of the displacement $\mathbf{u}_P$ of the point P Cesaro's known formula [1], [2: pp. 63—64]:

$$\mathbf{u}_P = \mathbf{u}_0 + \int_h \mathbf{U} \cdot \mathbf{t}\, ds = \mathbf{u}_0 + \psi_0 \cdot (\mathbf{r}_P - \mathbf{r}_0) + \int_h \mathbf{B} \cdot \mathbf{t}\, ds\,, \tag{2.9}_1$$

$$\mathbf{B} = b_{km}\, \mathbf{g}^k\, \mathbf{g}^m = [a_{km} + (a_{km;l} - a_{lm;k})\, \mathbf{g}^l \cdot (\mathbf{r}_P - \mathbf{r})]\, \mathbf{g}^k\, \mathbf{g}^m\,. \tag{2.9}_2$$

Since in the formula (2.9), P may be any of the points of the solid, the displacement field of the solid may be produced with the aid of Cesaro's formula in the knowledge of the strain field $\mathbf{A}(x)$, provided at a point (for example in point O) the displacement $\mathbf{u}_O$ and the rotation tensor ψ_0 are known. From the formula (2.9) one can read that with one single strain field infinitely many displacement fields are associated which, however, may only differ from each other in the forms of the rigid-body-like motion of the solid described by the expression $\mathbf{u}_O + \psi_O \cdot (\mathbf{r}_P - \mathbf{r}_O)$. The displacement field produced with

the aid of Cesaro's formula are, according to the conditions of normalization (1.5), single-valued.

2.3 Having the knowledge of the strain field, $\mathbf{u}_O$ and ψ_O Cesaro's formula yields the displacement of the point P, independent of the selection of the curve h between the points O and P, provided that in the region of the curves h which may come into question, for the integrand of the integral, along the curve $(2.9)_1$ the equation

$$\mathbf{B} \times \nabla = \mathbf{B}_{,n} \times \mathbf{g}^n = b_{km;\,n}\, \varepsilon^{mns}\, \mathbf{g}^k\, \mathbf{g}_s = 0 \tag{2.10}$$

will be satisfied. Having in mind the derivative

$$(\mathbf{r}_P - \mathbf{r})_{,n} = -\mathbf{r}_{,n} = -\mathbf{g}_n$$

as well as the rule of derivation of a product, from the condition (2.10) the substitution of the right-hand side of Eq. $(2.9)_2$ yields

$$b_{km;\,n}\, \varepsilon^{mns} = [a_{km;\,n} + (a_{km;\,l} - a_{lm;\,k})(-\delta^t_n)]\, \varepsilon^{mns} + \\ + (a_{km;\,ln} - a_{lm;\,kn})\, \varepsilon^{mns}\, \mathbf{g}^l \cdot (\mathbf{r}_P - \mathbf{r}) = 0\,.$$

From the fact, that the first term of the expression obtained identically yields zero, it follows that also the second term must be equal to zero. This condition will be satisfied independent of the vector $(\mathbf{r}_P - \mathbf{r})$ provided the equation

$$(a_{km;\,ln} - a_{lm;\,kn})\, \varepsilon^{mns} = 0$$

is true.

The fulfilment of the above equation, i.e., of the equation

$$\varepsilon^{akl}\, \varepsilon^{bmn}\, a_{km;\,ln} = 0 \qquad x \in V \tag{2.11}$$

to be produced from the former is the condition of the single-valued field of displacements $\mathbf{u}(x)$ in the region V in case of given $\mathbf{u}_O$, ψ_O and $\mathbf{A}(x)$. Eqs (2.11) are the so-called Saint-Venant's compatibility equations (or the conditions of integrability).

On the basis of Eq. $(2.5)_1$

$$\varepsilon^{akl}\, a_{km;\,l} = \frac{1}{2}\, \varepsilon^{akl}\, \psi_{kl;\,m}$$

may be written, thus, in case of the validity of $(2.5)_1$ one can make sure of the identical satisfaction of the compatibility equations (2.11):

$$\varepsilon^{akl}\, \varepsilon^{bmn}\, a_{km;\,ln} = \frac{1}{2}\, \varepsilon^{akl}\, \varepsilon^{bmn}\, \psi_{kl;mn} \equiv 0\,. \tag{2.12}$$

In other words, Eqs $(2.5)_1$ may also be considered as compatibility equations (or as conditions of integrability).

In the knowledge of the field of displacements, for the strain field formed according to Eq. (2.3), the compatibility equations are identically fulfilled. On the contrary (i.e., without the knowledge of the displacements), the compatibility equations represent conditions made against the strain field.

2.4 Let us denote the so-called incompatibility tensor by the terms

$$\mathbf{E} = e^{ab}\,\mathbf{g}_a\,\mathbf{g}_b\,;\quad e^{ab} = \varepsilon^{akm}\,\varepsilon^{blp}\,a_{kl;mp}\,. \tag{2.13}_1$$

The incompatibility tensor is symmetrical, its six scalar coordinates are as follows (g being the determinant of the metric tensor):

$$\begin{aligned}
g\,e^{11} &= a_{22;33} + a_{33;22} - 2\,a_{23;23},\\
g\,e^{22} &= a_{33;11} + a_{11;33} - 2\,a_{31;31}\,,\\
g\,e^{33} &= a_{11;22} + a_{22;11} - 2\,a_{12;12}\,,\\
g\,e^{12} &= a_{23;31} + a_{31;23} - a_{33;12} - a_{12;33}\,,\\
g\,e^{23} &= a_{31;12} + a_{12;31} - a_{11;23} - a_{23;11}\,,\\
g\,e^{31} &= a_{12;23} + a_{23;12} - a_{22;31} - a_{31;22}
\end{aligned} \tag{2.13}_{2-7}$$

Let the Saint-Venant's compatibility equations be fulfilled, so the incompatibility tensor vanishes, $e^{ab} = 0$. Eqs $e^{ab} = 0$ are not independent of each other, because from the interpretation of $(2.13)_1$ it is easy to recognize that Bianchi's formula will be identically satisfied:

$$\mathbf{E}\cdot\nabla \equiv \mathbf{0},\quad e^{ab}_{\cdot\cdot;b} = \varepsilon^{akm}\,\varepsilon^{blp}\,a_{kl;mpb} \equiv 0\,. \tag{2.14}$$

The consequences of identity (2.14) are dealt with under paragraph 4.

2.5 Let us now dwell on the role and structure of the stress functions.

In connection with the three-dimensional problems, in the Cartesian system of coordinates, stress functions have been introduced by MAXWELL [8] and MORERA [9] each in a different way. BELTRAMI [5] pointed out that Maxwell's and Morera's stress functions are only special cases of a generalized solution. GWYTHER [11] adapted Beltrami's generalized solution to the orthogonal system of curvilinear coordinates. The tensorial (invariant) form of the stress functions has separately been established by B. FINZI [12] and V. I. BLOKH [13]. M. E. GÜRTIN [14] verified that the production of the stress field according to Beltrami is complete only in case where the external systems of forces applied at each closed boundary surface of the continuum are in themselves separately in equilibrium. In another case, Beltrami's solution must be completed by a biharmonic solution.

For the case of couple stresses, the stress functions were found by W. GÜNTHER [15], then, Günther's solution was generalized by D. E. CARLSON [16] who also furnished a complete solution for the case of the couple stresses.

In this paper only the case of the continuum bounded by a single closed surface, without couple stresses is investigated. For this case, by existence of the potential of the volume system of forces, the complete solution is Eq. (2.15).

Assuming that the stress functions are the scalar coordinates of the symmetrical tensor field

$$\mathbf{F}(x) = f_{xy}\, \mathbf{g}^x\, \mathbf{g}^y$$

to be differentiated sufficiently several times, and the scalar field $\Phi(x)$, according to the relationship

$$q^k = -\Phi_{,s}\, g^{sk}$$

is the potential of the volume system of forces, the equilibrium Eqs (1.1) will identically be satisfied if the stress tensor is calculated according to the formula

$$t^{kl} = \varepsilon^{kxp}\, \varepsilon^{lyq}\, f_{xy;\,pq} + \Phi g^{kl}\,. \tag{2.15}$$

In case of a non-potential volume system of forces, in lieu of the second term at the right-hand side of the formula (2.15) a so-called biharmonic equilibrium stress field should be applied [14].

FINZI has already drawn the attention to the fact that in the formula (2.15) the value of t^{kl} does not change if instead of f_{xy} another stress function

$$f'_{xy} = f_{xy} + \frac{1}{2}(w_{x;y} + w_{y;x}) \tag{2.16}$$

is introduced by making use of an arbitrarily chosen vector field $\mathbf{w}(x) = w_k \mathbf{g}^k$. On the basis of Eq. (2.16) it is possible to set the three coordinates of f_{xy} equal to zero by conveniently selecting the coordinates w_k. Let us designate the three non-zero coordinates by f'_{XY} and the three coordinates equated to zero by f'_{AB} (also the subscripts written in capitals may take the values 1, 2, 3, however, in not every possible variation). For the latter case the equation

$$f'_{AB} = f_{AB} + \frac{1}{2}(w_{A;\,B} + w_{B;\,A}) = 0 \tag{2.17}$$

is valid. The subscripts A, B should be chosen in such a way that Eq. (2.17) should have solutions to all the three coordinates of vector $\mathbf{w}$ at the values of f_{AB} considered as being arbitrary.

Thus, according to Eqs (2.16) and (2.17), also in a completely generalized case is sufficient to assume such a symmetrical tensor field $\mathbf{F}(x)$ as the tensor of the stress functions of which only three independent scalar coordinates are not equal to zero. For example, in case of Maxwell's solution f_{11}, f_{22} and f_{33}, and in case of Morera's solution $f_{12} = f_{21}$, $f_{23} = f_{32}$ and $f_{31} = f_{13}$ are the non-zero coordinates. The feasible structures of the stress functions are investigated by BLOKH [13].

Below, at a possible selection of the stress functions XY are the pairs of subscripts of the non-zero coordinates of the tensor $\mathbf{F}$, and AB that of the zero coordinates of the tensor $\mathbf{F}$.

3. Condition of the single-valued displacements to be produced on the surface by making use of Cesaro's formula

3.1 Be the surface $\bar{S}$ described by the equation $x^k = x^k(\xi^1, \xi^2)$, in which ξ^1 and ξ^2 are the parameters of the surface, fully situated inside the region V or at its boundary S (Fig. 3).

According to the formula (2.9) the Cesaro's formula gives a single-valued displacement field on the surface $\bar{S}$ if for any closed curve h to be considered on the surface $\bar{S}$, the equation

$$\oint_h \mathbf{B} \cdot \mathbf{t}\, ds = 0$$

is true.

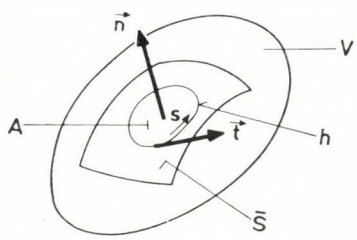


Fig. 3

Hence, by transforming by integration according to Stokes's theorem (by designating the surface region bounded by the curve h with the symbol A, and the transposed tensor with the symbol T) we have

$$\mathbf{0} = \oint_h \mathbf{t} \cdot \mathbf{B}^T\, ds = \int_A (\mathbf{n}\, x\, \nabla) \cdot \mathbf{B}^T\, dA = -\int_A (\mathbf{B}\, x\, \nabla) \cdot \mathbf{n}\, dA\,,$$

which means that in case of an arbitrary h and an arbitrary A the condition

$$(\mathbf{B} \times \nabla) \cdot \mathbf{n} = \mathbf{0}\,, \qquad x \in \bar{S} \tag{3.1}$$

should be valid. By carrying out the calculations according to paragraph 2.3, from (3.1) one obtains

$$\mathbf{E} \cdot \mathbf{n} = 0\,, \qquad x \in \bar{S} \tag{3.2}_1$$

$$e^{ab}\, n_b = \varepsilon^{akm}\, \varepsilon^{blp}\, n_b\, a_{kl;\,mp} = 0\,, \quad x \in \bar{S} \tag{3.2}_2$$

for the single-valued displacements to be produced on the surface $\bar{S}$.

The conditions (3.2) also remain valid with respect to the whole boundary S of the solid.

4. Independent compatibility conditions of the strain field

4.1 According to Bianchi's identity (2.14), the compatibility equations $e^{ab} = 0$ are not independent of each other.

WASHIZU [10] pointed out with the aid of Bianchi's formula that in the Cartesian system of coordinates by satisfying the conditions

$$e^{12} = e^{23} = e^{31} = 0\,, \qquad x \in V \tag{4.1}_1$$

and

$$e^{11} = e^{22} = e^{33} = 0\,, \qquad x \in S \tag{4.1}_2$$

also the equations

$$e^{11} = e^{22} = e^{33} = 0, \qquad x \in V \tag{4.1}_3$$

will be satisfied, i.e., inversely, by satisfying the conditions

$$e^{11} = e^{22} = e^{33} = 0, \qquad x \in V \tag{4.2}_1$$

and

$$e^{12} = e^{23} = e^{31} = 0, \qquad x \in S \tag{4.2}_2$$

also the equations

$$e^{12} = e^{23} = e^{31} = 0, \qquad x \in V \tag{4.2}_3$$

will be satisfied. Therefrom ensues that it is sufficient to consider either Eq. $(4.1)_1$ or Eq. $(4.2)_1$ as being field equations or also to complete the boundary conditions with either $(4.1)_2$ or $(4.2)_2$.

In the following, the conclusions to be drawn from Bianchi's formula will be treated in a general way by assuming an arbitrary system of curvilinear coordinates.

4.2 For this purpose let us compose the integral of the scalar product of Bianchi's identity (2.14) and the arbitrary, differentiable vector field $\mathbf{v}(x) = v_k \mathbf{g}^k$ in region V. Partial integration yields

$$0 = \int_V v_a\, e^{ab}_{\ldots;b}\, dV = \int_S v_a\, e^{ab}\, n_b\, dS - \int_V v_{a;b}\, e^{ab}\, dV \tag{4.3}$$

in which n_b is the external normal of boundary S.

Prescription of the condition of single-valued displacements (3.2) on the boundary S and also accounting for the symmetry of e^{ab} yields the equation:

$$\int_V v_{(a;\,b)}\, e^{ab}\, dV = 0. \tag{4.4}$$

The subscripts in brackets designate the symmetrical part of the gradient $v_{a;b}$.

Considering that all six coordinates of the symmetrical tensor field $v_{(a;b)}$ cannot be selected in an arbitrary way, from (4.5), it does not follow that e^{ab} vanishes at every point of V. However, three such pairs of subscripts may be selected in several ways, also depending on the system of coordinates which enable us to find a solution with respect to $v_a(x)$ of the equation

$$v_{(A;\,B)} = \frac{1}{2}(v_{A;\,B} + v_{B;\,A}) = \alpha_{AB}(x) \tag{4.5}$$

even in case of the three, completely arbitrary functions $\alpha_{AB}(x)$.

Making use of the vector field $v_a(x)$ selected according to (4.5), Eq. (4.4) may be written in the form

$$\int_V \left[\sum_{(AB)} \alpha_{AB}\, e^{AB} + \sum_{(XY)} v_{(X;\,Y)}\, e^{XY} \right] dV = 0. \tag{4.6}$$

In Eq. (4.6), the coordinates designated with the pairs of subscripts AB and XY, also taking into account the symmetry of e^{ab}, yield all of the coordinates of e^{ab}. The designations (AB) and (XY) under the summation sign mean that the summation is to be carried out according to the pairs of subscripts.

Hence, provided

$$e^{XY} = 0, \qquad x \in V, \tag{4.7}$$

is valid, considering that $\alpha_{AB}(x)$ is arbitrarily selected, from (4.6) it follows

$$e^{AB} = 0 \qquad x \in V. \tag{4.8}$$

By comparing Eqs (2.17) and (4.5), from the result (4.8) obtained by the assumption (4.7), the following statements may be made:

— assuming the coordinates $e^{XY} = e^{YX}$ of the incompatibility tensor to be zero which have the same pairs of indices as the three coordinates $f_{XY} = f_{YX}$ which, in the tensor of the stress functions may be selected to be different from zero, without invalidating the generality of the solution in the applied system of coordinates, and

— accepting the condition of the single-valued displacements, that $e^{ab}n_b = 0$ at the boundary of the solid,

— yield that the coordinates $e^{AB} = e^{BA}$ of the incompatibility tensor of the other three pairs of superscripts are equal to zero.

In short, in case of the fulfilment of conditions

$$e^{XY} = e^{YX} = 0, \qquad x \in V \tag{4.9}$$

and

$$e^{ab}\, n_b = 0\,, \qquad x \in S \tag{4.10}$$

also Eqs

$$e^{AB} = e^{BA} \qquad x \in V \tag{4.11}$$

are true; hence, it follows that Eqs (4.9) and (4.10) are equivalent with the compatibility conditions (2.11) meaning that these equations are the necessary and sufficient independent compatibility conditions of the single-valued displacements.

The XY pairs of superscripts may be selected in as various ways and in the same way as the non-zero coordinates of the tensor of the stress functions (see paragraph 2.5).

It is convenient to call the boundary conditions (4.10) "compatibility boundary conditions".

The results obtained by presenting all the possible cases, determining the compatibility boundary conditions and by the generalization of the results to arbitrary systems of curvilinear coordinates, mean the completion of WASHIZU's results mentioned in paragraph 4.1.

5. Field equations and boundary conditions of the linearized elastostatics with stresses

5.1 On the basis of the material presented in the preceding sections it can be stated that with the boundary-value problems of the linearized elastostatics formulated to the stress field the following field equations and boundary conditions are associated:

Field equations:

$$t^{kl}_{..;l} + q^k = 0\,, \qquad x \in V, \tag{5.1}_1$$

$$\varepsilon_{klm}\, t^{kl} = 0\,, \qquad x \in V, \tag{5.1}_2$$

$$a_{pq} = c_{pqkl}\, t^{kl}\,, \qquad x \in V, \tag{5.2}$$

$$e^{XY} = \varepsilon^{Xkm}\, \varepsilon^{Ylp}\, a_{kl;mp} = 0, \; x \in V. \tag{5.3}$$

(The pairs of superscripts XY may be selected in as many ways and in the same way as the tensor of the stress functions may be constructed.)

Boundary conditions:

$$t^{kl}\, n_l = p^k\,, \qquad x \in S_t, \tag{5.4}_1$$

$$u_k = \tilde{u}_k, \qquad x \in S_u, \tag{5.4}_2$$

$$e^{ab}\, n_b = 0\,, \qquad x \in S\,. \tag{5.5}$$

(The coordinates of the pairs of superscripts XY and AB yield all of the coordinates of e^{ab}.)

The boundary condition which is equivalent with $(5.4)_2$, and relating to the stress field, further the simpler boundary condition, which is equivalent at the boundary section S_u with that of the (5.5) will be established in paragraph 5.3.

It is to be noted that the compatibility field equations and consequently the compatibility boundary conditions may be selected in several ways, for the very same problem. It should always conveniently be applied to the nature of the problem.

5.2 The boundary-value problem of the linearized elastostatics formulated in paragraph 5.1 and completed by the conditions (1.5) produces a single-valued stress field $\mathbf{T}(x)$ and a displacement field $\mathbf{u}(x)$.

For proving this statement it is sufficient by considering that in case of fulfilment of the compatibility conditions (5.3) and (5.5) and of the conditions of normalization (1.5), Cesaro's formula puts a single-valued field $\mathbf{u}(x)$ to the field $\mathbf{T}(x)$. Thereby, the verification of the unique solution of the boundary-value problem has been led back to the known verification of the unique solution of the boundary-value problem under (1.1)—(1.4).

5.3 The boundary-value problem treated in paragraph 5.1 will only really be formulated by stresses in case where its boundary condition $(5.4)_2$ related to the displacement field of the boundary section S_u will be transformed into a boundary condition related to direct stresses.

For an easier understanding let us introduce the surface system of coordinates $\acute{x}^1$, $\acute{x}^2$, $\acute{x}^3$ designated with prime, at the boundary section S_u defined by the equation $x^k = x^k(\acute{x}^1, \acute{x}^2)$ (in which $\acute{x}^1$, $\acute{x}^2$ are the parameters of the surface), and in its immediate surrounding, in the region V_u. $\acute{x}^3$ is measured along the normal unit vector $\mathbf{n}$ (Fig. 4). The covariant basic vectors of the new system of coordinates are as follows:

$$\acute{\mathbf{g}}_\alpha = \frac{\partial \mathbf{r}}{\partial x^k} \frac{\partial x^k}{\partial \acute{x}^\alpha} = d^k_{\cdot\alpha}\, \mathbf{g}_k\,; \quad d^k_{\cdot\alpha} = \frac{\partial x^k}{\partial \acute{x}^\alpha}\,, \tag{5.6}_1$$

$$\acute{\mathbf{g}}_3 = \mathbf{n}\,. \tag{5.6}_2$$

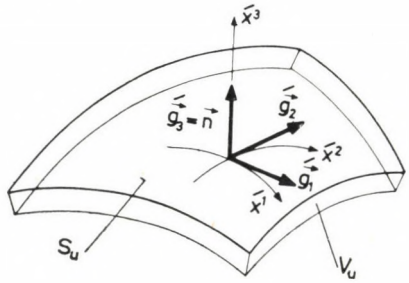


Fig. 4

(The subscripts and superscripts written in Greek letters can only take the values 1, 2.)

In the newly introduced surface system of coordinates, the boundary condition $(5.4)_2$ takes the following form:

$$\acute{u}_k(\acute{x}^1, \acute{x}^2, 0)\, \acute{\mathbf{g}}^k = \dot{\tilde{u}}_k(\acute{x}^1, \acute{x}^2)\, \acute{\mathbf{g}}^k; \qquad x \in S_u \tag{5.7}$$

In consequence of the identity (5.7) the gradients of the fields $\mathbf{u}$ and $\dot{\tilde{\mathbf{u}}}$ along S_u are equal

$$\acute{u}_{k;\beta} = \dot{\tilde{u}}_{k;\beta}, \qquad x \in S_u,$$

thus, on the basis of Eq. (2.2), and also considering Eqs (2.3) and (2.4) the equation

$$\acute{a}_{k\beta} + \acute{\psi}_{k\beta} = \dot{\tilde{u}}_{k;\beta}, \qquad x \in S_u,$$

and the equations

$$\acute{a}_{\alpha\beta} = \frac{1}{2}(\dot{\tilde{u}}_{\alpha;\beta} + \dot{\tilde{u}}_{\beta;\alpha}), \qquad x \in S_u, \tag{5.8}_1$$

$$\acute{\psi}_{\alpha\beta} = \frac{1}{2}(\dot{\tilde{u}}_{\alpha;\beta} - \dot{\tilde{u}}_{\beta;\alpha}), \qquad x \in S_u, \tag{5.8}_2$$

$$\acute{a}_{3\beta} + \acute{\psi}_{3\beta} = \dot{\tilde{u}}_{3;\beta}, \qquad x \in S_u, \tag{5.8}_3$$

respectively, are valid.

By taking into account Eq. (1.2) from $(5.8)_1$ one obtains directly the three scalar boundary conditions relating to the stress field and substituting Eq. $(5.4)_2$:

$$\acute{c}_{\alpha\beta pq}\, \acute{t}^{pq} = \frac{1}{2}(\dot{\tilde{u}}_{\alpha;\beta} + \dot{\tilde{u}}_{\beta;\alpha}) \qquad x \in S_u, \tag{5.9}_1$$

or, by returning to the stress tensor written in the original system of coordinates:

$$d^r_{.\alpha}\, d^s_{.\beta}\, c_{rsmn}\, t^{mn} = \frac{1}{2}(\dot{\tilde{u}}_{\alpha;\beta} + \dot{\tilde{u}}_{\beta;\alpha}), \qquad x \in S_u. \tag{5.9}_2$$

Eq. $(5.8)_{2,3}$ permits the simplification of the boundary condition (5.5) which have the form in the surface system of coordinates

$$\acute{e}^{a3} = 0, \qquad x \in S_u. \tag{5.10}$$

Namely, if at the surface section S_u the Eq.

$$\acute{\psi}_{kl;\gamma} = \acute{a}_{k\gamma;l} - \acute{a}_{l\gamma;k}, \qquad x \in S_u \tag{5.11}$$

according to $(2.5)_1$ will be fulfilled, so, according to (2.12) also the compatibility equation

$$\acute{e}^{a3} = \acute{\varepsilon}^{akm}\,\acute{\varepsilon}^{3\lambda\pi}\,\acute{a}_{k\lambda;\,m\pi} = \frac{1}{2}\,\acute{\varepsilon}^{akm}\,\acute{\varepsilon}^{3\lambda\pi}\,\acute{\psi}_{km;\,\lambda\pi} = 0 \tag{5.12}$$

will be satisfied. Thus, it is sufficient to assure the boundary condition (5.10) in lieu of (5.11).

By a formal calculation one can point out that in case of validity of $(5.8)_{1,2}$ the equation

$$\acute{\psi}_{\alpha\beta;\gamma} = \acute{a}_{\alpha\gamma;\beta} - \acute{a}_{\beta\gamma;\alpha}\,, \qquad x \in S_u$$

will identical be satified, thus there remains as boundary condition the equation

$$\acute{\psi}_{3\beta;\gamma} = \acute{a}_{3\gamma;\,\beta} - \acute{a}_{\beta\gamma;\,3}, \qquad x \in S_u.$$

By replacing Eq. $(5.8)_3$ one obtains

$$\acute{a}_{3\beta;\,\gamma} + \acute{a}_{3\gamma;\,\beta} - \acute{a}_{\beta\gamma;\,3} = \dot{\tilde{u}}_{3;\,\beta\gamma} \tag{5.13}_1$$

and with the consideration of (1.2) eventually yields the boundary condition

$$\begin{aligned}(\acute{c}_{3\beta pq}\,\acute{t}^{pq})_{;\gamma} + (\acute{c}_{3\gamma pq}\,\acute{t}^{pq})_{;\beta} - (\acute{c}_{\beta\gamma pq}\,\acute{t}^{pq})_{;\,3} + \acute{\Gamma}^{3}_{\beta\gamma}\,\acute{c}_{33pq}\,\acute{t}^{pq} = \\ = (\dot{\tilde{u}}_{3;\,\beta})_{,\gamma} - \acute{\Gamma}^{\sigma}_{3\gamma}\,\dot{\tilde{u}}_{\sigma;\,\beta} - \acute{\Gamma}^{\sigma}_{\beta\gamma}\dot{\tilde{u}}_{3;\,\sigma}\,, \qquad x \in S_u\,.\end{aligned} \tag{5.13}_2$$

Eq. (5.13) is symmetrical with respect to the subscripts β, γ, thus, indeed means only three scalar boundary conditions.

Satisfying the boundary conditions (5.9) and (5.13) assures the fulfilment of the boundary condition

$$\acute{u}_k = \dot{\tilde{u}}_k\,, \qquad x \in S_u\,.$$

5.4 Let us consider the following example to the boundary conditions on the boundary section S_u, which are given to the stresses.

Be the boundary section S_u of an arbitrarily loaded solid a plane surface. In the Cartesian system of coordinates in Fig. 5 $\mathbf{n} = \mathbf{g}_3$ the normal vector on S_u. The prescribed deplacement field should be arbitrary:

$$\mathbf{u}(x^1, x^2) = \tilde{u}_k(x^1, x^2)\,\mathbf{g}^k, \qquad x \in S_u\,.$$

At S_u determined according to the formula (5.8) satisfied the Eq.

$$a_{\alpha\beta}(x^1, x^2) = \frac{1}{2}\,(\tilde{u}_{\alpha,\,\beta} + \tilde{u}_{\beta;\alpha})\,,$$

whereby, with respect to the stress field, on the basis of $(5.9)_1$ the equation

$$c_{\alpha\beta pq}\, t^{pq} = a_{\alpha\beta}\ (x^1, x^2), \qquad x \in S_u\,. \tag{5.14}$$

is true.

One should observe that in (5.14) the coordinate $\overset{*}{u}_3$ does not enter, consequently, (5.14) cannot be a sufficient boundary condition. The missing boundary conditions are furnished by the equations $(5.13)_1$ assuring the single-valued displacements determined from the stresses which, written in detail, are as follows:

$$\begin{aligned} 2a_{31,1} - a_{11,3} &= \tilde{u}_{3,11}\,, \\ 2a_{32,2} - a_{22,3} &= \tilde{u}_{3,22}, \qquad x \in S_u \\ a_{31,2} + a_{32,1} - a_{12,3} &= \tilde{u}_{3,12}\,. \end{aligned} \tag{5.15}$$

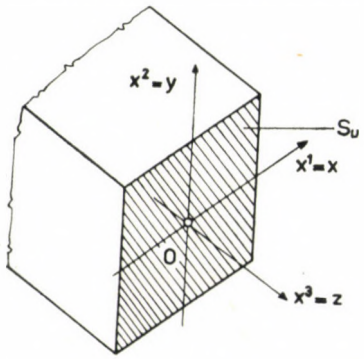


Fig. 5

5.5 The problem arises, that in connection with the known solutions of the elastostatics to the stress field why did the question not present itself under (5.3) of the selection of the compatibility equations as field equation and the need for the completion of the boundary conditions according to (5.5) which latter will be also demonstrated by the example of section 5.5.

The following answer could be given to the above question. In case of some of the solutions to the stress field of the linear elastostatics, completion of the boundary conditions and consideration of the selection of the compatibility equations as field equations are not needed because from the nature of the problem this latter obviously follows, and the compatibility boundary conditions will identically be fulfilled (for example, deformation in plane).

In other, more intricate instances, the authors avoid the problem by trying to find the solution beside the equilibrium equations by satisfying all six compatibility equations (considering, however, as a matter of course, that only three of them are independent). By proceeding in this way, the compatibility boundary conditions will identically be satisfied.

In case of the boundary-value problem formulated for the displacements, in conformity with the last paragraph of section 2.3, the compatibility boundary conditions are also identically fulfilled.

6. Conclusion

In an arbitrary system of curvilinear coordinates let us denote the field of displacements by $u_k(x)$, the strain field by $a_{kl}(x)$, and by

$$e^{ab}(x) = \varepsilon^{akm}\, \varepsilon^{blp}\, a_{kl;mp}$$

the incompatibility field.

6.1 By making use of Bianchi's formula being identically satified, the paper points out that in case of a solid bounded by a single, simply connected surface the fulfilment of the conveniently selected compatibility field equations

$$e^{XY} = 0\,, \qquad x \in V \tag{6.1}_1$$

and the compatibility boundary conditions

$$e^{ab}\, n_b = 0 \qquad x \in S \tag{6.1}_2$$

is the necessary and sufficient condition for producing by Cesaro's formula a single-valued displacement field from the field $a_{kl}(x)$ also satisfying the normalizing conditions (1.5). In the system of coordinates adopted, the pairs of indices XY may be selected in as many ways and in the same way as the non-zero coordinates of the tensor of the stress functions.

6.2 The compatibility boundary condition $(6.1)_2$ is the necessary and sufficient condition of the single-valued displacements on surface S.

6.3 The necessary and sufficient field equations and boundary conditions of the boundary-value problem of the linear elastostics formulated by stresses may be obtained in such a way that one adds to the constitutive equations and to the equilibrium equations three compatibility field equations made possible according to $(6.1)_1$, and one completes the boundary conditions (1.4) with the compatibility boundary condition $(6.1)_2$.

6.4 On the boundary section S_u where the prescription of the displacement field $\tilde{u}_k$ is the boundary condition, according to (5.9) and (5.13), the following boundary conditions are related to the stress field. If the constitutive equation being $a_{kl} = c_{klpq}\, t^{pq}$, the boundary conditions written with the coordinates $\acute{x}^1, \acute{x}^2$ and $\acute{x}^3$ variable on the surface and normal to the surface, respectively, are as follows:

$$\acute{a}_{\alpha\beta} = \frac{1}{2}\,(\acute{\tilde{u}}_{\alpha;\beta} + \acute{\tilde{u}}_{\beta;\alpha})\,, \qquad x \in S_u\,, \tag{6.3}_1$$

$$\overset{\prime}{a}_{3\alpha;\beta} + \overset{\prime}{a}_{3\beta;\alpha} - \overset{\prime}{a}_{\alpha\beta;3} + \overset{\prime}{\Gamma}{}^{3}_{\alpha\beta}\overset{\prime}{a}_{33} =$$
$$= (\overset{\vee}{u}_{3;\sigma}),_{\beta} - \Gamma^{\mu}_{3\beta}\overset{\vee}{u}_{\mu;\alpha} - \overset{\prime}{\Gamma}{}^{\mu}_{\alpha\beta}\overset{\vee}{u}_{3;\mu}, \qquad x \in S_u \qquad (6.3)_2$$

wherein $\overset{\prime}{\Gamma}{}^{m}_{ab}$ is the notation of Christoffel's symbols.

6.5 The separation of the unneeded compatibility conditions might have significance in the approximate calculations for determining the stress field.

7. Example

7.1 Let us write down in detail as an example, the equations and boundary conditions of the linearized elastostatics related to the stress field, in the Cartesian system of coordinates by assuming a solid of isotropic material bounded by a single, simply connected surface, and let us consider an actual boundary-value problem.

In this example $x^1 = x$, $x^2 = y$, $x^3 = z$ are the three coordinates, and x, y, z subscripts designating indices (they cannot obtain the values 1, 2, 3).

Let us denote the coordinates of the symmetrical stress tensor and the strain tensor as well as those of the density vector of the volume loading, respectively:

$$[a_{kl}] = \begin{bmatrix} \varepsilon_x & \frac{1}{2}\gamma_{xy} & \frac{1}{2}\gamma_{xz} \\ \frac{1}{2}\gamma_{yx} & \varepsilon_y & \frac{1}{2}\gamma_{yz} \\ \frac{1}{2}\gamma_{zx} & \frac{1}{2}\gamma_{zy} & \varepsilon_z \end{bmatrix}; \quad [t^{kl}] = \begin{bmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z \end{bmatrix}$$

$$[q^k] = [q_x \quad q_y \quad q_z] .$$

The shape of the equilibrium equations $(1.1)_1$ is

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} + q_x = 0, \ldots xyz,^{*} \qquad (7.1)_{1-3}$$

the shape of Hooke's law according to (1.2) is

$$E\varepsilon_x = \sigma_x - \nu(\sigma_y + \sigma_z), \quad \ldots xyz,$$
$$\frac{E}{2(1+\nu)}\gamma_{xy} = \tau_{xy}, \ldots xyz, \qquad (7.2)_{1-6}$$

wherein

E = Young's modulus;
ν = Poisson's ratio,

* The symbols *xyz* entering after the three points mean that the formula ahead is valid also in the case of a cyclic change of the subscripts *xyz*.

and finally, the coordinates of the incompatibility tensor according to $(2.13)_{2-7}$:

$$e^{xx} = \frac{\partial^2 \varepsilon_y}{\partial z^2} + \frac{\partial^2 \varepsilon_z}{\partial y^2} - \frac{\partial^2 \gamma_{yz}}{\partial y \, \partial z}, \ldots xyz, \tag{7.3}_{1-3}$$

$$e^{xy} = \frac{1}{2}\left(\frac{\partial^2 \gamma_{yz}}{\partial z \, \partial x} + \frac{\partial^2 \gamma_{zx}}{\partial y \, \partial z} - \frac{\partial^2 \gamma_{xy}}{\partial z^2}\right) - \frac{\partial^2 \varepsilon_z}{\partial x \, \partial y}, \ldots xyz. \tag{7.3}_{4-6}$$

The incompatibility coordinates (7.3) may be expressed by making use of Hooke's law with the coordinates of the stress, and then, also accounting for the equilibrium equations may further be independently developed. Thus, one obtains the transformed incompatibility coordinates denoted with overbars:

$$\bar{e}^{xx} = -\frac{E}{1+\nu} e^{xx} = \Delta\sigma_x - \frac{1}{1+\nu}\left(\frac{\partial^2 T_I}{\partial y^2} + \frac{\partial^2 T_I}{\partial z^2}\right) + \\ + \left(\frac{\partial q_x}{\partial x} - \frac{\partial q_y}{\partial y} - \frac{\partial q_z}{\partial z}\right), \ldots xyz, \tag{7.4}_{1-3}$$

$$\bar{e}^{xy} = -\frac{E}{1+\nu} e^{xy} = \Delta\tau_{xy} + \frac{1}{1+\nu} \frac{\partial^2 T_I}{\partial x \, \partial y} + \frac{\partial q_x}{\partial y} + \\ + \frac{\partial q_y}{\partial x}, \quad \ldots xyz, \tag{7.4}_{4-6}$$

wherein

$$T_I = \sigma_x + \sigma_y + \sigma_z \tag{7.5}$$

is the first scalar invariant of the stress tensor.

7.2 The six Saint-Venant's compatibility equations may be obtained by putting the incompatibility coordinates (7.3) (or their forms expressed by stresses) equal to zero:

$$e^{xx} = e^{yy} = e^{zz} = e^{xy} = e^{yz} = e^{zx} = 0, \tag{7.6}$$

and the Beltrami—Michell's compatibility equations may be established by setting the transformed incompatibility coordinates (7.4) equal to zero:

$$\bar{e}^{xx} = \bar{e}^{yy} = \bar{e}^{zz} = \bar{e}^{xy} = \bar{e}^{yz} = \bar{e}^{zx} = 0, \tag{7.7}$$

and, by making use of the first three equations thus obtained, one proceeds by further transformations.

It is to be noted that the usual form of Beltrami—Michell's compatibility equations may only be obtained in case where the validity of the equations $\bar{e}^{xx} = \bar{e}^{yy} = \bar{e}^{zz} = 0$ (or, $e^{xx} = e^{yy} = e^{zz} = 0$) has been assumed.

7.3 According to section 4.2, the necessary and sufficient conditions of the compatibility of the stress field may be obtained in such way that from among the incompatibility coordinates (7.3) expressed by stresses or from among the transformed incompatibility coordinates (7.4) selecting suitably three of them are set equal to zero, and considering them as compatibility field equations, and the compatibility boundary conditions according to (4.10) at the surface of the solid are prescribed:

$$n_x e^{xx} + n_y e^{yx} + n_z e^{zx} = 0, \ldots xyz ,$$

or

$$n_x \bar{e}^{xx} + n_y \bar{e}^{yx} + n_z \bar{e}^{zx} = 0, \ldots xyz \qquad (7.8)_{1-3}$$

wherein n_x, n_y, n_z define the external normal of the boundary.

The pairs of indices of the three compatibility field equations may be selected from among the six possible pairs of indices

$$xx,\ yy,\ zz,\ xy,\ yz,\ zx, \qquad (7.9)$$

on basis of section 4.2, according to the following prescription. According to Eq. (4.5) the six equations

$$\frac{\partial v_x}{\partial x} = \alpha_{xx}, \ldots xyz , \qquad (7.10)_{1-3}$$

$$\frac{1}{2}\left(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x}\right) = \alpha_{xy}, \ldots xyz , \qquad (7.10)_{4-6}$$

will be classified into groups of three (which means, in sum, twenty ways of grouping) then, it should be examined whether each of the groups of three equations have, in case of arbitrary α functions, solutions to the vector coordinates v_x, v_y, v_z. If so, the pairs of indices of functions α entering in the three equations in question should be cancelled from those included in (7.9). The remaining three pairs of indices give one of the possible variants of the pairs of indices of the compatibility field equations. Thus, for example, the system of equations

$$\frac{\partial v_y}{\partial y} = \alpha_{yy} , \quad \frac{1}{2}\left(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x}\right) = \alpha_{xy} ,$$

$$\frac{1}{2}\left(\frac{\partial v_z}{\partial x} + \frac{\partial v_x}{\partial z}\right) = \alpha_{zx}$$

have a solution for v_x, v_y, v_z also in case of arbitrary α_{yy}, α_{xy}, α_{zx}, i.e., the equations

$$e^{zz} = e^{xx} = e^{yz} = 0 \qquad (7.11)$$

are possible compatibility field equations.

A detailed consideration shows that the compatibility field equations may be selected in the following 17 ways in the Cartesian system of coordinates (written in the following, for convenience, with the incompatibility coordinates transformed):

$$\bar{e}^{xx} = \bar{e}^{yy} = \bar{e}^{zz} = 0\,,$$

$$\bar{e}^{xy} = \bar{e}^{yz} = \bar{e}^{zx} = 0\,,$$

$$\bar{e}^{xx} = \bar{e}^{yy} = \bar{e}^{xy} = 0, \ldots xyz\,,$$

$$\bar{e}^{xx} = \bar{e}^{yy} = \bar{e}^{yz} = 0, \ldots xyz\,,$$

$$\bar{e}^{xx} = \bar{e}^{yy} = \bar{e}^{zx} = 0, \ldots xyz\,,$$

$$\bar{e}^{xx} = \bar{e}^{yz} = \bar{e}^{zx} = 0, \ldots xyz\,,$$

$$\bar{e}^{xx} = \bar{e}^{xy} = \bar{e}^{yz} = 0, \ldots xyz\,. \tag{7.12}$$

K. WASHIZU [10] gave the possibilities included in the first two lines, all the other possibilities being pointed out by the present paper.

7.4 Consider the actual boundary-value problem related to a homogeneous orthogonal prism of isotropic material, represented in Fig. 6.

