ACTA TECHNICA

ACADEMIAE SCIENTIARUM HUNGARICAE

ADIUVANTIBUS
I. BARTA, E. BÖLCSKEI, P. O. GESZTI, A. LÉVAI

REDIGIT
M. MAJOR

TOMUS 61 FASCICULI 1—2

AKADÉMIAI KIADÓ, BUDAPEST
1968

ACTA TECHN. HUNG.

ACTA TECHNICA

A MAGYAR TUDOMÁNYOS AKADÉMIA MŰSZAKI TUDOMÁNYOK OSZTÁLYÁNAK KÖZLEMÉNYEI

SZERKESZTŐSÉG ÉS KIADÓHIVATAL: BUDAPEST V., ALKOTMÁNY U. 21.

Az Acta Technica német, angol, francia 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
Budapest V., Nádor u. 7.

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Az Acta Technica előfizetési ára kötetenként belföldre 120 forint, külföldre 165 forint. Megrendelhető a belföld számára az „Akadémiai Kiadó"-nál (Budapest V., Alkotmány utca 21. Bankszámla 05-915-111-46), a külföld számára pedig a „Kultúra" Könyv- és Hírlap Külkereskedelmi Vállalatnál (Budapest I., Fő utca 32. Bankszámla: 43-790-057-181) vagy annak külföldi képviseleteinél és bizományosainál.

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

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

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

Acta Technica
Nádor u. 7.
Budapest V.
Ungarn

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

Abonnementspreis pro Band: 165 Forint. Bestellbar bei dem Buch- und Zeitungs-Außenhandels-Unternehmen »Kultúra« (Budapest I., Fő utca 32. Bankkonto Nr. 43-790-057-181) oder bei seinen Auslandsvertretungen und Kommissionären.

ACTA TECHNICA

Tomus 61

INDEX

Csáki, F.—Gyürki, J.: Some Questions of the System Identification by Stochastic Signals — Einige Probleme der Systemidentifikation mit stochastischen Signalen — Quelques problèmes de l'identification d'un système par signaux stochastiques — *Чаки, Ф.—Дьёрки, Й.:* Некоторые проблемы идентификации системы, выполненной при помощи стохастических сигналов 227

Gádor, L.: Circuits with Impedances Having Z-Characteristics — Aus Impedanzen mit Z-Kennlinien aufgebaute Stromkreise — Circuits avec impédances à caractéristiques Z — *Гадор, Л.:* Цепи с импедансами имеющими Z-характеристику ... 155

Gádor, L.: Quelques points de vue pour le réglage de la tension du réseau — Some Points of View for the Regulation of the Mains Voltage — Einige Gesichtspunkte für die Regelung der Netzspannung — *Гадор, Л.:* Некоторые принципы регулирования сетевого напряжения 177

Gertler, J.: Some Concepts on the Programming of Direct Digital Process Control Systems — Einige Probleme der Programmierung unmittelbarer Digital-Prozess Steuersysteme — Quelques problèmes de la programmation des systèmes de commande digitale directe des processus — *Гертлер, Я.:* Некоторые вопросы программирования систем непосредственного цифрового управления процессами ... 55

Хатвани, Й.—Кардош, П.—Барат, Я.: Применение цифровых аналогов в устройствах промышленной автоматики — *Hatvany, J.—Kardos, P.—Barát, J.:* The Use of Counting Technique Instruments in Industrial Automatic Equipments — Anwendung zähltechnischer Mittel bei industriellen Automatik-Anlagen — L'emploi des techniques opérationnelles dans les systèmes d'automatisation industrielle... 31

Horváth, F.: Tooth Root Thickness between Trochoidal Root Profiles on Involute-Shaped Gears — Zahnwurzeldicke einer Evolventenverzahnung an der von Trochoiden begrenzten Zahnwurzel — L'épaisseur de base des dentures à développante, limitée par des trochoïdes — *Хорват, Ф.:* Толщина корня зуба, ограниченного трохоидами, в случае эвольвентного зацепления 439

Horváth, Z.: The Determination of Equilibrium and Kinetics in the Processes Occurring in Oxidation-Smelting — Kinetische und Gleichgewichtsuntersuchung der Vorgänge beim oxydierenden Schmelzen — Examen de l'équilibre et de la cinétique des phénomènes accompagnant la fusion oxydante — *Хорват, З.:* Равновесные и кинетические исследования процессов при окислительной плавке 413

Huszár, I.: Berechnung des Kräftespiels von Schwalbenschwanzkommutatoren — Calculation of the System of Forces of Arch-Bound Commutators — Calcul du jeu de forces des collecteurs à queue d'aronde — *Хусар, И.:* Расчет работы коллекторов в виде ласточкина хвоста 1

Kemény, A. P.: Effect of Structural Irregularities on Avalanche Breakdown in $p - n$ Step Junctions — Der Einfluß der strukturellen Irregularitäten von $p - n$ Übergängen auf die volumetrische Lawinendurchbruchspannung — L'influence des irrégularités structurales des jonctions $p - n$ abruptes sur la tension de rupture en avalanche spatiale — *Кемень, А. П.:* Воздействие нерегулярности структуры абруптных $p - n$ переходов на объемное напряжение лавинного пробоя 101

Korach, M.—Fülöp, J.: Un effet thermique des fours-tunnel — A Heat Effect in Continuous Furnaces — Ein wärmetechnischer Effekt der Tunnelöfen — *Корах, М.—Фюлёп, Я.:* Один теплотехнический эффект туннельных печей 137

Линковский, Ж. Б.: Средняя интенсивность отказов на заданном интервале времени работы аппаратуры — *Linkovski-Conde, J.:* The Mean Intensity of Instrument Breakdown in a Given Time Interval — Durchschnittliche Intensität des Schadhaftwerdens von Instrumenten in einem gegebenen Zeitintervall — L'intensité moyenne des pannes des instruments dans un intervalle de temps donné..... 67

Линковский, Ж. Б.: Долговечность и теория надежности — *Linkovski-Conde, J.:* Life and Theory of Reliability — Lebensdauer und Theorie der Verlässlichkeit — Durée des service et théorie de la sécurité de fonctionnement..... 219

Lipka, I.: Geometrie der Abwälzfräser mit Grundschnecken evolventenförmiger Ausführung — The Geometry of Hobs with Involute Basic Worm — Géométrie des fraises vis-mère à vis de base à développante — *Липка, И.:* Геометрия обкаточных червячных фрез с эвольвентным червяком 261

Lipka, I.: Geometrie der Schneckenwälzfräser mit Grundschnecken schleifenevolventischer Ausführung — The Geometry of Generating Cutters with Convolute Basic Worm — Géométrie des fraises vis-mère à vis de base convolutée — *Липка, И.:* Геометрия обкаточных фрез с конволютным червяком 371

Lőrinczy, A.—Németh-Sallay, M.—Szép, I. C.—Tihanyi, J.: The Effect of Quinone on the Surface Properties of Germanium — Einfluß von Chinon auf die Oberflächeneigenschaften von Germanium — L'influence de la quinone sur les qualités de surface du germanium — *Лёринци, А.—Немет-Шаллаи, М.—Сеп, И. Ц.—Тихани, Й.:* Воздействие хинона на поверхностные свойства германия 189

Pranab Kumar De: A Theoretical Approach to the Determination of the Amount of Wear during Sliding between Surfaces — Theoretische Bestimmung des Betrages des Abriebs zwischen gleitenden Oberflächen — Détermination théorique approchée de l'usure entre surfaces de frottement — *Пранаб Кумар Де.:* Теоретический приближенный метод определения степени износа при трении между поверхностями 429

Шомло, Я.: Исследование устойчивости цифрового регулятора — *Somló, J.:* Investigation of the Stability of a Digital Controller — Untersuchung der Stabilität eines digitalen Reglers — Examen de la stabilité d'un régulateur digital... 75

Szőke, B.: The Proper Conduction of the Flux of Mechanical Energy with Special Regard to the Self-Balanced Movable Clutch — Zweckmäßige Leitung des Energieflusses; die bewegliche, selbstausgleichende Wellenkupplung — Conduite efficace de flux d'énergie et accouplement auto-équilibrant mobile — *Сёке, Б.:* Целесообразное проведение потока энергии и самоуравновешивающаяся подвижная муфта сцепления 277

Tevan, Gy.: Calculation of the Coefficient of Self-Induction and of the Impedance of Cylindrical Coils with Open Iron Core — Berechnung des Selbstinduktionskoeffizienten bzw. der Impedanz von zylindrischen Spulen mit offenem Eisenkern — Calcul du coefficient des self-induction et de l'impédance des bobines cylindriques à noyau de fer ouvert — *Теван, Дь.:* Расчет коэффициента самоиндукции или же импеданса цилиндрических обмоток с незамкнутым сердечником 201

Tóth, L.: Distribution of Stresses Acting on the Rolle Surface in Strip Rolling — Verteilung der Druckspannung an der Walzenoberfläche beim Auswalzen von dünnem Stahlblech — Répartition des pressions à la surface d'un rouleau au laminage des feuillards *Тот, Л.:* Распределение напряжений, действующих на поверхность валков, при прокатке лент 351