Be the Young's modulus and Poisson's ratio respectively

$$E, \nu; \tag{7.13}$$

the system of volume forces applied on the prism:

$$q_x = q_y = 0 \quad \text{and} \quad q_z = Ay\,, \tag{7.14}$$

wherein A = const,

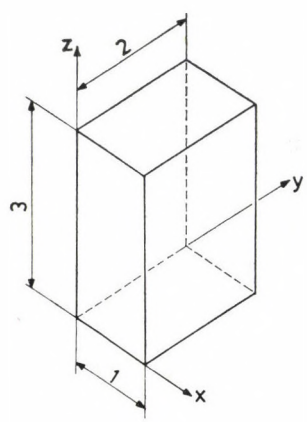


Fig. 6

and the dynamic boundary conditions (B being equally constant), on the surfaces $x = 0$ and $x = 1$:

$$\sigma_x = \tau_{xy} = \tau_{zx} = 0 ;$$

on the surface $y = 0$:

$$\tau_{xy} = \sigma_y = \tau_{yz} = 0 ;$$

on the surface $y = 2$:

$$\tau_{xy} = \sigma_y = 0 , \quad \tau_{yz} = -2A ;$$

on the surfaces $z = 0$ and $z = 3$:

$$\tau_{yz} = -\frac{A}{2} y^2 , \quad \sigma_z = B \sin \frac{\pi}{2} y . \tag{7.15}$$

The applied volume and surface systems of forces are in equilibrium in ensemble.

According to the known theorems of the theory of elasticity, from the data (7.13) to (7.15) the stresses

$$\sigma_x, \sigma_y, \sigma_z, \qquad \tau_{xy}, \tau_{yz}, \tau_{zx}$$

of the solid are uniquely ensuing as the functions of the position coordinates. According to the statement in paragraph 6.3 the following equation should be satisfied:

- as field equations
  - the equilibrium equations $(7.1)_{1-3}$ and
  - from among the equations under (7.12) three of the compatibility equations of some of the lines.
- further, as boundary conditions
  - the dynamic conditions (7.15) and
  - the compatibility conditions $(7.8)_{1-3}$.

For convenience sake, instead of trying to find the solution, it should be investigated as described above, whether the functions

$$\sigma_x = \sigma_y = \tau_{xy} = \tau_{zx} = 0 ,$$
$$\sigma_z = B \sin \frac{\pi}{2} y , \quad \tau_{yz} = -\frac{A}{2} y^2 \tag{7.16}$$

could be the solutions for the boundary-value problem considered.

By way of replacements it can be proved that the equilibrium equations $(7.1)_{1-3}$ and the dynamic boundary conditions (7.15) will be fulfilled, therefore, the answer may be given by considering the compatibility conditions.

The transformed coordinates of the incompatibility tensor, with the functions (7.16) are as

$$\begin{aligned}
\bar{e}^{xx} &= \frac{1}{1+\nu} B \frac{\pi^2}{4} \sin \frac{\pi}{2} y , \\
\bar{e}^{yy} &= 0 , \\
\bar{e}^{zz} &= -\frac{\nu}{1+\nu} B \frac{\pi^2}{4} \sin \frac{\pi}{2} y , \\
\bar{e}^{xy} &= 0 , \\
\bar{e}^{yz} &= -A + A = 0 , \\
\bar{e}^{zx} &= 0 .
\end{aligned} \qquad (7.17)_{1-6}$$

One of the possible ways to satisfy the compatibility conditions is the investigation of the field equations

$$\bar{e}^{xx} = \bar{e}^{yy} = \bar{e}^{zz} = 0 \qquad (7.18)$$

entering in the first line of the equations (7.12) and boundary conditions (7.8). However, the field equations (7.18) will only be satisfied, in conformity with the expressions (7.17), if $B = 0$. As a matter of course, in this case, also the boundary conditions (7.8) will be fulfilled.

Another possible way is to examine the field equations to be found in the second line of (7.12)

$$\bar{e}^{xy} = \bar{e}^{yz} = \bar{e}^{zx} = 0 \qquad (7.19)$$

and also the boundary conditions (7.8). In conformity with $(7.17)_{4-6}$ the field equations (7.19) are satisfied, wherefore, the problem of the compatibility will be decided by the compatibility boundary conditions. Considering the expressions (7.17) the boundary conditions according to (7.8) will be fulfilled

— on the faces $y = 0$ and $y = 2$,
— but at the faces $x = 0$ and $x = 1$, and
— at the faces $z = 0$ and $z = 3$ only in case of $B = 0$.

The very same result, $B = 0$ will be obtained if from among the possible compatibility field equations presented under (7.12) (in sum from 17 possibilities) any three and the compatibility boundary conditions (7.8) will be taken as a basis.

The result $B = 0$ means that the functions (7.16) are not the solutions to the boundary-value problem considered.

REFERENCES

1. CESARO, E.: Sulle formole del Volterra, fondamentali nella teoria della distorsioni elastiche. *Rend. Accad. Napoli*, **12** (1906), 311—321
2. ЛУРЧЕ, А. И.: Теория упругости, Наука, Москва 1970
3. SOKOLNIKOFF, I.: Mathematical Theory of Elasticity. 2nd ed., McGraw-Hill Book Company, New York 1956
4. MALVERN, L. E.: Introduction to the Mechanics of a Continuous Medium. Prentice-Hall Inc. Englewood Cliffs, New Jersey 1969
5. BELTRAMI, E.: Osservazioni sulla nota precedente (del socio Morera). *Rend. Lincei*, Ser. **5**, 1 (1892), 141—142
6. MICHELL, J. H.: On the Direct Determination of Stresses in an Elastic Solid, with Applications to the Theory of Plates. *Proc. Lond. Math. Soc.*, **31** (1899), 100—124
7. SOUTHWELL, R. V.: Castigliano's Principle of Miniumm Strain-energy and the Conditions of Compatibility for Strain. S Timoshenko 60th Anniversary Volume, The MacMillan Co., 1938
8. MAXWELL, J. C.: On Reciprocal Diagrams in Space and their Relation to Airy's Function of Stresses. *Proc. Lond. Math. Soc.* (1), 2 (1865—1869), 58—60
9. MORERA, G.: Soluzione generale delle equazioni indefinite dell'equilibrio di un corpo continuo. *Rend. Lincei*, Ser. **5**, 1 (1892), 137—141
10. WASHIZU, K.: A Note on the Conditions of Compatibility. *Journal of Mathematics and Physics* **36** (1958), 306—312
11. GWYTHER, R. R.: The Formal Specification of the Elements of Stress in Cartesian, and in Cylindrical and Spherical Polar Coordinates. *Mem. Manchester Phil. Soc.*, **56** (1912), No. 10
12. FINZI, B.: Integrazione delle equazioni indefinite della meccanica dei sistemi continui. *Rend. Lincei*, Ser. **6**, 19 (1934), 578—584, 620—623
13. БЛОХ, В. И.: Функции напряжений в теории упругости. Прикл. Мат. Мех., 14 (1950), 415—422.
14. GURTIN, M. E.: A Generalization of the Beltrami Stress Functions in Continuum Mechanics. *Arch. Rat. Mech. Anal.*, **13** (1963), 321—329
15. GÜNTHER, W.: Zur Statik und Kinematik des Cosseratschen Kontinuums. *Abh. Braunschweig. Wiss. Ges.*, **10** (1958), 195—213
16. CARLSON, D. E.: On Günther's Stress Functions for Couple Stresses. *Quartl. Appl. Math.*, **25** (1967), 139—146

Über die mit Spannungen aufgeschriebenen Feldgleichungen und Randbedingungen der linearen Elastostatik. — Die notwendige und hinreichende Bedingung der Einwertigkeit des von dem Verformungsfeld herstellbaren Verschiebungsfeldes, d.h. die sechs Sainte-Venantschen Kompatibilitätsgleichungen sind voneinander nicht unabhängig. Zur Beseitigung der Abhängigkeit gab im Jahre 1957 K. WASHIZU eine Lösung unter Berücksichtigung der Bianchischen Identität. Auch mit Hilfe der Bianchischen Identität vervollständigt die vorliegende Abhandlung die Lösung von K. WASHIZU aufgrund der Einwertigkeit des am Rand herstellbaren Verschiebungsfeldes (d.h., der Kompatibilitätsrandbedingung), zu allen möglichen Fällen und gleichzeitig verallgemeinert die Ergebnisse auf beliebige krummlinige Koordinatensysteme. Auf diese Weise besteht das Grundgleichungssystem der linearen Elastostatik in Zusammenhang mit den sechs Spannungskoordinaten von drei Gleichgewichtsgleichungen und von drei, zweckmäßig ausgewählten Kompatibilitätsgleichungen, jedoch müssen die üblichen Randbedingungen mit den drei Kompatibilitätsrandbedingungen ergänzt werden.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (3—4), pp. 247—258 (1980)

A HEURISTIC METHOD FOR SPEEDING UP MANUAL OPTIMIZATION OF BOOLEAN FUNCTIONS

A. VARGA,* P. ECSEDI-TÓTH** and F. MÓRICZ*

[Manuscript received April 26, 1978]

A simple heuristic procedure is presented to speed up manual optimization of Boolean functions (truth-functions). The procedure uses a tree representation of the Boolean functions and so it is geometric rather than algebraic in character. The method can easily be applied to totally determined Boolean functions as well as to partial ones, containing not more than 10 variables. Some illustrative examples are also included. The authors are indebted to the referee for his valuable remarks and suggestions.

Introduction

First we shall briefly describe the tree constructing procedure. A more detailed presentation can be found in [5].

If no confusion occurs, then a formula φ and the Boolean function represented by φ will not be distinguished. The set of truth-values is denoted by I; i.e. $I = \{\text{TRUE, FALSE}\} = \{0, 1\}$. Let a Boolean function φ of n variables ($n > 0$) be given. By a selection function for φ we mean an ordering of the variables occurring in φ. If a selection function for φ is fixed, then we can obtain a binary tree representation of φ by the following Construction 1 below.

If a function φ is attributed to a vertex of the tree to be constructed, then we denote the vertex by φ, too. If a function φ is expressed by a formula also containing constants (i.e. truth-values), then φ can be simplified either to a constant or to a constant-free formula by some well-known laws of the propositional calculus (such as computation laws, idempotency laws and the law of double negation). This simplification is thought to be performed, whenever it is possible. We denote by $S(\varphi)$ the variable chosen by the selection function from among the (not necessarily effective) variables of φ.

Construction 1. Let us assign a directed tree to a function φ so that it is the smallest tree fulfilling the three conditions *(i)—(iii)*:

(i) The beginning vertex of the tree is φ.

* Bolyai Institute, Szeged state university, Aradi vértanúk tere 1. 6720 Szeged, Hungary

** Research group on mathematical logic and theory of automata, Hungarian Academy of Sciences, Somogyi u. 7, 6720 Szeged, Hungary

(ii) If a vertex of the tree is a function ψ having $t(>0)$ variables, then there are exactly two outgoing edges from ψ, labelled by $S(\psi)$ and its negation $\overline{S(\psi)}$. The end vertices of these edges are the functions ψ_1, ψ_0 obtained from ψ by substituting TRUE and FALSE for $S(\psi)$, respectively. ψ_1 and ψ_0 are functions depending, at most, on $t-1$ variables.

(iii) The vertices TRUE and FALSE are not adjacent to any outgoing edge.

A more detailed explanation of the construction and a list where the referred laws can be found are given in [1].

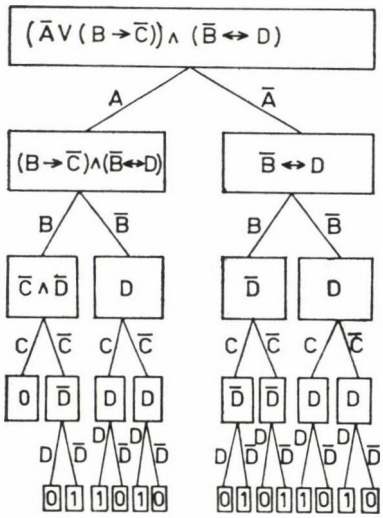


Fig. 1

We give a simple example to illustrate this procedure. Let the formula φ be $(\bar{A} \vee (B \to \bar{C})) \wedge (\bar{B} \leftrightarrow D)$ and the selection function $S = \langle A, B, C, D \rangle$ be the lexicographical ordering of the variables. The tree generated by Construction 1 is indicated in Fig. 1.

If a Boolean function φ and a selection function S are given, then the (uniquely determined) tree produced by this process will be denoted by φ_S. By the φ_S^1-tree (resp. φ_S^0-tree) we shall mean the tree that can be obtained from φ_S by deleting every edge which does not occur in a path leading from the beginning vertex to a TRUE vertex (resp. FALSE vertex) (Fig. 2).

By a path (in a φ_S-tree) we mean a sequence of literals corresponding to the edges of a path from the top to the bottom of the tree. For example, the paths of the φ_S-tree, given in Fig. 1, are

$$\langle A, B, C \rangle, \langle A, B, \bar{C}, D \rangle, \langle A, B, \bar{C}, \bar{D} \rangle, \langle A, \bar{B}, C, D \rangle, \langle A, \bar{B}, C, \bar{D} \rangle,$$

$\langle A, \bar{B}, \bar{C}, D\rangle$, $\langle A, \bar{B}, \bar{C}, \bar{D}\rangle$, $\langle \bar{A}, B, C, D\rangle$, $\langle \bar{A}, B, C, \bar{D}\rangle$, $\langle \bar{A}, B, \bar{C}, D\rangle$,
$\langle \bar{A}\ B, \bar{C}, \bar{D}\rangle$, $\langle \bar{A}, \bar{B}, C, D\rangle$, $\langle \bar{A}, \bar{B}, C, \bar{D}\rangle$, $\langle \bar{A}, \bar{B}, \bar{C}, D\rangle$, $\langle \bar{A}, \bar{B}, \bar{C}, \bar{D}\rangle$.

It is quite easy to check [5] that every path $p = \{A_1, \ldots, A_n\}$ of a given φ_S^i-tree can be associated to the minterm $A_1 \wedge \ldots \wedge A_n$ $(i = 1)$ and the maxterm $\bar{A}_1 \vee \ldots \vee \bar{A}_n$ $(i = 0)$ of φ. All other notions and notations not explained before can be found in any textbook on switching theory (e.g. [1, 3]) or in [2], [5]).

This paper is to show the heuristics proposed in the practical work, hence we leave the (very easy) proofs of lemmata to the reader. The case of partially defined Boolean functions is not presented here. The interested reader is referred to our paper [4].

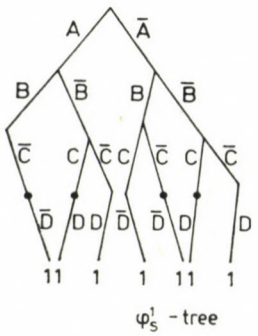


Fig. 2

1. The minimizing heuristics

1.1. In [5] we introduced the notion of the choice-point. We recall here the definition.

Definition. Let a formula φ and a selection function S be given. By a ***choice-point*** we mean a node of the φ_S^i-tree ($i \in I$, fixed) from which exactly two edges start.

The following lemma, stated and also proved in [5], shows that there is a relation between the notions of redundancy and choice-point. The proper understanding of this connection, however, is a future aim of our work. We believe that if this point is clear then we can prove that the heuristics presented here gives an irredundant formula as a result.

Lemma ([5]). Let a formula φ and a selection function S be given. If φ is redundant, then there exists a choice-point in the φ_S^1-tree and there exists a choice-point in the φ_S^0-tree.

By this Lemma the task of finding an irredundant form for a Boolean function can be composed in the following way: try to eliminate the choice-points from the tree as far as possible, by using simple tree-transforming rules. However, in general the choice-point-free version of the φ_S^i-tree cannot be achieved, for the converse of the lemma is not true. The aim of this paper is to find simple rules for eliminating choice-points from a φ_S^i-tree.

1.2 Definition. Let a formula φ and a selection function S be given. Let p_1 and p_2 be two paths in the φ_S-tree ($i \in I$, fixed). p_1 and p_2 are said to be *amicable paths* if

(i) the number of literals in p_1 is equal to the number of literals in p_2; and

(ii) p_1 and p_2 have the same but one literal, the exceptional literals, say A in p_1 and B in p_2, form a complementary pair, i.e. $\bar{A} = B$.

Note that, by definition, two amicable paths come to the same truth-value as an end vertex.

Lemma 1. Let a formula φ and a selection function S be given. Let p_1 and p_2 be amicable paths in the tree of φ obtained by Construction *1* and assume that the last common vertex in p_1 and p_2 is ψ. If the literal immediately following ψ is denoted by A_i in p_i ($i = 1, 2$), then $A_1 = \bar{A}_2$.

(Cf. Fig. 3, where α_1 and α_2 are the sequences of literals identical in p_1 and p_2.)

The following two tautologies are well-known:

$$\big((A \wedge B) \vee (\bar{A} \wedge B)\big) \leftrightarrow B\,, \tag{1}$$

$$\big((A \vee B) \wedge (\bar{A} \vee B)\big) \leftrightarrow B\,. \tag{2}$$

These tautologies, however, provide us with a simple but powerful means for eliminating certain choice-points.

Rule 1. Let a formula φ and a selection function S be given. Let

$$p_1 = \langle B_1, \ldots, B_k, A, B_{k+1}, \ldots, B_n \rangle$$

and

$$p_2 = \langle B_1, \ldots, B_k, \bar{A}, B_{k+1}, \ldots, B_n \rangle$$

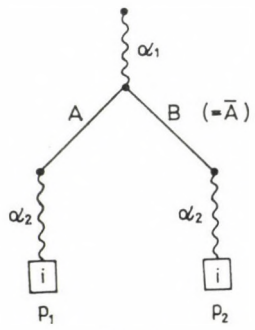


Fig. 3

be amicable paths in the φ_S^i-tree ($i = 0, 1$). Then p_1 and p_2 can be merged into one path

$$p = \langle B_1, \ldots, B_k, B_{k+1}, \ldots, B_n \rangle .$$

1.3. Definition. Let a formula φ and a selection function S be given. Let p_1 and p_2 be two paths in the φ_S^i-tree ($i \in I$, fixed). p_1 and p_2 are said to be ***quasi-amicable*** if

(i) $l_1 \leq l_2$, where l_i is the length of p_i ($i = 1, 2$),

(ii) there are $l_1 - 1$ literals in p_1 which occur in p_2,

(iii) the negation of the remaining "exceptional" literal of p_1 is present in p_2.

We note that this relation is not symmetrical.

Lemma 2. Let a formula φ and a selection function S be given. Let p_1 and p_2 be quasi-amicable paths in the tree of φ obtained by Construction 1, and assume that the last common vertex in p_1 and p_2 is ψ. If the literal immediately following ψ is denoted by A_i in p_i ($i = 1, 2$), then $A_1 = \bar{A}_2$. (Cf. Fig. 4, where $\alpha_1, \alpha_2, \alpha_3$ are sequences of the literals.)

The following two formulae are well-known tautologies of the propositional calculus:

$$(A \wedge B) \vee (\bar{A} \wedge B \wedge C) \leftrightarrow (A \wedge B) \vee (B \wedge C) , \tag{3}$$

$$(A \vee B) \wedge (\bar{A} \vee B \vee C) \leftrightarrow (A \vee B) \wedge (B \vee C) . \tag{4}$$

Our second simplification rule is based on these two tautologies:

Rule 2. Let a formula φ and a selection function S be given. Let

$$p_1 = \langle B_1, \ldots, B_k, A, B_{k+1}, \ldots, B_n \rangle$$

and

$$p_2 = \langle B_1, \ldots, B_k, \bar{A}, B_{k+1}, \ldots, B_n, C_1, \ldots, C_l \rangle$$

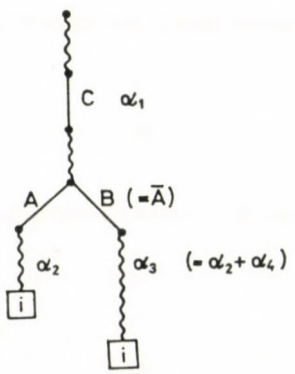


Fig. 4

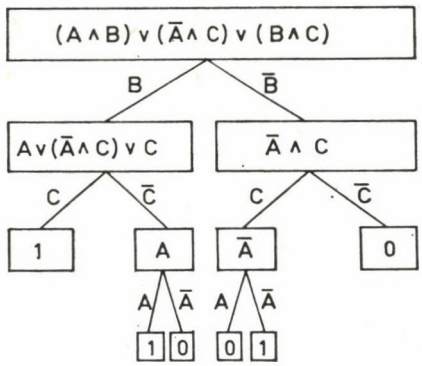


Fig. 5

be two quasi-amicable paths in the φ_s^i-tree ($i = 0, 1$). Then p_2 can be substituted by path

$$p = \langle B_1, \ldots, B_k, B_{k+1}, \ldots, B_n, C_1, \ldots, C_l \rangle .$$

Two quasi-amicable paths can be "simplified" by omitting the literal of the choice-point *from the longer* one, as the previous lemma and the tautologies (3), (4) tell us. Note, however, that after application of Rule 2, one-rootedness of the original tree will no longer hold; i.e. some "empty" edges may occur in the resulting tree (cf. the examples in Ch. 2).

1.4. These two simplifying rules are useful but, as the following example shows, not enough to obtain an irredundant form of an arbitrary Boolean function.

Let φ be given by the formula $(A \wedge B) \vee (\bar{A} \wedge C) \vee (B \wedge C)$, and fix the order of the selection by $\langle B, C, A \rangle$. The φ_S-tree is indicated in Figure 5.

It is easy to check that, by using the above two rules anyway, we obtain a formula at the same complexity level as the initial formula φ. E.g. by applying Rule *2* to the quasi-amicable paths p_1 and p_3 (Fig. 6), we get p_3 and $p_1' =$

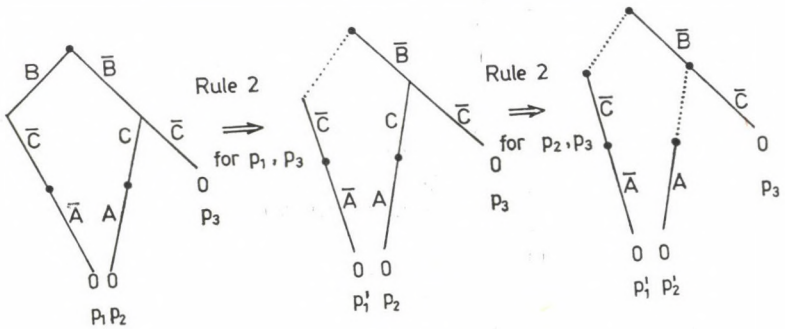


Fig. 6

$= \langle \bar{A}, \bar{C} \rangle$; from p_3 and p_2 we obtain p_3 and $p_2' = \langle \bar{B}, A \rangle$, again by Rule *2* (no possibility of employing Rule *1* at present). Considering the associated maxterms we have

$$\varphi = (B \vee C) \wedge (B \vee \bar{A}) \wedge (C \vee A) \qquad \text{(from } p_3, p_2', p_1', \text{ in order)}.$$

However, φ is not irredundant. In fact, $(A \vee C) \wedge (\bar{A} \vee B)$ is an irredundant form of φ.

In this example *we have not enough choice-points to eliminate*. For example, if we have another choice-point (e.g. in the φ_S^0-tree), then the desired form $(A \vee C) \wedge (\bar{A} \vee B)$ can be achieved (Fig. 7) by applying Rule *1* to the paths p_1 and p_3'', then p_2 and p_3'.

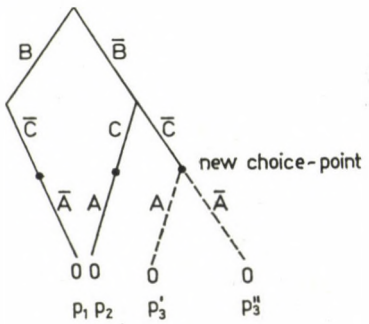


Fig. 7

Thus, in certain cases we can be in need of a "complicating" rule, as well, in order to increase the number of choice-points. The importance of this rule is in the fact that some local complication of the tree can lead to a more global simplification in the future. Of course, extreme care should be paid to this complication, and this is the crucial point why our algorithm is of heuristical nature only. The creation of new choice-points is, at the moment, on a trial-and-error basis.

The following two formulae are tautologies in the propositional calculus:

$$A_1 \wedge A_2 \leftrightarrow \big((A_1 \wedge B \wedge A_2) \vee (A_1 \wedge \bar{B} \wedge A_2)\big), \tag{5}$$

$$A_1 \vee A_2 \leftrightarrow \big((A_1 \vee B \vee A_2) \wedge (A_1 \vee \bar{B} \vee A_2)\big). \tag{6}$$

Our "complicating" rule is based on these two tautologies:

Rule 3. Let a formula φ of n variables and a selection function S be given. Let $p = \langle A_1, \ldots, A_m \rangle$ $(m < n)$ be a path in the φ_S^i-tree $(i = 0, 1)$ and let B be a variable of φ not occurring in p. Then two new paths

$$p_1 = \langle A_1, \ldots, A_j, B, A_{j+1}, \ldots, A_m \rangle$$

and

$$p_2 = \langle A_1, \ldots, A_j, \bar{B}, A_{j+1}, \ldots, A_m \rangle$$

can be substituted for p (Fig. 8).

Note that in practical applications of Rule *3*, it is always satisfactory to restrict ourselves to the case when the new choice-point is created at the place of a truth-value. Theoretically this can be justified by the "commutativity laws" of the 2-element Boolean algebra.

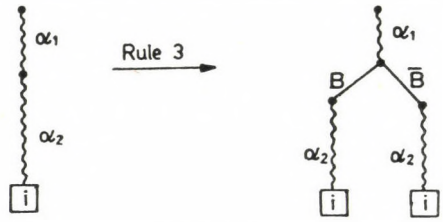


Fig. 8

1.5. Let a φ_S^i-tree be given. Consider the following sets of paths:

$IS = \{p \mid p$ is a path of φ_S^i but there is no path q in the φ_S^i-tree so that p and q are amicable$\}$, ("isolated paths");

$SIS = \{p \mid p \in IS$ and there exists at least one variable in φ, say A, so that neither A nor $\bar{A}$ occurs in $p\}$, ("short isolated paths").

Then, we can proceed with the following algorithm.

(i) Find the elements of *IS*;

(ii) Find the subset *SIS*;

(iii) Create a new choice-point at the place of the truth-value of an element of *SIS* if both paths, newly made, are amicable with some ***old*** paths. If this is not the case (e.g. only one of the new paths is amicable with an old one), then the creation of the choice-point is prohibited in this path. Make a trial with another element of *SIS*;

(iv) Apply Rule *1* as many times as can be;

(v) Apply Rule *2* going from the longest paths of the φ_S^i-tree upward to the shorter ones; if Rule *1* can be applied to a result of the application of Rule *2*, then use it again;

(vi) If no further simplification is possible, then write down the remaining paths. Finally, consider the associated minterms ($i = 1$) or maxterms ($i = 0$) and conjunct or disjunct them, respectively.

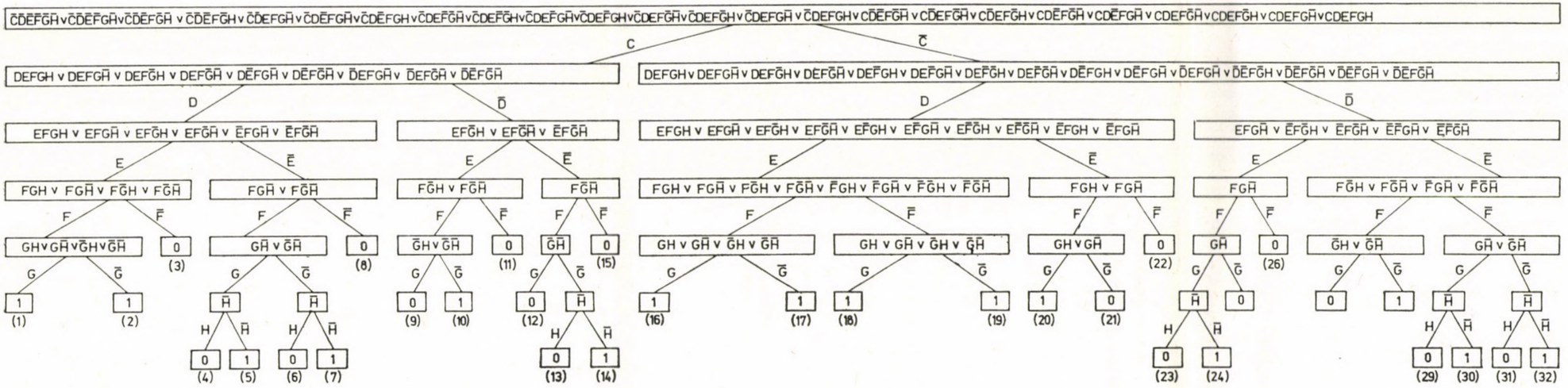

C
D
E
F
G
H
(1)
(2)
(3)
(4)
(5)
(6)
(7)
(8)
(9)
(10)
(11)
(12)
(13)
(14)
(15)
(16)
(17)
(18)
(19)
(20)
(21)
(22)
(23)
(24)
(26)
(29)
(30)
(31)
(32)

Fig. 9

2. Examples

In this section we present two examples. The formulas are given in full disjunctive normal forms. Conjunction is denoted by simple concatenation. We use "decimal coding" for the formulas (cf. [3]). The φ_S-trees of these formulas are indicated in Fig. 9 and Fig. 11, respectively. The selection functions can be read off these trees.

2.1. Let φ be

$$\bigvee_6 (0, 2, 4, 5, 14, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 36, 44, 45, 52, 54, 60, 61, 62, 63).$$

The φ_S^1-tree will be simplified. We need not use Rule *3*. Apply Rule *1* for the paths (16), (17) first, then for (18), (19); let us denote the results by (16, 17) and (18, 19), respectively. Employ Rule *1* again for (16, 17) and (18, 19). We obtain the path (16, 17, 18, 19) consisting of the literals $\bar{C}, D, E$. Similarly, from the paths (1), (2) we have C, D, E, F; from (7), (5) we have $C, D, \bar{E}, F, H$; and finally from (30), (32) we have $\bar{C}, \bar{D}, \bar{E}, \bar{F}, \bar{H}$. The tree of this intermediate state is indicated in Figure 10.

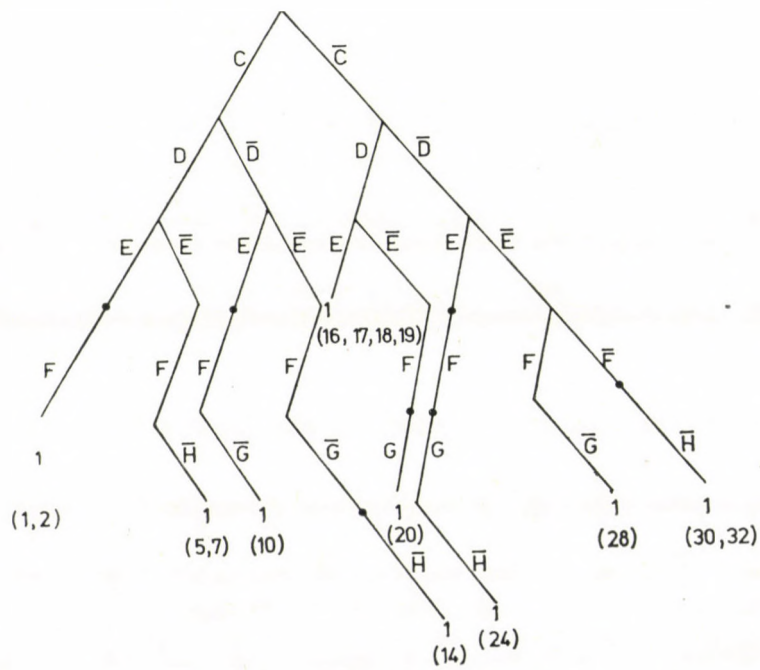


Fig. 10

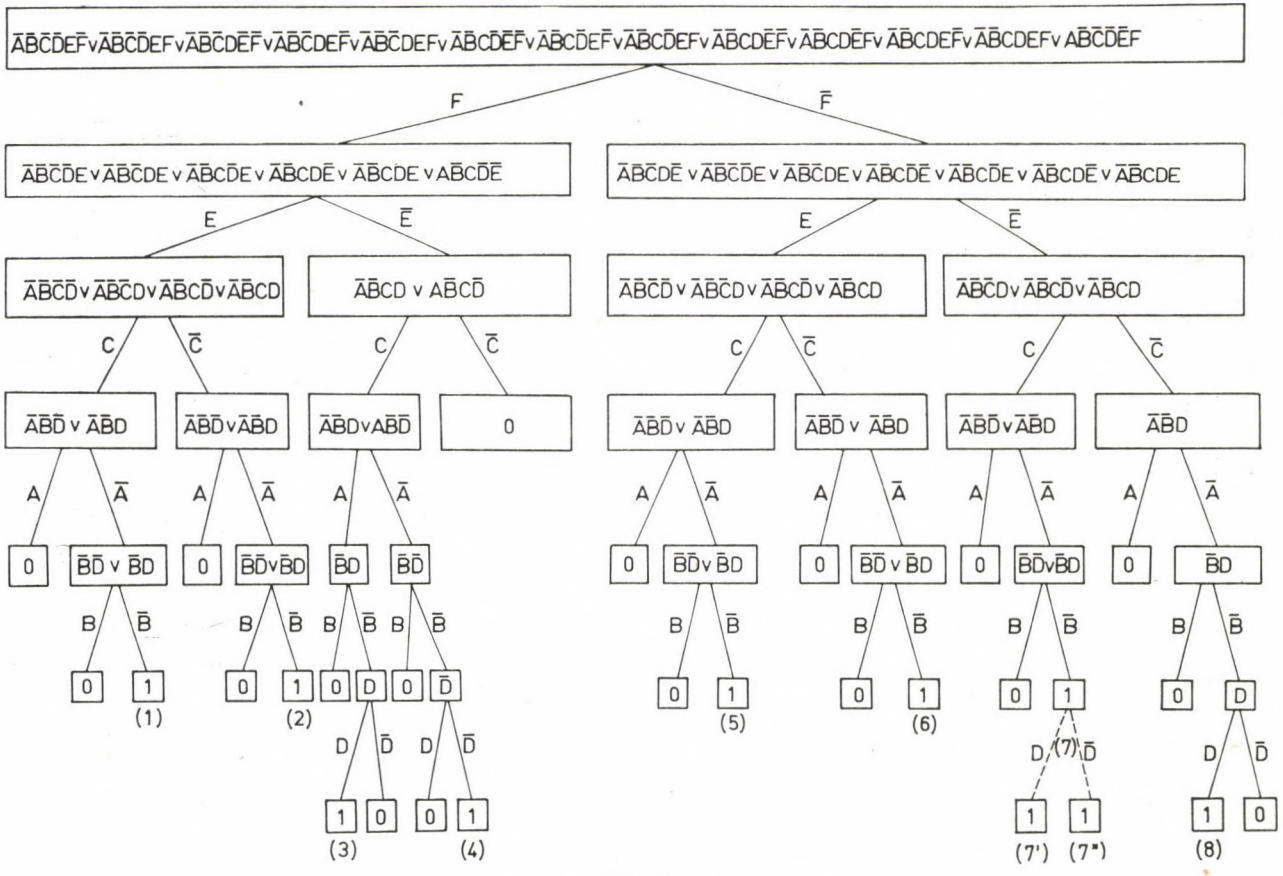

ĀB̄C̄DĒ∨ĀB̄C̄DE∨ĀB̄CD̄Ē∨ĀB̄CDE∨ĀBC̄DE∨ĀBC̄DEF∨ĀBCD̄Ē∨ĀBCD̄E∨ĀBCDĒ∨ĀBCDE∨ABC̄D̄Ē
F
F̄
ĀB̄C̄DE∨ĀB̄C̄DE∨ĀB̄CD̄E∨ĀB̄CDE∨ĀBC̄DE
ĀB̄C̄DĒ∨ĀB̄C̄DE∨ĀB̄CD̄E∨ĀB̄CDĒ∨ĀB̄CD̄E∨ĀB̄CDE∨ĀBC̄DE
E
Ē
E
Ē
ĀB̄C̄D̄∨ĀB̄C̄D∨ĀB̄CD̄∨ĀB̄CD
ĀB̄CD∨ĀBC̄D̄
ĀB̄C̄D̄∨ĀB̄C̄D∨ĀB̄CD̄∨ĀB̄CD
ĀB̄CD∨ĀBC̄D̄∨ĀB̄CD
C
C̄
C
C̄
C
C̄
C
C̄
ĀB̄D̄∨ĀB̄D
ĀB̄D̄∨ĀB̄D
ĀB̄D∨ĀBD̄
0
ĀB̄D̄∨ĀB̄D
ĀB̄D̄∨ĀB̄D
ĀB̄D̄∨ĀB̄D
ĀB̄D
A
Ā
A
Ā
A
Ā
A
Ā
A
Ā
A
Ā
A
Ā
0
B̄D̄∨B̄D
0
B̄D̄∨B̄D
B̄D
B̄D̄
0
B̄D̄∨B̄D
0
B̄D̄∨B̄D
0
B̄D̄∨B̄D
0
B̄D
B
B̄
B
B̄
B
B̄
B
B̄
B
B̄
B
B̄
B
B̄
0
1
0
1
0
D
0
D̄
0
1
0
1
0
1
0
1
(1)
(2)
(5)
(6)
D
D̄
D
D̄
D
(7)
D̄
D
D̄
1
0
0
1
1
1
1
0
(3)
(4)
(7')
(7")
(8)

Fig. 11

Rule *2* is applied for

(a) the result of (5), (7), and the result of (1), (2); $\bar{E}$ can be deleted from the former;

(b) the paths (14), (10); $\bar{E}$ can be deleted from the first;

(c) the paths (10) and the result of (1), (2); $\bar{D}$ can be deleted from (10);

(d) the result of point *(b)* and the result of point *(a)*; $\bar{D}$ can be deleted from the path obtained from (14) in *(b)*;

(e) the paths (20) and (16, 17, 18, 19); $\bar{E}$ can be deleted from (20);

(f) the paths (24) and the result obtained in *(e)* from (20); $\bar{D}$ can be deleted from (24);

(g) the result of (1), (2) and (16, 17, 18, 19); C can be deleted from the former.

We arrived at the following formula:

$$\bar{C}\bar{D}\bar{E}\bar{F}\bar{H} \vee \bar{C}\bar{D}\bar{E}\bar{F}\bar{G} \vee \bar{C}EFG\bar{H} \vee \bar{C}DFG \vee CF\bar{G}\bar{H} \vee CEF\bar{G} \vee \vee CDF\bar{H} \vee \bar{C}DE \vee DEF.$$

This is irredundant.

2.2 Let the formula φ be

$$\bigvee_{6} (2, 3, 4, 6, 7, 8, 10, 11, 12, 13, 14, 15, 41).$$

The φ_S^1-tree will be simplified (Fig. 11). The application of Rule *3* can be observed at (7), an element of the set SIS, the path (7′) is amicable with path

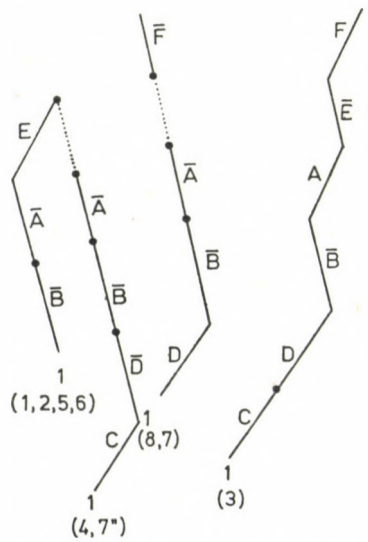


Fig. 12

(8) and (7″) is amicable with path (4). Observe that the creation of any other choice-point is prohibited, so we apply Rule *1* first for the paths (1) and (2); then for (5) and (6); for the results (1, 2) and (5, 6) of the previous steps; then for (7′) and (8); for (7″) and (4); respectively. The intermediate state is indicated in Fig. 12. We apply Rule *2* for (1, 2, 5, 6) and (4, 7″); $\bar{E}$ can be deleted from the latter; for (1, 2, 5, 6) and (8, 7′); from which $\bar{E}$ can be deleted. There is no other possibility for applying Rule *2* or Rule *1*. Finally, we get the following irredundant form:

$$A\bar{B}CD\bar{E}F \vee \bar{A}\bar{B}C\bar{D} \vee \bar{A}\bar{B}D\bar{F} \vee \bar{A}\bar{B}E\,.$$

REFERENCES

1. Ádám, A.: Truth Functions and the Problem of their Realization by Two-terminal Graphs, Akadémiai Kiadó, Budapest, 1968
2. McCluskey, E. J.: Minimization of Boolean functions, *Bell System Tech. J.* **35**, 1417—1444 (1956)
3. McCluskey, E. J.: Introduction to the Theory of Switching Circuits, McGraw-Hill, New York—London—Sydney—Toronto, 1965
4. Móricz, F.—Varga, A.—Ecsedi-Tóth, P.: A Method for Minimizing Partially Defined Boolean Functions, *Acta Cybernetica* **4**, 283—290 (1979)
5. Varga, A.—Ecsedi-Tóth, P.—Móricz, F.: On a Connection between Algebraic and Graph-Theoretic Minimization of Truth Functions, Proceedings of the International Colloquium on Algebraic Methods in Graph Theory, Szeged, 1978, Akadémiai Kiadó and North Holland Company, to appear

Heuristischee Methode zur Beschleunigung der Optimierung von Boole-Funktionen- — Der Aufsatz bringt eine einfache heuristische Methode zur Beschleunigung der manuellen Optimierung von Boole-Funktionen (Wahrheitsfunktionen). Das Verfahren funktioniert an einer Baumrepräsentation der Boole-Funktionen, daher hat es eher geometrischen als algebraischen Charakter. Es kann gleicherweise für komplette oder partialisch definierte Boole-Funktionen mit höchstens 10 Unbekannten angewendet werden. Die Arbeit enthält auch einige illustrative Beispiele.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (3—4), pp. 259—273 (1980)

OPTIMUM DESIGN FOR BENDING AND ULTIMATE SHEAR STRENGTH OF HYBRID I-BEAMS

J. FARKAS*
DOCTOR TECHN. SCI.

[Manuscript received May 28, 1979]

The paper presents the approximate analytical optimization of welded hybrid I-beams subject to bending. In the objective function the different material costs of web and flanges are taken into account. The following design constraints are considered: local buckling constraint of web and that of compressed flange, constraints of maximum stress and deflection, respectively. The dead weight and the lateral buckling is not considered. The obtained simple formulae are applied to the comparative calculations which show e.g. that the material cost of hybrid beams optimized with respect to stress constraint is by 15÷25% less than that of homogeneous ones. Further, limit shear stresses are determined below where the effect of shear can be neglected.