Verő, J.: The Yield Strength in Homogeneous Alloys of α-Iron — Einfluß der gelösten Legierungselemente auf die Fließgrenze des α-Eisens — L'influence des éléments d'alliage dissous sur la limite d'écoulement du fer α — *Верэ, Й.:* Воздействие растворенных легирующих элементов на условный предел текучести α-железа 243

Wawra, H.: Zur Frage der Konzentrationsabhängigkeit der Poisson-Zahl — The Problem of Poisson's Ratio as a Function of Concentration — Le problème de la constante de Poisson en fonction de la concentration — *Вавра, Х.:* К вопросу зависимости от концентрации числа Пуассона 401

Winter, E.—Bitó, J.: The Emission Dependence of the Length of the Cathode Space and of the Positive Column — Die Emissionsabhängigkeit der Länge der kathodischen Räume und der positiven Säule — La dépendance de l'émission des espaces cathodiques et de la longueur de la colonne positive — *Винтер, Э.—Бито, Я.:* Эмиссионная зависимость длин катодных пространств и положительных столбцов 239

ACTA TECHNICA

ACADEMIAE SCIENTIARUM HUNGARICAE

ADIUVANTIBUS
I. BARTA, E. BÖLCSKEI, P. O. GESZTI, A. LÉVAI

REDIGIT
M. MAJOR

TOMUS 61

AKADÉMIAI KIADÓ, BUDAPEST
1968

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 1—30 (1968)

BERECHNUNG DES KRÄFTESPIELS VON SCHWALBENSCHWANZKOMMUTATOREN

I. HUSZÁR

KANDIDAT DER TECHNISCHEN WISSENSCHAFTEN
GANZ ELEKTROTECHNISCHE WERKE, BUDAPEST

[Eingegangen am 18. Februar, 1967]

Das mechanische Modell des Schwalbenschwanzkommutators kann durch zwei eindimensionale Federsysteme dargestellt werden. Bei dem einen werden die radialen Kräfte und Deformationen, bei dem anderen die axialen berücksichtigt. Die zur Bestimmung der Federkonstanten der einzelnen Federelemente dienenden Beziehungen werden in der Arbeit angegeben. Durch Klärung der Berechnungsmethode des Federsystems können die den verschiedenen, beim Kommutatorbetrieb vorkommenden Zuständen (Montage, Rotation, Erwärmung, rotierender und warmer Zustand) entsprechenden Kräfte bestimmt werden. Die auf diese Weise bestimmten Beanspruchungen ergeben die zur Dimensionierung der Kommutatorkonstruktion erforderlichen Daten.

Bezeichnungen

Kräfte

B Gewölbekraft;
C Zentrifugalkraft;
N Normalkraft, einseitig;
p Gewölbedruck;
P Resultierende des Gewölbedruckes;
q Radialdruck;
Q Radialkraft;
R Radialkraft, einseitig;
T Axialkraft.

Verschiebungen

x, ξ Axialverschiebung;
X Axialüberdeckung;
y, η Radialverschiebung;
Y Radialüberdeckung.

Sonstige Charakteristiken

$a, b, d, D, e, h, k, l, L, r, s$ geometrische Maße;
K Lamellenzahl;
α Neigungswinkel des Schwalbenschwanzes;
φ Richtungswinkel;
χ Zentriwinkel der Lamellenhälfte.

Indizes für Teile

g Preßring;
l Segmentisolierung;
m Isolierring, Kappe;
s Lamelle.

Indizes für Zustände

0 Montage;
Δt Erwärmung;
c Rotation;
ϑ stehender warmer Zustand;
ω rotierender kalter Zustand;
$\omega\vartheta$ Betriebszustand.

I. Bedingungen der Berechnung

Bild 1 zeigt den Aufbau des Schwalbenschwanzkommutators. Zwischen den aus Kupfer gefertigten Lamellen (s) befinden sich Mikanit-Segmentisolierungen (l). Auf den an den beiden Seiten der Lamelle befindlichen geneigten, d.h. auf den gesamten Lamellenkörper bezogen, konischen Flächen sind die isolierenden Mika-Kappen (m) angebracht. An den beiden Seiten ist dann noch je ein Preßring (g) vorgesehen. Werden diese mit den Preßschrauben (a) zusammengezogen, kann eine entsprechende Druckkraft auf den Lamellenkörper ausgeübt werden [1—3].