Symbols

A	cross-sectional area
b	distance between transverse stiffeners, Fig. 3
C_a	material cost
c	distance of the plastic hinges in flanges, Fig. 3
c_f, c_g	material cost parameters for flanges and web, resp.
E	modulus of elasticity
F_g	resultant of web membrane stresses
f	a distance, Fig. 3
h	depth of web
I_x	moment of inertia
I_0	required moment of inertia
K_0	required section modulus
k_T	factor of local shear buckling, Eq. (33)
M_H	limit bending moment
M_u	maximum moment to be carried by girder, Eq. (46)
M_p	plastic moment of girder assuming fully effective web, Eq. (45)
M_f, M_g	Eq. (45)
M_p^*, M_{pf}	Eq. (44)
N	axial force, Fig. 1
n_h	safety factor
s	width of flanges
T	shear force
T_{Fg}	shear force to produce yielding of web, Eq. (41)
T_S, T_C, T_B	limit shear forces, Fig. 4 Eqs (42), (47), (48)
v	flange thickness
v_g	web thickness
w	maximum deflection
β	web slenderness parameter, Eq. (8)
γ	specific weight
δ	flange slenderness parameter, Eq. (17)

* Prof. Dr. J. FARKAS, Nehézipari Műszaki Egyetem (Technical University) H-3515 Miskolc-Egyetemváros, Hungary

$\eta = c_g/c_f$	material cost ratio
ν	Poisson's ratio
$\xi = \sigma_{Fg}/\sigma_{Hf}$	Eq. (2)
σ_{Hg}, σ_{Hf}	limit stress of web and flanges, resp.
σ_{Fg}, σ_{Ff}	yield stress of web and flanges, resp.
σ_t^F	tension field web membrane stress
τ_{Fg}	yield shear stress of web material
τ_{kr}	critical shear stress
Θ	inclination of tension field to flange, Fig. 3
Θ_d	inclination of web panel diagonal, Fig. 3

1. Introduction

The application of steels having increased yield stress may result in significant saving in weight and cost of metal structures. 370-steels (tensile strength 370 MPa) are mainly replaced by 520-steels (tensile strength 520 MPa, yield stress 340—400 MPa), but in some countries other weldable high strength structural steels are developed as well, e.g. C 60/45, C 70/60, C 85/75 steels in USSR [24, 28], HW 70, 80, 90 steels in Japan, A 514 steel (yield stress 690 MPa) in USA, etc. More detailed informations about high strength steels used in metal structures can be found for example in [12] or in [24].

The use of high strength steels is economical mainly in structural parts subjected to static tension (storage tanks, pressure vessels, trusses) or compression in plastic buckling region (columns, trusses). In structures with active displacement constraints (radiotelescopes, machine structures) or compressed in elastic buckling region (slender bars) it is not economical to increase the yield stress. High-strength steels are less economical in structures loaded with pulsating loads, because the fatigue strength of welded joints does not increase in proportion to the yield stress, although there are some new techniques to improve the fatigue strength of welded joints (e.g. [22]).

In structures subject to bending, the hybrid I-beams, with flanges made of steels of higher strength than web, are more economical than the homogeneous ones. Design problems of hybrid girders were investigated in many papers, for example in [4, 6, 7, 13, 14, 16, 21, 25, 26, 27].

This paper presents the analytical optimum design of hybrid I-beams loaded for bending. In the objective function the cost difference between the two types of steels used in flanges and web, respectively, is taken into account. In more intricate cases, when in the objective function the effect of dead weight and the fabrication costs are also involved, mathematical programming methods can be used. SZABÓ [23] has used the backtrack programming for this purpose. This discrete programming method may be advantageous for optimum design of welded structures with few unknowns [11].

2. Optimum design for minimum material cost of hybrid I-beams subject to pure bending

2.1 Characteristics of the cross-section optimized on the basis of stress constraint

The stress distribution in the limit state of a hybrid I-beam is shown in Fig. 1. Note, that we do not use the fully plastic stress distribution, because as was shown in [9], the beams designed elastically are more economical than those designed plastically.

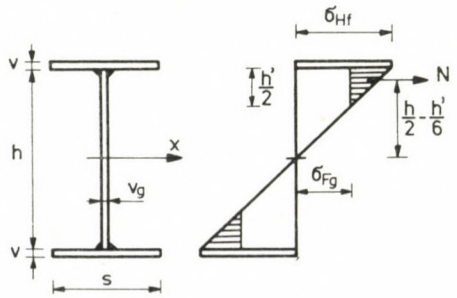


Fig. 1. Limit state stress distribution in a hybrid I-girder loaded in bending

Here is

$$N = (\sigma_{Hf} - \sigma_{Fg})\, h' v_g/4 = \sigma_{Hf}(1 - \xi)^2\, h v_g/4 \tag{1}$$

because

$$\frac{h'}{h} = \frac{\sigma_{Hf} - \sigma_{Fg}}{\sigma_{Hf}} = 1 - \xi;\ \xi = \sigma_{Fg}/\sigma_{Hf}. \tag{2}$$

The limit bending moment is approximately

$$M_H = \sigma_{Hf}\left(\frac{v_g h^2}{6} + svh\right) - 2\,N\left(\frac{h}{2} - \frac{h'}{6}\right). \tag{3}$$

On the basis of (1) and (2), neglecting the dead weight, the ***stress constraint*** takes the form as follows:

$$\frac{M_H}{\sigma_{Hf}} = svh + \frac{v_g h^2}{12}(3\xi - \xi^3) \geq K_0\,. \tag{4}$$

K_0 is the required section modulus. The cross section area is

$$A = hv_g + 2sv. \tag{5}$$

From (5) the area of flanges is

$$2sv = A - hv_g\,. \tag{6}$$

Inserting (6) into (4), the stress constraint may be expressed as

$$A \geq \frac{2K_0}{h} + \frac{v_g h}{6}(6 - 3\xi + \xi^3) . \tag{7}$$

The local buckling constraint of unstiffened web is

$$\frac{v_g}{h} \geq \beta = \sqrt{\frac{12(1-\nu^2)n_h \sigma_{Hg}}{23{,}9\pi^2 E}} . \tag{8}$$

For 370-steels it is $\sigma_{Hg} = 2 \cdot 10^4$ N/cm^2; $\nu = 0{,}3$; $E = 2{,}1 \cdot 10^7$ N/cm^2, with a safety factor of $n_h = 1{,}08$ one obtains $\beta = 1/145$. For other steels (8) takes the form

$$\frac{v_g}{h} \geq \frac{1}{145}\sqrt{\frac{\sigma_{Hg}}{\sigma_{H370}}} . \tag{8a}$$

Treating the constraints (7) and (8) as actives, we get

$$A = \frac{2K_0}{h} + \frac{\beta h^2}{6}(6 - 3\xi + \xi^3) . \tag{9}$$

The material cost of the unit length of the beam is

$$C_a = c_g \gamma v_g h + 2c_f \gamma s v \tag{10}$$

Using (6) and (9) we get

$$\frac{C_a}{c_f \gamma} = \frac{2K_0}{h} + \frac{\beta h^2}{6}(6\eta - 3\xi + \xi^3) . \tag{11}$$

The condition $dC_a/dh = 0$ gives

$$h_{\text{opt}} = \sqrt[3]{\frac{6K_0}{\beta(6\eta - 3\xi + \xi^3)}} . \tag{12}$$

For a homogeneous beam it is $\xi = \eta = 1$ and (12) takes the known form [8]

$$h_{\text{opt}} = \sqrt[3]{\frac{3K_0}{2\beta}} . \tag{13}$$

On the basis of (11) and (12) the minimal cost can be expressed as

$$\frac{C_{a.\text{min}}}{c_f \gamma} = \sqrt[3]{4{,}5\, K_0^2 \beta(6\eta - 3\xi + \xi^3)} . \tag{14}$$

By using (9) and (12) one obtains for the cross-sectional area of the beam at minimal cost

$$A_{c.\min} = \frac{\beta h^2}{2}(4\eta + 2 - 3\xi + \xi^3)\,, \tag{15}$$

and, according to (6) and (15), the area of flanges is

$$2sv = A - \beta h^2 = \frac{\beta h^2}{2}(4\eta - 3\xi + \xi^3)\,. \tag{16}$$

The local buckling constraint for compressed flange

$$\frac{v}{s} \geq \delta = \frac{1}{30}\sqrt{\frac{\sigma_{Hf}}{\sigma_{H370}}} \tag{17}$$

can be treated as active. On the basis of (16) and (17)

$$s = \frac{h}{2}\sqrt{\frac{\beta}{\delta}(4\eta - 3\xi + \xi^3)} \quad \text{and} \quad v = \delta s = \frac{h}{2}\sqrt{\beta\delta(4\eta - 3\xi + \xi^3)}\,. \tag{18}$$

The approximate formula of moment of inertia for the section of minimal cost is

$$I_x = \frac{v_g h^3}{12} + \frac{svh^2}{2} = \frac{\beta h^4}{24}(12\eta + 2 - 9\xi + 3\xi^3) =$$

$$= (12\eta + 2 - 9\xi + 3\xi^3)\sqrt[3]{\frac{3}{32\beta}\left(\frac{M_{\max}}{\sigma_{Hf}}\right)^4 \frac{1}{6\eta - 3\xi + \xi^3}}\,. \tag{19}$$

2.2 Characteristics of the cross section optimized on the basis of deflection constraint

The deflection constraint is

$$I_x = \frac{v_g h^3}{12} + \frac{svh^2}{2} \geq I_0\,. \tag{20}$$

The required moment of inertia I_0 can be calculated approximately, considering only the elastic deformations in the whole beam. By using (6) and (20) we get

$$A \geq \frac{4\,I_0}{h^2} + \frac{2v_g h}{3}\,. \tag{21}$$

If the deflection constraint is active, the actual max. stress in the web (σ) is less than σ_{Hg} and the exact formula of web buckling would be

$$\frac{v_g}{h} \geq \beta \sqrt{\frac{\sigma}{\sigma_{Hg}}} .$$

Instead of this exact formula we use (8). According to (6) and (21) the material cost can be expressed as

$$\frac{C_a}{c_f \gamma} = \frac{4I_0}{h^2} + \frac{\beta h^2}{3}(3\eta - 1) . \tag{22}$$

The condition $dC_a/dh = 0$ gives

$$h'_{\text{opt}} = \sqrt[4]{\frac{12 I_0}{\beta (3\eta - 1)}} . \tag{23}$$

Using (21) and (23), the cross-sectional area is

$$A'_{c.\min} = \frac{\beta h^2}{3}(3\eta + 1) = (3\eta + 1)\sqrt{\frac{4 I_0 \beta}{3(3\eta - 1)}} . \tag{24}$$

The area of flanges is

$$2sv = \frac{\beta h^2}{3}(3\eta - 2) , \tag{25}$$

and the minimal cost may be expressed as

$$\frac{C'_{a.\min}}{c_f \gamma} = \sqrt{\frac{16}{3}(3\eta - 1)\beta I_0} . \tag{26}$$

2.3 Comparative calculations

The economy of hybrid I-beams can be demonstrated by means of the above derived approximate formulae. Let us use for the calculations the data of the following two hybrid beams (stresses in MPa):

1) 52/37-beam; steels according to the Hungarian Standard MSZ 6280:

eb 37B, $\sigma_{Hg} = 200$; $\sigma_{Fg} = 240$; flanges 52C,

$\sigma_{Hf} = 280$; $\xi = 0{,}86$; $\eta = 0{,}901$;

2) 85/38-beam: steels according to the Soviet Standard [28]:

$$\text{web C 38; } \sigma_{Hg} = 210; \ \sigma_{Fg} = 230; \text{ flanges C 85/75 ;}$$
$$\sigma_{Hf} = 530; \ \xi = 0{,}434; \ \eta = 0{,}617 .$$

a) ***Comparison of weights*** of hybrid and homogeneous beams, respectively, optimized on the basis of the stress constraint (for homogeneous beam it is $\xi = \eta = \sigma_{Hg}/\sigma_H = 1$). Using (15) and (12), we get

$$\frac{A_{\text{hyb}}}{A_{\text{hom}}} = \frac{4\eta + 2 - 3\xi + \xi^3}{\sqrt[3]{4(6\eta - 3\xi + \xi^3)^2}} \left(\frac{\sigma_{Hg}}{\sigma_{Hf}}\right)^{2/3} . \tag{27}$$

For 52/37 and 85/38-beam (27) gives the values of 0,805 and 0,602, resp., thus, hybrid beams are by 20—40% lighter than the homogeneous ones.

b) ***Comparison of material costs.*** Using (14) one obtains

$$\frac{C_{a.\text{hyb}}}{C_{a.\text{hom}}} = \frac{1}{\eta} \left(\frac{\sigma_{Hg}}{\sigma_{Hf}}\right)^{2/3} \sqrt[3]{\frac{6\eta - 3\xi + \xi^3}{4}} . \tag{28}$$

For 52/37 and 85/38-beam (28) gives the values of 0,845 and 0,746, respectively, thus, hybrid I-beams assure 15÷25% savings in material cost.

c) ***Comparison of deflections.*** Assuming, that $\beta_{\text{hyb}} = \beta_{\text{hom}}$, according to (19) we get

$$\frac{w_{\text{hyb}}}{w_{\text{hom}}} = \frac{I_{x.\text{hom}}}{I_{x.\text{hyb}}} = \frac{\sqrt[3]{2}\,(6\eta - 3\xi + \xi^3)^{4/3}}{12\,\eta + 2 - 9\xi + 3\xi^3} \cdot \left(\frac{\sigma_{Hf}}{\sigma_{Hg}}\right)^{4/3} . \tag{29}$$

For 52/37 and 85/38-beam (29) gives the values of 1,48 and 2,53 respectively, thus, deflections of hybrid I-beams are larger than those of homogeneous ones, therefore, it is advisable to examine whether the cross section optimized on the basis of stress constraint does fulfil the deflection constraint, i.e. whether $I_x \geq I_0$. I_x can be calculated with (19).

d) ***Comparison of material costs*** of hybrid and homogeneous beams, respectively, optimized on the basis of ***deflection constraint.*** Using (26) we get

$$\frac{C'_{a.\text{hyb}}}{C'_{a.\text{hom}}} = \frac{1}{\eta} \sqrt{\frac{3\eta - 1}{2}} . \tag{30}$$

For 52/37 and 85/38-beam (30) gives the values of 1.024 and 1.057 resp., thus, if the deflection constraint is active, hybrid beams are uneconomical.

3. Ultimate shear strength of homogeneous and hybrid I-beams optimized for bending

In calculating I-beams loaded for bending and shear, designers can follow two ways as follows:

1. Use of interactive formulae for web buckling valid for all values of shear stress. This method is prescribed, for instance, by the Hungarian Standard MSZ 15024. The optimization of homogeneous I- and box beams on the basis of this principle is treated in [10].

2. Determine a limit curve for shear stresses or an interaction curve for shear forces; shear stresses below these limit values can be neglected. This method simplifies the calculations and, therefore, it is advantageous for the purposes of optimization. Here we calculate these limit shear stress values for homogeneous and hybrid I-beams without and with transverse stiffeners, respectively.

3.1 Beams with unstiffened web

According to CARSKADDAN [3], the critical buckling stress of the simply supported web of a hybrid I-beam can be calculated with the following formula:

$$\tau_{kr} = \frac{5{,}34\ \pi^2 E}{12(1-\nu^2)} \left(\frac{v_g}{h}\right)^2 [1 - 1{,}61(1-\xi)^2]. \tag{31}$$

Fig. 2 shows the limit curves of shear stresses for beams with web made of steel 37B calculated with (31), taking into account a safety factor of $n_h = 1{,}08$.

Note that BRONOWICKI and FELTON [2] have eliminated the effect of shear stress in the optimization of homogeneous I-beams prescribing a constraint as follows:

$$T/(hv_g) \leq 0{,}4\ \tau_{kr}\ .$$

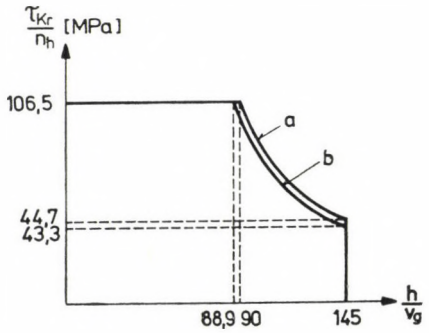


Fig. 2. Limit shear stress curves for unstiffened beams; *a)* homogeneous I-beam made of 97B steel; *b)* 52C/37B hybrid I beam, $n_h = 1{,}08$; $\tau_{Hg} = 115$ MPa; $\xi = 0{,}86$

DELESQUES [5] has proposed for homogeneous girders with web made of 360-steel an empirical formula

$$\tau_{kr}\,[\mathrm{MPa}] = 16 + 8889\frac{v_g}{h} \quad \text{for} \quad 67 < \frac{h}{v_g} < 167$$

and for slendernesses $167 < h/v_g < 333$ has given an interaction formula. CHONG [4] has optimized hybrid beams on the basis of active shear stress constraint according to (31). This method is not correct, because the constraint of shear stress is not active in all cases.

3.2 Beams with transverse stiffeners

There are several methods to predict the collapse load of plate girders considering the post-buckling behaviour of web and the influence of flexural rigidity of flanges [15]. Here we shell construct, on the basis of the paper of ROCKEY, EVANS and PORTER [18], the interaction diagrams for optimized homogeneous and hybrid beams.

According to Basler's theory [1], the web buckles when the critical shear stress τ_{kr} is reached and any additional load is supported by tensile membrane stress σ_t (Fig. 3a, b, c). Failure occurs when the buckling stress τ_{kr} and the

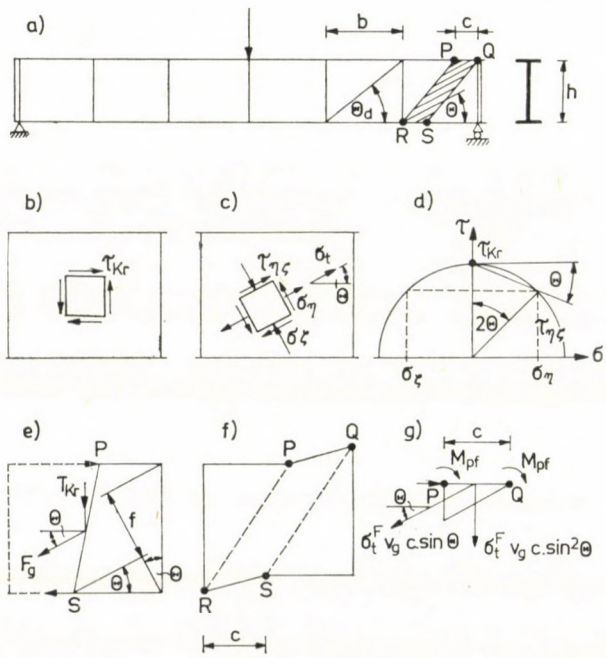


Fig. 3. Determination of the collapse load of a transversely stiffened plate girder; *a)* the simply supported girder; *b)* direction of critical shear buckling stress in the end panel; *c)* stress components in direction of tension membrane stress field; *d)* stress components in Mohr's circle; *e)* forces acting in the section P—S; *f)* the collapse mechanism; *g)* equilibrium of the flange element

membrane stress σ_t^F produces yielding in the web. Hinges formed in the flanges together with yield zone in the web form a plastic collapse mechanism (Fig. 3f). ROCKEY and ŠKALOUD [17] have considered the effect of flexural rigidity of flanges, too. They have assumed that the inclination of the tension field is $\Theta = \Theta_d$. ROCKEY, EVANS and PORTER [18, 19] have developed a refined method, which determines the inclination angle Θ by an iteration procedure, taking the initial value of $\Theta = 2\Theta_d/3$ and iterating until $T_{S\max}$ is reached. For design purposes it is recommended to take $\Theta = 2\Theta_d/3$.

The stress components in the direction of Θ in the stress state characterized by τ_{kr} according to the Mohr's circle (Fig. 3a, b, c, d) are as follows:

$$\sigma_\eta = \tau_{kr} \sin 2\Theta; \quad \sigma_\zeta = -\tau_{kr} \sin 2\Theta; \quad \tau_{\eta\zeta} = \tau_{kr} \cos 2\Theta \tag{32}$$

where

$$\tau_{kr} = \frac{k_T \pi^2 E}{12(1-\nu^2)} \left(\frac{v_g}{h}\right)^2; \quad k_T = 5{,}34 + 4(b/h)^2 \quad \text{for} \quad b/h \geq 1,$$
$$k_T = 4{,}0 + 5{,}34(b/h)^2 \quad \text{for} \quad b/h < 1. \tag{33}$$

The yield criterion relating to the stress components of $\sigma'_\eta = \sigma_\eta + \sigma_t^F$, σ_ζ and $\tau_{\eta\zeta}$ can be written in the form

$$\sigma_\eta'^2 + \sigma_\zeta^2 - \sigma'_\eta \sigma_\zeta + 3\tau_{\eta\zeta}^2 = \sigma_{Fg}^2. \tag{34}$$

From (34) we get

$$\sigma_t^F = -\frac{3}{2}\tau_{kr} \sin 2\Theta + \sqrt{\sigma_{Fg}^2 + \tau_{kr}^2 \left[\left(\frac{3}{2} \sin 2\Theta\right)^2 - 3\right]}. \tag{35}$$

From the equation of equilibrium (Fig. 3e) it can be written that the ultimate shear load is

$$T_S = F_g \sin\Theta + T_{kr} \tag{36}$$

where

$$F_g = \sigma_t^F v_g (2c \sin\Theta + f); \quad f = h\cos\Theta - b\sin\Theta \tag{37}$$

thus

$$F_g = \sigma_t^F v_g \sin\Theta (2c + h\cot\Theta - b). \tag{38}$$

The equilibrium of the flange element according to Fig. 3g gives

$$2M_{pf} = \sigma_t^F v_g c \sin^2\Theta \cdot \frac{c}{2}, \tag{39}$$

thus

$$c = \frac{2}{\sin\Theta} \sqrt{\frac{M_{pf}}{\sigma_t^F v_g}}. \tag{40}$$

Introducing

$$T_{Fg} = \tau_{Fg}\, hv_g = \frac{\sigma_{Fg}}{\sqrt{3}}\, hv_g \tag{41}$$

by means of (35), (36), (38), (40) and (41) one obtains

$$\frac{T_S}{T_{Fg}} = \frac{\tau_{kr}}{\tau_{Fg}} + \sqrt{3}\sin^2\Theta\left(\cot\Theta - \frac{b}{h}\right)\frac{\sigma_t^F}{\sigma_{Fg}} + 4\sqrt{3}\sin\Theta\sqrt{\frac{\sigma_t^F}{\sigma_{Fg}}\, M_p^*} \tag{42}$$

where

$$\frac{\sigma_t^F}{\sigma_{Fg}} = -\frac{\sqrt{3}}{2}\frac{\tau_{kr}}{\tau_{Fg}}\sin\Theta + \sqrt{1 + \left(\frac{\tau_{kr}}{\tau_{Fg}}\right)^2 (0{,}75\sin^2 2\Theta - 1)}, \tag{43}$$

and

$$M_p^* = M_{pf}/(\sigma_{Fg} h^2 v_g)\,; \quad M_{pf} = \sigma_{Ff}\, sv^2/4\,. \tag{44}$$

The second and third term in (42) expresses the effect of a membrane tension field and the flexural rigidity of the flanges, respectively.

The characteristic points of the interaction diagram shown in Figure 4 can be calculated by means of the following formulae:

$$M_p = \sigma_{Ff}\, svh + \sigma_{Fg}\, h^2 v_g/4 = M_f + M_g\,, \tag{45}$$

M_u can be obtained from (4)

$$M_u = M_H = \sigma_{Hf}\left[svh + \frac{v_g h^2}{12}(3\xi - \xi^3)\right]. \tag{46}$$

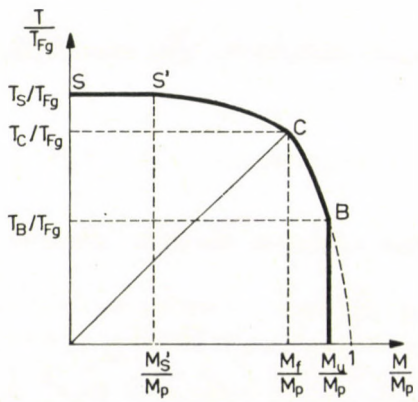


Fig. 4. Interaction diagram for I-girders with transverse stiffeners

For the point S' it is $T'_S = T_S$ and $M'_S = T_S b$, i.e. the maximum bending moment in the end panel of a simply supported beam, with a restriction that $M'_S \leq 0{,}5\, M_f$. For T_C the following empirical formula was given [18]:

$$\frac{T_C}{T_{Fg}} = \frac{\tau_{kr}}{\tau_{Fg}} + \frac{\sigma_t^F}{\sigma_{Fg}} \sin\left(\frac{4\Theta_d}{3}\right)\left[2 - \left(\frac{b}{h}\right)^{1/8}\right]\left(0{,}544 + \frac{36{,}8\, M_{pf}}{M_f}\right). \quad (47)$$

M_f is the contribution of the flanges to the plastic moment of resistance of the girder, as defined by (45). The section B—C of the interaction diagram may be represented by a simple parabola, therefore, it is

$$\frac{T_B}{T_C} = \sqrt{\frac{M_p - M_u}{M_g}}. \quad (48)$$

The simple relationships of optimal cross sections obtained in point 2.1 make it possible to derive the actual interaction diagram for important special cases. Let us determine the interaction diagram for two special cases as follows.

Case 1. Homogeneous I-girder optimized on the basis of stress constraint with data as follows: $h/v_g = 145$; steel 370, stresses in MPa: $\sigma_F = 240$; $\sigma_H = 200$; $\tau_F = 0{,}577$; $\sigma_F = 138.5$; $\xi = \eta = 1$; distance of stiffeners $b = 1{,}5\, h$.

With Eq. (33) $k_T = 7.12$; $\tau_{kr} = 64{,}3$; $\Theta_d = \text{arc tan}\,(1.5) = 33{,}7°$; $\Theta = 22.5°$; with (43) we get $\sigma_t^F/\sigma_{Fg} = 0{,}7763$.

By means of (17) and (18), with $\delta = 1/30$ one obtains $s = h\sqrt{\beta/(2\delta)} = 0.3216\, h$ and $v = h\sqrt{\beta\delta/2} = h/93{,}27$. Using these values, M_p^* can be calculated with (44): $M_p^* = 1{,}34 \cdot 10^{-3}$. Eq. (42) gives

$$T_S/T_{Fg} = 0{,}4645 + 0{,}1803 + 0{,}0854 = 0{,}7302,$$

these three terms numerically show the effect of τ_{kr}, σ_t^F and M_{pf}, respectively. Further, Eq. (47) gives $T_C/T_{Fg} = 0{,}8626$. Since $T_C > T_S$, it can be taken that $T_C = T_S$.

With (45) $M_p = 3\sigma_F h^2\, v_g/4$; $M_f/M_p = 0{,}6667$; using (46) $M_u = 2\sigma_H\, h^2\, v_g/3$, thus $M_u/M_p = 8\sigma_H/(9\sigma_F) = 0{,}7407$ and with (48) we get $T_B/T_C = \sqrt{3 - 8\sigma_H/(3\sigma_F)} = 0{,}8819$ and $T_B/T_{Fg} = 0{,}7302 \cdot 0{,}8819 = 0{,}6440$.

The interaction diagram is shown in Fig. 5. It can be seen that, in a cross-section loaded in $M_{\max} = 0{,}7407\, M_p$ ($\sigma_{\max} = \sigma_H$), the effect of shear can be neglected, if

$$\tau \leq \tau_{HB} = \frac{T_B}{T_{Fg}} \cdot \frac{\tau_{Fg}}{n_h} = 0{,}6440 \cdot \frac{138{,}5}{1{,}08} = 82{,}6 \text{ MPa}.$$

In cross sections mainly loaded in shear it should be

$$\tau \leq \tau_{HS} = \frac{T_C}{T_{Fg}} \cdot \frac{\tau_{Fg}}{n_h} = 93{,}6 \text{ MPa}.$$

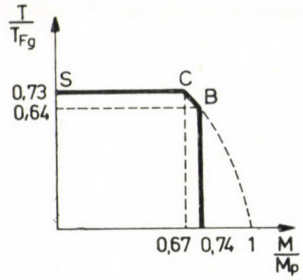


Fig. 5. Interaction diagram for optimized homogeneous I-beams; steel 370; $h/v_g = 145$; $s/v = 30$; $b/h = 1{,}5$; $\tau_{HS} = 93{,}6$ MPa; $\tau_{HB} = 82{,}6$ MPa

Case 2. Hybrid I-girder optimized on the basis of stress constraint with data as follows (stress in MPa): flanges made of steel 52C, web of 37B, thus $\sigma_{Fg} = 240$, $\sigma_{Hf} = 280$, $\sigma_{Ff} = 360$; $\xi = 0{,}86$; $\eta = 0{,}901$; with (17) $\delta = 1/25$.

By using Eq. (18), we get $s = 0{,}2675\ h$ and $v = h/93{,}46$. With (44) $M_p^* = 1{,}665 \cdot 10^{-3}$. Eq. (42) gives

$$T_S/T_{Fg} = 0{,}4645 + 0{,}1803 + 0{,}0854 \sqrt{1{,}665/1{,}34} = 0{,}7400\ .$$

Using Eqs (44) and (45) one obtains $M_{pf}/M_f = v/(4h) = 2{,}675 \cdot 10^{-3}$. Eq. (47) gives $T_C/T_{Fg} = 0{,}8643$. Since $T_C > T_S$, we take $T_C = T_S$. According to (45) $M_p = 0{,}8725\sigma_{Fg}h^2v_g$ and $M_g = \sigma_{Fg}h^2v_g/4$. By means of Eq. (12)

$$M_u = M_{Hf} = \sigma_{Hf}\frac{v_g h^2}{6}(6\eta - 3\xi + \xi^3) = 0{,}577\sigma_{Hf}h^2 v_g\ .$$

With Eq. (48) $T_B/T_C = 0.8929$ and $T_B/T_{Fg} = 0{,}74 \cdot 0{,}8929 = 0{,}6607$. Further, $M_u/M_p = 0{,}7715$ and $M_f/M_p = 0{,}7135$.

The interaction diagram is shown in Fig. 6. In cross sections loaded in $M = 0.7715\ M_p$ ($\sigma_{\max} = \sigma_{Hf}$), the effect of shear can be neglected, if

$$\tau \leq \tau_{HB} = \frac{T_B}{T_{Fg}} \cdot \frac{\tau_{Fg}}{n_h} = 0{,}6607 \cdot \frac{138{,}5}{1{,}08} = 84{,}7 \text{ MPa}\ .$$

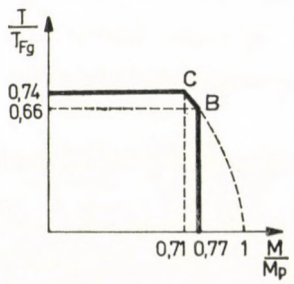


Fig. 6. Interaction diagram for optimized hybrid I-beams; steels 52C/37B; $h/v_g = 145$; $s/v = 25$; $b/h = 1{,}5$; $\tau_{HS} = 94{,}9$ MPa; $\tau_{HB} = 84{,}7$ MPa

In cross sections loaded mainly in shear, it should be

$$\tau \leq \tau_{HS} = \frac{T_C}{T_{Fg}} \cdot \frac{\tau_{Fg}}{n_h} = 94{,}9 \text{ MPa}.$$

REFERENCES

1. Basler, K.: Strength of Plate Girders in Shear. — *J. Struct. Div. Proc. ASCE* **87** (1961) No. ST7, 151—179
2. Bronowicki, A. J.—Felton, L. P.: Optimum Design of Continuous Thin-walled Beams. — *Int. J. Numer. Meth. Eng.* **9** (1975), 711—720
3. Carskaddan, P. S.: Shear Buckling of Unstiffened Hybrid Beams. — *J. Struct. Div. Proc. ASCE* **94** (1968) No. ST8, 1965—1990
4. Chong, K. P.: Optimization of Unstiffened Hybrid Beams. — *J. Struct. Div. Proc. ASCE* **102** (1976) No. ST2, 401—409
5. Delesques, R.: Résistance des ames de poutres sans raidisseurs intermédiaires. — *Construction Métallique* **11** (1974) No. 2, 5—19
6. Design of Hybrid Steel Beams. — *J. Struct. Div. Proc. ASCE* **94** (1968) No. ST6, 1397—1426
7. Farkas, J.: Optimum Design of Crane Bridge and Runway Girders of Box Cross-Section (in Hungarian). — *Proc. 5th Conf. Materials Handling*, Budapest, 1967. MTESZ-KAB.
8. Farkas, J.: Festigkeitseigenschaften von geschweißten, auf Biegung optimal bemessenen I- und Kastenträgern. — *Acta Techn. Hung.* **66** (1969), 427—439
9. Farkas, J.: Economy of Welded Unstiffened I-Section Rods. — *Regional Colloquium on Stability of Steel Structures, Budapest*, 1977. Final Report, Budapest, 1978, 107—112
10. Farkas, J.: Minimization of the Cross-Section Area of Welded Unstiffened Plate and Box Girders Subjected to Bending and Shear. — *Acta Techn. Hung.* **87** (1978), 295—306
11. Farkas, J.—Szabó, L.: Optimum Design of Welded Beams by Means of Backtrack Programming Method. — *Acta Techn. Hung.* (to be published)
12. Kato, B.—Okumura, T.: Structural Behaviour Including Hybrid Construction. — *10th Congress Int. Assoc. Bridge and Struct. Engrs* (IABSE) Tokyo 1976. Introductory Report. Zürich 1976, 199—223
13. Kunihiro, T. et al.: Study of Hybrid Girders. — *10th Congress IABSE Tokyo* 1976. Preliminary Report, Zürich 1976, 409—414
14. Maeda, Y. et al.: Structural Behaviour of Hybrid Girders in Bending. Application to Actual Bridges. — *10th Congress IABSE Tokyo* 1976. Preliminary Report, Zürich, 415—420
15. Manual on Stability of Steel Structures. — European Convention for Constructional Steelwork, 1976
16. Nethercot, D. A.: Buckling of Welded Hybrid Steel I-Beams. — *J. Struct. Div. Proc. ASCE* **102** (1976), No. ST3, 461—474
17. Rockey, H. C.—Škaloud, M.: The Ultimate Load Behaviour of Plate Girders Loaded in Shear. — *Structural Engineer* **50** (1972), 29—47
18. Rockey, K. C.—Evans, H. R.—Porter, D. M.: A Design Method for Predicting the Collapse Behaviour of Plate Girders. — *Proc. Inst. Civ. Engrs.* Part 2. **65** (1978) Mar., 85—112
19. Rockey, K. C.—Valtinat, G.: Zur Traglastberechnung vollwandiger Biegeträger mit Vertikalsteifen. — »Integration von Maschinen- und Stahlbau. Mainz, Kraußkopf-Verlag, 1978.« 95—112
20. Rondal, J.— Maquoi, R.: Optimization of Unstiffened Hybrid I-Beams with Stability Constraints. — *Proc. Regional Coll. Stability of Steel Structures*, Budapest 1977, 373—382
21. Schilling, Ch. G.: Optimum Properties for I-Shaped Beams. — *J. Struct. Div. Proc. ASCE* **100** (1974) No. ST12, 2385—2401
22. Shimada, W. et al.: Improvement of Fatigue Strength in Fillet Welded Joint by CO_2 Soft Plasma Arc Dressing on Weld Toe. — IIW-doc. XIII-881-78. *Int. Inst. Welding*, 1978
23. Szabó, L.: Optimum Design of Welded Hybrid I-Beams. — Ph. D. Dissertation (in Hungarian), Miskolc, Techn. Univ., 1978

24. Жербин, М. М.: Высокопрочные строительные стали. Киев, Будивельник, 1974.
25. Клинов, С. И.: Оптимальные параметры бистальных балок. Сборник Трудов Ленингр. Инж.—Строит. Инст. 1973. № 84, 51—62
26. Куницкий, Л. П.: Оптимизация и экономика сплошных балок из двух классов сталей. «Сопротивление материалов и теория сооружений. Вып. 32. Киев, Будивельник 1978.» 17—25
27. Сикало, П. И.: Практические методы оптимизации сечений бистальных сварных двухтавровых балок минимального веса. «Металлические конструкции и испытания сооружений. Сб. тр. Ленингр. инж—строит. инст. № 1 (134) 1977.» 79—87
28. Строительные Нормы и Правила. Стальные конструкции. СНиП II—В.3.72. Москва, Стройиздат, 1974.

Optimalbemessung auf Biegung und Querkraftstragfähigkeit von hybriden I-Trägern. — Der Artikel behandelt die analytische Optimierung der auf Biegung beanspruchten hybriden (mit Gurtblechen aus höherfester Stahl) geschweißten I-Trägern. In der Zielfunktion wurden die verschiedenen Werkstoffpreise der Steg- und Gurtblechstähle berücksichtigt. Es wurde mit der Beulungsbedingung des Steges und des gedrückten Gurtblechs, ferner mit der Bedingung der maximalen Spannung bzw. Durchbiegung gerechnet. Das Eigengewicht und die Kippung wurde nicht berücksichtigt. Die abgeleiteten einfachen Formeln wurden auf einige Vergleichsberechnungen angewendet, die zeigten, daß z. B. die Werkstoffkosten von den auf Spannungsbedingung optimisierten Hybridträgern um 15÷25% niedrigere sind, als die Kosten von homogenen Trägern. Es wurden die Grenzschubspannungen bestimmt, unter denen der Schub vernachlässigbar ist.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (3—4), pp. 275—284 (1980)

THE THEORY OF RELATIVITY AS SEEN BY THE ENGINEER

Z. HENNYEY*

[Manuscript received, July 27, 1979]

Physics is the most important underlying discipline of the engineering sciences. The way of thinking built upon the notions of absolute space and time, bases of classical physics, means for the man of engineering the stable foundation of his own science. No wonder, if he does not get to friendly terms with the new ideas of the theory of relativity. There is indeed a contradiction between the traditional logics of classical physics and the logics of the theory of relativity and to bridge it is a hard task. To bridge this gap is the aim of this study.

1. Introduction

The difficulties of understanding the theory of relativity arises fundamentally from the fact that this theory rejects the notion of absolute space and thus, as a matter of fact, it removes the ground from the conventional way of thinking. In the following an attempt is made to bridge the "logical vacuum" arising from this fact.

Let us take step by step the fundamentally new ideas of the theory of relativity and see whether they can be interpreted when the axiom of absolute space and time is retained. As the first step we shall mention the equivalence of material and energy. Considering it as objective reality we get to the generalization of Newton's mechanics. The next ideas are the Lorentz contraction and the time dilatation. They have turned up in connection with the Michelson experiment and are considered by the theory of relativity as virtual phenomena. We shall, however, regard them as physical realities.

We arrive at the same formulas as the theory of relativity, but we find them valid in absolute space only! On the other hand, the theory of relativity says — and this is its most important statement verified by experiment — that the laws of physics are uniformly valid in all systems of inertia. Here opens the contradiction which may be bridged by a new interpretation of the Lorentz transformation.

* Z. HENNYEY, H-1068 Budapest, Gorkij fasor 32. Hungary.

2. The generalization of Newton's theory

The equivalence of material and energy expressed by Einstein's famous equation ($E=c^2m$) can easily be interpreted in classical physics.* We must assume that like material, the energy has an inert mass. Consequently, for energy in motion the laws of motion of material systems are valid.

Energy appears most often in connection with material. When a material point of mass m_0 is in motion, together with it moves its kinetic energy too. Thus, the entire inert mass m of the mixture material-energy moving with velocity $\mathbf{v}$ consists of two parts: the *resting mass* m_0 representing the material and the inert mass m_e of the kinetic energy: $m = m_0 + m_e$. The magnitude of the latter is given by the equation $m_e = E_{kin}/c^2$.

The kinetic energy depends upon velocity, consequently when the material-energy complex of mass m moves with changing velocity, its mass will also change. This mass is minimum when the system is resting (then $m = m_0$). Let us observe that the axiom of the equivalence of material and energy *requires the existence of absolute space.* (Hardly can such a definition be accepted of the state of rest according to which it means rest with respect to the observer.)

Classical physics did not take into account the inert mass of energy. Thus, it is conceivable that its results hold only for that special case when the inert mass of energy can be neglected.

Let us follow the fundamental ideas of classical mechanics *without neglecting the mass of energy*.

Let a material point of resting mass m_0 move in the absolute space in constant field $\mathbf{F}$. According to Newton's axiom, the force $\mathbf{F}$ acting on this point changes its impulse during time dt as follows

$$d(m\mathbf{v}) = \mathbf{F}\,dt\,. \tag{1}$$

This differential equation is not sufficient to determine both time functions $m(t)$ and $\mathbf{v}(t)$. One equation more is needed and this is given by the energy axiom. The total mass changes during time dt because its velocity and thus also its kinetic energy changes. The change of the kinetic energy is, however, equal to the work performed by force $\mathbf{F}$ during time dt: $dE_{kin} = \mathbf{F}\cdot\mathbf{v}\,dt$. In these terms the lacking differential equation can be written as

$$dm = (\mathbf{F}\cdot\mathbf{v}/c^2)\,dt\,. \tag{2}$$

The above two differential equations may be regarded as the basis of Newton's heory in its generalized form (completed with the energy axiom).

* This was interpreted in this way by PLANCK (1903).

In classical mechanics the trajectory of mass point has been obtained by integrating the differential equation $m\,\mathrm{d}v = \mathbf{F}\,\mathrm{d}t$. We want to trace back the general case to this special one.

Let us rewrite equation (1), by expanding the differential of the product, into the form $m\,\mathrm{d}\mathbf{v} = \mathbf{F}\,\mathrm{d}t - v\mathrm{d}m$. Now in terms of (2) we eliminate $\mathrm{d}m$:

$$m\,\mathrm{d}\mathbf{v} = (\mathbf{F} - c^{-2}\,\mathbf{v}\,(\mathbf{v}\cdot\mathbf{F})\,\mathrm{d}t\,. \tag{3}$$

This differs from the special case in such a way that on the right side beside the force appears a component named the *relativistic* component of force $\mathbf{F}$:

$$\mathbf{F}_{\mathrm{rel}} = c^{-2}\,\mathbf{v}(\mathbf{v}\cdot\mathbf{F}) = \left(\frac{v}{c}\right)^2 \mathbf{F}_{\mathrm{tg}}\,. \tag{4}$$

Here v is the absolute value of vector $\mathbf{v}$ and $\mathbf{F}_{\mathrm{tg}}$ is the tangential component of force (parallel with velocity). The complementary part of the relativistic component is the "accelerating" component which we denote as $\mathbf{F}_{\mathrm{ac}}$. According to the expansion $\mathbf{F}_{\mathrm{rel}} - \mathbf{F}_{\mathrm{ac}}$, differential equation (1) can also be divided into"components":

$$\begin{aligned} m\,\mathrm{d}\mathbf{v} &= \mathbf{F}_{\mathrm{ac}}\,\mathrm{d}t \\ \mathbf{v}\,\mathrm{d}m &= \mathbf{F}_{\mathrm{rel}}\,\mathrm{d}t \\ \hline \mathrm{d}(m\mathbf{v}) &= \mathbf{F}\,\mathrm{d}t \end{aligned} \tag{5}$$

This expansion is very suggestive. The accelerating component of force changes the velocity while the relativistic one changes the mass. The latter is, according to (4), part of the tangential component and in the extreme case $\mathrm{v} = \mathrm{c}$ coincides with the total tangential component. Thus, in this extreme case the "relativistic-accelerating" decomposition of the force just agrees with the "tangential-normal" one.

The differential equation system (5) is an alternative form of the system (1)—(2) which is the basis of the generalized Newton theory. The advantage of the form (5) is the symmetry appearing in variables m and v. We have seen that in the extreme case $v = c$ the components of force can be interchanged and (5) can be written as

$$m\,\mathrm{d}\mathbf{v} = \mathbf{F}_{\mathrm{norm}}\,\mathrm{d}t \quad \text{and} \quad \mathbf{v}\,\mathrm{d}m = \mathbf{F}_{\mathrm{tg}}\,\mathrm{d}t\,.$$

These equations lead to the very important discovery that the force acting upon the mass point moving with velocity of light *cannot modify the absolute value of velocity*. This follows from the first equation of (5) according to which the acceleration is due to the normal component of force, consequently only the direction of velocity changes but not its magnitude.

The increase of inertia of the moving material point is related to the increase of its kinetic energy and this depends only on the absolute value of velocity. Thus, a relation between variables m and $\mathbf{v}$ must exist and has to be determined now. Let us start out from the basic equations

$$\mathrm{d}(m\mathbf{v}) = \mathbf{F}\mathrm{d}t \tag{1}$$

and

$$c^2\,\mathrm{d}m = \mathbf{F}\cdot\mathbf{v}\,\mathrm{d}t\,. \tag{2}$$

By means of the first equation we eliminate the product $\mathbf{F}\mathrm{d}t$ from the second equation and thus we obtain the following equation:

$$c^2\,\mathrm{d}m = \mathbf{v}\cdot\mathrm{d}(m\mathbf{v})\,. \tag{6}$$

Here appear only the variables m and v, consequently this differential equation is suitable for finding the relation existing between them. After accomplishing the simple calculations and separating the variables we obtain the following differential equation

$$\frac{\mathrm{d}m}{m} = \frac{1}{2}\,\frac{\mathrm{d}(v^2)}{c^2 - v^2}\,. \tag{7}$$

The solution of this differential equation can be written into the form:

$$m_0 = m\sqrt{1 - \left(\frac{v}{c}\right)^2} \tag{8}$$

which is well known from the theory of relativity. This equation gives the material part of total mass m moving with velocity v.

In the above we dealt with mass m moving with the velocity of light (it must evidently be understood as a total mass). It has been established that in this state of motion m can be neither accelerated nor slowed down. The force acting on it can modify only its trajectory. In terms of relation (8) it can be added that the material part of the total mass m moving with the velocity of light is zero. It is a pure energy; the inert mass of light energy can easily be recognized in m moving with velocity of light. An interesting interpretation of the energy of light is found here: this energy is nothing but the *kinetic energy of photon.*

Finally, let us determine the kinetic energy of the total mass m moving with velocity v. Using (8) in the relation $E_{kin} = c^2(m - m_0)$, we obtain

$$E_{kin} = c^2 m\left(1 - \sqrt{1 - \left(\frac{v}{c}\right)^2}\right)\,. \tag{9}$$

This is a formula well known from the theory of relativity, too.

Let us go further and according to our program proeeed with the Lorentz contraction and Einstein's dilatation.

3. The Michelson effect

The most important experimental basis of the theory of relativity is given by the experiments of MICHELSON. Their negative results produce the illusion that we find ourselves in absolute space where the propagation of light is isotropic, i.e. the light propagates with the same velocity in all directions. According to classical physics this cannot be true. But then the Michelson effect, the illusion of absolute space must have some explanation.

This explanation is given by the interpretation of the Lorentz contraction and the time dilatation as objective physical phenomena. Both phenomena are, in the theory of relativity, consequences of the Lorentz transformation, i.e. they follow from the postulation of the equivalence of inertia systems. We follow the inverse way: let us pronounce as postulate the Lorentz contraction and the time dilatation and then arrive at the Lorentz transformation.

It must be assumed that in the material system moving with velocity v_0 in direction of axis x (which is, of course, the velocity in the absolute space) the following physical phenomena occur:

- every body contracts parallel to axis x (x contraction);
- every physical phenomenon slows down in relation to its rate in absolute space;
- the magnitude of both the contraction and slowing down is determined by the same factor $f = \sqrt{1-(v/c)^2}$.

These phenomena in moving material systems "falsify" the measurement of length and time because our yard-sticks shortened in position parallel to axis x show a larger size of x and our slowed down clocks show shorter lengths of time than they are in reality. *For us, however, these measurements mean the reality.*

We take into account this "falsification" of nature by multiplying the coordinates x and t of the Galilei transformation in the moving system by an adequate factor. Thus we obtain the modified Galilei transformation

$$x' = (x - v_0 t)/f\,, \quad y = y'\,, \quad z' = z\,;$$
$$t' = t \cdot f \tag{10}$$

Let us call this transformation shortly *virtual*; indeed, so we perceive the distance x and times t measured in the moving system by physical tools. The virtual transformation reminds us of the Lorentz transformation, it deviates from that only in the time.

Returning to the Michelson effect, let us examine how much is the velocity of light in various directions when measuring it by physical tools in our moving system. Let us consider, e.g., the beam of light departing from the

origin of our coordinate system and propagating along the path making an angle φ' to axis x' in plane x', y' with the velocity *w measured by us.* Then the motion of the point of light is described by the following equations:

$$y' = w \cos \varphi \cdot t \,,$$
$$y' = w \sin \varphi \cdot t \,. \tag{11}$$

Let this motion be transformed into the absolute space by the virtual transformation given by (10). By simple calculus we obtain

$$x = (v_0 + f^2 w \cos \varphi) t \,,$$
$$y = fw \sin \varphi \, t \,.$$

Hence, the components x and y of the absolute velocity of light can be determined. The sum of their squares must be c^2. Writing it down we get a quadratic equation for w with the surprisingly simple solution

$$w = \frac{c}{1 + \frac{v_0}{c} \cos \varphi} \,. \tag{12}$$

(This solution is rational because the quadratic equation could be transformed into a pure quadratic form.) Thus, if we could measure the velocity of light in our moving system by some tool we would get a result depending on both its direction and our own velocity.

Substantially, the measuring arrangement of Michelson compared those intervals of time during which the beam covered the perpendicular arms of identical length. Let the length of the arm be k and the angle of the position of one of the arms φ. According to (12) the time required to cover this distance is

$$T(\varphi) = \frac{k}{c} \left(1 + \frac{v_0}{c} \cos \varphi \right) \,. \tag{13}$$

The direction of the back course is $\varphi + 180°$, thus in the above formula $\cos \varphi$ changes its sign. The sum of partial times of the forward and backward runs is 2k/c. So, the total time depends neither of the direction of the beam nor our own velocity v_0.

According to the above discussion, the measurement of Michelson is insensitive to the anisotropy of the propagation of light. The substantially faulty consequences are nevertheless practically *true.* The experiments verify the fundamental principle of the theory of relativity, the equivalence of the inertia systems. The laws of physics are described in our moving system of coordinates by the same equations as in the absolute space (let us think of the Maxwell equations). *What is the explanation for this?*

4. Theory of relativity

Einstein started with the assumption based on the Michelson experiment that all inertia systems are equivalent. Although we have found that classical physics was right rejecting this conclusion, it has a great practical importance because it represents a necessary idealization expressed in Lorentz transformation. We must conclude that the equivalence of inertia systems, i.e. the validity of Maxwell's equations in all inertia systems, is a very good approximation and it is true with such an accuracy that its error cannot be demonstrated by terrestrial experiments.

This statement can be translated into the language of mathematics as follows:

— *The Lorentz transformation is an approximation of the virtual transformation* representing the physical reality, which is favoured by symmetry properties. *Its error is negligible in terrestrial dimensions.*

For the analysis of this error let us write side by side both transformations, omitting the trivial coordinates y and z:

VIRTUAL	LORENTZ
$x' = (x - v_0 t)\frac{1}{f}$	$x'' = (x - v_0 t)\frac{1}{f}$
$t' = t \cdot f$	$t'' = (t - v_0 x/c^2)\frac{1}{f}$

(14)

where

$$f = \sqrt{1 - (v_0/c)^2}\,.$$

In space coordinates the transformations agree, only the time coordinates differ. At first sight this difference seems to be very substantial. In the formulas x' and t' the contraction of x and the dilatation of t appear explicitly from the position of factor f, but in the entirely symmetric formulas x'' and t'' it seems as if there were contractions in both coordinates. This misleading picture (that is also important from the engineer's viewpoint) will be corrected if coordinates x'' and t'' are expressed by virtual coordinates in terms of the above equations:

$$x'' = x' \quad \text{and} \quad t'' = t' - \frac{v_0}{c^2} x'\,. \tag{15}$$

In these formulas the closeness of the two transformations appears quite clearly: the rate of contraction of x and the dilatation of t coincide exactly, only the origo of the time scale t'' is shifted depending on distance.

This shift of the starting point is the error of the Lorentz transformation which is negligible in terrestrial scales (with small values of x).

We arrived at the end of our experiment. We have built up the completed theory of Newton which by an easily followable conservative way of thinking arrives at the Lorentz transformation. As a matter of fact this way is longer, of course, since Einstein departs from the Lorentz transformation, but for the engineering thinking this certainly is compensated by an easy understanding.

5. Deflection of light

The shift of the starting point of time scale given by the Lorentz transformation depends on distance. It is obvious to seek the error of Lorentz transformation in a phenomenon of "celestial size". Such a phenomenon is the deflection of the beam of light passing near to the Sun. An opportunity to measure it is seldom given, it can be measured only in case of total eclipse of the Sun at some points of the Earth. Owing to this the evaluation of the results of measurements cannot be regarded as concluded yet. This is also shown in PATHRIA's book "The Theory of Relativity" issued in 1974 which dedicates a chapter to the evaluation and reevaluation of the results obtained in measurements of eclipses of the Sun between 1919 and 1952. He gives these results in the following table:

Table 1

Date	Deflection (in seconds of arc)
29 May 1919	2,07 ± 0,09
21 September 1922	1,83 0,11
9 May 1929	1,96 0,08
19 June 1936	2,68 0,37
20 May 1947	2,20 0,38
25 February 1952	1,43 0,18

As the tolerances given in the table show, these results differ substantially from the value of 1,75 seconds of arc prognosticated by EINSTEIN. (There is only one measurement in which this value is within the tolerance range.) This is a direct refutation of isotropy in the propagation of light in inertia systems. Of course, if, knowing that the equivalence of inertia systems is approximative, we can regard this table from other aspects, then we must establish that the error of this approximation which seems to be undetectable (with our tools of to-day) in terrestrial dimensions, is small on celestial scale, too.

According to our theory set forth above, however, the table of PATHRIA has been expectable, if the solar system does not rest in the absolute space. This is only a proof that the light deflection is not independent of the momentaneous position of Sun and Earth in the space. This position in the space is repeated every year, consequently the deflection must also show a yearly periodicity. Let us rearrange the table of PATHRIA according to the dates:

Table 2

Date	Deflection (in seconds of arc)
25 February	1,25—1,61
9 May	1,88—2,04
20 May	1,82—2,58
29 May	1,93—2,16
19 June	2,31—3,05
21 September	1,72—1,94

The order in Table 2 is quite surprising. From these few data, of course, laws of the change of deflection can hardly be concluded. At any rate, it would be interesting to know the results of measurements after 1952 since they would undoubtedly widen our knowledge of the laws of deflection. For us, however, these sparse data of measurements have a proving force, since we sought only for empiric proof of the anisotropy of light propagation.

6. Conclusion

We have arrived from Newton's theory to the Lorentz transformation using the simple and traditional way of thinking of classical physics. The practical equivalence of systems of inertia is the most important principle of modern physics. This principle affirms that physical laws valid in the absolute space can be considered valid in all systems of inertia and thus "illusion of absolute space" is justified.

EINSTEIN started out from the Lorentz transformation; to be more exact, from the *exact* equivalence of systems of inertia. His method has the advantage that it leads directly to the most important principle of modern physics. Its disadvantage is that just his considering the equivalence at an exact feature has led to serious logical contradictions and to the famous paradoxes of the theory of relativity, to the indigestible thoughts on the mutual contraction and dilatation (observers moving near to each other "see contract each other" — but what is the reality?).

The most important advantage of the generalized Newton's theory is its clear structure; it does not require a concept of physical world deviating from the usual way of thinking as e.g. the curved space (it is difficult to consider a straight line as being curved; with respect to what the curvature is?).

Classical physics have been considered by the theory of relativity as an approximation valid at small velocities. We think this is true rather in the reverse order now. Einstein's theory of relativity can be regarded by classical physics as an approximation valid at *small distances*.

REFERENCE

PATHRIA, R. K.: The Theory of Relativity. Pergamon Press, Hindustan, pp. 225—233. 1974

Die Relativitätstheorie mit den Augen des Technikers gesehen. — Die wichtigste Grundwissenschaft der technischen Wissenschaften ist die Physik. Für den Techniker bedeutet die auf die absoluten Raum-Zeit-Begriffe der klassischen Physik aufgebaute Anschauungsweise und Denkungsart die stabile Grundlage seiner eigenen Wissenschaft. Kein Wunder daher, wenn er sich mit den neuartigen Gedanken der Relativitätstheorie schwer anfreundet. In der Tat besteht zwischen der traditionellen Logik der klassischen Physik und der relativitäts-theoretischen Logik ein Abgrund, dessen Überbrückung keine leichte Aufgabe ist. Die vorliegende Studie versucht diesen Abgrund zu überbrücken.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (3—4), pp. 285—290 (1980)

IMPROVABLE BRACKETING OF THE CIRCULAR FREQUENCIES OF A STRAIGHT ROD WITH CHARACTERISTICS VARYING ALONG ITS LENGTH, PERFORMING FLEXURAL OSCILLATIONS

PART II

Á. BOSZNAY*
DOCTOR OF TECHN. SCI.

GY. RICHLIK*—GY. TÓTH*

[Manuscript received April 27, 1979]

The paper solves the problem with the aid of the Poincaré—Rayleigh—Ritz and Fichera's methods and points out some problems arising in connection with the numerical solution.

1. Introduction, aims

The previous paper having the same title, [1] briefly resumes a possible solution of the above problem, i.e. the improvable upper bounds are calculated by the Poincaré—Rayleigh—Ritz finitization and then the Fichera's formula is used for the determination of the lower bounds. In order to facilitate the practical application the authors think it necessary to attract, with the aid of a numerical example, the attention to some problems arising during the application. Subsequently [1] Part I will be considered and the references referring to it will carry a "I" before the number of the equation.

2. The example

The object of the investigations: a rod free at one end, fixed at the other one, of diameter linearly varying along its length, of circular cross section, provided the mass distribution is homogeneous and Young's modulus is constant. The "own system of coordinates" and the geometrical parameters of the rod are shown at Fig. 1. With their aid the diameter d, the area A of the cross section and the necessary cross-sectional moment of inertia I_z are

$$d(x) = \frac{2}{L}\left[LR - (R - r)x\right], \tag{1}$$

* Prof. Dr. Á. Bosznay Chair of Technical Mechanics, Faculty of Electrical Engineering, Technical University, Budapest, Hungary

$$A(x) = \frac{\pi}{L^2}[LR - (R - r)x]^2, \tag{2}$$

$$I_z(x) = \frac{\pi}{4L^4}[LR - (R - r)\,x]^4. \tag{3}$$

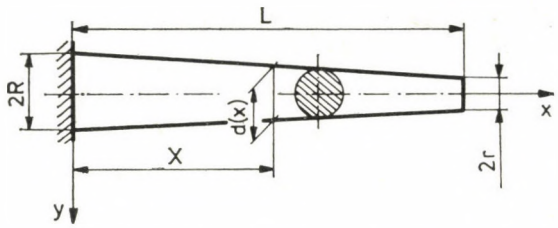


Fig. 1. The "own system of coordinates" of the rod and its determining geometrical parameters

Knowing these functions and provided $\varrho = \text{const}$ and $E = \text{const}$, the operator symbols defined in I./(4) have been calculated. The system of boundary conditions I./(6) is, in the present case, equivalent to

$$v(0) = 0, \qquad v'(0) = 0 \tag{4}$$

geometrical and

$$v''(L) = 0 \qquad v'''(L) = 0 \tag{5}$$

dynamic boundary conditions. (The apostrophs signify differentiation with respect to x.)

3. Choice of the basis functions $\varphi_n(x)$

In the following ***only*** that partial problem will be examined in connection with the question, whether it is sufficient in the example to select these functions from the admissible class of functions (fulfilling only the geometrical boundary conditions).

For the applicability of the larger class of functions it is sufficient if the problem fulfils a special condition of definiteness [2], which was discovered by COLLATZ and named K-definiteness from the initial of the name KAMKE [4].

In order to examine the K-definiteness it is sufficient in the present case to examine the scalar product $(\mathcal{A}\varphi, \varphi)$ in the numerator of the Rayleigh quotient.

Twice integrating partially gives

$$(\mathcal{A}\varphi, \varphi) = \int_0^L EI_z(x)\varphi''^2(x)\,dx + [R(\varphi)]_0^L \tag{6}$$

where

$$R(\varphi) = [\varphi(x)\,(EI_z(x)\varphi''(x))' - EI_z(x)\varphi'(x)\varphi''(x)]_0^L \tag{7}$$

is the Dirichlet's boundary member.

According to KAMKE [2] the definiteness can be examined in the present example if the second and third differential quotient can be eliminated from (7) with the dynamic boundary conditions (5). In our case the members containing $\varphi''(0)$ and $\varphi'''(0)$ cannot be eliminated by (5), thus the K-definiteness cannot be investigated.

Therefore, according to the Kamke principle, the permitted functions cannot be used. The φ_n functions (see I.) must be selected from amongst the set of comparative functions.

4. The φ_n functions used

According to the numerical test of the authors, more accurate bounds are obtained if instead of Hermite polynomials the d^4/dx^4 operator eigenfunctions are used, provided with the boundary conditions (4) and (5). For a similar problem MICHLIN has also used them [3]. They are

$$\varphi_n = a_{1n} \sin \beta_n x + a_{2n} \cos \beta_n x + a_{3n} \sinh \beta_n x + a_{4n} \cosh \beta_n(x)$$
$$n = 1, 2, \ldots \tag{8}$$

where the β_n-s are the positive roots of the transcendent equation

$$\cosh \beta L \cdot \cos \beta L + 1 = 0\,, \tag{9}$$

the a_{jn}-s ($j = 1, 2, 3, 4$) must fulfil the following system of equations which is a consequence of the boundary conditions (4) and (5):

$$\left[\begin{array}{c|c|c|c} 0 & 1 & 0 & 1 \\ \hline \beta_n & 0 & \beta_n & 0 \\ \hline -\beta_n^2 \sin \beta_n L & -\beta_n^2 \cos \beta_n L & \beta_n^2 \operatorname{sh} \beta_n L & \beta_n^2 \operatorname{ch} \beta_n L \\ \hline -\beta_n^3 \cos \beta_n L & \beta_n^3 \sin \beta_n L & \beta_n^3 \operatorname{ch} \beta_n L & \beta_n^3 \operatorname{sh} \beta_n L \end{array}\right] \left[\begin{array}{c} a_{1n} \\ \hline a_{2n} \\ \hline a_{3n} \\ \hline a_{4n} \end{array}\right] = \boldsymbol{0}$$

Here the rank of the coefficient matrix is 3, because e.g. the subdeterminant of the third order belonging to the lower left corner is

$$\beta_n^3 \cosh \beta_n L + \beta_n^3 \cos \beta_n L\,,$$

or also considering the characteristic equation (9) it has the value

$$\beta_n^3 \left(\cosh \beta_n L - \frac{1}{\cosh \beta_n L}\right)$$

which is only zero for $\beta_n = 0$, but this case has been excluded in the foregoing.

Consequently there is one free unknown amongst the a_{jn}-s; one possibility is e.g. the selection $a_{1n} = 1$; another possibility is to normalize φ_n by the prescription $(\mathcal{A}\varphi_n, \varphi_n) = 1$. E.g. MICHLIN calculates in this way [3].

The obvious fact that the upper bounds obtainable in this way are not influenced by the choice of the free unknown has been reflected also in our numerical investigations. For the calculations $a_{4n} = 1$ has been used, therefore

$$a_{2n} = -a_{4n} = -1\,,$$

and

$$a_{1n} = -a_{3n} = \frac{\cos \beta_n L + \cosh \beta_n L}{\sin \beta_n L + \sinh \beta_n L}\,.$$

5. Calculation of the lower bounds

The lower bounds can be calculated with the aid of formula I./(12) if first the values of the integrals in the orthogonal invariants I./(13) have been determined. For this purpose one must first know the operators defined in I./(14). Considering the concrete case, the constants

$$E_l \doteq E\,, \qquad \varrho_l \doteq \varrho$$

$$A_l \doteq A(L) \doteq r^2 \pi\,, \qquad I_{zl} \doteq I_z(L) = \frac{r^4 \pi}{4} \tag{10}$$

defining the "lower base problem" (the prismatic rod-satisfying the geometrical boundary conditions of the original problem — of circular cross-section, the diameter equal to the smallest diameter in the original problem or smaller than it, with the same material constants as the original rod) the operators $\mathcal{A}_l$, $\mathcal{B}_l$, $\mathcal{C}_l$ and $\mathcal{M}_l$ can be established according to Part I. The operator $\mathcal{G}_{jl}$ is the integral operator formed with the Green function as a kernel of the "lower base problem". To form the kernel [4] let us start from the function

$$\eta_l(x) = a_1 \pm b_1 + (a_2 \pm b_2)\, x + (a_3 \pm b_3)\, x^2 + (a_4 \pm b_4)\, x^3 \tag{11}$$

where the "+" sign is valid if $x \leq \xi$ and "−" if $x \geq \xi$ and $x, \xi \in [O, L]$. The values a_i, b_i $(i = 1, 2, 3, 4)$ independent of x can be determined with the aid of the boundary conditions (4), (5) if it is considered that at the point $x = \xi$ $\eta_l(x)$, $\eta_l'(x)$, $\eta_l''(x)$ are continuous, $\eta_l'''(x)$ has the discontinuity $1/EI_{zl}$. (Mechanically the so constructed Green function gives the displacement at an arbitrary point $x \in [O, L]$ of the axis of the rod of the "lower basic problem" under the action of a unit force parallel to the y axis and acting on $x = \xi$, $\xi \in [O, L]$. Considering, that amongst the quantities a_i, b_i $(i = 1, 2, 3, 4)$ may also be quantities dependent on ξ, using the notations $G_l(x, \xi)$ instead

of $\eta_l(x)$,

$$G_l(x, \xi) = \begin{cases} \dfrac{1}{6EI_{zl}} (3\xi x^2 - x^3) , & \text{if} \quad x \leq \xi , \\ \dfrac{1}{6EI_{zl}} (3\xi^2 x - \xi^3) , & \text{if} \quad x \geq \xi . \end{cases} \tag{12}$$

From (12) it can be seen that for a rod of any length L the same Green function is valid. The values of the definite integrals I/(14) can be calculated and the inversion of the designated matrix can be carried out on a computer, inasmuch as the linearly independent systems $\eta_s(x)$ $(s = 1 \ldots k)$ and $\zeta_t(x)$ $(t = 1 \ldots k_2)$ were previously selected. (Let us remark that it is unnecessary to construct two linearly independent systems of functions, even the system of equations (8) can be used.) After these already I./(13) is available in the form of a concrete numerical value and from I./(12) n concrete lower bounds can be gained.

Resuming the results obtained so far, it can be stated that an algorhitm well adaptable to a computer is available for calculating the improvable bounds of the circular frequencies of the chosen problem. I./(12) shows that only by using upper bounds approaching the real solution well can it be hoped to obtain lower bounds approaching the real solutions well enough. (The increased accuracy rule refers to the first several upper bounds.)

6. Numerical example

In order to recognize the efficiency of the described method, in the table are summarized the results for concrete geometrical parameters ($R =$ $= 20$ mm, $r = 15$ mm, $L = 10^3$ mm). For the sake of comparison the lines marked $\alpha^2_{ldi}\, \varrho/E$ and $\alpha^2_{udi}\, \varrho/E$ contain the results for prismatic rods of $2r$ and $2R$ diameter, respectively.

Table

	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$
$\alpha^2_{udi} \frac{\varrho}{E}$	$1{,}235 \cdot 10^{-3}$	$4{,}851 \cdot 10^{-2}$	$3{,}807 \cdot 10^{-1}$	1,462	3,997
$\alpha^2_{ui} \frac{\varrho}{E}$	$1{,}029 \cdot 10^{-3}$	$4{,}327 \cdot 10^{-2}$	$3{,}062 \cdot 10^{-1}$	1,143	3,217
$\alpha^2_{li} \frac{\varrho}{E}$	$1{,}017 \cdot 10^{-3}$	$2{,}983 \cdot 10^{-3}$	$7{,}305 \cdot 10^{-2}$	$8{,}857 \cdot 10^{-2}$	$9{,}317 \cdot 10^{-2}$
$\alpha^2_{ldi} \frac{\varrho}{E}$	$6{,}948 \cdot 10^{-4}$	$2{,}730 \cdot 10^{-2}$	$2{,}140 \cdot 10^{-1}$	$8{,}232 \cdot 10^{-1}$	2,248

The calculations showed that by using the Poincaré—Rayleigh—Ritz method, thus increasing the number of basis functions until 5 the value of the upper bounds improves. Increasing the number of basis functions above 5 the value of the determinant ordered to the coefficient matrix in the eigenvalue problem I./(9) rapidly approached zero and correspondingly for the squares of the resulting eigenfrequencies negative numbers were obtained, too.

When computing the lower bounds $k_1 = k_2 = 1$ and $\eta(x) = \zeta(x) = 6x^2 - 4x^3 + x^4$ was chosen because of the limited computer hours made possibilities for us. The investigations have shown that by choosing suitably the "lower base problem" the numerical errors in the calculation of the lower bounds can be considerably reduced. The prescribed integrations were carried out by modified Romberg procedure. The inversion I./(14) was unnecessary because of the choice $k_1 = k_2 = 1$.

The computer work was done on the CDC 3300 computer of the SzTAKI Institute of the Hungarian Academy of Sciences. The frame of the program system was written in SIMULA 67 program language. The solution of the transcendent equations (9) and of the eigenvalue problem I./(9) were obtained with a FORTRAN program of double machine accuracy.

REFERENCES

1. Kovács, M.—Richlik, Gy.—Takács, F.—Tóth, Gy.: Improvable Bracketing of the Circular Frequencies of a Straight Rod with Characteristics Varying along its Length, Performing Flexural Oscillations (in Hungarian). *Műszaki Tudomány*. **54** (1977), 1—2
2. Kamke, E.: Differentialgleichungen, Lösungsmethoden und Lösungen. Akad. Verlagsges. Geest u. Portig K.-G., Leipzig 1956, 217
3. Michlin, S. G.: Variationsmethoden der mathematischen Physik. Akademie-Verlag, Berlin 1962, 338
4. Collatz, L.: Eigenwertaufgaben mit technischen Anwendungen. Akad. Verlagsges. Geest u. Portig K.-G., Leipzig 1949, 58, 202

Verbesserungsfähige Einschließung der Eigenkreisfrequenzen eines Biegeschwingung untervorfenen geraden Stabes mit verändrelichem Querschnitt. — Die Verfasser lösen das Problem mit Hilfe der Poincaré—Rayleigh—Ritzschen Methode und des Verfahrens von lenken die Aufmerksamkeit auf einige Probleme in Zusammenhang mit der numerischen Fichera, und Lösung hin.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (3—4), pp. 291—301 (1980)

OPTIMAL DESIGNING OF PIPE NETWORKS BY INTEGER PROGRAMMING

M. GRÓSZ*

[Manuscript received September 16, 1978]

In the paper the author deals with the extension of existing water networks and the design of new networks. For the optimal solution of these problems a nonlinear integer programming model is presented. It differs from models known from literature inasmuch as a) the pipe diameters can assume only discrete values; b) the water quantities fed from the sources are not constants given beforehand, but parameters depending on the network; c) minimalization of the investent costs and of the operating expenses is aimed at. The definition of the problem is based on graph theory and network flow theory. In order to reduce the possible number of pipe diameter combinations a special so-called minimal tree has been built up. The nonlinear integer programming problem has been solved by the method of penalty functions. Knowing the minimal tree the discrete diameter combination near the accurate optimum have been sought for which provides, together with the corresponding water quantity a local optimum.

1. Introduction

The present paper deals with the extension of existing water networks and the planning of new networks. During the last few years many paper have been published on this problem, which fact reflects its practical importance.

A good review of the results is given by CENEDESE and MELE [3]. Part of the authors have already assumed discrete variables for the pipe diameters and the investment costs besides [1, 3, 11] they have also striven at minimizing the operating costs [5]. But the majority of authors do not consider the quantity of water fed from the sources (water tower and pump), but considers this as being constant.

In this paper a possible non-linear integer programming model for designing pipe networks is presented where the pipe diameters can assume only discrete values (i.e. they can be selected only from the available pipes) and the real quantity of feed from the sources is considered. Application of the integer programming model became necessary because the results of the continuous model do not satisfy practical requirements.

Optimal designing of pipe network shall mean the following: those pipe diameters belonging to the branches of the network with open branches and

* M. GRÓSZ, Csalit u. 9, H-1025 Budapest, Hungary

with a closed-loop are selected from a given set of pipe diameters where the investment costs and the operating costs are minimum and in addition to the fulfilment of hydraulic conditions the limiting conditions for the nodal pressures are complied with, too. The paper does not deal with optimization of the network outlay, nor with the schedule of the construction of the new network, the prognostication of consumption or the maintenance costs. But the problem is formulated in such a general way that it is suitable for solving other network problems, besides those of water networks.

2. Definition of the network problems

Let us introduce a fictive node into the network [9] and connect it with the source nodes (fictive branches). In the fictive node the fed-in quantity is made equal to the total consumption of the network. The fictive nodal point permits to consider the real water quantities delivered by the water tower and the pumps. This means that the quantities fed-in by the various sources need not be fixed beforehand, but that they may change. Depending on the characteristic of the pump and for several water towers the water levels developing in them can be calculated.

Subsequently the water network is considered a connected graph $D(K, I)$ where K is the set of the edges of the graph. I, the set of the nodes of the graph.

On the edges D of the graph are assumed arbitrary directions (Fig. 1). Let us introduce the sign rule for the nodes according to Fig. 2, and establish the two Kirchhoff laws:

$$p_i - p_j = f_k(d_k) \cdot q_k \cdot |q_k|, \quad (k = 1, \ldots, n;\ i, j = 1, 2, \ldots, m) \tag{1}$$

$$\sum_{k \in K_i} q_k - \sum_{x \in K_i} q_k = a_i, \quad (i = 1, 2, \ldots, n) \tag{2}$$

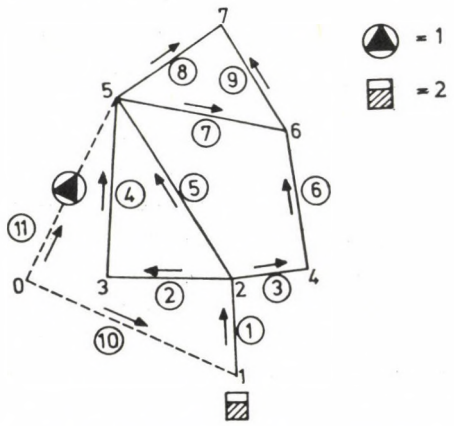


Fig. 1. (1 = pump; 2 = water tower)

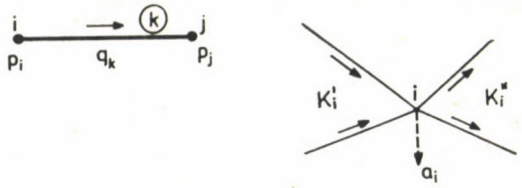


Fig. 2

where

p_i — the node pressure in the i-th node,
q_k — the delivered (flowing water quantity in the edge (i, j) (with sign according to the direction)
$f_k(d_k) = \dfrac{8{,}27 \cdot 10_7 \cdot r_k \cdot l_k}{d_k^5}$
d_k — the pipe diameter belonging to the edge (i, j)
K_i', K_i'' — the sets of incoming and of outgoing edges belonging to the i-th node
r_k — roughness
l_k — length of the edge (i, j)
a_i — consumption in the i-th node
$n = |I|$
$m = |K|$. (the sign $|\ .\ |$ signifies the rank of the set, the number of elements of the set.)

Now let us define the network problems with the aid of edge-node incidence matrix:

$$G = \{g_{ki}^t\} = \begin{cases} 0, \text{ if the } k\text{-th edge does not belong to the } i\text{-th node} \\ 1, \text{ if the } k\text{-th edge runs into the } i\text{-th node} \\ -1, \text{ if the } k\text{-th edge emerges from the } i\text{-th node} \end{cases}$$

By this a sign rule for the vertices has been laid down.

From the theory of network flows it is known [7] that the rank of matrix $G = m(n - 1)$. It is also known that matrix G is always of complete column rank, that means that it is sufficient to establish the nodal law for $n - 1$ nodes.

The two Kirchhoff laws and the limiting conditions are in matrix form:

$$G \cdot p = s\,, \tag{3}$$

$$G^T \cdot q = a\,, \tag{4}$$

$$p \geq p^{\min}\,, \tag{5}$$

where

$s = \{s_i\} = \{f_i q_i |q_i|\}$,
a — the vector of nodal consumptions
$p^{\min}$ — the vector of minimum pressures permitted in the nodes.

It is also known from the theory of network flows [7] that to the non-singular submatrix of matrix G corresponds to a connected tree and conversely.

The number of trees can be given on the graph D as follows:

$$t = \det |G^T \cdot G| .$$

From these trees is constructed a uniquely connected minimum tree. A connected minimum tree means that any vertex of the graph D can be reached along the tree by the shortest route from the nearest source.

The construction of the minimum tree is needed because minimization of the operating costs is strived for so that the majority of the water quantities is transported on the shortest route as far as possible and that this is taken into consideration when choosing the pipe diameters. But this beginning does not influence the result of the optimization. On the other hand, knowing the minimum tree, the matrix G is partitioned into two parts and accordingly Eq. (4) is rewritten:

$$[GR^T \quad GS^T]_x \begin{bmatrix} q' \\ q'' \end{bmatrix} = a$$

where, expressing the transported quantities belonging to the branches of the tree

$$q' = [GR^T]^{-1} (a - GS^T \cdot q'')$$

For one component

$$q'_i = \beta_i - \sum_{t \in N} \alpha_{it} \cdot q''_t \qquad (i \in B) \tag{6}$$

where

$$[GR^T]^{-1} a = \{\beta_i\} \qquad (i \in B)$$

$$[GS \cdot GR^{-1}]^T = \{\alpha_i\} \qquad (i \in B, t \in N)$$

N — the set of the indices of the edges not belonging to the minimum tree (or the set of the indices of the corresponding variables q''_t),
B — the set of indices of the variables q_i belonging to the tree.

Eq. (6) shows that the water quantities belonging to the complementary of the tree uniquely define the water quantities belonging to the tree.

The following chapter presents a possible method for constructing a minimum tree.

3. Construction of the minimum tree

For constructing the minimum tree the solution of the shortest path problem known from literature is used.

To every edge of the tree is assigned a distance z_{ij}, which may only assume a positive value. Let us determine the shortest path from the i-th vertex to

the other vertices of the graph. Then with the aid of the Dijkstra algorithm such a tree is constructed which contains all the shortest paths from the i-th vertex to the other vertices of the graph [7]. The i-th vertex is called the ***root*** of the tree.

The value z_{ij} assigned to the fictive edge is taken as being zero and for the root the fictive node is chosen. The tree constructed by the Dijkstra algorithm is at the same time a minimum tree, any arbitrary node being accessible from the fictive node by the shortest path along the tree, the length of the fictive edge being zero.

To determine the minimum tree the values z_{ij} assigned to the edges are required, therefore, let us revert to the selection of the parameters z_{ij}. For a network situated on a flat land the parameters z_{ij} are equal to the distances between nodes (l_k). In hilly country the parameters z_{ij} must be determine taking into account the height differences between the two nodes.

The Dijkstra algorithm used for constructing the minimum tree a further edge to the tree is connected with each step. Before deciding whether the edge in question belongs to the tree or not, the parameters z_{ij} of the edges must be calculated with which the tree can be continued:

$$z_{ij} = \alpha_{ij}(\Delta h_{ij}) \cdot l_k \tag{7}$$

where

Δh_{ij} — the geodetical height difference between the nodes i and j.

The parameter α_{ij} is chosen according to practical experience and it influences the construction of the minimum tree, but not the solution of the problem, the minimum tree only helps to produce a good approximation.

Now all knowledge necessary for constructing the minimum tree is available and thus the mathematical model of the problem can be dealt with.

4. The mathematical model of the problem

Knowing the minimum tree and correspondingly the decomposition of the matrix G the Kirchhoff law (3) is rewritten as:

$$GR \cdot p = s', \tag{8}$$

$$GS \cdot p = s''. \tag{9}$$

From (8) the node pressure can uniquely determined:

$$p = GR^{-1} \cdot s', \tag{10}$$

and if this is substituted into (9)

$$GS \cdot GR^{-1}s' = s''. \tag{11}$$

The system of Eq. (11) is but the loop law. (11) is rewritten so that the values q' are introduced from (6):

$$\sum_{i \in B} \alpha_{it} f_i(d_i) \cdot \gamma_i \left(\beta_i - \sum_{t \in N} \alpha_{it} q_t''\right)^2 = f_t(d_t) q_t'' \cdot |q_t''| \qquad (t \in N) \tag{12}$$

where

$$\gamma_i = \operatorname{sign}\left(\beta_i - \sum_{t \in N} \alpha_{it} q_t''\right).$$

So a nonlinear system of equations has been obtained where only the water quantities carried along the edge belonging to the complementary of the tree and the pipe diameters belonging to the edges of the complete graph are unknown.

Now let us try to write down the limiting conditions (5) with the aid of the same unknowns. Eq. (10) the node pressure drops, therefore, knowing the permitted node pressure drop (5) can be written as follows:

$$GR^{-1} \cdot s' \leq \Delta p^{\max},$$

or

$$\sum_{j \in B} u_{ij} \gamma_j \cdot f_j(d_j) \left(\beta_j - \sum_{t \in N} \alpha_{it} q_t''\right)^3 \leq \Delta p_i^{\max} \quad (i \in I/J) \tag{13}$$

where

GR^{-1} = $\{u_{ij}\}$
J — the set of such nodes where a source is placed,
$\Delta p_i^{\max}$ — the maximum pressure drop permitted at the i-th node.

Condition (13) can be interpreted only for the nodes not containing the source, besides it is not absolutely necessary to express it for every node.