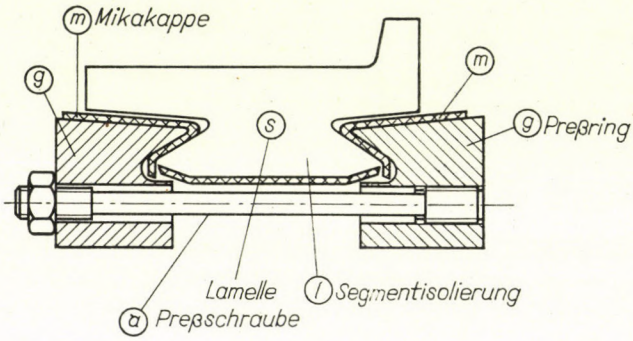


Bild 1. Schwalbenschwanzkommutator

Das Zusammendrücken der Preßringe kann auch dadurch erreicht werden, daß diese auf eine rohrförmige Nabe montiert werden. An einem Ende derselben stützt eine Schulter den Preßring (diese kann aus einem Stück gefertigt oder auch angeschweißt sein), am anderen Ende wird dieser mit einer Mutter angespannt. Der Kommutator kann auch so gebaut sein, daß zwischen den Preßring und die Stützfläche der Nabe noch eine besondere Feder, mit nahezu rechteckigem Querschnitt eingeschaltet wird, die mit der einen Seite auf dem äußeren, und mit der anderen auf dem inneren Umfang aufliegt und so die nötige weiche Federung gewährleistet. Auf alle diese Lösungen bezieht sich unsere Ableitung, unsere Feststellungen in Bezug auf die Schrauben müssen natürlich sinngemäß angewandt werden.

Die Abstützung des fertigen Kommutators auf der Nabe darf das Kräftespiel des Kommutators nicht beeinflussen. Das heißt, sie soll gegenüber Belastungen in Radialrichtung entsprechend weich sein und soll die axialen Verschiebungen nicht hindern. Im folgenden wird vorausgesetzt, daß die Fassung des Kommutators ideal ist, d. h. von der Wirkung der Nabe wird abgesehen.

Der Querschnitt des Kommutators ist symmetrisch. Dies gilt im großen und ganzen auch für den Längsschnitt, wenn wir das mit Fahnen versehene

Ende und die etwas abweichende Form der Preßringe, z. B. die zur Auswuchtung gefertigte Nut, nicht in Betracht ziehen. Im weiteren setzen wir also voraus, daß sowohl der Längs-, als auch der Querschnitt des Kommutators symmetrisch seien.

Wir rechnen auch nicht mit den Ungenauigkeiten der Herstellung und setzen für den gesamten Kommutator eine vollkommene Geometrie voraus.

Bei den Berechnungen wird angenommen, daß die einzelnen Werkstoffe — Kupfer, Stahl, Mikanit — ideal homogen, isotrop und linear elastisch bzw. an der Fließgrenze ideal plastisch sind.

Die Eigengewichte der einzelnen Elemente werden gegenüber den auftretenden Kraftwirkungen vernachläßigt.

Eine sehr wesentliche Bedingung ist, daß die Schwalbenschwanzflächen als glatt betrachtet werden, d. h. bei dem Preßring und der Kappe bzw. bei dem konischen Teil des Lamellenkörpers rechnen wir mit einer reibungslosen Verbindung. Dementsprechend bilden die hier auftretenden Kräfte ein zur Oberfläche senkrechtes, verteiltes System. In Wirklichkeit tritt auf den Berührungsflächen Reibung auf. In Anbetracht jedoch der Unsicherheit in der Berechnung der Reibungskraft, ist es zweckmäßig, diese in erster Annäherung zu vernachläßigen. Die so für ideale Verhältnisse gewonnenen Beziehungen können aber für Reibungsgrenzfälle verhältnismäßig leicht abgeändert werden, wie darüber im Rahmen einer späteren Arbeit berichtet werden soll.

Im Sinne des vorangehenden kann angenommen werden, daß die Symmetrie auch bei den einzelnen Berührungen vollkommen ist. Infolgedessen sind die auftretenden Oberflächendrücke ideal kreissymmetrisch, bzw. werden sie in Längsrichtung als in gleicher Weise verteilt betrachtet. Wir betrachten also die Konstruktion auch vom Gesichtspunkt der Kräfte gesehen als symmetrisch.

II. Kräfteverhältnisse im idealen Kommutator

Die Tatsache, daß die Kräfte in Längs- und auch in Querrichtung als symmetrisch angesehen werden können, bietet eine bedeutende Vereinfachung. Als weitere Näherung kann auch noch vorausgesetzt werden, daß die verteilte Belastung auf den Berührungsflächen — zwischen den Lamellen, auf der Auflagefläche des Preßrings — gleichförmig verteilt ist.

Laut obigen Voraussetzungen kann man den Kräfteplan der idealen Kommutatorkonstruktion aufzeichnen, und zwar können die einzelnen Elemente gesondert dargestellt werden (Bild 2).