When stating the problem it has been mentioned that optimum pipe diameters are needed for a given pipe diameter set. In order to write down this in a simple mathematical way, let us introduce a variable x_{ij}:

$$x_{ij} = \begin{cases} 1, \text{ if along the } i\text{-th edge the } j\text{-th pipe diameter type occurs,} \\ 0 \text{ in the opposite case.} \end{cases}$$

As along each edge only one pipe diameter may occur, so the variables x_{ij} must comply with the following condition:

$$\sum_{j=1}^{j^*} x_{ij} = 1, \qquad (i \in K).$$

With the aid of the variable x_{ij} let us write the functions d_i and $f_i(d_i)$:

$$d_i = \sum_{j=1}^{j^*} \delta_{ij} x_{ij},$$

$$f_i(d_i) = h_i \sum_{j=1}^{j^*} \frac{x_{ij}}{\delta_{ij}^5}, \tag{14}$$

where

δ_{ij} — the j-th pipe diameter permitted along the i-th edge,
$h_i = 8{,}27 \cdot 10^7 \cdot r_i \cdot l_i$,
j^* — the number of permitted pipe diameters.

Now let us turn to the objective function. According to CEMBROWICZ [2] the investment cost of the network is

$$\sum_{i \in K} \sum_{j=1}^{j^*} l_i \cdot c_i \cdot d_{ij}^{1/W} \cdot x_{ij} \,, \tag{15}$$

where

$$1 \leq \frac{1}{w} \leq 1{,}5$$

c_i — the cost of construction 1 m of the i-th pipe diameter (Ft/m).

The total operating cost of the network is composed of two factors: the operating costs of the pump stations and the cost of energy loss required for overcoming the network resistance. The first factor is proportionate with the product of the working pressure of the pump and the fed-in water quantity, while the second factor is proportionate to the sum of the absolute value of the pressure drops.

Now the objective function $C(x, q)$ will be, taking into account the total operating and investment costs of the network (14):

$$C(x, q'') = k_1 \sum_{i \in B} \sum_{j=1}^{j^*} h_i \left(\beta_i - \sum_{t \in N} \alpha_{it} q_t''\right)^2 \delta_{ij}^{-5} x_{ij} + k_1 \sum_{i \in N} \sum_{j=1}^{j^*} h_i (q_i'')^2 \delta_{ij}^{-5} x_{ij} + \tag{16}$$

$$+ k_2 \sum_{i \in K_{sz}} \left(\beta_i - \sum_{t \in N} \alpha_{it} q_t''\right) \cdot B_i(q'') + k_3 \sum_{i \in K} \sum_{j=1}^{j^*} l_i \cdot c_i \cdot \delta_{ij}^{1/W} \cdot x_{ij} \to \min.$$

Based on the foregoing the limiting conditions are:

$$\sum_{i \in B} \sum_{j=1}^{j^*} h_i \gamma_i \alpha_{it} \left(\beta_i - \sum_{t \in N} \alpha_{it} q_t''\right)^2 \cdot \delta_{ij}^{-5} \cdot x_{ij} = \sum_{j=1}^{j^*} h_t \cdot q_t'' \, |q_t''| \cdot \delta_{tj}^{-5} \cdot x_{tj} \quad (t \in N) \tag{17}$$

$$\sum_{k \in B} \sum_{j=1}^{j^*} u_{ik} \cdot \gamma_k \cdot h_k \left(\beta_k - \sum_{t \in N} \alpha_{kt} q_t''\right)^2 \delta_{kj}^{-5} x_{kj} \leq \Delta p_i^{\max} \qquad (i \in I/J) \tag{18}$$

$$\sum_{j=1}^{j^*} x_{ij} = 1 \qquad (i \in K) \tag{19}$$

$$q_i'' \leq q_i^{\max} \qquad (i \in K_{sz}) \tag{20}$$

where

k_1, k_2, k_3 — weighting parameters (constants),
$q_i^{\max}$ — the maximum output of the pump,
$B_i(q'')$ — polynomial obtained from the characteristic of the pump.

It depends on the order of magnitude of the weighting parameters, whether greater importance is ascribed to the investment costs or to the operating costs of the planned network.

The first and the second member of the objective function represent the sum of the absolute values of the pressure drops along the edges belonging to the tree and to the complementary of the tree. The third member gives the operating costs of the pump stations, while the fourth provides the investment costs.

In the sense of the construction of the minimum tree the fictive edges belong to the tree. The length of the fictive edges is taken as being zero, the correspoding $h_{kj} = 0$. Therefore, for the first member of the objective function it is enough to carry out the addition for the set B/K_{sz}.

The condition (17) describes the loop law as a function of the variables $q_i''(i \in N)$ and x_{ij}. Conditions (18) and (20) state the permitted pressure drops and accordingly the upper limit for the feeding by the pump.

5. Methods for the solution of the problem

For solving the problem (16)÷(20) two methods can be proposed. The ***first method*** is an algorithm consisting of two steps.

In the first step by some heuristic method a permitted solution must be found, i.e. such a diameter combination which complies with conditions (17)÷(20). The heuristic method is based on the following idea: knowing the minimum tree and starting everywhere from the smallest permitted pipe diameter, the nonlinear system of equations (17) is solved. For the obtained variables q_i'' conditions (18) and (20) are checked (condition (19) is fulfilled automatically). If condition (18) is not fulfilled for some node, so for the shortest path leading from the root of the tree (source) to the node another diameter combination is chosen. In order to reduce the number of combinations and to minimize the objective function it is assumed that along the edges belonging to the tree constructed on the base of the parameters ordered to the edge Z_{ij}, the water transportation is smaller than along the edges belonging to the complementary of the tree. By this we have tried to obtain as good as possible initial estimate.

If for some pump the condition (20) is not fulfilled, to reduce the pipe diameters along the edges connected to this node was tried. Thus, by engineering considerations and with the aid of a heuristic method after a finite number of steps, either a permitted solution is obtained, or it becomes clear that the conditions cannot be complied with (no solution exists for such conditions). If there is a permitted solution, then as a second step with the aid of the implicit enumeration method [6] to reduce the value of the objective function was

tried (using as a starting limit the objective function value of the permitted solution).

The number of combinations being $(j^*)^m$ and for large problems this algorithm being practically unuseable in this form, it is more suitable to use here a simplified enumeration algorithm, which again is based on the knowledge of the minimum tree and the principles connected with it. With this algorithm a local optimum is obtained after a finite number of steps and practically after an acceptable time.

The *second method* considers as a first step the pipe diameters as continuously changing. In this case it is suitable to re-write the problem (16)÷(20) as follows:

$$F(d, q) = k_1 \sum_{i \in B} h_i \left(\sum_{t \in N} \alpha_{it} q_t'' \right)^2 d_i^{-5} + k_1 \sum_{i \in N} h_i (q_i'')^2 d_i^{-5} + k_2 \sum_{i \in K_{sz}} B_i(q'') \times$$

$$\times \left(\beta_i - \sum_{i \in N} \alpha_{it} q_t'' \right)^2 + k_3 \sum_{i \in K} l_i \cdot c_i d^{1/W} \to \min,$$

$$\sum_{i \in B} \alpha_{it} \cdot \gamma_i \cdot h_i \left(\beta_i - \sum_{t \in N} \alpha_{it} q_t'' \right)^2 \cdot d_i^{-5} = h_t q_t'' |q_t''| \cdot d_t^{-5} \qquad (t \in N) \quad (PF)$$

$$\sum_{i \in B} u_{ij} \cdot \gamma_j \cdot h_j \left(\beta_j - \sum_{t \in N} \alpha_{jt} q_t'' \right)^2 \cdot d^{-5} \leq \Delta p^{\max} \qquad (i \in I/J)$$

$$d_i^{\max} \geq d_i \geq d_i^{\min} \qquad (i \in K)$$

$$\beta_i - \sum_{t \in N} \alpha_{it} q_t'' \leq q_i^{\max} \qquad (i \in K_{sz}).$$

In the (PF) problem written in this way the number of unknowns is considerably reduced $(2m - n + 1)$. In the (PF) problem generally the objective function as well as the set of permitted solutions is non-convex in the general case, therefore, the attainment of the global optimum is not guaranteed and so one can speak only of a local optimum.

In the first step the local optimum of the problem (PF) is determined if one exists at all. For this the method of penalty functions are used and thus the problem is reduced to a series of unconstrained extremum value problems for the solution of which efficient methods are already known from literature [4]. For increasing the efficiency of these methods we start out from an estimated system of diameters and the nonlinear system of equations is solved. From the diameters thus obtained and the variables $q_t''(t \in N)$ the starting solution for the second step is obtained.

If in the vicinity of the local optimum of the problem (PF) the problem (16)÷(20) has not even one permitted solution, that does not mean that the problem (16)÷(20) has no solution at all. In such a case we try to find a permitted solution by a heuristic method. If this method does not give a result either then the problem (16)÷(20) has no solution, in the opposite case a local optimum is sought for with the aid of the enumeration algorithm.

With the second step we try to "discretize" the diameters obtained from the optimal solution, i.e. that solution is sought for, where the diameters can assume only the discrete values in store. This part of the calculation is also supported by the minimum tree and the principles connected with it.

The diameter values obtained from the optimum solution of the (PF) problem are discretized according to whether the corresponding edge belongs to the minimum tree or not. If it belongs there, then that diameter value is selected falling nearest to the obtained value. If it does not belong to the minimum tree, from those two diameters of the set between which the obtained diameter is situated, the smaller one is selected. For the diameter series thus selected let us resolve the system of nonlinear equations and then calculate the values of the objective function and its gradient. The latter can be calculated for the unknown diameters only. From among the components of the gradient vector, the one the largest absolute value, is selected and considering the sign the pipe diameter belonging to this component of the gradient is changed for the neighbouring diameter belonging to the set. After that the nonlinear system of equations is again solved and the limiting conditions are checked. If these latter are complied with the value of the objective function is calculated. If the value of the objective function has diminished, the gradients are calculated and the steps just decribed are repeated. If the obtained objective function value has not decreased, the change of the first step is omitted and the above investigation is repeated with the next size gradient component.

With this enumeration algorithm a local optimum is obtained in a finite number of steps (maximum 2^m); although it is not always the global optimum it practically supplies a convenient solution. It can be seen that by the latter method the number of diameter combinations has been successfully reduced from $(j^x)^m$ to 2^m.

In the next chapter the algorithm of the second method will be described.

Description of the algorithm

1. The (PF) problem is solved by the method of penalty functions.

2. If a solution exists for the (PF) problem, then it is marked (d^0, q^0) and step 3. follows, in the opposite case step 18. follows.

3. The corresponding discrete diameter combinations are produced:

a) if the i-th edge belongs to the minimum tree and

$$\delta_{ij1} \leq d_i^0 \leq \delta_{ij2} \qquad (i \in K)$$

then $$d_i^0 - \delta_{ij1} < \delta_{ij2} - d_i^0 \quad \text{follows} \quad d_i^1 = \delta_{ij1}$$

in the opposite case $$d_i^1 = \delta_{ij2}$$

b) if the i-th edge does not belong to the minimum tree, then

$$\delta_{ij1} = d_i^1 .$$

4. $k = 1$ and $T = K$ (set T is identical with set K).

5. For the pipe diameter combination produced the nonlinear system of equations is solved, for a starting solution the vector (d^k, q^0) is taken, the solution is marked $y_k = (d^k, q^k)$.

6. If for y_k the constraints are not fulfilled than 9 follows, in the opposite case the calculations are continued with 7.

7. The value of the objective funtion $C_k = F(d^k, q^k)$ is calculated for the point y_k.

8. If $k = 1$ so $Z = C_1$, in the opposite case 8 follows.

9. If $C_k < Z$, 9 follows, in the opposite case 15 follows.

10. The gradient of the objective function is calculated for d_i at the point y_k.

$$\nabla F_i(d^k, q^k) \qquad (i \in T)$$

11. The following expression is sought for:

$$\max \mid \nabla F_i(d^k, q^k) \mid = t_{i*}.$$

12. $T = T/\{i^*\}$.

13. If $d_i^k = \delta_{ij1}$, so $d_i^{k+1} = \delta_{ij2}$, in the opposite case $d_i^{k+1} = \delta_{ij1}$.

14. $k = k + 1$, $T = K/\{i^*\}$ and we return to 5.

15. $T = T/\{i^*\}$.

16. If $T \neq \emptyset$ (not empty), the calculation is followed by 12, in the opposite case by 17.

17. If $T = \emptyset$ and a discrete diameter combination has not been found, then follows 18, in the opposite case the permitted solution corresponding to the limit Z is a local optimum of the problem (16)—(20) and 20 follows.

18. By a heuristic method a permitted solution is sought for. If that is found the continuation is 19, in the opposite case the problem (16)—(20) has no solution and 20 follows.

19. With the simplified enumeration algorithm a local optimum is sought for.

20. End.

REFERENCES

1. Artina, S.: The Use of Mathematical Programming Techniques in Designing Hydraulic Networks
2. Cembrowicz, R. G.: Optimierung von Rohrnetzen. In the book: W. Zielke, Elektronische Berechnung von Rohr- und Gerinneströmungen, Erich Schmidt Verlag, 1974
3. Cenedese, A.—Mele. P.: Optimal Design of Water Distribution Networks, *Journal of the Hydraulics Division*, **104,** No. HY2, February, 1978
4. Fiacco, A. V.—McCormick: Nonlinear Programming: Sequential Unconstrained Minimization Techniques. Wiley, New York 1968
5. Garbai, L.—Molnár, L.: Optimization of Urban Public Utility Networks by Discrete Dynamic Programming. Colloquia Mathematica Societatis János Bólyai, Progress in Operations Research, Eger 1974
6. Garfinkel, R. S.—Nemhauser, G. L.: Integer Programming, John Wiley, New York 1972
7. Hu, T. C.: Integer Programming and Network Flows. Addison-Wesley Publishing Co., 1970
8. Jacoby, L. S.: Design of Optimal Hydraulic Networks. *Journ. Hydr. Div. Proc. ASCE*, **94,** HY3 (1968)
9. Korte, J. W.—Vielhaber, H.: Ein Beitrag zur Berechnung von Wasserversorgungsnetzen I—IV. *Gas- und Wasserfach* **108** (1967), H. 8., 14., 24., 28
10. Kruskal, J. B.: On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem. *Proc. Amer. Math. Soc.* **7** (1956) 48—50
11. Lam, C. F.: Discrete Gradient Optimization of Water System, *Journal of the Hyraulics Division* **99,** No. HY6, June 1973
12. Molnár, L.: Design of Radial Pipe Networks. *Periodica Polytechnica* **19,** No. 3
13. Watanda, T.: Least-Cost Design of Water Distribution System. *Journal of the Hydraulics Division*, **99,** No. HY9, 1973 Sept.

Optimalentwurf von Rohrnetzen mit Hilfe von ganzzahliger Programmierung. Die Erweiterung von vorhandenen, und der Entwurf von neuen Wasserleitungsnetzen werden behandelt. Zur Optimallösung der Probleme der Konstruktion von Rohrnetzen wird ein nichtlineares ganzwertiges Programmierungsmodell dargestellt. Zur Abfassung des Problems basiert man auf die Graphtheorie und auf die Netzströmungstheorie. Zur Abminderung der Zahl der möglichen Kombinationen wird ein Spezialbaum, der sog. Minimalbaum aufgestellt. Die nichtlineare, ganzwertige Programmierungsaufgabe wird mit Hilfe eines Abrechnungsalgorithmus gelöst. Bei einer anderen Behandlungsmethode des Problems werden die Rohrdurchmesser als kontinuierlich Veränderliche betrachtet, sodann wird die nichtlineare Programmierungsaufgabe mit Hilfe der Straffunktionsmethode gelöst, und bei Kenntnis des Minimalbaums wird eine in der Nähe des exakten Optimums befindliche diskrete Durchmesserkombination gesucht, die mit der Wassermenge q'' ein lokales Optimum liefert.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (3—4), pp. 303—319 (1980)

ELECTROMAGNETIC WAVE PROPAGATION IN MOVING MEDIA WITH SPECIAL REGARD TO FREQUENCY-SHIFTS

"ANOMALOUS" FREQUENCY-SHIFTS IN ASTRONOMY

PART III. EXAMPLES OF APPLICATION

CS. FERENCZ*

[Manuscript received August 2, 1978]

In the first part after a comprehensive review of problems which appear to be fundamental in the study of electromagnetic wave propagation in moving media it is shown how to eliminate the observed discrepancy and the possible ways of analyzing the electromagnetic wave propagation in moving media are discussed. One of the discussed methods, developed under the name of relativistic ray tracing, is described in detail in the second part. It is shown how to determine the signal spectrum and the so-called basic effect for which expressions formulated in cases thought to be most important. The weighting functions and the proofs of their form are given. Finally in the *third part* results of earlier similar analyses are utilized for explaining "anomalous" frequency shifts and other associated wave propagation phenomena (line broadening, polarisation rotation, etc.) observed in the solar system. It is suggested that similar apparently anomalous frequency shifts as well as other associated phenomena observed in galactic and extragalactic wave propagation could be interpreted in the same way.

In this section applications of the above results to some of the reshifts of non-cosmological origin will be briefly reviewed, mentioning, if necessary for completion, also other application fields. Finally perspective applications will be suggested. Since part of the applications has been already summarized in reports [17, 18], details of the interpretations will not be discussed only referred to.

III/1 Solar limb effect [16]

The so-called solar limb effect has been an open problem for about half a century, until 1970. This phenomenon is observed on the extreme limb of the Sun where the redshift of the Fraunhofer lines exceeds the gravitational redshift of 0.636 km/s [2, 3, etc]. The frequency variation within the Sun can be accounted for by the motion of the source, that is by the motion of the absorption site [59]. By making use of expressions (74) and (77) the total measured curve can be predicted. In this case

$$-\frac{\widehat{\Delta f}}{f_0} = \frac{\Delta f_{\text{grav}}}{f_0} - \frac{\Delta f_{\text{str}}}{f_0} + \frac{v \sin \Theta}{c} \, |p(n_s)| \left[1 - e^{-\sigma_0 \int_{vp} N ds}\right] \tag{86}$$

*Dr. Cs. FERENCZ, Puskin u. 24, H-1088 Budapest, Hungary

where Θ is the angle of the radius pointing to the solar surface, to the straight line from the Sun to Earth, Δf_{grav} is the gravitational effect and Δf_{str} is the already known effect produced by the direct motion of the source (absorption site). The substitution of $v \sin \Theta$ for v in (77) running over the surface of the Sun (along the equator, etc.) becomes trivial on seeing the foregoing. Several models have been proposed for Δf_{str}, two of the known models have been used:

a) For simple outflow:

$$\frac{\Delta f_{\text{str}}}{f_0} \cong \frac{v}{c} \cos \Theta \tag{87a}$$

b) For Schröter's model:

$$\frac{\Delta f_{\text{str}}}{f_0} \cong \frac{v}{c} \beta_{\odot} \cos \Theta \tag{87b}$$

where $v = 0.42$ km/s is the proposed value and $\beta_{\odot}$ is explained in [59].

In the calculations the density model was chosen to be an average value of $N_{\max} \cong 10^9$ cm^{-3}. In Fig. 13 it can be seen that the theoretical model (86) is in good agreement with the measured values of the limb effect [2]. The full and open circles represent values measured on the East and on the West limb, respectively. Curve (1) was calculated by making use of (87a) for $v = 0.3$ km/s, while for the calculation of curves (2) and (3) expression (87b) was used. For curve (2) $v|p(n_e)| = 0.3$ km/s, for curve (3) $v|p(n_e)| = 0.57$ km/s. In Fig. 14 curve (2) predicted from model (86)—(87a) is to be seen as turned by 9° and curve (2) predicted from model (86)—(87b) as turned by 6—7° because of an outflow similar to a helical arm.

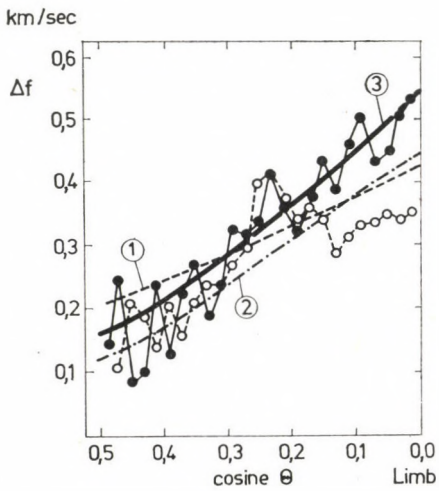


Fig. 13

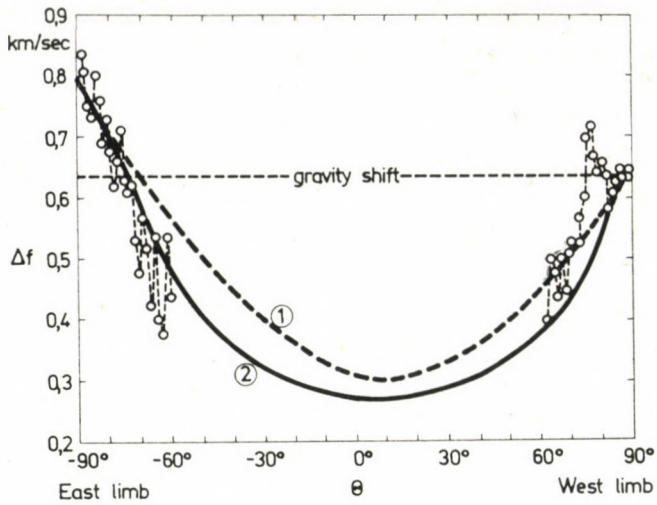


Fig. 14

The most important conclusion which can be drawn is that the total picture including the observed redshift effect is accurately described by (86). The redshift in excess of the gravitational effect is due to the radial outflow of the solar atmosphere.

It is of interest to note, in addition, that (86) does not contain an explicit frequency dependence, it varies only in a slight way with n_e while it has to reflect at the same time the change in the interaction in other ways, too (scattering, change in scintillation, etc.). It can be understood from Fig. 11 that the line broadens in the direction towards the limb. Thus, a method has become available for detailed calculations of this effect in the future. The explanation of the starting (central) lines is assumedly a general problem of solar physics which cannot be specifically related to the limb effect. The present results thus give an explanation of every aspect of the observations [67]. (For the evaluation of Fig. 11, it must be known that in this qualitative diagram the interaction free — effect free — state corresponds to $n = 0$.)

III/2 Occultation redshift and propagation "anomalies" due to propagation in the solar corona [13, 14, 15, 20]

It follows from the above considerations that both the calculation models (74) and (73)—(73a) will necessarily predict propagation anomalies in the inhomogeneous corona expanding at high velocity. Owing to the low density in the corona (a slight departure of the refraction index from 1) and to the rapid decrease in density with increasing distance, the observation of an

anomaly is to be expected only within a distance of a few sun radius and it must be strongly dependent on the average extension area, density and outflow velocity of the corona at the given period. In addition, transient redshift effects should be caused by the high velocity outbursts of matter from the Sun (flare, etc.) if precise measurements are carried out.

The given effect has been observed with both a natural radiation source (Taurus-A) [4] and a space probe (Pioneer-6) [5]. Interpretation studies were made by making use of the conventional corona models [56, 57, 60, etc.] as reported in [13, 15].

Note: In the early investigations [13] the result (74) for α obtained from the total calculation model of the radial outflow has not been available yet. The difficulties caused by the value determined from the secondary effect $\sim v^2/c^2$ were later [15] eliminated by using expression (74).

The total picture model is

$$\frac{\tilde{\Delta f}}{f_0} \cong \left\{ -|p(n_{ie},\beta)| \frac{v_r}{c} + \left[q(n_{ie},\beta) \frac{v_r^2}{c^2} - \right. \right. \\ \left. \left. - |p_{\parallel}(n_e, v_i)| \frac{\tilde{v}_r^2}{c^2} \right] \right\} \left[1 - e^{-\sigma_0 \int_{rp} N ds} \right] \tag{88}$$

where the notations are in agreement with those used in (73a), (74), (82), v_r is the radial outflow, $\tilde{v}_i$ the velocity average of the high velocity quasi-plane parallel outbursts.

In the calculation the average electron density was assumed to vary as r^{-3} over the given heliocentric distances (3—10 $R_\odot$), thus for $3R_\odot$ we calculated with $N_0 = 2 \times 10^5$ cm^{-3} [56, 57, 60]. This gives for α values of $\sim 10^{-6} \geq 10^{-7}$. In inactive periods of solar cycle it may decrease below 10^{-7}. The value of σ_0 was determined by fitting (88) to values measured at a single point. This point was that of $\Delta f_{\max}$ in Sadeh's measurements at each time at which the radiation source Taurus A gives the best approximation to the Sun [13]. Calculating with this value, (88) reproduced the values obtained in both the Taurus A [4] and the Pioneer 6 [5] measurements to an accuracy within the errors of the measurements.

Figures 15.a—b show the data obtained from the Taurus A occultation experiment in 1967 (a) and the combined 1967—68 (b) data as published [4, etc.] along with the curves predicted from (88) where the heliocentric distance is R and the approximation takes the form

$$N(R) \cong \frac{N(3R_\odot) \cdot (3R_\odot)^i}{R^i} \tag{89}$$

with $i = 2, 3, 4$ in accordance with the above considerations. The residual of the measurements will be discussed later. An open problem is, in fact, the basic

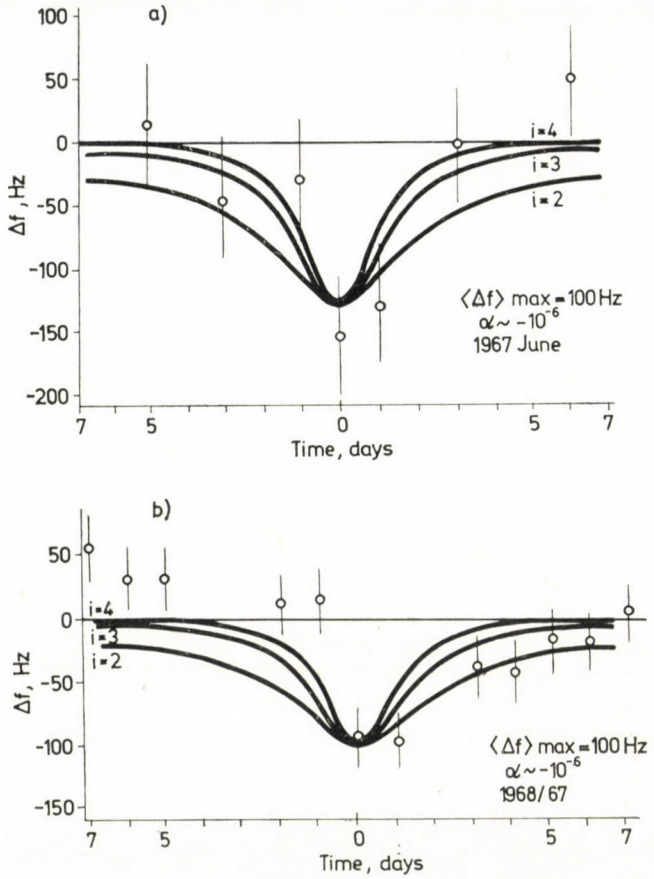


Fig. 15

line of the measurements since the error of the measurements is large. Another difficulty is that the aim was to find a strictly constant effect, and therefore, the 1967 and 1968 results have been published in a combined form. (The measurements were carried out on 21 cm.)

Figure 16 shows the values measured with Pioneer-6 in 1968 (S-band) [5] along with the theoretical curve. The agreement seems to be good. Here (89) was evaluated with the density curve for $i = 3$, only. In this case the earlier (Fig. 15) predicted results had to be fitted into the measured values by taking into account the heliocentric distance. It is of interest to see that the same type of transient effect produced by solar activity, the flares, is quite apparent from the data with low error of measurements. The W-type polarization transients necessarily appearing in the case of inhomogeneous steep fronts [7] — denoted by *IPT* in Fig. 16 — and their uniform interpretation [20] as well as the transient line broadenings, in agreement with the results of the spectrum

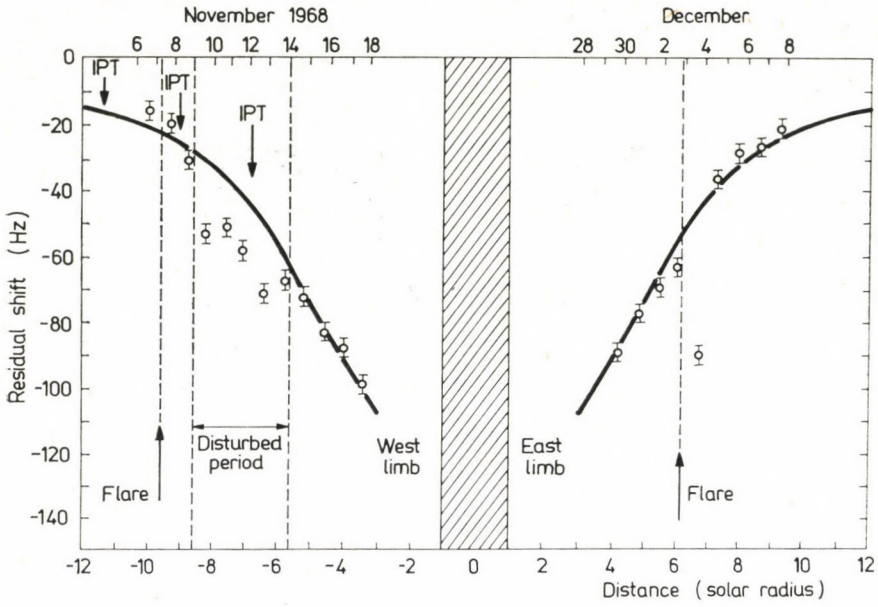


Fig. 16

calculations discussed in section II [8], are another and additional confirmations of the theory. (The delays between the IPT's and $(\widetilde{\Delta f}/f_0)_{\max} - s$ are a consequence of the delay between the pressure and velocity maximums in the streams near the Sun [70] under occultation.) It can be seen from (83) that we are now at the initial stage of $[1 - \exp(-\sigma_n n)]$ and that according to Fig. 11 a broadening by a factor of ~ 10 is to be expected for this transient increase in α and n. A reported broadening [8] in agreement with this prediction is shown in Fig. 17a. As many factors can be responsible for broadening, a detailed analysis of this phenomenon would be beyond the scope of this paper. However, it can be stated that the average value of the measured line broadening corresponds to that expected from Fig. 11 (Fig. 17b [68]).

It has to be noted that the curves predicted for the limb effect by calculations made up to the limb of the Sun ($1R_\odot$) using the known density models, fit well to the redshifts measured on the limb of the Sun within the allowed fluctuations in both cases (II.1 and II.2).

The most important conclusion is that a redshift caused by moving inhomogeneous media appears in the corona which is greatly dependent on the parameters of the latter. It has to be noted that this effect can become immeasurable due to a small decrease in the outbursts, that is, if the higher power terms predominate in the density profile

$$N(R) \cong \sum_i \frac{N_{0i}}{r^i} \quad \text{(in our case } i \geq 3\text{)}$$

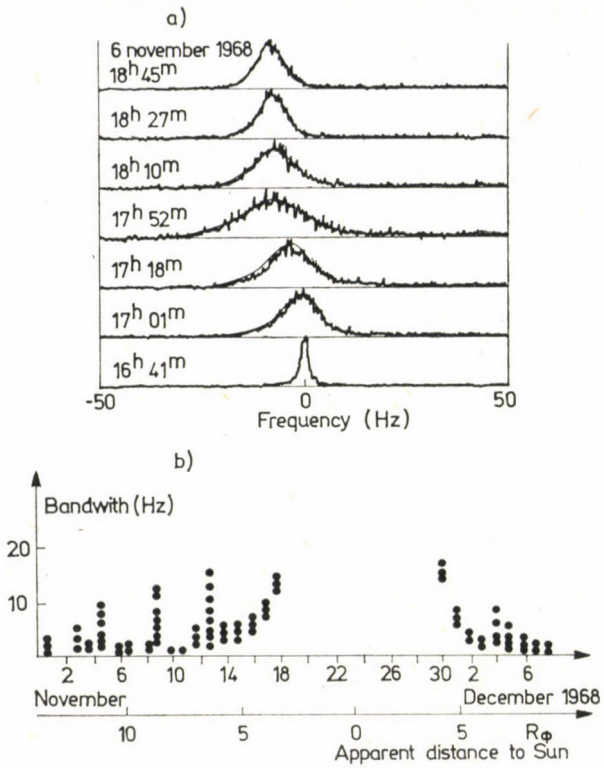


Fig. 17

and that the effect substantially increases with increasing outbursts (flares, etc.). For this reason the results reported in [61] (measurements on 18 cm) are not thought to be contradictions to the theory — as will be shown later. It is thought that if the results of present interpretation are utilized, occultation technique, could become a useful method for the study of the corona. The only objection made to the present interpretation [12, 68] will not be dealt with here since it has already been refuted with relevant references in sections I and II.

III/3 Rotating, inhomogeneous media [15, 18]

In this case expression (80) lends itself to the interpretation of observations as it shows that it changes its sign due to changes in the direction of the velocities on the opposite sides of the "rotating atmosphere". For lack of measurement data of sufficient accuracy no detailed curve fitting could be made because of the strong "geometry" dependence.

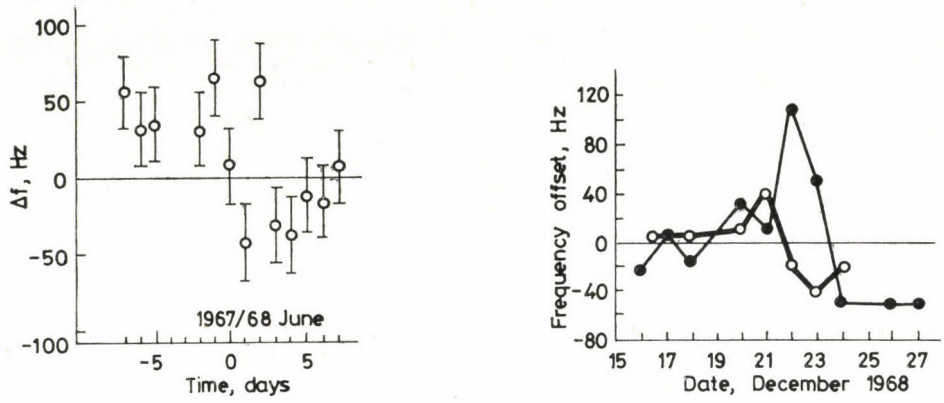


Fig. 18

Nevertheless, some illustrative results are available. The residuals after integration [13, 18] of the Taurus A measurements [4] show the effect of solar rotation of the expected order of magnitude of a few 10^{-8} [13]. A result in agreement with this value was obtained from the measurements reported in [61]. Both results are shown in Fig. 18 with *a)* the Taurus A residuals and *b)* the 1968 occultation experiments on the W28S radiation source.

It is also apparent from Fig. 18 that the effect which can be linearly approximated in the transient stage of decreasing tendency in the time of total occultation which is lower by an order of magnitude as compared with the similar frequency shift of the Pioneer-6 oscillator [5] and thus the effect was automatically suppressed by the mean oscillator frequency determination technique used in the experiment.

It is important to add that the rotation and internal flows of the dense planetary atmospheres [15] render the interpretation (81) and not (80) of the occultation measurements highly problematical [58, 62, 63] if, it is used as usual. For this reason it has already often been proposed [15, 18] that these effects should be taken into account.

III/4 Effect of the terrestrial atmosphere [17, 18, 19]

In this case it is useful to separate the effects occuring in the "traversing"-type of propagation e.g. satellite to Earth as described by (79) from those arising in the "reflected"-type propagation e.g. from Earth to Earth as described by (78).

a) Satellite-Earth relation:

According to calculations [18] the effect described by (79) for signals from satellite to a terrestrial receiver station greatly depends on the angle of

incidence. For perpendicular incidence ($\vartheta = 90°$) the expected effect is

$$\frac{\tilde{\Delta f}}{f_0} \lesssim 10^{-11}$$

It increases if the incidence is oblique e.g. for $\vartheta = 60°$, depending on the structure and state of the ionosphere, it can be

$$\frac{\tilde{\Delta f}}{f_0} \sim 10^{-9} \div 10^{-12}$$

This means that possibly the effect has to be taken into account in extremely precise observation, but generally it can be ignored. An estimation has been made with diffraction index values which hold in the wavelength range from *m* to *dm*.

b) Terrestrial transmitter-receiver relation:

In this case the effect is excepted to occur as described by (78), thus, it has the form

$$\frac{\tilde{\Delta f}}{f_0} \cong - A(n_g, \vartheta, \text{etc}) \frac{v}{c} \cos\vartheta \left[1 - e^{-\sigma_0 \int_{rp} N \cdot s}\right]. \tag{90}$$

Measurements were carried out [1] in accordance with this arrangement and the results were successfully interpreted by making use of a method which can be considered to be a treatment of a special case in terms of the here discussed comprehensive theory. The effect was properly attributed to the contribution from tropospheric winds. In the following it will be shown how these experimental results can simply be reproduced in terms of (90).

The experimental arrangement is shown in Fig. 21 which explains at the same time the meaning of the angle measured with respect to the horizontal axes in Figs 19 and 20 and shows that ϑ in (90) is understood to be this angle.

It has to be taken into account in this analysis that v/c is extremely small. It is known that $v/c \sim 10^{-8}$ and that in (90) it can be taken to be constant along the different paths of propagation [1]. It can also be seen that the observed effect is throughout of the same order of magnitude. (The measurement was made by evaluating the difference in frequency after runs to and fro using 960 MHz in one and 810 MHz in the other direction.) In addition, A is taken to be independent of ϑ in the very restricted range of ϑ. However, since the ray path diverges from a straight line, that is, the path is strongly curved with respect to the inhomogeneous medium (reduced to ϑ the value of ϑ_{eff} increases) [1], the coefficients are treated as an unknown constant in the expansion for the small ϑ. It is an allowed approximation that

$$\sigma_0 \int_{rp} N ds \cong \sigma_0 \bar{N} \int_{rp} ds = C_N s(\vartheta)$$

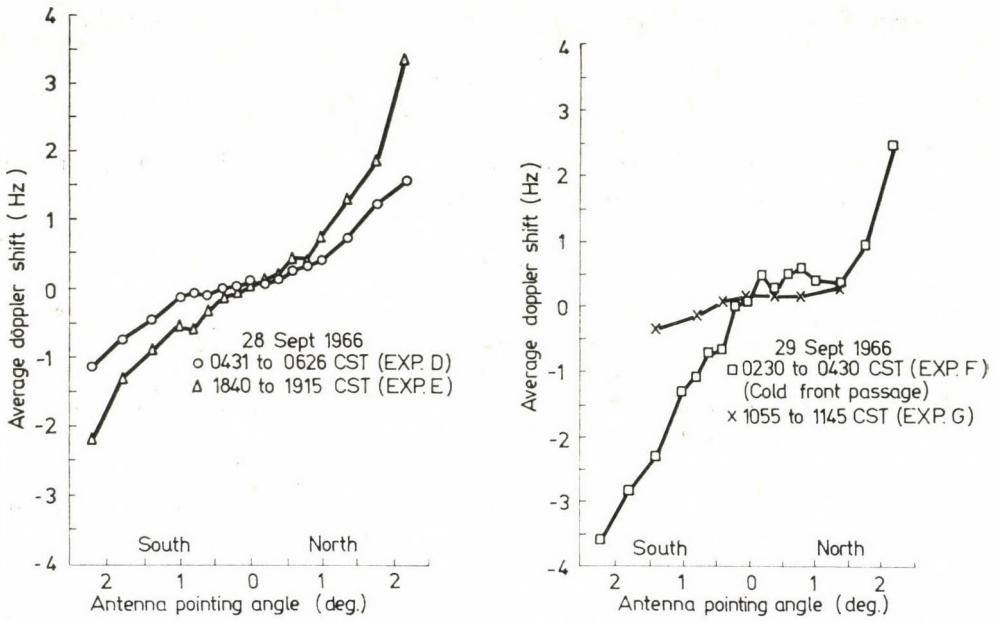

Average doppler shift (Hz)
28 Sept 1966
o 0431 to 0626 CST (EXP. D)
Δ 1840 to 1915 CST (EXP. E)
South
North
Antenna pointing angle (deg.)
Average doppler shift (Hz)
29 Sept 1966
□ 0230 to 0430 CST (EXP. F)
(Cold front passage)
× 1055 to 1145 CST (EXP. G)
South
North
Antenna pointing angle (deg.)

Fig. 19

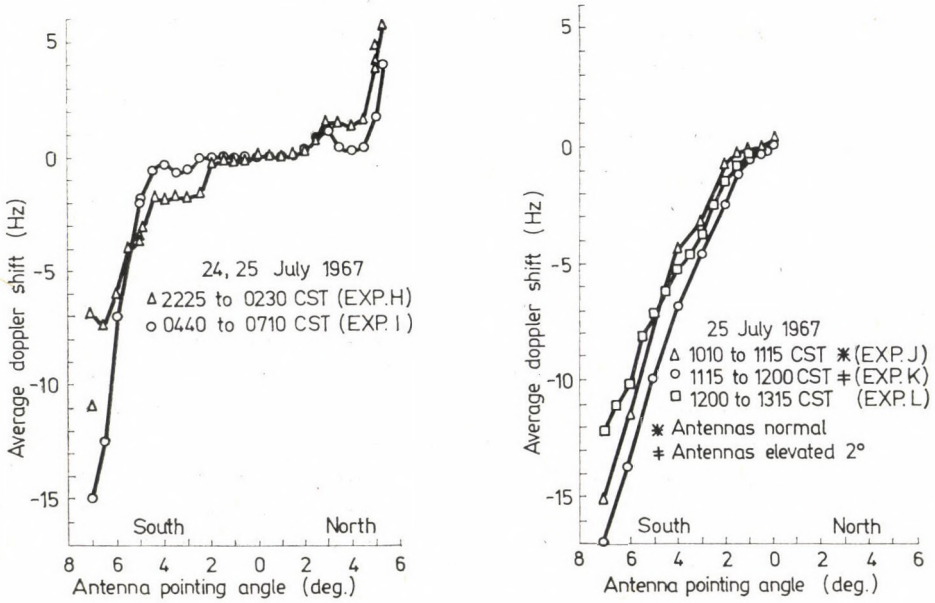

Average doppler shift (Hz)
24, 25 July 1967
Δ 2225 to 0230 CST (EXP.H)
o 0440 to 0710 CST (EXP. I)
South
North
Antenna pointing angle (deg.)
Average doppler shift (Hz)
25 July 1967
Δ 1010 to 1115 CST ✱ (EXP. J)
o 1115 to 1200 CST ǂ (EXP. K)
□ 1200 to 1315 CST (EXP. L)
✱ Antennas normal
ǂ Antennas elevated 2°
South
North
Antenna pointing angle (deg.)