Auf eine einzige Lamelle — die Hälfte der beiderseitig befindlichen Segmentisolierung inbegriffen — wirkt das über die Ringe und durch die Kappe übertragene normale und über den gesamten Umfang als gleichmäßig

verteilt angenommene Kräftesystem q, dessen Resultierende für eine Lamelle mit N_s bezeichnet wird:

$$N_s = l_m(s'_s + s'_l)\, q, \tag{1}$$

wobei $s'_s + s'_l$ die mittlere, im Bogen gemessene Breite der Lamelle an der Stelle der Lastübertragung ist. Dabei ist die radiale Komponente aufgrund des Vektordiagramms

$$R_s = N_s \cos\alpha, \tag{2}$$

die axiale Komponente ist hingegen

$$T_s = N_s \sin\alpha = R_s \tan\alpha, \tag{3}$$

wo α der Neigungswinkel des Schwalbenschwanzes ist. Die Resultierende der an beiden Seiten wirkenden Kräfte N_s ist

$$Q_s = 2\, R_s. \tag{4}$$

Auf den Seitenflächen mit der Oberfläche A_s tritt der als gleichmäßig verteilt angenommene Gewölbedruck p mit der Resultierenden P_s auf:

$$P_s = A_s p. \tag{5}$$

Die Resultierende der an den beiden Seiten auftretenden Kräfte P_s, d.h. die Gewölbekraft, kann aus dem Vektordiagramm berechnet werden:

$$B_s = 2\, P_s \sin\chi. \tag{6}$$

Schließlich wirkt auch die verteilte Zentrifugalkraft, deren Resultierende für eine Lamelle C_s ist, die Zentrifugalkraft der beiden halben Segmentisolierung inbegriffen. Aus dem das Gleichgewicht der gesamten Lamelle darstellenden Vektordiagramm kann man anschreiben:

$$Q_s - C_s = B_s. \tag{7}$$

Die Belastung und das Gleichgewicht der Segmentisolierung sind aus dem Bild ersichtlich (mit Gewölbedruck p).

Bild 2 zeigt auch das Kräftespiel des Lamellenkörpers. Äußere Belastung bilden jetzt das auf die einzelnen Lamellen wirkende Kräftesystem N_s sowie die Zentrifugalkräfte C_s. An je einer Seite wird die Resultierende der Kräfte N_s die Kraft

$$T = KT_s = R \tan\alpha \tag{8}$$

sein, ausgedrückt durch die Lamellenzahl K. Gleichzeitig bilden diese für den gesamten Lamellenkörper ein Gleichgewichtssystem, nachdem die Resultie-