Fig. 20

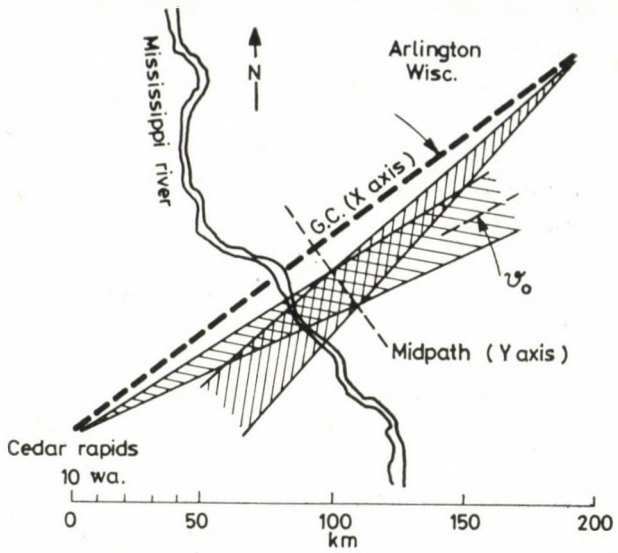


Fig. 21

and it is apparent from Fig. 21 that

$$C_N s(\vartheta) \cong \frac{C_{rp}}{\sin \vartheta} \tag{91}$$

where C_{rp} is a constant uniformly characteristic of propagation path, density etc.

On substituting (91) into (90) and expanding the latter in a power series of ϑ, we get

$$\frac{\tilde{\Delta f}}{f_0} \cong C_1 \vartheta + C_2 \vartheta^3 \tag{92}$$

where C_1 and C_2 depend on A, N, v/c, etc. It is obvious that the curves shown in Figs 19 and 20 as well as all other measured curves reported in [1] are approximated by expression (92). The coefficients C_1; C_2 obtained by fitting and the results of the detailed analysis of the propagation path (Figs 1, 11 etc. in [1]) show a good agreement. The interpretation in terms of the comprehensive theory is simpler and more general.

III/5 Brief review of other fields

In the foregoing a complete set of phenomena (i.e. the solar system) was discussed in terms of the defined aspects. These results suggest that the total application field of the theory should be explored and that its utilization should be attempted in the interpretation of problems which are still open [17]. There

are known "problematical" frequency shifts in some areas of astronomy (a non-complete compilation is given in [55]). There is still a lot of work to be done in this field. Now we must content ourselves with pointing out some possibilities.

a) High velocity flows:

Into this category we may classify the outbursts of materials from explosions in the Universum, the high velocity outbursts from highly active galactic cores, transfer of material between colliding gallaxies, etc. In a crude approximation their general characteristic can be taken to be, besides the high velocity, a central expansion or contraction or the simultaneous appearance of both.

In this case the basic effect described by (74) is expected to predominate. Thus, in extragalactic phenomena of this type the frequency should shift as compared with the average gravitational frequency determined by spacetime metrics more to the red on the "limb" of the phenomenon, if it expands, and more to the blue in the central parts. It is possible to see objects with simultaneous red and blue shifts. (The measurements around the Sun and the arrangement of the solar limb effect can be taken as analogy.) If the flow pattern is more complex a corresponding red-blue ordering can be obtained.

Some anomalous observations in astronomy K-effect in our Galaxy, K-effect in the Magellan Cloud, etc.) seem to be interpretable in an obvious way. The integration of (86) over the total visible stellar surface gives some broadened spectral line exhibiting a change in the average frequency. If the outflowing amount of material is small a blue-shift should appear, on the other hand, for a large amount of material the change of $|p(n_\varepsilon)|$ and of $s(\epsilon, n, \ldots)$ in (84) should cause a predominant redshift.

Interesting effects can be envisaged in pulsing systems, in double, triple, etc. systems with internal material flow.

b) Quasars:

It seems that a high velocity flow of large amounts of material is due to occur in these objects. The explanation of their spectra still seems to be an open problem.

A possible way for isolating the flow (expansion) effect seems to be the construction of a model based on the considerations under points II/4 a.3 and a.5 relying on measurements of line broadening and of possibly other parameters related to flow velocity and on possibly random measurements of distance, related not only to the spectrum. In terms of this model (69) could be used for analyzing the effect produced by the $\tilde{\Delta f}/f$ caused by the flow. This would provide a better model and lead to a better approximation to the quasar arrangement, etc.

It is precisely the importance of these analyses which is thought to justify this detailed summary. Of course, other (technical, etc.) application fields could be envisaged.

IV. Summary, conclusions

IV/1. A study of electromagnetic wave propagation in moving media has been presented as follows:

On summarizing the results of investigations into this field carried out date, it is shown that a discrepancy seems to arise between the transformations of the "velocities" of a moving plane and of a moving point from one type of inertial system into another one. This problem is necessarily encountered in the propagation of electromagnetic waves in moving media when the analysis must simultaneously cover the plane wave and the medium consisting of mass points.

A review of the reported analyses reveals that by them, even by summing up the discrepancy could not be eliminated. A solution of the problem is then proposed with reference to [25].

The reviewed analyses are critically discussed. On showing how the motion of a physical phenomenon is described either in terms of velocity or in terms of propagation vector and how these two descriptions are related to each other, it can be seen that neither an ideal plane, nor an ideal point phenomenon can be well described in either of these terms and that they behave in either way.

The proposed elimination of the discrepancy is general, it is formulated in terms of geometry and it holds for any phenomenon.

It is stated that the proposed model of a physical phenomenon is expected to be invariant with respect to inertial system only if the model is complete.

It follows from the above that for the description of actual, moving physical phenomena one has to find a formalism which transforms at the limit to terms of propagation vector for plane phenomena and to terms of velocity for point phenomena. It is suggested that the simultaneously particle- and wave-like behaviour of physical phenomena predictable in terms of geometry may entail further consequences. This seems to be a promising way for further investigations.

A method is presented for the study of complete interference patterns and the necessity of using the "relativistic" ray tracing procedure is shown.

Understanding that the stationary state and monochromatic are in fact ensured by true plane waves, the relativity of their maintenance in real physical systems is seen.

The importance of the problem whether the electromagnetic field is open or closed and the difficulties this may cause in practice are pointed out. At the same time a possibility is indicated for further investigations of energy impulse tensor forms.

IV/2. The "relativistic ray tracing method" is presented in detail.

First the basic physical problem is discussed. It is shown that in the case of a *moving medium* a stationary excitation in one of the inertial systems represents, even for arrangements in steady state, a non-stationary excitation changing with time into another inertial system. Two kinds of analytical methods are presented, a simple procedure used in the present study and a more accurate one for the analysis of this signal relatively stationary in time.

The relativistic ray tracing method is described in detail. The "basic effect" and the "weighting function" are defined. For illustration a block diagram is given.

The spectrum of the signal generated by the moving set of mirrors is described, in general, and in a special case, giving detailed results for the special case.

It is shown how to determine the average frequency variations caused by moving inhomogeneous media. The functions describing the basic effect are formulated for co-moving inhomogeneous media of plane parallel structure; non-co-moving, inhomogeneous media of plane parallel structure; expanding radial flow; co-moving inhomogeneous media of non-strictly plane parallel structure and for rotating inhomogeneous media of spherical symmetry. The weighting function is constructed from matching results of physical considerations and spectrum calculations.

IV/3. Examples of application are presented:

Interpretation of the solar limb effect.

Uniform interpretation of the anomalous redshifts (frequency variations) in solar occultation experiments also showing their relation to the solar limb effect.

The effect of the rotation of the Sun around its axis is analyzed and it is pointed out that the analogue of this effect should be taken into account in the occultation study of fast-rotating planets with dense atmospheres.

Effects occurring in the terrestrial atmosphere and their fitting to the uniform interpretation are shown.

It is suggested that this interpretation method could be successful in explaining phenomena (K-effect, etc.) within the galaxy and anomalies observed in galaxies-galactic groups and in the study of high velocity drifts and of quasars.

IV/4. The proposed theory permits the "anomalous redshifts" to be trivially and uniformly explained in a way which simultaneously confirms the extensively supported Big-Bang theory and, of course, it is consistent with our present notions of physics (it seems that it helps, also in elucidating the concepts of "plane" and "point").

IV/5. It seems possible in the light of present notions to group the effects inducing changes in experimental investigations and concerned with space research and astronomy.

a) Possible causes of frequency shifts: movement of signal source, movement of receiver, space-time geometry (gravitation), moving inhomogeneous medium.

b) Possible causes of line broadening: rotation of signal source etc. (direct motions) turbulence, thermal convection, flowing medium, quantum-physical causes (ratio of spontaneous to induced emission, etc.).

c) Possible causes of change in polarization: anisotropic medium, change of polarization in signal source, traversal of inhomogeneity, absorption sensitive to polarisation.

This grouping is neither complete nor final, nevertheless, it seem to be useful.

*

Author is indebted to geophysicist GY. TARCSAI for his collaboration in interpretations and model calculations and to colleagues GY. PÁL astronomer and F. KÁROLYHÁZY physicist for advice and critical discussions.

REFERENCES

1. W. P. BIRKEMEIER, H. S. MERRIL, D. H. SARGEANT, D. W. THOMSON, C. W. BEAMER, G. T. BERGMANN: Observation of Wind-Produced Doppler-Shifts in Tropospheric Propagation; Radio Science; **3,** 309, 1968
2. M. G. ADAM: Interferometric Measurements of Solar Wavelength and an Investigation of the Einstein Gravitational Displacement; Mon. Not. Roy. Astron. Soc; **108,** 446, 1948
3. L. A. HIGGS: The Solar Red-Shift; Mon. Not. Roy. Astron. Soc; **121,** 421, 1960
4. D. SADEH, S. H. KNOWLES, B. S. YAPLEE: Search for a Frequency Shift of the 21 centimeter Line from Taurus A Near Occultation by the Sun; Science; **159,** 307, 1968
5. R. M. GOLDSTEIN: Superior Conjuction of Pioneer 6; Science; **166,** 598, 1969
6. J. A. JACOBS, T. WATANABE: Doppler Frequency Changes in Radio Waves Propagating Through a Moving Ionosphere, Radio Science, **1,** 257, 1966
7. M. K. BIRD: Coronal Transient Events Observed with S-Band Faraday Rotation Measurements during Solar Occultation; Space Res., **XVI.** 711, Akademie-Verlag, Berlin, 1976
8. R. WOO, F. C. YANG, A. ISHIMARU: Structure of Density Fluctuations Near the Sun Deduced from Pioneer-6 Spectral Broadening Measurements; The Astrophys. J; **210,** 593, 1976
9. J. W. KIEREIN, B. M. SHARP: Compton Effect Interpretation of Solar Red-Shift; Solar Phys; **3,** 450, 1968
10. P. MERAT, J. C. PECKER, J. P. VIGIER: Possible Interpretation at an Anomalous Redshift Observed at the 2292 MHz Line Emitted by Pioneer-6 in the Close Vicinity of the Solar Limb; Astron. Astrophys; **30,** 167, 1974
11. A. A. CHASTEL, J. F. HEYVAERTS: Perturbations of Pioneer 6 Telemetry Signal During Solar Occultation; Nature; **249,** 21, 1974
12. A. A. CHASTEL, J. HEYVAERTS: Broadening and Anomalous Shift of Pioneer VI Telemetry Line an Effect of Coronal Inhomogeneities Useful for Diagnostics; Astron. Astrophys; **51,** 171, 1976
13. Cs. FERENCZ, GY. TARCSAI: A New Experimental Possibility of Investigating the Solar Corona: Frequency Measurements on Radio Sources When Occulted by the Sun; Planet. Space Sci; **18,** 1213, 1970
14. Cs. FERENCZ, GY. TARCSAI: Redshift during Pioneer-6 Solar Occultation — Unexplained or Predicted; Nature; **252,** 615, 1974
15. Cs. FERENCZ, GY. TARCSAI: Frequency Shift Effects due to Atmospheric Motions in Interplanetary Occultational Measurements; Space Res., **XVI.** 705, Akademie-Verlag, Berlin, 1976
16. Cs. FERENCZ, GY. TARCSAI: Theoretical Explanation of the Solar Limb Effect; Planet. Space. Sci; **19,** 659, 1971
17. Cs. FERENCZ, GY. TARCSAI: Interaction of Gravitational and Electromagnetic Fields or Another Effect?; Nature; **233,** 404, 1971

18. Cs. FERENCZ, GY. TARCSAI: Refraction Effects due to Moving Media in Doppler Measurement; Space Res; **XII.** 595, Akademie-Verlag, Berlin, 1972
19. Cs. FERENCZ: Electromagnetic Wave Propagation in Inhomogeneous Linear Media; Thesis submitted for the degree of Candidate of Sciences, Budapest, 1970, Library of the Hung. Acad. Sci. (in Hungarian)
20. Cs. FERENCZ: Electromagnetic Wave Propagation in Inhomogeneous Media: The Analysis of the Rotation of Polarization, and the Application of the Principle of Modified Ray Tracing; Part I—II., Acta Techn. Acad. Sci. Hung. **86,** 363, 1978; **87,** 263, 1978
21. D. MIDDLETON: A Statistical Theory of Reverberation and Similar First-Order Scattered Fields-Part. III; IEEE Trans. on Inf. Theory; **IT-18,** 35, 1972
22. Cs. FERENCZ: Electromagnetic Wave Propagation in Inhomogeneous Media: Strong and Weak Inhomogeneities: Acta Techn. Acad. Sci. Hung., **85,** 433, 1977
23. Cs. FERENCZ: Electromagnetic Wave Propagation in Inhomogeneous Media: Method of Inhomogeneous Basic Modes; Acta Techn. Acad. Sci. Hung., **86,** 79, 1978
24. Cs. FERENCZ: Electromagnetic Wave Propagation: The Analysis of the Group Velocity; Acta Techn. Acad. Sci. Hung., **86,** 169, 1978
25. Cs. FERENCZ: A Geometric Resolution of the Contradiction between the Propagation of Electromagnetic Plane Waves in Moving Dielectrics and the Einsteinian Addition of Velocities; Acta Techn. Acad. Sci. Hung., **84,** 147, 1977
26. Cs. FERENCZ: Permittivity of Inhomogeneous Media; Proc. of the Fourth Coll. on Micr. Com; **ET-10,** Akadémiai Kiadó, Budapest, 1970
27. I. E. TAMM: Osznovi Teorii Elektritsestva, Izd. "Nauka", (in Russian) Moscow, 1966
28. J. J. BRANDSTATTER: An Introduction to Waves, Rays and Radiation in Plasma Media; McGraw-Hill Book. Co. Inc. New York, 1963
29. K. NOVOBÁTZKY: The Theory of Relativity; Tankönyvkiadó (in Hungarian), Budapest, 1951
30. A. EINSTEIN: Ann. der Physik; **17,** 891, 1905. („Válogatott Tanulmányok" — in Hungarian, Gondolat, Budapest, 1971)
31. J. L. SYNGE: Relativity, the Special Theory; North-Holland Publ. Co. Amsterdam, 1965
32. M. V. LAUE: Die Relativitätstheorie, I; Vieweg und Sohn; Braunschweig, 1955
33. H. OTT: Zum Energie-Impuls Tensor der Maxwell-Minkowskischen Elektrodynamik; Ann. der Physik, (6), **11,** 33, 1952
34. F. BECK: Die Allgemeingültigkeit des Trägheitsgesätzes der Energie in der Planckschen Fassung; Z. für Physik; **134,** 136, 1953
35. G. MARX: Relativistische Elektrodynamik der Magnete; Acta Phys. Acad. Sci. Hung., **II,** 67, 1952
36. G. MARX: Das Elektromagnetische Feld in Bewegten Anisotropen Medien; Acta Phys. Acad. Sci. Hung., **III,** 75, 1953
37. H. ARZELIES, DE J. HENRY: Milieux Conducteurs on Polarisables en Mouvement; Travaux de l'Institut Scientifique Chérifien, Série Sciences Physiques; **No. 5,** Rabat, 1959
38. E. SCHMUTZER: Relativische Physik (Klassische Theorie): B. G. TEUBNER Verlagsgesellschaft, Leipzig, 1968
39. R. P. FEYNMANN, R. B. LEIGHTON, M. SANDS: The Feynmann Lectures on Physics, II; (in Hungarian) Műszaki Könyvkiadó, Budapest, 1970
40. L. D. LANDAU, E. M. LIPSIC: Theoretitseskaia Fizika II; Teoria Polia; (in Hungarian) Tankönyvkiadó, Budapest, 1976
41. W. P. ALLIS, S. J. BUCHSBAUM, A. BERS: Waves in Anisotropic Plasmas; M.I.T. Press, Cambridge, Mass., 1963
42. I. V. LINDELL: On the Definiteness of the Constitutive Parameters of a Moving Anisotropic Medium; Proc. IEEE; **60,** 638, 1972
43. J. A. ARNAUD, A. A. M. SALEH: Theorems for Bianisotropic Media; Proc. IEEE; **60,** 639 1972
44. J. A. KONG, D. K. CHENG: Wave Reflections from a Conducting Surface with a Moving Uniaxial Sheath; IEEE Trans. on Ant. and Prop; **AP-16,** 577, 1968
45. J. A. KONG, D. K. CHENG: Reflection and Refraction of Electromatnetic Waves by a Moving Uniaxially Anisotropic Slab; J. of Appl. Phys; **40,** 2206, 1969
46. S. W. LEE, Y. T. LO: Radiation in a Moving Anisotropic Medium; Radio Science; **1,** 313, 1966
47. J. A. KONG, D. K. CHENG: Modified Reciprocity Theorem for Bianisotropic Media; Proc. IEEE; **117,** 349, 1970
48. D. CENSOR: Ray Tracing in Weakly Nonlinear Moving Media; J. Plasma Phys; **16,** 415, 1976

49. M. E. KARPUK, W. G. TIEDERMAN: Effect of Finite-Size Probe Volume Laser Doppler Anemometer Measurements; IAAA Journal; **14,** 1099, 1976
50. K. G. BUDDEN: Radio Waves in the Ionosphere; Cambridge at the Univ. Press; 1966
51. L. B. FELSEN: Rays, Modes and Equivalent Networks; Proc. of the Fourth Coll. on Micr. Com; **ET-9,** Akadémiai Kiadó, Budapest, 1970
52. R. CRISTESCU, G. MARINESCU: Introduction to the Theory of Distributions and its Applications; (in Hungarian) Műszaki Könyvkiadó, Budapest, 1969
53. Cs. FERENZ: Wave Propagation in Inhomogeneous Linear Media; Acta Techn. Acad. Sci. Hung. **68,** 215, 1970
54. R. S. LAWRENCE, D. J. PASAKONY: A Digital Ray-Tracing Program for Ionospheric Research; Space Res. **II,** 258, North-Holland Publ. Co; Amsterdam, 1961
55. J. C. PECKER: Possible Explanations of Non Cosmological Redshifts; College de France Institute d'Astrophysique de Paris (preprint) 1976
56. A. ZIRIN: The Solar Atmosphere; Blaisdell, Waltham, Mass., 1966
57. D. E. BILLINGS: A Guide to the Solar Corona; Academic Press, New York, 1966
58. G. FJELDBO: Mariner 10 and Pioneer 10 and 11 Radio Occultation Measurements of Planetary Ionospheres; COSPAR XVIII. Plen. Meet; VII, **4.5,** Varna, 1975
59. E. H. SCHRÖTER, Zur Deutung der Rotverschiebung und der Mitte-Rand-Variation der Fraunhoferlinien bei Berücksichtigung der Temperaturschwankungen der Sonnenatmosphere; Zeitschrift für Astrophysik; **41,** 141, 1957
60. E. N. PARKER: Coronal Expansion and Solar Corpuscular Radiation; Proc. of the Plasma Space Sci. Symp; 99, Reidel, Dortrecht, 1965
61. I. A. BALL, D. F. DICKINSON, A. E. LILLEY, H. PENFIELD, T. I. SHAPIRO: Search for an Effect of the Sun on the Frequency of 18-Centimeter Radiation; Science; **167,** 1955, 1970
62. V. R. ESHLEMAN: Jupiter's Atmosphere: Problem and Potential of Radio Occultation; Science, **189,** 876, 1975
63. A. G. SOLOVJEV, A. D. KUZMIN, A. N. KAZANTSEV: "Pioneer 10" and Radio Astronomical Measurements of Jupiter; COSPAR XVIII. Plen. Meet; VII. **4.2,** Varna, 1975
64. M. OHKUBO: Reflection and Transmission of Electromagnetic Waves by Anisotropic Plasma Half-Space Moving in Longitudinal Direction; IEEE Trans. on Ant. and Prop; **AP-19,** 569, 1971
65. K. DAVIES: The Measurements of Ionospheric Drifts by Means of a Doppler Shift Technique; J. of Geophys. Res; **67,** 4909, 1962
66. K. DAVIES: Doppler Studies of the Ionospheric Effects of Solar Flares; Proc. of the Int. Conf. on the Ionosphere, 1962; 76, Chapman and Hall Ltd; London, 1963
67. M. G. ADAM, P. A. IBBETSON, A. D. PETFORD: The Solar Limb Effect: Observations of Line Contours and Line Shifts; Mon. Not. Roy. Astr. Soc., **177,** 687, 1976
68. A. A. CHASTEL: A Critical Analysis of the Explanation of Red-Shifts by a New Field; Astron. Astrophys., **53,** 67, 1976
69. G. L. JAMES: Geometrical Theory of Diffraction for Electromagnetic Waves; IEE Electromagnetic Waves Series 1; Peter Peregrinus Ltd. London, 1976
70. L. F. BURLAGA: Magnetic Fields, Plasmas and Coronal Holes: The Inner Solar System; SI. **1.2.** XXI. COSPAR Plen. Meeting, Innsbruck, 1978

Fortpflanzung von elektromagnetischen Wellen in bewegten Medien mit besonderer Berücksichtigung der Frequenzänderungen (Anomalistische Frequenzverschiebungen in der Astronomie). III. Anwendungsbeispiele. — Im dritten Teil der Arbeit wird schließlich aufgrund der Ergebnisse der vorhergehenden Untersuchungen eine einheitliche Erklärung für die im Sonnensystem beobachteten »anomalistischen« Frequenzverschiebungen und die daran anschließenden sonstigen Wellenfortpflanzungserscheinungen gegeben (Linienverbreiterung, Polarisationsdrehung, usw.). Es wird wahrscheinlich gemacht, daß auf die gleiche Weise auch die galaktischen und extragalaktischen »anomalistischen« Frequenzverschiebungen (sowie sonstige, anschließende Wellenausbreitungserscheinungen) erklärt werden können.

Acta Technica Academiae Scientiarum Hungaricae, Tomus 90 (3—4), pp. 321—349 (1980)

GENERAL ANALYSIS OF MONOCHROMATIC SIGNALS PROPAGATING ALONG INHOMOGENEOUS TRANSMISSION LINES

PART II

I. ÁRKOS, Mrs. FERENCZ*

[Manuscript received September 8, 1978]

The *first part* of this paper gives a short summary of fundamental relations and common methods for the investigation of inhomogeneous transmission lines. Then using the method of inhomogeneous basic modes it shows a process for determining the wave pattern along the line which seems to be generally sufficiently useful. Based on this, one may have different categories of transmission lines. Finally by using a further method too, the way of determining the wave pattern and the different line categories is shown. The *second part* of the paper gives the manner of determining the group velocity by using a perturbation process based on the method of inhomogeneous basic modes. In this case homogeneous, inhomogeneous and within strictly linear and dispersive lines are analysed. Finally an example is shown for determining the wave pattern and group velocity by using the method of inhomogeneous basic modes.

6. Group velocity on the transmission line

Besides determining the wave pattern, it is very interesting to investigate the propagation of energy, that is, the group velocity of the signal under propagation both in free space and along transmission lines. This is the first step of determining the delay-time, and of analyzing the propagation of transients or modulations.

6.1 Conditions for defining the average energy

As is known [1, 2, 8, etc.], the energy equation coming from Maxwell's equations takes the form of

$$\oint_A \bar{S}\, d\bar{A} = -\int_V \left(\bar{H} \frac{\partial \bar{B}}{\partial t} + \bar{E} \frac{\partial \bar{D}}{\partial t} \right) dV, \tag{6.1}$$

where the Poynting vector is

$$\bar{S} = \bar{E} \times \bar{H}.$$

* Ferenczné Árkos Ilona, H-1088 Budapest, Puskin u. 24.

It can be seen, that the result of the energetic analysis depends generally on geometry and time too, apart from the characteristics of the electromagnetic field. For a more general investigation the stochastic analysis of the momentary Poynting vector $(\bar{s}\,\bar{r}, t)$ may provide more information on the oscillation, drifting in space of the energy. However, group velocity cannot be spoken about, for the results depend on the interval of integration, too.

The problem will be somewhat easier in case of transmission lines, as was settled in the beginning, that the average values of the different parameters are unambiguously defineable (Part 2).

Further, a detailed analysis using wave propagation showed [8] the conditions of determining average energy, independently of the surface along which it is "measured" and of time, and the propagation of which may already be defined and therefore group velocity can be determined.

The momentary value of Poynting's vector is

$$\bar{s}(\bar{r}, t) = \bar{E}(\bar{r}, t) \times \bar{H}(\bar{r}, t) .$$

The average in time of the energy flowing through a surface $\Delta\bar{A}$ during a time interval $\Delta t = t_2 - t_1$ is

$$\widetilde{W} = \frac{1}{\Delta t}\left[\int_{t_1}^{t_2}\int_{\Delta A} \bar{s}(\bar{r}, t)\, d\bar{A}\right] dt$$

or (6.2)

$$\widetilde{W}^0 = \int_{\Delta A}\left[\frac{1}{\Delta t}\int_{t_1}^{t_2} \bar{s}\,(\bar{r},\ t)\, dt\right] d\bar{A} .$$

The average mentioned above may be considered as unambiguous in that case when the integrals in t and $\bar{A}$ can be commutated, that is, when

$$\widetilde{W} = \widetilde{W}^0 \tag{6.3a}$$

and in an unambiguous and sufficiently small neighbourhood of a point $(\bar{r}_0, t_0)$

$$\frac{\partial \widetilde{W}}{\partial(\Delta t)} \simeq 0 \quad \text{and} \quad \operatorname{grad}_{\Delta\bar{A}} \widetilde{W} = \text{const.} \tag{6.3b}$$

are valid.

The validity of these conditions will be investigated now. An electromagnetic field of the form of (4.4) will be chosen which in case of transmission lines has the form of (4.8) or (4.21), as was seen earlier. These are stationary solutions, that is

$$\frac{\partial \bar{F}_0}{\partial t} \equiv 0 .$$

Therefore, with the aim of determining v_g, $\bar{F}_0$ may only be perturbed by satisfying the condition:

$$0 \cong \frac{\partial \bar{F}_{\text{opert}}}{\partial t} \ll \frac{\partial \bar{F}}{\partial t} . \tag{6.4}$$

The detailed analysis of satisfying the conditions (6.3) can be found in [8]. It is verified there, that the average energy and related group velocity are defineable when $\Delta\bar{A}$ and Δt can be chosen so that

$$\frac{\partial \left[\int_{\Delta A} \bar{f} \, (E_{0i} \, H_{0j}) \, d\bar{A} \right]}{\partial t} \cong 0 \tag{6.5}$$

and

$$\frac{\partial \bar{f} \, (E_{0i} \, H_{0j})}{\partial t} \cong 0 ,$$

where $\bar{f}$ is a linear function and

$$\text{grad} \left(\int_{\Delta t} \sin^2 \omega_0 \, t \, dt \right) \cong \text{grad} \left(\int_{\Delta t} \sin \omega_0 \, t \cos \omega_0 \, t \, dt \right) \cong 0 .$$

The perturbation can be considered as permissible when (6.5) holds. In that case (6.3) is automatically valid.

Further, conditions for the validity of Eqs (6.3b) are:

— the Δt time interval must be so large, that the

$$\int_{\Delta t} \sin^2 \omega_0 \, t \, dt , \int_{\Delta t} \cos^2 \omega_0 \, t \, dt \quad \text{and} \quad \int_{\Delta t} \sin \omega_0 \, t \cos \omega_0 \, t \, dt$$

integrals may be substituted by a linear function of Δt with the expected accuracy:

— analogous condition issues, for $\Delta\bar{A}$, where the role of $\omega_0 t$ will be taken over by $\varphi(x)$;

— the amplitude defined along $\Delta\bar{A}$ can be taken as constant;

— in the perturbed case the same condition must be fulfilled for the Δt time interval;

— the $\Delta\bar{A}$ surface may be approximated by a plane.

Now an average energy may define the value of which per unit time is — using complex field components —:

$$\bar{W} \cong \frac{1}{2} \Delta\bar{A} \left[(\text{Re} \, \bar{E}_0 \times \text{Re} \, \bar{H}_0) + (\text{Im} \, \bar{E}_0 \times \text{Im} \, \bar{H}_0) \right] . \tag{6.6}$$

The average value of the Poynting vector is:

$$\tilde{\tilde{S}} \cong \frac{1}{2} \operatorname{Re}(\bar{E} \times \bar{H}^*) = \operatorname{Re}(\$) \tag{6.7}$$

where * means the complex conjugate and

$$\widetilde{W} = \operatorname{Re}\left[\int_{\Delta A} \$ \, d\bar{A}\right] \tag{6.8}$$

In comparing these conditions with Chapter 2. and 4.1, it can be seen that they do not include any unusual restrictions. (In our case they hold if u and i may be defined.)

Using these results let now the propagation of energy be analyzed. First the known methods will be shown, than using the most expedient one, the analysis will be carried out for transmission lines.

6.2 *The usual treatment of group velocity*

The energy propagation can be traced by the moving of an actual change in $\bar{S}$ in space-time. In the usual treatment a wave packet will be investigated with a well determined beginning and end [1]. The purely monochromatic signals are essentially excluded.

The usual form of the wave packet in a homogeneous medium is

$$\bar{F}(\bar{r}, t) = e^{j(\omega_0 t - \bar{k}_0 \bar{r})} \int_{\bar{k}_0 - \Delta\bar{k}_0}^{\bar{k}_0 + \Delta\bar{k}_0} F_0(\bar{k}, \bar{k}_0) \, e^{j\left\{(\bar{k} - \bar{k}_0)\left[\left(\frac{\partial\omega}{\partial\bar{k}_0}\right)t - \bar{r}\right]\right\}} d\bar{k} \tag{6.9}$$

where $\bar{k}$ is the wave vector ($\varphi = \bar{k}\bar{r}$),

$$\begin{aligned} \omega(\bar{k}) &= \omega(\bar{k}_0) + (\bar{k} - \bar{k}_0)\left(\frac{\partial\omega}{\partial k}\right)_0 + \dots \\ \omega(\bar{k}_0) &\equiv \omega_0 , \end{aligned} \tag{6.10}$$

and the spectrum of the wave packet is chosen to be so narrow, that the linear approximation (6.10) will suffice. Then the phase of the "group" described by (6.9) will be given by

$$e^{j(\omega_0 t - \bar{k}_0 \bar{r})}$$

which can be brought in front of the integral, and the moving of the amplitude uniquely characterizing the value of $\tilde{\tilde{S}}$ will be given by the equation

$$(\bar{k} - \bar{k}_0)\left[\left(\frac{\partial\omega}{\partial t}\right)_0 t - \bar{r}\right] = \text{constant}. \tag{6.11}$$

Hence the velocity of the propagation of the "constant amplitude surfaces", that is, the group velocity is given by

$$\frac{d\bar{r}}{dt} = \left(\frac{\partial\omega}{\partial\bar{k}}\right)_0 = \bar{v}_g. \tag{6.12}$$

It can be seen, that if $\omega(\bar{k}) \sim \bar{k}$, then — as is well known —

$$\bar{v}_g = \bar{v}_f,$$

where

$$\bar{v}_f = \omega_0 \left(\frac{\partial\varphi}{\partial s}\right)^{-1}$$

is the phase velocity and s is the arc length along the phase path.

This definition in itself is correct. However, it does not give a general answer to the problem of the velocity of energy propagation within the domain of validity of the phenomenological description, where in the majority of cases we are dealing with continuous and frequently strictly monochromatic signals. The definition comes from the assumed form of the solution and not from the basic equations (Maxwell's equations, telegraph equations, etc.).

Therefore, in the following neither this nor the similar methods (stationary phase theorem, etc. [1]) will be used. However, a lot of cases are known where the method treated in (6.9)—(6.12) is expedient. These cases are generally inhomogeneous in that manner, that the question is not the propagation of energy but the dispersion of the energy-packet, and for these investigations the usual methods are useful.

6.3 *The investigation of energy propagation by using perturbation*

This method is based on the method of inhomogeneous basic modes and its applications [8]. As for those cases which were chosen for analysis, the conditions of Chapt. 6.1 hold automatically; it is known, that the energy and the signal amplitude are connected mutually and unambiguously.

The analysis of the group velocity can be executed in such a way [8], that the amplitude of the solution of Maxwell's equations obtained by the method of inhomogeneous basic modes will be perturbed. The perturbed solution must also satisfy Maxwell's equations, and the perturbation must satisfy the conditions shown earlier. Thus,

$$\begin{aligned}
\bar{F}_{\text{pert}} = \bar{F}^{\delta} &= \\
&= \sum_{n=1}^{N} \left(C_n + \delta_n(x,t)\right) e^{-j\varphi_{0n}(x)}\, \bar{F}_{0n}\, e^{j(\omega_0 t - \varphi_n(x))} = \\
&= \sum_{n=1}^{N} C_n(x)\left(1 + f_n(x,t)\right) e^{-j\varphi_{0n}(x)}\, \bar{F}_{0n}\, e^{j(\omega_0 t - \varphi_n(x))} = \\
&= \sum_{n=1}^{N} \left(1 + f_n(x,t)\right) \bar{F}_n(x) = \sum_{n=1}^{N} \bar{F}_n^{\delta}(x)
\end{aligned} \tag{6.13}$$

where

$$|f_n(x,t)| = \frac{|\delta_n(x,t)|}{\sqrt{A_n^2(x) + B_n^2(x)}} \ll 1, \quad C_n(x) = A_n(x) + jB_n(x).$$

If it is considered now, that

$$\bar{F} = \sum_{n=1}^{N} \bar{F}_n(x)$$

is a solution and $F_n(x)$-s are inhomogeneous basic modes, the $f_n(x,t)$ may be determined (or $\bar{\delta}_n = \bar{a}_n + j\bar{b}_n$), and the group velocity:

$$|\bar{v}_{gn}| = \frac{\left|\frac{\partial f_n(x,t)}{\partial t}\right|}{|\operatorname{grad} f_n(x,t)|}. \tag{6.14}$$

The method and its application can be found in [8] in details together with the cases, when non-amplitude perturbations are used, namely, the

$$\varphi_{cn}(x) + \delta_n(x,t)$$

phase-perturbation,

$$\omega_0 + \delta_n(x,t)$$

frequency-perturbation and

$$\bar{F}_n(x) + \bar{\delta}_n(x,t)$$

polarization-perturbation.

It can be simply justified, that

— the $e^{-j(\varphi_{cn}+\delta_n)}$ phase-perturbation is the same as the

$$(1 - j\delta_n)\, e^{-j\varphi_{cn}(x)} \tag{6.15}$$

amplitude perturbation;

— the polarization-perturbation divides the signal into two independent parts, i.e.: we have besides the original mode a new one and the group velocities of the two modes are not necessarily identical;

— the frequency-perturbation can be rewritten into the form

$$e^{j(\omega_0+\delta_n)t} \simeq (1 + j\delta_n t)\, e^{j\omega_0 t} \tag{6.16}$$

and because of t will not be a permissible perturbation, therefore, a fundamentally new investigation will be necessary with non-monochromatic signals.

However, in the case of transmission lines polarization will not be dealt with and because of (6.16) the frequency-perturbation will be omitted.

For the sake of interpretability of the following it seems to be important to analyze (6.13) and (6.15) perturbations in detail. In more general cases this results in a generalized form of (6.14).

6.4 *Investigation of the complex perturbations*

It follows from the detailed analysis of [8], that the (6.14) form of the group velocity is exact whenever δ, δ_n, f and f_n are real and relate to the amplitude. However, it seems to be difficult to interpret (6.14) in complex cases. Let the possibility now be investigated for calculating the group velocity when the perturbation is complex.

Let the original, unperturbed solution or a component of it be

$$F = F_0 \, e^{j(\omega_0 - \varphi)} \, . \tag{6.17}$$

The phase velocity belonging to this can easily be obtained, however, now φ may be complex as well. So in one-dimensional case it follows trivially, that

$$\frac{1}{v_f} = \frac{1}{2\omega_0} \left(\frac{\partial \varphi}{\partial x} + \frac{\partial \varphi^*}{\partial x} \right), \tag{6.18}$$

where * means the complex conjugate, as earlier, or by using the notations of (5.3):

$$\frac{1}{v_f} = -\frac{j}{2\omega_0} (\gamma - \gamma^*) \, . \tag{6.19}$$

In other not one-dimensional cases it elementarily can be seen, that

$$\frac{1}{|v_f|} = \frac{|\overline{K} + \overline{K}^*|}{2\omega_0}$$

where $\overline{K} = \operatorname{grad} \varphi$.

The perturbed solution is:

$$F_\delta = (1 + f) \, F_0 \, e^{j(\omega_0 t - \varphi)}$$

where f, in general, may be complex:

$$f = f_0 \, e^{j\psi_f} \, , \tag{6.20}$$

and

$$\begin{aligned} F_\delta &= (1 + f_0 \, e^{j\psi_f}) \, F_0 \, e^{j(\omega_0 t - \varphi)} = \\ &= (1 + f_{\mathrm{Re}} + j f_{\mathrm{Im}}) \, F_0 \, e^{j(\omega_0 t - \varphi)} \simeq \\ &\simeq (1 + f_{\mathrm{Re}}) \, e^{-j(\varphi + \delta\varphi)} \, F_0 \, e^{j\omega_0 t} \, , \end{aligned} \tag{6.21}$$

and it follows, that

$$\delta_\varphi \cong -\frac{f_{\mathrm{Im}}}{1+f_{\mathrm{Re}}} \cong -f_{\mathrm{Im}}\,. \tag{6.22}$$

It will be noticed here, that from (6.21) it follows, that in general — in contrast to [8] — it is not true that the energy and the changes of it can be directly connected to the amplitude of the signal or the changes of this for determining v_g. When f is known, Eq. (6.21) is not sufficient to determine v_g, though the relation is unique.

Let the result of energy analysis be used — Eqs (6.6), (6.7) and (6.8) — or that form of them which relate to the transmission lines. Now

$$\bar{\mathcal{S}}_1 = \frac{1}{2}\bar{E}\times\bar{H}^*\,,$$

or

$$\mathcal{S} = \frac{1}{2}U\cdot I^*$$

In non-perturbed case

$$\mathcal{S} = \frac{1}{2}U_0(x)I_0(x)\,e^{j(\varphi_i-\varphi_u)} = \mathcal{S}_0\,.$$

In perturbed case:

$$\mathcal{S}_{\mathrm{pert}} = \mathcal{S}_0(1+f_{0u}e^{j\psi_{fu}})(1+f_{0i}\,e^{-j\psi_{fi}}) = \mathcal{S}_0(1+f_s)\,, \tag{6.23}$$

where f_s describes the energy perturbation. From (6.23) the propagation of energy variations (signal) can already be given. Let two cases be underlined:

α) the current- and voltage-perturbations are (or can be taken as) identical:

$$f_{0u}\,e^{j\psi_{fu}} = f_{0i}\,e^{j\psi_{fi}} = f_0\,e^{j\psi_f} \tag{6.24}$$

Now the effectively propagating energy — (6.7) — and the "complex energy" will be perturbed by the same function:

$$f_s = ff^* + (f+f^*) = f_0(f_0 + 2\cos\psi_f)\,, \tag{6.25}$$

that is, if $ff^* \ll f+f^*$:

$$f_s \cong f+f^* = 2f_0\cos\psi_f\,,$$

which is real. The group velocity can be obtained from this

$$v_g = \frac{-\dfrac{\partial f_s}{\partial t}}{\dfrac{\partial f_s}{\partial x}} = -\frac{f_0\dfrac{\partial f_0}{\partial t} + \cos\psi_f\cdot\dfrac{\partial f_0}{\partial t} - f_0\sin\psi_f\cdot\dfrac{\partial\psi_f}{\partial t}}{f_0\dfrac{\partial f_0}{\partial x} + \cos\psi_f\cdot\dfrac{\partial f_0}{\partial x} - f_0\sin\psi_f\cdot\dfrac{\partial\psi_f}{\partial x}}\,, \tag{6.26}$$

that is

$$v_g \cong -\frac{\cos\psi_f \cdot \dfrac{\partial f_0}{\partial t} - f_0 \sin\psi_f \dfrac{\partial \psi_f}{\partial t}}{\cos\psi_f \cdot \dfrac{\partial f_0}{\partial x} - f_0 \sin\psi_f \dfrac{\partial \psi_f}{\partial x}} . \tag{6.27}$$

In any case if $\psi_f = \text{const}$

$$v_g = -\frac{\partial f_0/\partial t}{\partial f_0/\partial x} .$$

is valid.

Analoguously to (6.26) $|v_g|$ can be determined in case of "free-space propagation" as well, which is necessary to interprete the results of calculating f, as seen in [8].

$\beta)$ $f_u \neq f_i$, then

$$(1 + f_s) = [1 + f_{0u} \cdot f_{0i} \cdot e^{j(\psi_{fu} - \psi_{fi})} + \\ + f_{0u}\, e^{j\psi_{fu}} + f_{0i}\, e^{-j\psi_{fi}}] .$$

Now it follows, that

$$\mathcal{S} = (1 + f_{s\mathrm{Re}} + j f_{s\mathrm{Im}})\, \mathcal{S}_0 ,$$

the effective power — within the validity of the real part (see (6.7)) — is:

Hence the effective perturbance to be used when defining and calculating v_g is:

$$f_{\mathrm{seff}} = f_{s\mathrm{Re}} - f_{s\mathrm{Im}} \frac{\mathrm{Im}\ \mathcal{S}_0}{\mathrm{Re}\ \mathcal{S}_0} . \tag{6.28}$$

Thus v_g is the following:

$$v_g = -\frac{\partial f_{\mathrm{seff}}/\partial t}{\partial f_{\mathrm{seff}}/\partial x} . \tag{6.29}$$

NB: Often in practice the interpretation of these relations is difficult. This is caused by the fact that $\bar{\varphi}$ and $\mathcal{S}$ are originated in [8] from fields of the form $\exp[j(\omega_0 t - \varphi)]$. The perturbing function $f(\bar{r}, t)$ or $f(x, t)$, however, have some other, not exactly defined form.

These investigations will be restricted so that an *average* eliminates the appearance of the fluctuation of the energy in $\mathcal{S}$ as a function of space and time in v_g. In more general cases $\mathcal{S}_{\mathrm{pert}}$ and $\mathcal{S}_{\mathrm{pert}}$ must be defined taking into account $f(\bar{r}, t)$ or $f(x, t)$, respectively.

6.5 *The analysis of line parameters by perturbation*

It can be seen from (2.1) and Chapt. 2, that the introduction of the line parameters is always bound to certain conditions. In this paper it was assumed that the parameters in (2.6) and (2.7) are defineable and the basic equations are (2.8) and (2.9). It can be seen that the parameters are independent of the signal, that is of u and i, respectively.

Further it was assumed that these parameters do not depend on the amplitude and phase of the signal, that is, for monochromatic solutions the basic equations are linear. However, the parameters even in this case depend on the geometry of the line and the characteristics of the medium [2]. So the dispersion of the medium even in linear cases can be taken into account if the propagation of the perturbed signal will be investigated. Several cases of medium analysis — given in [8] — will be used here. So two groups of linear lines will be obtained, namely:

— strictly linear lines
— linear, dispersive lines.

The first group consists of lines the parameters of which either do not depend on the signal parameters in the interval of the investigation (i.e. the investigated frequency band) or the dependence is completely negligeable within the accuracy we want.

The second group consists of lines, the parameters of which (Z_T, etc.) depend only on the frequency of the signal (within the given accuracy) in that way, that Z_T for the non-perturbed signal and $Z_{T\,\mathrm{pert}}$ for the perturbed signal suffice the

$$\begin{aligned} \mathbf{Z}_T &= \mathbf{Z}_T(\omega_0), \\ \mathbf{Z}_{T\mathrm{pert}} &= \mathbf{Z}_T(\omega_{0\mathrm{pert}}) \end{aligned} \tag{6.30}$$

and

$$\omega_{0\mathrm{pert}} = \omega_0 \left(1 + \frac{\partial f / \partial t}{j\omega_0}\right)$$

conditions. Relation (6.30) does not mean anything else than to use the conditions valid for ϵ, μ, $\varkappa$ and ν medium parameters in [8].

If geometric changes occur as a result of the signal, a further investigation is necessary. This will not be dealt with here.

It is known [2, etc.], that in a lossless transmission line filled with vacuum or linear medium:

$$L = L_0 \quad \text{or} \quad L = \mu_r L_0 ,$$

and

$$C = C_0 \quad \text{or} \quad C = \varepsilon_r C_0 ,$$

where L_0 and C_0 depend only on the geometry of line.

In the case of strictly linear lines the medium is also strictly linear, that is:

$$\frac{\partial \varepsilon}{\partial \omega_0} \equiv 0 \, .$$

So

$$\varepsilon_{\text{pert}} = \varepsilon = n^2 \tag{6.31}$$

where n is the refraction coefficient and therefore

$$\frac{\partial \mathbf{Z}_T}{\partial \omega_0} \equiv 0 \, , \quad \text{i.e. } \mathbf{Z}_T = \mathbf{Z}_{T\text{pert}} \, .$$

If the medium is dispersive in the line, the results of [8] may be applied. In isotropic ionized gases (isotropic plasmas):

$$\varepsilon \cong 1 - \frac{\omega_p^2}{\omega_0^2} = 1 - e$$

and

$$\varepsilon_{\text{pert}} = 1 - \frac{\omega_p^2}{\left(\omega_0 - j \dfrac{\partial f}{\partial t}\right)^2} \cong \varepsilon - j2 \, e \, \frac{\partial f / \partial t}{\omega_0} \, . \tag{6.32}$$

NB: In the cases when the medium in the line is anisotropic, the definition and interpretation of the line parameters is a further problem to be investigated. This exceeds the volume of this paper.

Further, a number of cases can be analysed with the detailed deduction of L, C, R and G for F and F_δ. So in the following the notations: $\mathbf{Z}_{T\,\text{pert}}$, $L_{T\,\text{pert}}$, L_{pert}, R_{pert}, etc can be used in the adequate sense.

Example: A line filled with ionized, isotropic gas (i.e. tempered isotropic plasma):

$$\mu \cong 1 \, ,$$
$$\varkappa = \nu = 0 \, ,$$

ϵ can be obtained according to (6.32), so:

$$L_{\text{pert}} \equiv L = L_0 \, ,$$
$$C = \varepsilon \, C_0 \, , \tag{6.30a}$$
$$C_{\text{pert}} = C - j2e \, \frac{\partial f / \partial t}{\omega_0} \, C_0 \, ,$$

and if the line is ideal:

$$R = G = 0.$$

There is a further "dispersive" case worth mentioning. The component materials of the line are not dispersive. The geometry of the line (L_0, C_0, etc.) are independent of signal within the interval of analysis. This is the most frequent case in the practice i.e. the case of the *simple transmission lines with attenuation*.

It is known that in attenuated lines L_T and C_T can be given by (2.6) and (2.7), respectively. In monochromatic cases — $\exp(j\omega_0 t)$ —:

$$L_T = L_0 + R \frac{1}{i(x)\ e^{j\omega_0 t}} \int i(x)\, e^{j\omega_0 t}\, dt = L - j\frac{R}{\omega_0} \tag{6.33}$$

and similarly:

$$C_T = C - j\frac{G}{\omega_0}\,. \tag{6.34}$$

These show, that L_T and C_T are dispersive. Let the value of them be determined in the most simple perturbed case, when the current is

$$i_{\text{pert}} = (1 + f) \cdot i(x) \cdot e^{j\omega_0 t}.$$

Then

$$\frac{1}{i_{\text{pert}}} \int i_{\text{pert}}\, dt \cong \frac{1}{j\omega_0}\left[1 - \frac{\partial f/\partial t}{j\omega_0}\right] = \frac{1}{j\omega_{\text{pert}}} \tag{6.35}$$

assuming that $\dfrac{\partial^2 f}{\partial t^2} \ll \dfrac{\partial f}{\partial t}$, which can be accepted as a consequence of the restriction in Chapter (6.1). So

$$\begin{aligned} L_{T\text{pert}} &= L_T + \frac{R}{\omega_0^2}\frac{\partial f}{\partial t} \\ \text{and} \qquad & \qquad (6.36) \\ C_{T\text{pert}} &= C_T + \frac{G}{\omega_0^2}\frac{\partial f}{\partial t}\,. \end{aligned}$$

It should be noted that this dispersion is also linear because (6.36) corresponds to (6.30). However, the character of (6.36) differs from the dispersive cases caused by the medium which can be seen in comparing (6.30a) and (6.36).

Because of (6.36) and its applicational results (shown later) — which are generally considered as theoretical examples — it is necesarry to analyse R and C_T in concrete tasks. More exactly, because of the physical processes giving R and C_T [17] it must be investigated, whether or not

$$R(t) \quad \text{and} \quad G(t)$$

may be defined for a given

$$\bar{E}(t) \quad \text{and} \quad \bar{H}(t).$$

If the answer is yes, then using the so defined $i(t)$ and $u(t)$ — in this case it can be generally assumed, that they exist and have the form of $\exp(j\omega_0 t)$ — it seems to be expedient to investigate in fact

$$L_T = L + \frac{1}{i(t)} \int R(t)\, i(t)\, dt$$

and (6.37)

$$C_T = C + \frac{1}{u(t)} \int G(t)\, u(t)\, dt$$

and trivially to define $R(\omega_0)$ and $G(\omega_0)$, respectively. It must be controlled, by what conditions it holds that

$$\frac{\partial R}{\partial \omega_0} \equiv \frac{\partial G}{\partial \omega_0} \equiv 0,$$

as (6.36) may be used in this case. If this last restriction will not hold, the relation between $R(\omega_0)$ and $R_{\text{pert}}(\omega_{0\text{pert}})$, etc. should be analysed, and (6.35) may be used independently of this. In this remaining cases, one may start from (6.37).

In case f and $\partial f/\partial t$ being complex, the relation between $\mathbf{Z}_{T\text{pert}}$ and $\mathbf{Z}_T$ will be a different problem. This was not excluded in the examinations executed, but going deeper into the problem may result in interesting details, e.g. the different properties of AM and PM propagation, etc.

Now we have all the means together to execute the investigation of group velocity using the chosen method.

6.6 *The analysis of group velocity using the method of inhomogeneous basic modes*

Let us start on the basis of Chapter 4.2 and at the beginning for the sake of generality let it be assumed, that

$$\mathbf{Z}_{T\text{pert}} \neq \mathbf{Z}_T .$$

Let the perturbation *be*

$$a_n + \delta_n = a_n(1 + f_n),$$

which will do for investigating v_g.

These mean as a consequence of Eq. (4.8) that by using the $\bar{F}_n \begin{pmatrix} u_n \\ i_n \end{pmatrix}$ inhomogeneous basic modes the perturbations of the current and voltage are identical. This, however, — see Chapter 4.5, etc. — does not generally involve the identity of perturbing functions of the resultant current or voltage or of the resultant propagating modes.

Let us start from Eq. (2.9) and Eq. (6.13), the second giving the perturbed form of the signal. It is known, that the unperturbed solution suffices Eqs (4.11) and (4.12). Evolving now Eq. (2.9) in the manner shown in Chapter 4.2, and taking into consideration that

$$\bar{F}_\delta = \sum_n (1 + f_n)\, \bar{F}_{0n}\, e^{j(\omega_0 t - \varphi_n)} = \sum_n \bar{F}_{\delta 0n}\, e^{j(\omega_0 t - \varphi_n)} = \\ = \sum_n \bar{F}_{\delta n} = \sum_n (1 + f_n)\, \bar{F}_n \tag{6.38}$$

one has

$$\sum_n \left[\nabla_{F\delta 0n} - j \frac{\partial \varphi_n}{\partial x} \cdot \mathbf{1} \right] \cdot \bar{F}_{\delta n} = \\ = - \sum_n \left[\frac{\partial \mathbf{Z}_{T\text{pert}}}{\partial t} + \mathbf{Z}_{T\text{pert}} \nabla^t_{F\delta 0n} + j\omega_0 \mathbf{Z}_{T\text{pert}} \right] \bar{F}_{\delta n} \tag{6.39}$$

where

$$\nabla_{F\delta 0n} = \begin{bmatrix} \dfrac{\partial \ln u_{\delta 0n}}{\partial x} & 0 \\ 0 & \dfrac{\partial \ln i_{\delta 0n}}{\partial x} \end{bmatrix} \cong \nabla_{F0n} + \frac{\partial f_n}{\partial x} \cdot \mathbf{1}$$

and (6.40)

$$\nabla_{F\delta 0n} = \begin{bmatrix} \dfrac{\partial \ln u_{\delta 0n}}{\partial t} & 0 \\ 0 & \dfrac{\partial \ln i_{\delta 0n}}{\partial t} \end{bmatrix} = \frac{\partial f_n}{\partial t} \cdot \mathbf{1} .$$

It was kept in mind that the solution is originally stationary and strictly monochromatic, and the transmission line does not vary in time, that is

$$\frac{\partial \mathbf{Z}_T}{\partial t} = 0$$

or, even in a dispersive case is stationary, that is

$$\frac{\partial \mathbf{Z}_{T\text{pert}}}{\partial t} = 0 . \tag{6.41}$$

The last restriction is non-trivial, but it will be required now, for only this type of line can be considered as stationary, linear, that is, which will not vary in time.

If follows from Eq. (6.39), that

$$\sum_n \left[\nabla_{F0n} - j \frac{\partial \varphi_n}{\partial x} \cdot \mathbf{1} + \frac{\partial f_n}{\partial x} \cdot \mathbf{1} \right] (1 + f_n) \bar{F} = \\ = - \sum_n \left(\frac{\partial f_n}{\partial t} + j\omega_0 \right) \mathbf{Z}_{T\text{pert}} (1 + f_n) \bar{F} . \tag{6.42}$$

Hence, f_n can be determined for the different tasks. During the process Eqs (4.10) and (4.12) may be taken into consideration when necessary, and it can be also used when f_n and its derivatives are small, consequently their product results in second order small terms. In the knowledge of f_n we can calculate v_{gn} and the resulting f and v_g.

It should be noted here, that elementary perturbations are investigated, that is

$$\frac{\partial f_n}{\partial x} \to 0 \quad \text{and} \quad \frac{\partial f_n}{\partial t} \to 0 ,$$

the quotient of them, however, — which arises when describing the *propagation* of perturbation — will be finite.

In this way — as it is suggested in [8] — non-elementary perturbations can also be investigated. Then the previous restrictions naturally will not hold. Therefore, in case of non-elementary perturbations, especially when we have a dispersive case, one has to start from Eq. (2.1), and it is not certain that equivalent f perturbing functions can be found for u and i. (This is generally valid for [8] too, though this problem is not emphasized there.

Therefore, in case of non-elementary perturbations pure amplitude, phase, polarization, and frequency perturbations in field-calculations cannot be spoken about. The detailed analysis of this question does not belong to the aim of this paper.)

An important special case is the following:

In many cases it may be correct to assume that all modes of the elementary perturbation are identical, that is

$$f_n \equiv f.$$

This may be the consequence of the structure of the solution and the boundary conditions.

In other cases the inhomogeneous basic modes are independent — orthogonal, uncoupled — ones. This will be obtained e.g. for homogeneous lines.

So Eq. (6.42) can be further evolved:

$$\left(\frac{\partial f}{\partial x}\cdot\mathbf{1}-j\omega_0\,\mathbf{Z}_T\right)\sum_n \bar{F}_n=-\left(\frac{\partial f}{\partial t}+j\,\omega_0\right)\mathbf{Z}_{T\,\mathrm{pert}}\sum_n \bar{F}_n\,, \tag{6.43}$$

or in case of a simple mode:

$$\left[\left(\frac{\partial f}{\partial x}-j\,\frac{\partial \varphi}{\partial x}\right)\mathbf{1}+\left(\frac{\partial f}{\partial t}+j\omega_0\right)\mathbf{Z}_{T\,\mathrm{pert}}\right]\bar{F}=0\,. \tag{6.44}$$

If f is real, now v_g can immediately be given or the differential equation for $\partial f/\partial x$ and $\partial f/\partial t$ can be solved. Writing Eq. (6.44) for the different components, evolving them u or i in knowing that $u\neq 0$ or $i\neq 0$, respectively, omitting the terms being small in second order and taking into consideration Eq. (4.11) it can be obtained, that

$$\frac{\partial f}{\partial x}+\frac{\omega_0}{\partial\varphi/\partial x}L_{T\,\mathrm{pert}}C_{T\,\mathrm{pert}}\frac{\partial f}{\partial t}+\frac{j}{2}\,\frac{\omega_0^2}{\partial\varphi/\partial x}(L_{T\,\mathrm{pert}}C_{T\mathrm{pert}}-L_T\,C_T)=0. \tag{6.45}$$

It can be seen, that this will be correct, if

$$C_T\,(L_{T\mathrm{pert}}-L_T)-L_T\,(C_{T\mathrm{pert}}-C_T)\to 0\,,$$

that is, it correlates with the earlier restrictions for the perturbation.

Thus, for concrete problems the energy propagation can be investigated by solving Eqs (6.42), (6.43) or (6.45).

6.7 *The analysis of the strictly linear case*

It was mentioned above that in a nondispersive case $\mathbf{Z}_{T\,\mathrm{pert}}=\mathbf{Z}_T$. Then it follows from Eq. (6.42), that

$$\sum_n\left[\left(\nabla_{F0n}-j\,\frac{\partial\varphi_n}{\partial x}\mathbf{1}+j\omega_0\,\mathbf{Z}_T\right)+\left(\frac{\partial f_n}{\partial x}\mathbf{1}+\frac{\partial f_n}{\partial t}\mathbf{Z}_T\right)\right](1+f_n)\bar{F}_n=0.$$

However, as Eq. (4.10) must be fulfilled

$$\sum_n\left[\nabla_{F0n}+\left(\frac{\partial f_n}{\partial x}\mathbf{1}+\frac{\partial f_n}{\partial t}\mathbf{Z}_T\right)\right](1+f_n)\,\bar{F}_n=0\,.$$

If furthermore it may be assumed that $f_n\equiv f$, it follows from Eq. (6.43), that

$$\left(\frac{\partial f}{\partial x}\mathbf{1}+\frac{\partial f}{\partial t}\cdot\mathbf{Z}_T\right)\sum_n\bar{F}_n=0\,. \tag{6.46}$$

Again considering Eq. (4.10) it can be seen, that if

$$v_g = -\frac{\partial f/\partial t}{\partial f/\partial x},$$

then

$$v_g \equiv v_f .$$

In other cases, when f is complex, the process must be executed according to Chapter 6.4.

In a number of practical cases, till ψ_f does not depend on x — Eq. (6.20) — the result is $v_g \equiv v_f$.

If there is only one mode, it follows from Eq. (6.45), that

$$\frac{\partial f}{\partial x} + \frac{\omega_0 L_T C_T}{\partial \varphi/\partial x} \frac{\partial f}{\partial t} = 0 .$$

This means because of Eq. (4.18), that

$$\frac{\partial f}{\partial x} = -\sqrt{L_T C_T}\, \frac{\partial f}{\partial t}$$

so

$$-\frac{\partial f/\partial t}{\partial f/\partial x} = \frac{1}{\sqrt{L_T C_T}},$$

from which it follows in most of cases and in every homogeneous cases, that

$$v_g \equiv v_f .$$

6.8 *The analysis of the dispersive lines*

Generally in dispersive cases Eqs (6.42), (6.43) and (6.45) hold. Let us now look at several applications in a homogeneous case, when there is only one mode for illustrating the earlier ones.

Let us start from

$$\frac{\partial f}{\partial x} = -\frac{\omega_0}{\partial \varphi/\partial x}\left[L_{T\,\mathrm{pert}}\, C_{T\,\mathrm{pert}} \frac{\partial f}{\partial t} + j\frac{\omega_0}{2}\left(L_{T\mathrm{pert}}\, C_{T\mathrm{pert}} - L_T C_T\right)\right]. \quad (6.45\mathrm{a})$$

a) Non-attenuated line filled with non-magnetized, tempered, isotropic plasma:

The line parameters will be given by Eqs (6.37). Using them it follows, that

$$L_T = L_0,\ C_T = C,\ L_{T\mathrm{pert}} = L_0,\ C_{T\mathrm{pert}} = C_{\mathrm{pert}} .$$

Further

$$L_{T\,\mathrm{pert}}\,C_{T\,\mathrm{pert}} = L_0\,C_{\mathrm{pert}} = L_0\,C_0\left[\varepsilon - j\,2\,e\,\frac{\partial f/\partial t}{\omega_0}\right]$$

where ε and e are given in Eq. (6.32). So

$$L_{T\,\mathrm{pert}}\,C_{T\,\mathrm{pert}} - L_T\,C_T = -j2e\,\frac{\partial f/\partial t}{\omega_0}\,L_0\,C_0$$

and from here

$$\frac{\partial f}{\partial x} = -\frac{\omega_0}{\partial\varphi/\partial x}\left[L_0\,C_0\,\varepsilon\,\frac{\partial f}{\partial t} + e\,L_0\,C_0\,\frac{\partial f}{\partial t} - j2e\,L_0\,C_0\,\frac{\left(\frac{\partial f}{\partial t}\right)^2}{\omega_0}\right].$$

Neglecting the terms small in second order and taking $\partial\varphi/\partial x$ from Eq. (4.18), one gets

$$\frac{\partial f}{\partial x} = -\frac{\partial f}{\partial t}\,\frac{1}{\sqrt{\varepsilon\,L_0\,C_0}}\,L_0\,C_0 = -\frac{\partial f}{\partial t}\,\frac{\sqrt{L_0\,C_0}}{\sqrt{\varepsilon}},$$

and so it follows that

$$v_g = -\frac{\partial f/\partial t}{\partial f/\partial x} = \frac{\sqrt{\varepsilon}}{\sqrt{L_0\,C_0}}\,. \tag{6.47}$$

It is known, that

$$v_f = \frac{1}{\sqrt{\varepsilon}}\,\frac{1}{\sqrt{L_0\,C_0}}\,,$$

therefore,

$$v_g\,v_f = \frac{1}{L_0\,C_0} = v_0^2\,, \tag{6.48}$$

where v_0 is the signal velocity on the lines "filled" with vacuum.

The effect of the dispersive medium parameters ($\epsilon, \varkappa, \mu$, and ν) can also be similarly investigated in other cases.

It is important to mention that our results are valid for strictly monochromatic, continuous signals.

b) Simple attenuated lines:

In this case the line parameters can be determined from Eqs (2.6), (2.7) and (6.36). So

$$L_T = L - j\,\frac{R}{\omega_0}\,,\quad L_{T\mathrm{pert}} = L_T + \frac{R}{\omega_0^2}\,\frac{\partial f}{\partial t}\,,$$

$$C_T = C - j\,\frac{G}{\omega_0}\,,\quad C_{T\mathrm{pert}} = C_T + \frac{G}{\omega_0^2}\,\frac{\partial f}{\partial t}\,.$$

Again neglecting the terms being small in second order and developing Eq. (6.45), one gets

$$-\frac{\partial f}{\partial x} = \frac{\partial f}{\partial t} \frac{LC - j\dfrac{LC + RC}{2\omega_0}}{\sqrt{\left(LC - \dfrac{RG}{\omega_0^2}\right) - j\dfrac{LG + RC}{\omega_0}}}. \tag{6.49}$$

It is known, that the phase-velocity [2] is in this case

$$\frac{1}{v_f^2} = \left(\frac{\mathrm{Re}\,\partial\varphi/\partial x}{\omega_0}\right)^2 = \frac{1}{2}\left(LC - \frac{RG}{\omega_0^2}\right) + \frac{1}{2}\sqrt{\left(LC + \frac{RG}{\omega_0^2}\right)^2 + \left(\frac{RC - LG}{\omega_0}\right)^2}.$$

However, the group velocity may be determined according to Eq. (6.26) as f is complex. In homogeneous case Eq. (6.49) may be simply solved as for the solution of the equation

$$-\frac{\partial f}{\partial x} = \bar{\mathfrak{K}} \cdot \frac{\partial f}{\partial t} = (\mathrm{Re}\ \bar{\mathfrak{K}} + j\ \mathrm{Im}\ \bar{\mathfrak{K}})\ \frac{\partial f}{\partial t} \tag{6.50}$$

can be sought for in the form of $f = F(x)\,G(t)$, in the known manner. The f obtained — for $\bar{\mathfrak{K}} = \mathrm{const.}$ — gives v_g according to Eq. (6.26). It is known [11], that e.g. f can take the following form

$$f = A_0\, e^{\alpha(t - \mathrm{Re}\,\bar{\mathfrak{K}}\cdot x) + \beta\,\mathrm{Im}\,\bar{\mathfrak{K}}\cdot x} \cdot e^{j\beta(t - \mathrm{Re}\,\bar{\mathfrak{K}}\cdot x) - j\alpha\,\mathrm{Im}\,\bar{\mathfrak{K}}\cdot x} \tag{6.51}$$

where α and β are constant. Now f_0 and ψ_f can easily be obtained

$$-\frac{\partial fs/\partial x}{\partial fs/\partial t} = \mathrm{Re}\,\bar{\mathfrak{K}} + \mathrm{Im}\,\bar{\mathfrak{K}}\,\frac{\beta\cos\psi_f + \alpha\sin\psi_f}{\beta\sin\psi_f - \alpha\cos\psi_f} = \mathrm{Re}\,\bar{\mathfrak{K}} - \mathrm{Im}\,\bar{\mathfrak{K}}\cdot\mathrm{tg}\,\psi_{\mathrm{ekv}}, \tag{6.52}$$

where ψ_{ekv} means the following:

$$\psi_{\mathrm{ekv}} = \mathrm{arctg}\,\frac{\beta\cos\psi_f + \alpha\sin\psi_f}{\alpha\cos\psi_f - \beta\sin\psi_f}.$$

Here arises the problem for the first time which was mentioned in Chapter 6.4, that is, the definition of $\bar{\varphi}$ when the character of $[1 + f(x, t)]\cdot\bar{F}$ is non-sinusoidal. However, as Eq. (6.52) is simple, v_g can be obtained by averaging. (In more complicated cases the definition of $\bar{\varphi}_{\mathrm{pert}}$ and f_s must be revised.)

So we have

$$v_g = \frac{1}{\mathrm{Re}\,\bar{\mathfrak{K}}} \neq v_f$$

where

$$\operatorname{Re}\overline{\mathfrak{K}} = \operatorname{Re}\frac{LC - j\dfrac{LG + RC}{2\omega_0}}{\sqrt{\left(LC - \dfrac{RG}{\omega_0^2}\right) - j\dfrac{LG + RC}{\omega_0^2}}}\,. \tag{6.53}$$

This result is important, because this case is generally not considered as dispersive.

6.9 *Perturbation of the resultant wave pattern*

It is self-evident that not only the solution obtained by using the method of inhomogeneous basic modes can be investigated in this way — though it proved to be very useful — but e.g. the wave pattern obtainable in the manner shown in Chapter 5. The perturbation of this resultant wave pattern does not necessarily result in "propagation". Depending on the boundary conditions and the line parameters it may occur that the perturbation simply "dissolves" in the pictures. For the sake of generality let this case be shortly surveyed too.

It is assumed that the solution is known in the form of Eq. (5.3). Perturbed this one has

$$\begin{aligned} u_f &= U_0(1 + f_u)\, e^{j\omega_0 t - \int_0^x \gamma_u(x)\,dx}\,, \\ i_f &= I_0(1 + f_i)\, e^{j\omega_0 t - \int_0^x \gamma_i(x)\,dx}\,. \end{aligned} \tag{6.54}$$

The fundamental solutions Eqs (5.3) satisfy Eqs (5.5) the perturbed solutions — Eqs (6.54) — satisfy the following equations:

$$\begin{aligned} \frac{\partial u_f}{\partial x} &= -L_{T\,\mathrm{pert}}\frac{\partial i_f}{\partial t}\,, \\ \frac{\partial i_f}{\partial x} &= -C_{T\,\mathrm{pert}}\frac{\partial u_f}{\partial t}\,. \end{aligned} \tag{6.55}$$

Developing this in the way shown in Chapter 5, taking Eqs (6.54) into consideration and neglecting the terms being small in second order equations will be obtained for f_u and f_i, meanwhile we shall use Eqs (5.6) and that

$$\frac{1}{1+f}\,\frac{\partial f}{\partial q} = \frac{\partial\, ln\,(1+f)}{\partial q} \cong \frac{\partial f}{\partial q}\,,$$

where q may play the role of x or t, respectively. The following equations (by using the notations $\partial C/\partial x = C'$ and $\partial L/\partial x = L'$) will be obtained:

$$\left(2\gamma_i + \frac{C'_{T\,\mathrm{pert}}}{C_{T\,\mathrm{pert}}}\right)\frac{\partial f_i}{\partial x} + 2j\omega_0 L_{T\,\mathrm{pert}} C_{T\,\mathrm{pert}} \frac{\partial f_i}{\partial t} =$$

$$= \gamma_i \left(\frac{C'_{T\,\mathrm{pert}}}{C_{T\,\mathrm{pert}}} - \frac{C'_T}{C'_T}\right) + \omega_0^2 (L_{T\,\mathrm{pert}} C_{T\,\mathrm{pert}} - L_T C_T)$$

and (6.56)

$$\left(2\gamma_u + \frac{L'_{T\,\mathrm{pert}}}{L_{T\,\mathrm{pert}}}\right)\frac{\partial f_u}{\partial x} + 2j\omega_0 L_{T\,\mathrm{pert}} C_{T\,\mathrm{pert}} \frac{\partial f_u}{\partial x} =$$

$$= \gamma_u \left(\frac{L'_{T\mathrm{pert}}}{L_{T\,\mathrm{pert}}} - \frac{L'_T}{L_T}\right) + \omega_0^2 (L_{T\mathrm{pert}} C_{T\,\mathrm{pert}} - L_T C_T),$$

respectively.

Now in knowing f_i and f_u, $f_{s\,\mathrm{eff}}$ may be determined according to Chapter 6.4, and if energy flow may be spoken about in the resulting wave pattern, v_g may be given.

It is worth mentioning, that in a non-dispersive case the equations are:

$$\left(2\gamma_i + \frac{C'_T}{C_T}\right)\frac{\partial f_i}{\partial x} + 2\, j\omega_0 L_T C_T \frac{\partial f_i}{\partial t} = 0$$

and (6.57)

$$\left(2\gamma_u + \frac{L'_T}{L_T}\right)\frac{\partial f_u}{\partial x} + 2\, j\omega_0 L_T C_T \frac{\partial f_u}{\partial t} = 0,$$

respectively.

In transmission lines of *homogeneous class* as in the dispersive, as in a strictly linear case

$$\gamma_u = \gamma_i = \gamma$$

and

$$f_u = f_i = f.$$

In a homogeneous case even more can be told. In homogeneous dispersive lines we have

$$2\gamma \frac{\partial f}{\partial x} + 2\, j\omega_0 L_{T\,\mathrm{pert}} C_{T\,\mathrm{pert}} \frac{\partial f}{\partial t} = \omega_0^2 [L_{T\mathrm{pert}} C_{T\,\mathrm{pert}} - L_T C_T] \qquad (6.58)$$

which is identical with (6.45)!

In homogeneous, non-dispersive case it follows, that

$$\frac{\partial f}{\partial x} + j\frac{\omega_0}{\gamma} L_T C_T \frac{\partial f}{\partial t} = 0. \qquad (6.59)$$

However, as for $\gamma = j\omega_0 \sqrt{L_T C_T}$, it follows, that

$$\frac{\partial f}{\partial x} = -\frac{\partial f}{\partial t}\sqrt{L_T C_T}\,.$$

This means, that $v_g = v_f$, which was foreseen.

That case of *ideal, homogeneous, dispersive lines* is also worth mentioning, when — in many cases but not always — the following approximation is acceptable (see chapter 6.5, and [8]):

$$L_{T\,\mathrm{pert}}\, C_{T\,\mathrm{pert}} \frac{\partial f}{\partial t} \cong LC \frac{\partial f}{\partial t}$$

and (6.60)

$$L_{T\,\mathrm{pert}}\, C_{T\,\mathrm{pert}} - L_T C_T \cong j \frac{L\Delta_C - C\Delta_L}{\partial t} \frac{\partial f}{\partial t}\,,$$

where Δ_L and Δ_C can be calculated by using (2.6) and (2.7); both of them may have negative values, too.

However, now an interesting but not general consequence follows from (6.58):

$$v_g = -\frac{\partial f/\partial t}{\partial f/\partial x} = \frac{\sqrt{LC}}{LC - \frac{1}{2}(L\Delta_C + C\Delta_L)}\,,$$

that is: (6.61)

$$v_g \cdot v_f = \frac{1}{LC - \frac{1}{2}(L\Delta_C - C\Delta_L)}\,.$$

Applying this to our example in Chapt. 6.8a, we get Eq. (6.46).

As a consequence, it is worth executing the analysis of the group velocity in both ways.

7. An example: Analysis of the exponential transmission lines using the method of inhomogeneous basic modes

For the sake of shortness and of being easily controllable, let the exponential transmission line be analysed, which is (partly) known.

The parameters of the line are given by Eqs (3.5). Using the notation of (3.8) and (4.32) and applying the method given in Chapter 4 for determining the *wave pattern* if follows that

$$f = \frac{d \ln z_0(x)}{dx} = \frac{d}{dx} \ln\, [z_{00}\, e^{-kx}] = -k,$$

$$f' = f'' = 0\,,$$

$$g = (-1)^{n-1} \Gamma(x) = \pm\gamma, \qquad n = 1, 2$$

$$\gamma = j\omega_0 \sqrt{L_0 C_0} = j\beta_0 \,.$$

$$z_{00} = \sqrt{\frac{L_0}{C_0}} \,.$$

From that it follows that the condition in (4.50) holds and Δ is constant.

Further

$$\mathcal{G}(x) = \frac{f'}{f} - 2g = \mp 2\gamma = G_{on} \,,$$

$$\mathcal{H}(x) = -\frac{f^2}{4} = -\frac{k^2}{4} = H_0 \,,$$

and from these

$$\Delta = \mathcal{G}_{0n}^2 - 4\mathcal{H}_0 = 4\gamma^2 + k^2 = \text{constant} \,,$$

that is

$$\Delta_{01} \equiv \Delta_{02} \equiv \Delta_0 .$$

It follows then — by using Eq. (4.52) —

$$v_n(x) = -\frac{1}{2}\left[\mathcal{G}_{0n} \pm 2\gamma \sqrt{1 - \frac{k^2}{4\gamma^2}}\right] \quad n = 1, 2 \,, \tag{7.1}$$

that is

$$v_1(x) = -j\beta_0 \left(-1 \pm \sqrt{1 - \frac{k^2}{4\beta_0^2}}\right) = -j(-\beta_0 \pm \beta) \,,$$

$$v_2(x) = -j\beta_0 \left(-1 \pm \sqrt{1 - \frac{k^2}{4\beta_0^2}}\right) = -j(\beta_0 \pm \beta) .$$

It can be seen that in this case the wave pattern is composed of four inhomogeneous basic modes which will be noted by the indices "$1\pm$" and "$2\pm$", respectively.

The Eq. (4.37) provides the solution:

$$\eta_{n\pm}(x) = C_{n\pm} e^{\frac{k}{2}x} \exp j[(-1)^n \beta_0 \pm \beta]\, x \cdot \exp. j\,[\omega_0 t - (-1)^n \beta_0 x] \,. \tag{7.2}$$

By using Eq. (4.23) we arrive at the resultant wave pattern

$$\begin{aligned} i(x,t) &= e^{\frac{k}{2}x + j\omega_0 t}\,[(C_{1-} + C_{2-})\, e^{-j\beta x} + (C_{1+} + C_{2+})\, e^{j\beta x}] = \\ &= e^{\frac{k}{2}x + j\omega_0 t}\,[I_0^+ e^{-j\beta x} + I_0^- e^{j\beta x}] \end{aligned} \tag{7.3}$$

and

$$\begin{aligned} u(x,t) &= e^{-\frac{k}{2}x + j\omega_0 t}\, z_{00}\,[(C_{2-} - C_{1-})\, e^{-j\beta x} + (C_{2+} - C_{1+})\, e^{j\beta x}] = \\ &= e^{-\frac{k}{2}x + j\omega_0 t}\,[u_0^+ e^{-j\beta x} + u_0^- e^{j\beta x}] \,. \end{aligned} \tag{7.4}$$

Further the relationship among the coefficients can be determined in the known way:

$$I_0^{\pm} = \frac{\frac{k}{2} \pm j\beta}{j\omega_0 L_0} U_0^{\pm} .$$

Let now the *group velocity* be examined. Taking the relations of (3.5) into consideration, now $Z_T = Z_{T\text{pert}}$, and the point to start from is Eq. (6.47). As the wave pattern in this case is composed of four inhomogeneous basic modes, the validity of Eqs (6.48) will not hold as the equivalence of the perturbant functions is not necessarily valid.

The Eq. (6.47) can be rewritten by taking into consideration that the perturbation is infinitesimally small:

$$\sum_{n\pm} \left[\nabla_{F0n\pm} f_{n\pm} + \frac{\partial f_{n\pm}}{\partial x} \cdot \mathbf{1} + \frac{\partial f_{n\pm}}{\partial t} \mathbf{Z}_T \right] \bar{F}_{n\pm} = 0 \tag{7.5}$$

where $\bar{F}_{n\pm}$ means the known inhomogeneous basic modes according to Eq. (4.23).

By using Eqs (7.2), (4.13) and (4.23), the Eq. (7.5) can be developed. It must be noted, that because of the manner of decomposing the signal when the method was introduced it is not permitted to carry out any reductions in Eq. (7.2), for the sake of $F_{0n\pm}$ and $\varphi_{n\pm}(x)$ to keep their form according to the definition. From the solution of Eq. (7.2) it follows that

$$\nabla_{F0n\pm} = \begin{bmatrix} -\frac{k}{2} + j(-1)^n \beta_0 \pm j\beta & 0 \\ 0 & \frac{k}{2} + j(-1)^n \beta_0 \pm j\beta \end{bmatrix}.$$

Developing Eq. (7.5) and after elementary algebraic transformations one has:

$$\begin{aligned}
&\frac{C_{1-}}{C_{2-}} \cdot \frac{k}{2} \cdot f_{1-} + j(\beta_0 - \beta) f_{2-} + \frac{\partial f_{2-}}{\partial x} + \sqrt{L_0 C_0} \frac{\partial f_{2-}}{\partial t} = 0 , \\
&\frac{C_{2-}}{C_{1-}} \cdot \frac{k}{2} \cdot f_{2-} - j(\beta_0 - \beta) f_{1-} + \frac{\partial f_{1-}}{\partial x} - \sqrt{L_0 C_0} \frac{\partial f_{1-}}{\partial t} = 0 . \\
&\frac{C_{1+}}{C_{2+}} \cdot \frac{k}{2} \cdot f_{1+} + j(\beta_0 - \beta) f_{2+} + \frac{\partial f_{2+}}{\partial x} + \sqrt{L_0 C_0} \frac{\partial f_{2+}}{\partial t} = 0 , \\
&\frac{C_{2+}}{C_{1+}} \cdot \frac{k}{2} \cdot f_{2+} - j(\beta_0 - \beta) f_{1+} + \frac{\partial f_{1+}}{\partial x} - \sqrt{L_0 C_0} \frac{\partial f_{1+}}{\partial t} = 0 .
\end{aligned} \tag{7.6}$$

After thoroughly analysing these partial differential equations it can be seen that a possible solution is

$$f_{n\pm} = -j^n A_{n\pm} e^{j(\Omega t + B_\pm x)} \tag{7.7}$$

which is limited.

Examining the possible solutions of Eqs (7.6), it follows that in relation (7.7):

a) the argument of the exponential function can only be imaginary;

b) the argument is a linear function of t and x;

c) the multiplicating factor of t is identical in the perturbant functions of all the basic modes, however, the multiplicating factor of x

$$B_{1+} = B_{2+} = B_+ \quad \text{and} \quad B_{1-} = B_{2-} = B_-;$$

d) and finally, when the amplitude of A_{2+} is real, the amplitude of $f_{1\pm}$ takes the form of $(-jA_{1\pm})$.

According to the original restrictions on the perturbation let be

$$\Omega \ll \omega_0 \quad \text{and} \quad A_{n\pm} \ll C_{n\pm}\,. \tag{7.8}$$

Then from Eq. (7.6) it follows that

$$(\beta_0 + \Omega\sqrt{L_0 C_0})^2 - (\beta \pm B_\pm)^2 = \frac{k^2}{4},$$

that is, neglecting the terms being small in second order

$$B_\pm = \pm\Omega \frac{\sqrt{L_0 C_0}}{\sqrt{1 - \dfrac{k^2}{4\beta_0^2}}} = \pm\Omega\gamma_0 \tag{7.9}$$

and

$$\frac{C_{2\pm}}{C_{1\pm}} \cdot \frac{A_{2\pm}}{A_{1\pm}} = \frac{2}{k}\left[(\beta_0 + \Omega\sqrt{L_0 C_0}) - (B_\pm \pm \beta)\right].$$

As in the equations no further restrictions can be found and

$$\max\, e^{j(\Omega t + B_\pm x)} = 1\,,$$

the original restrictions for the perturbation may be fulfilled taking (7.8) into consideration.

The equations (7.6) give restrictions only for the proportion of the amplitudes of the perturbant functions and the $B_\pm \ll \max(\beta_0, \beta)$ also holds (see (7.9)).