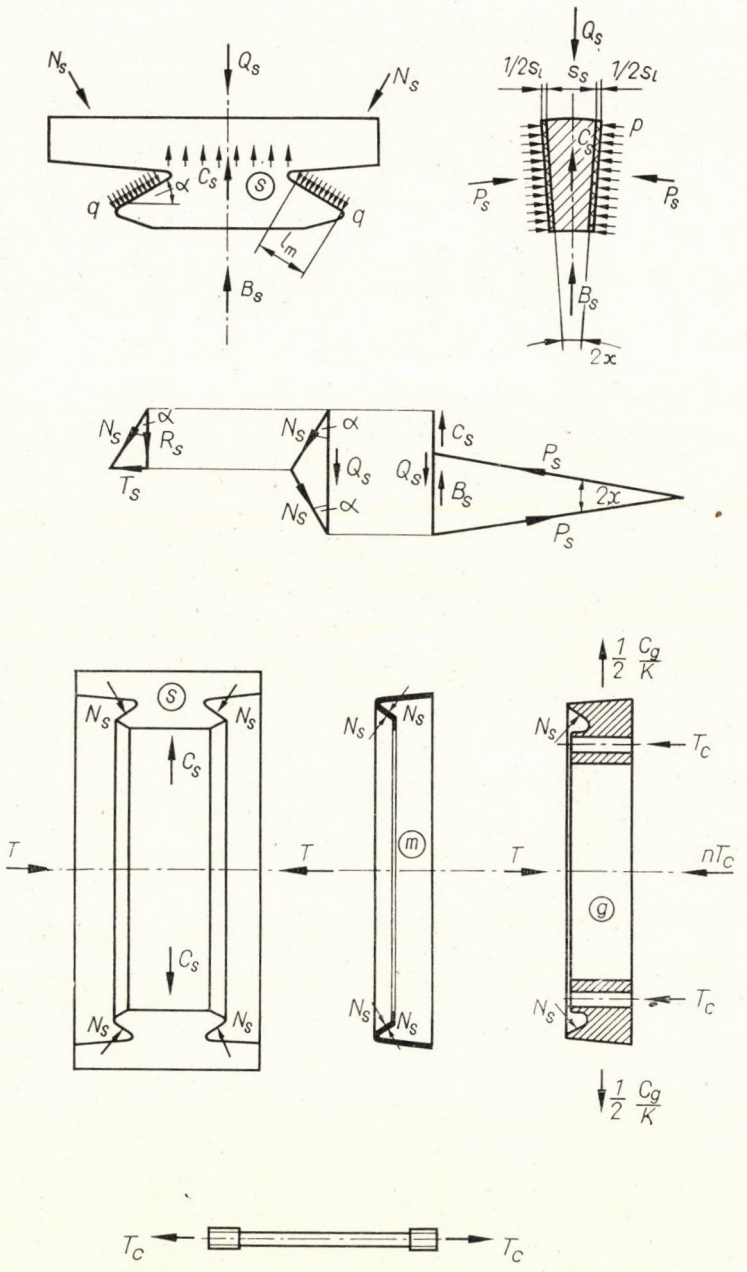


Bild 2. Kräfteverhältnisse der Kommutatorelemente

renden der an den beiden Seiten wirkenden Kräfte sich von einander bloß durch das Vorzeichen unterscheiden. Im Interesse der weiteren Ausführungen wird die algebraische Summe der Radialkomponente eingeführt und für den

gesamten Kommutatorkörper wird:

$$R = KR_s = Q/2\,. \tag{9}$$

Zu bemerken ist, daß sich dieses Radialkräftesystem im Gleichgewicht befindet. Vollständigkeitshalber wird auch die numerische Summe der Kräfte N_s definiert:

$$N = KN_s. \tag{10}$$

Ähnlich ist die numerische Summe des übrigens im Gleichgewicht befindlichen Kräftesystems Q_s

$$Q = KQ_s. \tag{11}$$

Auch das System der Zentrifugalkräfte befindet sich im Gleichgewicht. Bezüglich seiner numerischen Summe können wir schreiben:

$$C = KC_s. \tag{12}$$

Auch für die Gewölbekräfte wird die algebraische Summe eingeführt:

$$B = KB_s. \tag{13}$$

Schließlich können wir aufgrund der Ausdrücke (7), (11), (12) und (13) schreiben:

$$B = Q - C, \tag{14}$$

was als Gleichgewichtsgleichung des Lamellenkörpers betrachtet wird.

Das Bild stellt auch die Belastung der Kappe dar. In Anbetracht der kleinen Masse wird jedoch die Zentrifugalkraft vernachläßigt. Wie ersichtlich, ist die Belastung wieder ein Gleichgewichtssystem mit der Normalkraft N_s je Lamelle.

Die Belastung des Preßringes ist einesteils das System der je Lamelle übertragenen Kräfte N_s, deren Resultierende die Axialkraft T ist, anderenteils die eigene Fliehkraft, die an sich ein Gleichgewichtssystem bildet. Auf beide Ringe berechnet ist die gesamte Fliehkraft numerisch C_g, von der im Bild der auf eine Lamelle entfallende Wert $0{,}5\,C_g/K$ eingetragen ist. Bei den Berechnungen ist in C auch die Fliehkraft der Preßschrauben inbegriffen.

Schließlich tritt an den Schraubenstellen auch die Axialkraft T_c auf, deren Resultierende im Fall von n Preßschrauben

$$T = nT_c \tag{15}$$

ist. Dementsprechend wird das Gleichgewicht des Preßringes durch die Kräfte T und nT_c ausgedrückt.

Zuletzt wurde im Bild die Preßschraube dargestellt, die sich unter Wirkung der Zugkraft T_c im Gleichgewicht befindet. Die bei der Rotation in Berechnung zu ziehende Fliehkraft wird bei der Bestimmung des Kräftespiels mit dem Preßring zusammen in Betracht gezogen, es wird also vorausgesetzt, daß die Schrauben auf Zug beansprucht sind; bei der Kontrolle auf Festigkeit wird aber mit dieser als mit einer verteilten Kraft, bei der Biegebeanspruchung gerechnet.

III. Ersatz-Federsystem

Der Kommutator ist eine statisch unbestimmte Konstruktion, d. h. die Berechnung des Kräftespiels kann nur unter Berücksichtigung der Deformationen durchgeführt werden. Um einen besseren Überblick zu gewinnen, ist es zweckmäßig, den Kommutator durch ein einfacheres mechanisches Modell zu ersetzen, das eben vom Gesichtspunkt der Kräfte und der Deformationen aus, mit der ursprünglichen Konstruktion gleichwertig ist.

Da laut unseren Voraussetzungen der Kommutator sowohl vom Gesichtspunkt der Kräfte als auch von dem des Aufbaus aus vollkommen kreissymmetrisch ist, kann offenbar dasselbe auch von den Radialdeformationen gesagt werden. Die Festigkeitsverhältnisse sind also in einem beliebigen Querschnitt, entlang eines Radius mit irgendeinem Neigungswinkel identisch. In diesem Sinne kann das gesamte System durch einen Meridianschnitt ersetzt werden, in obigem Sinne sind nämlich die Festigkeitsverhältnisse der Meridianschnitte von der Orientierung der Meridianebene unabhängig.

Aufgrund des obigen Gedankenganges kann der Kommutator durch ein zweidimensionales mechanisches Modell ersetzt werden.

Aus dem auch in Axialrichtung symmetrischen Aufbau folgt gemäß dem vorangehenden auch gleich, daß es zweckmäßig ist, als Mittelpunkt des Modells den Origo der Konstruktion zu wählen.

Bei der Darstellung des Ersatzsystems wurde zum Ausdruck gebracht, daß die Verhältnisse einesteils in Radial-, anderenteils in Axialrichtung zu verfolgen sind. Gleichzeitig können sie unter Anwendung des Superpositionsprinzips getrennt behandelt werden. Demnach besteht das Modell aus zwei separaten Federsystemen: eines zur Beschreibung der radialen, das andere zur Beschreibung der axialen Festigkeitsverhältnisse.

Die Bestimmung des Federsystems kann verhältnismäßig einfach erfolgen. Eine grundlegende Bedingung ist, daß die Deformation der einzelnen konstruktiven Elemente und die der entsprechenden Federelemente bei identischen Verhältnissen übereinstimmen. Bild 3 stellt das Modell dar.

Im radialen Federsystem gilt der Lamellenkörper — unter Berücksichtigung des Zusammenbaus — als gedrückte Feder, die wir im Interesse des später folgenden gesondert als aus den Schwalbenschwanzteilen und aus dem Mittelteil zusammengesetzt betrachten (Bild 3a). Im Lamellenkörper sind

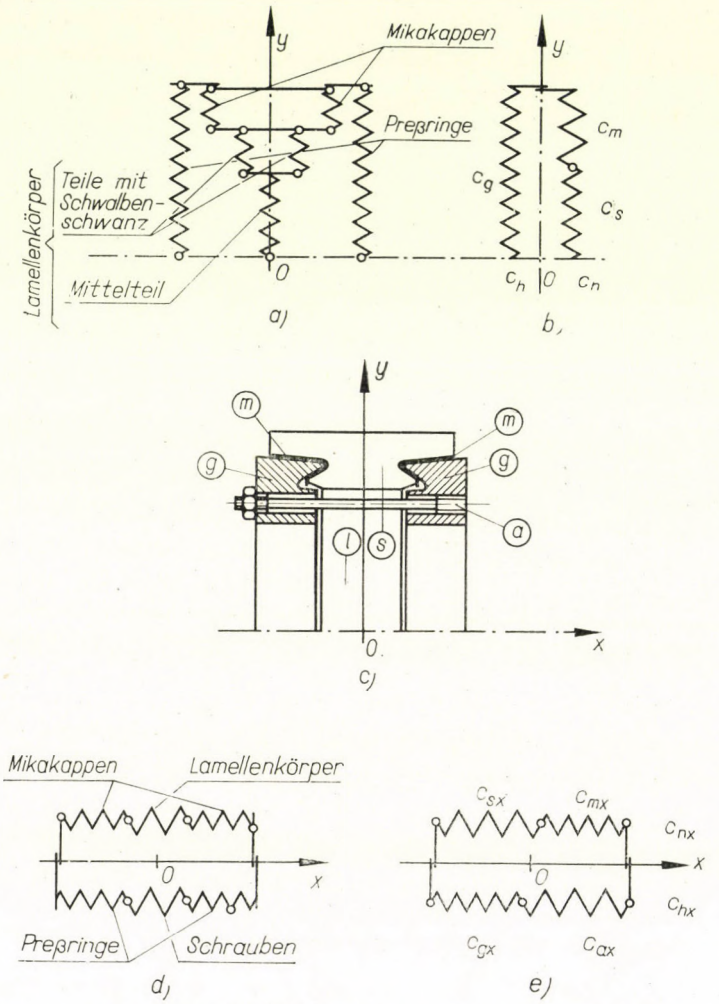


Bild 3. Mechanisches Modell des Schwalbenschwanzkommutators

natürlich auch die auf Druck beanspruchten Segmentisolierungen zwischen den Lamellen mit inbegriffen. Gleicherweise sind die beiden Mika-Kappen gedrückte Federn. Die beiden Preßringe sind auf Zug beansprucht. Radial nimmt das Schraubensystem nicht am Kräftespiel teil.