As the $f_{n\pm}$ solutions fulfil the conditions mentioned above (Chapt. 6.1 and [8]) they are usable for determining the group velocity:

$$f_{n\pm} = -j^n A_{n\pm} e^{j\Omega(t \pm \gamma_0 x)}. \tag{7.10}$$

The perturbed solutions are

$$i_{\text{pert}}(x,t) = i(x,t) + [I_0^+ \cdot f_{i0+} e^{j\Omega(t-\gamma_0 x)} e^{-j\beta x} + \\ + I_0^- f_{i0-} e^{j\Omega(t+\gamma_0 x)} e^{j\beta x}] e^{\frac{k}{2} x + j\omega_0 t},$$

and

$$u_{\text{pert}}(x,t) = u(x,t) + [U_0^+ f_{u0+} e^{j\Omega(t-\gamma_0 x)} e^{-j\beta x} + \\ + U_0^- f_{u0-} e^{j\Omega(t+\gamma_0 x)} e^{j\beta x}] e^{-\frac{k}{2} x + j\omega_0 t}$$

where $f_{i0\pm}$ and $f_{u0\pm}$ can be obtained from $f_{n\pm}$ by using Eqs (7.3) and (7.4).

Taking the relations (6.23), (6.26) and (6.29) into consideration the group velocity can be obtained. As $f_{n\pm}$, $f_{u0\pm}$, and $f_{i0\pm}$ are constant,

$$v_{g\pm} = \pm \frac{1}{\gamma_0}, \tag{7.11}$$

that is

$$|v_g| = \frac{1}{\sqrt{L_0 C_0}} \sqrt{1 - \frac{k^2}{4\beta_0^2}},$$

and v_{g+} and v_{g-} are the group velocities of the wave propagating along the $+x$ and $-x$ axis, respectively.

It is worth mentioning, that the phase velocity of waves propagating along the $\pm x$ axis is

$$v_f = \pm \frac{\omega_0}{\beta} = \frac{1}{\sqrt{L_0 C_0}} \frac{1}{\sqrt{1 - \frac{k^2}{4\beta_0^2}}},$$

that is

$$v_g \cdot v_f = \frac{1}{L_0 C_0} = \frac{1}{(\sqrt{L_0 C_0})^2} = v^2.$$

This example shows the uniformity and the usability of the procedure based on the inhomogeneous basic modes. This is an important adventage of the method itself.

8. Summary, conclusions

a) The ways of defining line parameters (R, L, C and G) are briefly surveyed for transmission lines with given restrictions and the basic equations are given in equivalent forms.

b) The "method of inhomogeneous basic modes" is applied at homogeneous, quasi-homogeneous and inhomogeneous

$$\left(\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}\right)$$

transmission lines for determining the wave pattern assuming the solution to take the form of

$$e^{j\omega_0 t} \quad (\omega_0 = \text{const.}).$$

c) With the aid of the "method of inhomogeneous basic modes" several groups of inhomogeneous transmission lines are defined: homogeneous class, polynomial class, exponential class, periodical class and others.

d) A possible method is given for the direct determination of the resultant wave pattern which is simultaneously useful for the classification of inhomogeneous lines.

e) Adapting results and methods of electromagnetic wave theory the group velocity may be defined and determined on transmission lines by an existing perturbing solution assuming even monochromatic (continuous) signals.

As a part of this the restrictions relating to the perturbation of signals and originating from e.m. wave theory are summarized in detail.

The commonly used definition, the way of calculating group velocity are briefly surveyed, together with their criticism.

f) The possibility of calculating the group velocity from perturbing function obtained in case of using the method of signal perturbation is given for real and complex perturbing functions which can be identical or different for current and voltage. The difficulties are shown in detail arising when the power (Poynting-vector) will be complex.

g) The method of calculating the line parameters valid for perturbing signal in case of dispersive transmission lines, and the definition of linear dispersive transmission lines are given.

h) The way of determining the group velocity is given using the method of inhomogeneous basic modes for linear, non-dispersive and dispersive transmission lines as well.

The group velocity is analysed in detail in a non-dispersive case and it is justified that in some cases and in any homogeneous case $v_g \equiv v_f$.

i) The analysis of group velocity is given for ideal transmission lines filled with tempered, isotropic plasma. The value of v_g and its relation to v_f is determined. This serves to illustrate the method of solution in more general cases.

Similarly v_g is determined in the case of a simple attenuated transmission line and its relation to v_f. It is important, that now $v_g = v_f$.

j) The way of investigation of energy propagation (or dissolve) is given with the aid of perturbing the resultant wave pattern. An applicational example is shown for homogeneous and homogeneous-class of lines.

k) It is an important result that the group velocity can be developed from the basic equations without any special assumptions. So, a way opens up for the applications of dispersive and for non-dispersive linear lines in analysing the delay time, transients, etc.

l) We give an applicational example in Chapt. 7.

m) Several experiences show that when generalizing the method for non-elementary perturbations, a perturbation will not be "clear" but simultaneously all the signal parameters will be perturbed. Assuming e.g. a "clear" amplitude perturbation the contradiction will not arise in case of assuming the perturbation and its derivatives being elementarily small. These difficulties may be avoided in non-elementary cases by perturbing all the signal parameters. This needs to be further investigated.

REFERENCES

1. Collin, R. E.: Grundlagen der Mikrowellentechnik; VEB Verlag Technik, Berlin 1973
2. Simonyi, K.: Theoretische Elektrotechnik; VEB Deutsches Verlag der Wissenschaften, Berlin 1968
3. Ebel, I. L.—List, F. D.: Izmerityel szaprotyivljenyija ballaszta tipa ISzB-1; Avtomatyika i Telemechanika („Szovjaz") **3**, 1967, No. **5**
4. Erdős L.: Ballasztellenállás vizsgálata rádióizotópos mérési módszerrel, tekintettel az ágyazat szennyezettségére és az aljak minőségi állapotára; Vasúti Tudományos Kutató Intézet, 421/1972
5. Erdős L.: A ballasztellenállás vizsgálati módszereinek továbbfejlesztése, automatizálása, Vasúti Tudományos Kutató Intézet, OP 12/1972
6. Ferencz, Cs: Elektromágneses hullámterjedés inhomogén, lineáris közegekben; kandidátusi értekezés, MTA Könyvtár, Budapest 1970
7. Ferencz, Cs.: Electromagnetic Wave Propagation in Inhomogeneous Media: Method of Inhomogeneous Basic Modes; *Acta Techn. Hung.*, **86** (1978), 79
8. Ferencz, Cs.: Electromagnetic Wave Propagation: The Analysis of the Group Velocity; *Acta Techn. Hung.*, **86** (1978), 169
9. Fodor Gy.—Simonyi K.—Vágó J.: Elméleti villamosságtan példatár; Tankönyvkiadó, Budapest 1967
10. Ferencz, Cs.: Electromagnetic Wave Propagation in Inhomogeneous Media: Strong and Weak Inhomogeneities; *Acta Techn. Hung.*, **85** (1977), 433
11. Kamke, E.: Differentialgleichungen, Lösungsmethoden und Lösungen; Akademische Verlagsgesellschaft Geest & Portig K.-G., Leipzig 1951
12. Baule, B.: Die Mathematik des Naturforschers und Ingenieurs, Band IV. Gewöhnliche Differentialgleichungen; S. Hirzel Verlag, Leipzig, 1953
13. Arnaud, J. A.—Saleh, A. A. M.: Theorems for Bianisotropic Media; *Proc. IEEE*, **60** (1972), 639
14. Ralston, A.: Bevezetés a numerikus analízisbe; Műszaki Könyvkiadó, Budapest 1969

15. BUDDEN, K. G.: Radio Waves in the Ionosphere; Cambridge at the University Press, 1966
16. HEALD, M. A.—WAHRTON, C. B.: Plasma Diagnostics with Microwaves; John Wiley & Sons, New York 1965
17. ÁRKOS, I.: Elektronfizika, Jegyzet, Budapest, 1967
18. ALLIS, W. P.—BUCHSBAUM, S. J.—BERS, A.: Waves in Anisotropic Plasma; M.I.T. Press, Cambridge, Mass. 1963
19. IDEMEN, M.: The Maxwell's Equation in the Sense of Distributions; *IEEE Trans. Ant. and Prop.*, **AP-21**, (1973), 736
20. FERENCZ, Cs.: Electromagnetic Wave Propagation in Inhomogeneous Media: The Analysis of the Rotation of Polarization and the Application of the Principle of Modified Ray Tracing; under publ. in *Acta Techn. Hung.*
21. SIMONYI, K.: Foundations of Electrical Engineering; Pergamon Press, 1963
22. REITER, GY.: Generalized Telegraphist's Equations for Waveguide of Varying Cross-sections. *Proc. of the Inst. of the Electrical Engineers*, Part B. Supplement, Convention on long distance transmission by waveguide. (1959) January, pp 54—57

Allgemeine Analyse von monochromatischen Signalen bei Fortpflanzung entlang inhomogener Übertragungsleitungen. II. — Der zweite Teil der Arbeit bringt das Verfahren zur Bestimmung der Gruppengeschwindigkeit durch Verwendung eines Perturbationsprozesses auf Grundlage der Methode der inhomogenen Grundmoden. In diesem Fall werden homogene, inhomogene und innerhalb strikt linearer Grenzen Dispersionslinien analysiert. Schließlich wird ein Beispiel gezeigt für die Bestimmung der Wellenform und Gruppengeschwindigkeit unter Benützung der Methode der inhomogenen Moden.

BOOK REVIEW

L. Gábor

ÉPÜLETSZERKEZETTAN (BAUKONSTRUKTIONSLEHRE), BAND 4

Tankönyvkiadó, Budapest 1979. Seiten, 234 Abbildungen, fachliterarisches Verzeichnis

Dieses Buch ist der 4. abschließende Band der Baukonstruktionslehre des Lebenswerkes des Professors L. GÁBOR. Dieser Band befaßt sich mit der Konstruktion von Türen, Toren, Fenstern, Glaswänden, Portalen, lichtdichten und Abschirmvorrichtungen, von eingebauten Kästen, Fußböden und Wandbekleidungen. Das Buch behandelt diese Konstruktionen, sowohl in architektonischer, als auch in konstruktiver Beziehung kritisch und führt charakteristische Ausführungsbeispiele dieser Konstruktionen an.

Inhaltlich ist das Buch völlig neuartig. Weder in der ungarischen, noch in der ausländischen Fachliteratur ist ein Fachbuch vorzufinden, das sich mit den ergänzenden Konstruktionen der Gebäude so ausführlich und im wissenschaftlich so wertvoller Weise befaßt. Die Sammlung, Systematisierung und kritische Wertung dieses weit verzweigten Stoffes verpflichtet die ungarische Bauwissenschaft und die Bauindustrie zu großem Dank und Anerkennung dem Verfasser gegenüber. Das Werk ist für die ungarische Architektur ein verläßlicher Wegweiser und bietet grundlegende Hilfe für den Unterricht an den technischen Universitäten und Hochschulen.

Im Text des Buches ist jede Einzelheit sorgfältigst behandelt, doch sind Zweck und Anordnung der ergänzenden Konstruktionen, sowie die günstigste Befriedigung der dem Zweck dienenden Bedingungen maßhaltend erörtert. Der Verfasser unterläßt es nicht die Aufmerksamkeit ständig auf die Befriedigung der baupolizeilichen, statischen, stoffkundlichen, lüftungstechnischen und sonstigen physikalischen und chemischen Anforderungen, sowie auf den Feuer- und Lärmeschutz, ferner auf fertigungstechnologische, sanitäre und ästhetische Erfordernisse zu richten. Auch erinnert der Autor stets an den Zusammenhang zwischen Ursache und Wirkung, an die Dauerhaftigkeit und Wirtschaftlichkeit, ferner an die Notwendigkeit Probleme in zweckdienlicher und vernünftiger Weise zu lösen, die sich aus nicht gleichgerichteten, sondern einander sogar entgegengesetzten Anforderungen ergeben.

Einen besonderen Wert verleihen dem Buch das mit musterhafter Sorgfalt und fachmännischer Genauigkeit gezeichnete reiche Bildmaterial, sowie dessen geschmackvolle und demonstrative Ausführung, die den Vorteil, der sich aus der Anwendung zweier Farben bietet, weitgehend ausnützt. In dieser Beziehung sind besonders die klaren und übersichtlichen Bilder hervorzuheben, welche die Fensterkonstruktionen in einer Art darstellen, die nur die Anwendung zweier Farben ermöglicht.

Zusammenfassend muß festgestellt werden, daß der vorliegende vierte Band des Werkes des Professors L. GÁBOR, den vorhergehenden drei Bänden gleich, ein hervorragendes Produkt der ungarischen technischen Fachliteratur darstellt und als Höhepunkt der während eines Menschenlebens geleisteten wertvollen wissenschaftlichen Arbeit des Autors zu werten ist.

P. Csonka

L. Palotás

THEORY OF MATERIALS FOR ENGINEERING CONSTRUCTIONS

1, GENERAL KOOWLEGDE ON MATERIALS

Akadémiai Kiadó, Budapest 1979

Prof. PALOTÁS, the Hungarian Academy of Sciences and its Publishing House fulfilled a long-standing and urgent requirement by the publication of this book. A two-volume book by the same author entitled *Building Materials* published under the auspices of the publishing House of the Hungarian Academy of Sciences in 1959—1961 went out of print in a few years time.

The first of the three volumes of the new work in question is subdivided into three main chapters.

In main chapter one the basic knowledge of the theory of the building materials and engineering constructions is summed up in short and, hereafter, the structural materials are treated beginning with the cultures of ancient times up to the present day. Author demonstrates with persuasive data, numerous figures and examples that the development of the building, in general, and architectural and engineering constructions in particular, cannot be separated from building materials: the growth in assortment of materials, the deepening of the knowledge on materials determined the methods of building construction and their development.

In this main chapter the development of the theory of building materials, their internal structures are described, by explaining those adjacent fields without the knowledge of which the solution of the day-to-day problems of the research studies or the utilization of the materials is unimaginable today.

Main chapter two summarizes the chemical, physical and mechanical properties of the structural materials as well as the measuring methods, the system and units of weights and measures, giving a clear idea of the dialectic unity of the materials, measuring devices, investigating person and material properties. The particular value of this chapter is the way in which the author distinguishes between the linearly elastic models, linearly elastic-linearly hardening, linearly elastic-perfectly plastic, rigid-perfectly plastic and rigid-linearly hardening materials, and explains and presents the fields of application of these models, at the same time giving a significant help to the designers of engineering structures in choosing the suitable loads and stress-strain models.

Main chapter three sums up the investigation and evaluation methods of the strength of structural materials. The methods and principles of static, sustained fatigue and dynamic tests, the equipments needed, their measuring range, exactitude, and applicability, are detailed. The stress-strain patterns, the manner of mechanical evaluation of the deformations and the associated limit-design theories basically determine the load capacity of the engineering constructions. This chapter truely mirrors the creative studies performed in this field, by the author, and gives information in an excellent way for the designer how he can draw sure conclusions from the measurement results of strains to the stress distribution in the structures. The mathematical evaluation gives a great help which is dealt with in detail by the author as a clause of this chapter, with in particular takes into account the mathematic-statistical methods.

The whole work verifies that the basic condition of the duration and economical construction of the projects is the suitable quality as regards the used material and design, execution and maintenance work. The essential condition of the good quality is the knowledge of the material, permanent checking and continuous testing. The Publishing House of the Hungarian Academy of Sciences, bringing to the fore the author's plentiful experiences obtained in his fifty-year's scientific tutorial and industrial activities, presenting the Hungarian civil engineer's society with a very useful, up-to-date, high-level book.

J. Ujhelyi

L. Palotás

THEORY OF MATERIALS FOR ENGINEERING CONSTRUCTIONS

2. WOOD, STONE, METALS AND BINDING MATERIALS

Akadémiai Kiadó, Budapest 1979, 586 pages, 337 figures, 98 tables, 230 references

This volume deals with the properties and testing methodes of wood, stone, metallic and binding materials as well as with the manufacturing of their products and their utilization. Its subject matter is comprehensive, the manner of treating is detailed and manysided in every field which has been promoted by outstanding co-workers chosen by the author, such as P. KERTÉSZ and S. VERESS, P. KERTÉSZ wrote the chapter dealing with stones, while S. VERESS was the writer of that of the metals, both with thorough grounding in their professions suitable to the whole of the work, at the same time lending to it a perfect uniformity in regard both to its content and form.

The main chapter which deals with *wood*, in ten subchapters treats the significance of the wood materials used for engineering constructions, the quality requirements, testing methods and the possible fields of application in Hungary.

The first subchapter deals with the utilization of wood through the history of mankind, with its phenomena of living as well as with its properties: with the anatomy and internal structural pattern. The second subchapter discusses the production of wood, its different products; the third one the kinds and characteristics of structural wood; the fourth subchapter details the statistical data of the wood economy management throughout the world and in Hungary. The unfavourable statistical data for the Hungarian wood economy management require a significant improvement and the augmentation of the wood stock, the full utilization of the properties of wood the acquirement of a detailed and wide ranged knowledge in this field, to which the book serves as a a great help.

Subchapters five, six and seven analyse the chemical, physical and mechanical properties of wood, respectively. The role of the chemical elements and compounds, the constituents, water content, swelling and shrinking properties, density and body density of woods, further their behaviour in connection thermal, electric and acoustic effects; hardness, wear resistance, friction characteristics, cleavability, compressive, tonsile, flexural, shear, torsional and impact strengths, strain behaviour under load, nail and screw bearing capacity. Author refers to the fact that the wood, due to its special structural characteristics, has been a valuable, very useful organic building material from ancient times up to our own days for a great number of different constructional purposes, whereby it is worth a particularly careful treatment.

In subchapters eight, and nine the defect and illnesses, respectively durability and preservation of the wood are described. From subchapter ten, the refinement of the wood, the substituting wood materials, testing of the substitution and improving of the wood materials and their behaviour may be recognized.

The main chapter entitled Natural building stones, treats in five subchapters the knowledge relating to stone materials.

The structure, classification and kinds of stone materials are shown in subchapter four. Herein the stones are characterized; the crystalline and amorphous rock constituents, the rock structures, the formations, classifications and the kinds of natural stones are thoroughly discussed. In subchapter two, information is given on the production, processing and treatment of stones. The mineral disclosed resources in turn by quarry, gravel and sand-pit operation and by processing into building materials.

In subchapter three, the physical-strength properties are presented. The concept of the quality of stones and the quality requirements in rocks are interpreted in a very interesting up-to-date way. The properties expressing the quality should be determined by tests; the tests to be performed constitute an order of qualification which may be the qualification of the products, of the rock, as well as the evaluation of the research of mineral resources. The reliability mostly depends on the effectivity of sampling. The most significant properties of stones are defined by petrographic characterization and by the investigation relating to mass distribution, the destructive and non-destructive strength tests, and by those performed to determine the hardness, wear, cleavability, processing, aggregate-strength and aggregate size distribution. Chapter four is closely attached to the third one and analyses the weathering and durability properties, the test and evaluation problems as well as those of stone protection and preservation. In connection with durability, the concept of the effective factor of variation of the model and rock behaviour has been introduced.

Fortunately, the main chapter of the natural building stones is written in the spirit of the Hungarian Standard series MSZ 18280—18296 and qualification system which came into force in our own days and, therefore, not only by its aspect but also by its content it is richer than other works published so far on this subject-matter. The observance of the standard prescriptions in a day-to-day obligation. He, who reports on the standards at the same time also by showing the scientific roots of the day-to-day problems, also diffuses the scientific way of thinking. All these are solved conveniently yet in moderate size.

Main chapter three entitled *Metals* deals with metals which are significant from the point of view of utilization in the structures of building industry. After the exposition of general problems this is followed by a detailed discussion of iron and steel, aluminium and other metals in 27 subchapters.

In six subchapters, the general properties of the metals are described i.e., the differences between metals and non-metallic materials, the division of metals according to their application and metallurgic viewpoints internal structures and crystallization, the structural characteristics of the different alloys and of their changes of state, equilibrium diagrams, the physical properties of metals, the meaning of corrosion. The most interesting subchapter is that which reads about alloys, because from the about 90 elements of metalloid and metallic nature, 4000 two-component, 118000 three-component and several millions of four-components alloy systems may be derived.

The informations in connection with iron and steel are given in 10 subchapters. The economic significance of iron is followed by the discussion of the constitutional diagrams of iron-carbon alloys which is a very significant subject matter, because the different kinds of iron used in practice are, in fact, carbon alloys. It is of great importance to know that the properties of steel may be improved by cold and warm working, heat treatment and incrustation. The welding of steel is discussed in detail.

The fusion and fusion-pressure welding procedures are systematized in a clearly constructed figure. A particular subchapter deals with the effects of welding, weldability and with the investigation of weldability, and also with the testing of welded connections. The mechanical properties of iron and steel are related to the test results and standard prescriptions. These properties are, in fact, tensile strength, work capacity, compressive, shear strength, hardness, torsional, impact and flexural strength, fatigue strength, relaxation, creep of steels, as well as the technological bending and repeated folding. The kinds of steels are classified into the following groups. Carbon steels: structural steels, weldable steels, hot-rolled steels and other types of carbon steels; cold-shaped steels: profiles, concrete reinforcement steel networks, twisted reinforcement steel bars, stressing wires; alloyed steels; weather-proof and corrosion-resisting steels; these are also grouped according to casts. The subchapters of iron and steel are closed by discussing the phenomenon of corrosion.

The second most significant metallic material for engineering structures, aluminium is treated in five subchapters. The production of aluminium from bauxite, the types and alloys of aluminium, its casting, processing, physical behaviour, heat treatment; welding, brazing and riveting of aluminium structures and their corrosion and protection are presented.

From among the other metals, the most significant properties of copper and its alloys, those of lead, tin, and zinc are detailed in separate chapters.

From the main chapter treating the metals, it may be stated that the reader gets acquainted with a high-level, up-to-date knowledge matter.

In the *main chapter* entitled: *Binding materials*, subsequent to a subchapter in treating of problems of general chapter; treats in four subchapters, subdivided according to their applications in the building industry, the binding materials utilized in the building industry.

The non-hydraulic binding materials hardening in air, are lime, gypsum and magnesia. The production of these binding materials from limestone, rock gypsum and rock magnesite, respectively, the mechanism of binding-assuring the needed strength, the utilization of the binding materials, their properties and test methods are written in detail in this part of the book.

The weakly hydraulic binding materials are the ancestors of cements which were already used in Roman times. Such are hydraulic lime made from marly limestone or from chalky clay; the Roman cement which is burnt from marl; the lime puzzuolana which is produced from a mixture of hydrated-lime and natural sour hydraulic additives.

The basic slags rich in lime, the santonine earth, puzzuolana, trasses, diatomaceous earth and the artificial supplementary binding materials fall into the group of hydraulic additives. Their common behaviour is, that ground into flourfine powder, it cannot harden by themselves, neither by mixing with water, however, under the effect of a catalytic agent, both in air and water. The hydraulic additives are treated in a detailed way suitable to their significance. In Hungary, these materials are put into circulation combined with cement or lime products.

A hydraulic binding material is the relatively young Portland-cement discovered 135 years ago, but since that time being one of the most significant building material of the modern building industry. Its importance may be assessed from the large amount that is utilized in our country, which is as much as 6.5 million tons. Portland-cement is grouped among the silicate cements which, together with the alumina cement constitute the hydraulic binding materials. The book describes the manufacturing of cement, outlines the chemical processes of the clinker formation, the chemical composition and mineral clinkers, as well as the hydration and structure of the cement stone. Hereafter, the physical and strength properties, the time of setting, degree of grinding fineness, constancy of volume, binding force, density, water-holding capacity, thermal expansion, thermal curability, and storage stability are discussed, including the test methods, the recent standard prescriptions and the test results obtained so far.

A survey on the types of cements and their field of application may be interesting for a large reading public, giving an outline of Portland-cements, alumina cements, swelling cements, sigma-cement.

In summary, it may be stated that the new volume, similarly to the first volume, accomplished the task as purpose, which has been set by the author referring to Galileo Galilei in the preface of the book which helps us to acquire knowledge. The Publishing House and the Printing Office of the Hungarian Scientific Academy contributed the attainment of this aim with its valuable work.

J. Ujhelyi

ZEMENT TASCHENBUCH 1979/80

Verein Deutscher Zementwerke, Bauverlag G.m.b.H., Wiesbaden—Berlin, 594 S., 75 Abb., 105 Tafeln

Dieses Taschenbuch erscheint in Zeiträumen von ein-zwei Jahren. Der vorliegende Band ist die 47 Ausgabe des Werkes.

Der erste Teil des Buches behandelt in den beiden ersten Kapiteln die quantitativen und qualitativen Kennzeichen der Bestandteile des Zements. Im dritten Kapitel wird ein guter Überblick der zeitgemäßen Zementfabrikation und über deren Bedarf an Rohmaterialien, sowie an technologischen Anforderungen geboten. Besondere Beachtung verdient die systematische Zusammenfassung der, den Bestimmungen des Umweltschutzes angepaßten Vorschriften der Zementfabrikation. Die physikalischen und mechanischen Kennzeichen der Zemente enthält das vierte Kapitel.

Der zweite und dritte Teil des Buches erfaßt das ganze Gebiet der derzeit gebräuchlichen Zuschlagstoffe, die es ermöglichen die Eigenschaften des Betons (Festigkeit, Raumgewicht, Wasserdichtheit, Abbindebeschleunigung, Färbung usw.) innerhalb weiter Grenzen zu bestimmen.

Der vierte und fünfte Teil behandeln die Bedingungen der Herstellung von normalen und Leichtbeton, sowie die diesbezüglichen Klassifikationsverfahren. Diese Kapitel enthalten auch Hinweise zur Schätzung der Festigkeit, der Formänderung und der rheologischen Eigenschaften des frischen und erhärteten Betons. Die Kapitel, die die Planung von Spezialbetonen (z. B. wasserdichten, wärme- und frostbeständigen, strehlendichten und verschleißfesten Betonen) behandeln, sind für Betontechnologen, die sich mit der Planung solcher Betone befassen, von besonderem Interesse.

Große Aufmerksamkeit verdient auch der sechste Teil des Werkes, der die neuesten Anwendungsmöglichkeiten des Zements zur Herstellung von Unterwasser- und Torkretbeton, von mit Glas- und Kunststoffasern verstärktem Beton und die Herstellung von unterschiedlichen Ansprüchen entsprechenden Sichtbetonen erörtert.

Das Handbuch schließt eine Zusammenfassung der in früheren Ausgaben angeführten Spezialbetone und deren Anwendungsgebiet, sowie die Sammlung der wichtigsten Normen und Bauvorschriften ab.

Das Buch ist ein nützlicher Behelf für Forscher und für die im Gange befindliche Modernisierung der ungarischen Vorschriften.

B. Goschy

INDEX

BOOK REVIEW

BUCHBESPRECHUNG

Printed in Hungary

A kiadásért felel az Akadémiai Kiadó igazgatója — Műszaki szerkesztő: Zacsik Annamária

A kézirat nyomdába érkezett: 1980. IV. 28. — Terjedelem: 15,05 (A/5) ív, 48 ábra, 1 melléklet

81.8284 Akadémiai Nyomda, Budapest — Felelős vezető: Bernát György

Acta Techn. Hung. **90** (3—4), pp. 187—200

KÉZDI, Á.—MLYNAREK, ZB.: *Static Penetration Test Results with Soils Having Slight or Medium Cohesion*

The paper analyses the influence of the phase composition: (volume percentages of solid, air and water) on the cone resistance during static penetration test. The analysis is based on data obtained in investigations using the ííboudler loam" in Great Poland Lowland (ííKonin clay" and ííRudnicze clay") and loess from Törökbálint, Hungary. The static souding test was carried out using the laboratory device. The function which furnishes the relationship between cone resistance and phase composition was determined numerically. The analysis of the cohesion and angle on internal friction effect has been done using statistical methods.

Acta Techn. Hung. **90** (3—4), pp. 201—220

SOARE, M. V.: *A proposal of Codification of Theories in the Theory of Structures*

Starting from the statement that there is no systematic classification of theories in the theory of structures, a codification is proposed based on four main factors and namely: the manner of expressing the equilibrium of the structure, the constitutive low of the material, the magnitude of deformation and the curvature expression. The paper is concluded with examples of codification of some problems in the theory of structures.

Acta Techn. Hung **90** (3—4), pp. 221—246

KOZÁK, I.: *Notes on the Field Equations with Stresses and on the Boundary Conditions in the Linearized Theory of Elastostatics*

The necessary and sufficient conditions of the single-valued displacements can be determined from the strain field, that is the six Saint-Venant's compatibility equations are not independent of each othert. To eliminate the interdependence, K. WASHIZU in 1957 has given a solution by allowing for Bianchi's identity. In this paper, on the basis of the condition of the single-valued displacements concering the boundary (i. e. on the base of the compatibility boundary condition) similarly with the aid of Bianchi's identity, the solution given by K. WASHIZU is extended to all feasible caves and, at the same time, the results are generalized to aay curvilinear coordinate systems. In this way, the set of the governing equations of the linearized theory of the elastostatics, in connection with the six stress coordinates, is composed of the three equilibrium equations and of three suitably selected compatibility equations, however, the customary boundary conditions chould be complemented by three compatibility boundary conditions.

Acta Techn. Hung. **90** (3—4), pp. 247—258

VARGA, A.—ECSEDI—TÓRH, P.—MÓRICZ, F.: *A Heurestic Method for Speeding up Manual Optimization of Boolean Functions*

A simple heuristic procedure is procented to speed up manual optimization of Boolean functions (truth-functions). The procedure uses a tree representation of thr Boolean functions and so it is geometric rather than algebraic is character. The method can easily be applied to totally determined Boolean functions as well as to partial ones, containing not more than 10 variables. Some illustrative examples are also included. The authors are induced to the referee for his valuable remarks and suggestions.

Acta Techn. Hung. **90** (3—4), pp. 259—274

FARKAS, J.: *Optimum Design for Bending and Ultimate Shear Strength of Hybrid I-Beams*

The paper presents the approximate analytical optimization of welded hybrid I-beams subject to bending. In the objective function the different material costs of web and flanges are taken into account. The following design constraints are considered: local buckling constraint of web and that of compressed flange, constraints of maximum stress and deflection, respectively. The dead weight and the lateral buckling is not considered. The obtained simple formulae are applied to the comparative calculations which show e. g. that the material cost of hybrid beams optimized with respect to stress constraint is by 15—25% less than that of homogeneous ones. Further, limit shear stresses are determined below where the effect of shear can be neglected.

Acta Techn. Hung. **90** (3—4), pp. 275—284

HENNYEY, Z.: *The Theory of Relativity as Seen by the Engineer*

Physics in the most important underlying discipline of the engineering sciences. The way of thinking built úpin the notions of absolute space and time, bases of classical physics, means for the man of ingineering the stable foundation of his own science. No wonder, if he does not get to fricendly terms with the new ideas of the theory of relativity. There is indeed a contribution between the traditional logics of classical physics and the logics of the theory of relativity and to bridge it is a hard task. To bridge this gap is the aim of this study.

Acta Techn. Hung. **90** (3—4), pp. 285—290

BOSZNAY, Á. et al.: *Improvable Bracketing of the Circular Frequences of a Straight Rod with Characteristics Varying along its Length, Performing Flexural Oscillations*

The paper solves the problem with the aid of the Poincaré—Rayleigh—Ritz and Fichera's methods and points out some problems arising in connection with the numerical solution.

Acta Techn. Hung. **90** (3—4), pp. 291—302

GRÓSZ, M.: *Optimal Designing of Pipe Networks by Integer Programming*

In the paper the author deals with the extension of existing water networks and the design of new netwirks. For the optimal solution of these problems a nonlinear integer programming model is presented. It differs from models known from literature inasmuch as a) the pipe diameters can assume only discrete values; b) the water quantities fed from the sources are not constants given beforehand, but parameter depending on the network; c) minimalization of the investent costs and of the operating expenses is aimed at the definition of the problems is based on graph theory and network flow theory. In order to reduce the possible number of pipe diameter combinations a special so-called minimal tree has been built ip. The nonlinear integer programming problem has been solved by the method of penalty functions. Knowing the minimal tree the discrete diameter combination near the accurate optimum have been sought for which prodives, togethe with the corresponding water quantity a local optimum.

Acta Techn. Hung. **90** (3—4), pp. 303—320

FERENCZ, CS.: *Electromagnetic Wave Propagation in Moving Media with Special Regard to Frequency-Shifts "Anomalous", Frequency-Shifts in Astronomy*

In the first part after a comprehensive review of problems which appear to be fundamental in the study of electromagnetic wave progapation in moving media 00 is shown how to eliminate the observing discrepancy and the possible ways of analyting the electromagnetic wave propagation in moving media are discussed. One of the discussed methods, developed under the name of relativistic ray tracing, is described in detail in the second part. It is show to determine the signal spectrum and the so-called basic effect for which expressions formulated in cases thought to be most important. The weighting functions and the proofs of their form are given. Finally, in the *third part* results of earlier similar analyses are utilized for explaining "anomalous" frequency shifts and other associated wave propagation phenomena (line broadening, polarisation rotation, etc.) observed in the solar system. It is suggested that similar apparently anomalous frequency shifts as well as other associated phenomena observed in galactic and extragalactic wave propagation could be interpreted in the same way.

Acta Techn. Hung. **90** (3—4), pp. 321—349

MRS, I. ÁRKOS: *General Analysis of Monochromatic Signals Propagating along Inhomogeneous Transmission Lines*

The *first part* of this paper gave a short summary of fundamental relations and commun methods for the investigation of inhomogeneous tranmission lines. Then using the method of inhomogeneous basic modes it shows a process for determining the wave pattern along the line which seems to be generally sufficiently useful. Based on this, one may have different categories of transmission lines. Finally by using a further method too, the way of determining the wave pattern and the different categories will be shown. The *second part* of the paper gives the manner of determining the group velocity by using a perturbation process based on the method of inhomogeneous basic modes. In this case homogeneous, inhomogeneous and within strictly linear and dispersive lines are analysed. Finally an example is shown for determining the wave pattern and group velocity by using the method of inhomogeneous basic modes.

The *Acta Technica* publish papers on technical subjects in English, French, German and Russian.

The *Acta Technica* appear in parts of varying size, making up one volume.

Manuscripts should be addressed to

Acta Technica
H-1051 Budapest
Münnich Ferenc u. 7.
Hungary

Correspondence with the editors and publishers should be sent to the same address

Orders may be placed with "Kultura" Foreign Trading Company (H-1389 Budapest 62, P.O.B. 149. Account No. 218-10990) or its representatives abroad.

Les *Acta Technica* paraissent en français, allemand, anglais et russe et publient des travaux du domaine des sciences techniques.

Les *Acta Technica* sont publiés sous forme de fascicules qui seront réunis en volumes.

On est prié d'envoyer les manuscrits destinés à la rédaction à l'adresse suivante:

Acta Technica
H-1051 Budapest
Münnich Ferenc u. 7.
Hongrie

Toute correspondance doit être envoyée à cette même adresse.

On peut s'abonner à l'Entreprise du Commerce Extérieur «Kultura» (H-1389 Budapest 62, P.O.B. 149. Compte courant No. 218-10990) ou chez représentants à l'étranger.

«*Acta Technica*» публикуют трактаты из области технических наук на русском, немецком, английском и французском языках.

«*Acta Technica*» выходят отдельными выпусками разного объема. Несколько выпусков составляют один том.

Предназначенные для публикации рускописи следует направлять по адресу:

Acta Technica
H-1051 Budapest
Münnich Ferenc u. 7.
Венгрия

По этому же адресу направлять всякую корреспонденцию для редакции и администрации.

Заказы принимает предприятие по внешней торговле «Kultura» (H-1389 Budapest 62, P.O.B. 149. Текущий счет № 218-10990) или его заграничные представительства и уполномоченные.

Reviews of the Hungarian Academy of Sciences are obtainable at the following addresses:

AUSTRALIA
C.B.D. LIBRARY AND SUBSCRIPTION SERVICE, Box 4886, G.P.O., *Sydney N.S.W. 2001*
COSMOS BOOKSHOP, 145 Ackland Street, *St. Kilda (Melbourne), Victoria 3182*

AUSTRIA
GLOBUS, Höchstädtplatz 3, *1200 Wien XX*

BELGIUM
OFFICE INTERNATIONAL DE LIBRAIRIE, 30 Avenue Marnix, *1050 Bruxelles*
LIBRAIRIE DU MONDE ENTIER, 162 Rue du Midi, *1000 Bruxelles*

BULGARIA
HEMUS, Bulvar Ruszki 6, *Sofia*

CANADA
PANNONIA BOOKS, P.O. Box 1017, Postal Station "B", *Toronto, Ontario M5T 2T8*

CHINA
CNPICOR, Periodical Department, P.O. Box 50, *Peking*

CZECHOSLOVAKIA
MAD'ARSKÁ KULTURA, Národní třida 22, *115 66 Praha*
PNS DOVOZ TISKU, Vinohradská 46, *Praha 2*
PNS DOVOZ TLAČE, *Bratislava 1*

DENMARK
EJNAR MUNKSGAARD, Norregade 6, *1165 Copenhagen*

FINLAND
AKATEEMINEN KIRJAKAUPPA, P.O. Box 128, *SF-00101 Helsinki 10*

FRANCE
EUROPÉRIODIQUES S.-A., 31 Avenue de Versailles, *78170 La Celle St.-Cloud*
LIBRAIRIE LAVOISIER, 11 rue Lavoisier, *75008 Paris*
OFFICE INTERNATIONAL DE DOCUMENTATION ET LIBRAIRIE, 38 rue Gay Lussac, *75240 Paris Cedex 05*

GERMAN DEMOCRATIC REPUBLIC
HAUS DER UNGARISCHEN KULTUR, Karl-Liebknecht-Strasse 9, *DDR-102 Berlin*
DEUTSCHE POST ZEITUNGSVERTRIEBSAMT, Strasse der Pariser Kommüne 3—4, *DDR-104 Berlin*

GERMAN FEDERAL REPUBLIC
KUNST UND WISSEN ERICH BIEBER, Postfach 46, *7000 Stuttgart 1*

GREAT BRITAIN
BLACKWELL'S PERIODICALS DIVISION, Hythe Bridge Street, *Oxford OX1 2ET*
BUMPUS, HALDANE AND MAXWELL LTD., Cowper Works, *Olney, Bucks MK46 4BN*
COLLET'S HOLDINGS LTD., Denington Estate, *Wellingborough, Northants NN8 2QT*
W.M. DAWSON AND SONS LTD., Cannon House, *Folkestone, Kent CT19 5EE*
H. K. LEWIS AND CO., 136 Gower Street, *London WC1E 6BS*

GREECE
KOSTARAKIS BROTHERS, International Booksellers, 2 Hippokratous Street, *Athens-143*

HOLLAND
MEULENHOFF-BRUNA B.V., Beulingstraat 2, *Amsterdam*
MARTINUS NIJHOFF B.V., Lange Voorhout 9—11, *Den Haag*
SWETS SUBSCRIPTION SERVICE, 347b Heereweg, *Lisse*

INDIA
ALLIED PUBLISHING PRIVATE LTD, 13/14 Asaf Ali Road, *New Delhi 110001*
150 B-6 Mount Road, *Madras 600002*
INTERNATIONAL BOOK HOUSE PVT. LTD., Madame Cama Road, *Bombay 400039*
THE STATE TRADING CORPORATION OF INDIA LTD., Books Import Division, Chandralok, 36 Janpath, *New Delhi 110001*

ITALY
EUGENIO CARLUCCI, P.O. Box 252. *70100 Bari*
INTERSCIENTIA, Via Mazzé 28, *10149 Torino*
LIBRERIA COMMISSIONARIA SANSONI, Via Lamarmora 45, *50121 Firenze*
SANTO VANASIA, Via M. Macchi 58, *20124 Milano*
D. E. A., Via Lima 28, *00198 Roma*

JAPAN
KINOKUNIYA BOOK-STORE CO. LTD., 17-7 Shinjuku-ku 3 chome, Shinjuku-ku, *Tokyo 160-91*
MARUZEN COMPANY LTD., Book Dpartment, P.O. Box 5056 Tokyo International, *Tokyo 100-31*
NAUKA LTD. IMPORT DEPARTMENT, 2-30-11 Minami Ikebukuro, Toshima-ku, *Tokyo 171*

KOREA
CHULPANMUL, *Phenjan*

NORWAY
TANUM-CAMMERMEYER, Karl Johansgatan 41—43, *1000 Oslo*

POLAND
WĘGIERSKI INSTYTUT KULTURY, Marszałkowska 80, *Warszawa*
CKP I W ul. Towarowa 28 *00-958 Warszawa*

ROMANIA
D. E. P., *Bucureşti*
ROMLIBRI, Str. Biserica Amzei 7, *Bucureşti*

SOVIET UNION
SOJUZPETCHATJ — IMPORT, *Moscow*
and the post offices in each town
MEZHDUNARODNAYA KNIGA, *Moscow G-200*

SPAIN
DIAZ DE SANTOS, Lagasca 95, *Madrid 6*

SWEDEN
ALMQVIST AND WIKSELL, Gamla Brogatan 26, *101 20 Stockholm*
GUMPERTS UNIVERSITETSBOKHANDEL AB, Box 346, *401 25 Göteborg 1*

SWITZERLAND
KARGER LIBRI AG, Petersgraben 31, *4011 Basel*

USA
EBSCO SUBSCRIPTION SERVICES, P.O. Box 1943, *Birmingham, Alabama 35201*
F. W. FAXON COMPANY, INC., 15 Southwest Park, *Westwood, Mass. 02090*
THE MOORE-COTTRELL SUBSCRIPTION AGENCIES, *North Cohocton, N. Y. 14868*
READ-MORE PUBLICATIONS, INC., 140 Cedar Street, *New York, N. Y. 10006*
STECHERT-MACMILLAN, INC., 7250 Westfield Avenue, *Pennsauken, N. J. 080110*

VIETNAM
XUNHASABA, 32, Hai 3a Trung, *Hanoi*

YUGOSLAVIA
JUGOSLAVENSKA KNJIGA, Terazije 27, *Beograd*
FORUM, Vojvode Mišiča 1, *21000 Novi Sad*

3. IV. 1981. HU ISSN 0001—7035 Index: 26.025