Bild 3b zeigt die einzelnen Federn miteinander verbunden bzw. durch ihre Resultierenden ersetzt, wo also die gemeinsame Federkonstante der

beiden Ringe c_g, die des gesamten Lamellenkörpers c_s und die der Mika-Kappe c_m ist.

Im axialen Federsystem gelten der Lamellenkörper und die Mika-Kappen als gedrückt. Unter Zug stehen die beiden Preßringe und die Schrauben (Bild 3d). Im Bild 3c sind wieder die Resultierenden eingezeichnet. Folglich ist c_{sx} die gemeinsame Federkonstante des gesamten Lamellenkörpers, c_{mx} die der beiden Mika-Kappen, c_{gx} die resultierende Federkonstante der beiden Ringe, c_{ax} die der sämtlichen Schrauben.

Mit Hilfe einer weiteren Vereinfachung wird es in beiden Federsystemen einfach mit resultierenden gezogenen und gedrückten Federn c_h und c_n sowie mit c_{hx} und c_{nx} gerechnet.

Für die Berechnung der Federkonstanten muß die Bezugsgrundlage vereinbart werden. Diese sei in Radialrichtung die Kraft Q und in Axialrichtung die Kraft T, wie diese die Ausdrücke (11) und (8) angeben. Dementsprechend sind die Federkonstanten durch die Ausdrücke

$$y = cQ \quad \text{und} \quad x = c_x T \tag{16}$$

definiert; y bedeutet die dem Radius entsprechende Maßänderung, x die gesamte radiale Längsänderug.

Die Zusammenhänge tragen der linearen Elastizität der einzelnen Teile Rechnung, d.h. der lineare Zusammenhang zwischen Kraft und Deformation wird als Grundlage angesehen.

IV. Berechnung des Federsystems

Die beiden Federmodelle sind voneinander nicht unabhängig. Die Berührungsflächen ergeben den Zusammenhang. Genauer gesagt, die Bedingung für die auf dem Lamellenkörper und auf der konischen Auflagefläche der Preßringe entstehenden Kräfte, nämlich die Reibungslosigkeit, bestimmt den Zusammenhang zwischen den beiden Systemen.

Eben deshalb werden zuerst die zufolge des Zusammenhanges auftretenden Deformationen eingehend untersucht, nachher werden die Federsysteme aufgrund der gewonnenen Resultate rechentechnisch getrennt. Das radiale und das axiale Federschema werden getrennt berechnet, auf Grund des Superpositionsprinzips wird die Resultierende dargestellt, das heißt, die auf den gesamten Kommutator bezogene Lösung.

Unsere Überlegungen werden für die einzelnen charakteristischen Zustände des Kommutators getrennt durchgeführt, dann werden die wichtigeren Betriebszustände verglichen.

V. Montage

Bild 4 stellt die Verbindung zwischen dem Preßring und dem Lamellenkörper im Einzelnen dar. Durch eine starke ununterbrochene Linie ist die zusammengebaute Lage dargestellt, das Berührungspunktpaar A und B ist eingezeichnet.

Wenn wir uns die Verbindung aufgehoben denken, schrumpfen die Ringe einerseits in Radialrichtung zusammen, andererseits nähern sie sich einander axial. Hört nämlich das bei der Berührung auftretende Kräftesystem N_s, d. h. die radiale Spannkraft R_s und die axiale Kraft T_s auf (Bild 2), nimmt das Federsystem seine dem unbelasteten Zustand entsprechende Form an. Ähnliches gilt auch für den Lamellenkörper, natürlich muß hier mit einer Formänderung in entgegengesetztem Sinn gerechnet werden, d. h. mit einer Ausdehnung nach beiden Richtungen. In dem Bild ist die diesem getrennt gedachten unbelasteten Zustand entsprechende Lage mit einer gestrichelten Linie angedeutet. Hier wurde auch das in zusammengebautem Zustand markierte Berührungspunktpaar eingezeichnet, das jetzt in die Lagen A_0 und B_0 gelangt.

Die tatsächliche Montage wird natürlich nicht der getrennten Lage entsprechend ausgeführt. Erst werden die Preßringe mit dem Lamellenkörper in Berührung gebracht (siehe dünne Linie in Bild 4b). In dieser Lage vor der Montage entsteht in den Schrauben noch keine Kraft, die Muttern placieren sich auf den Bolzen derart, daß sie eben die Ringe berühren (Bild 1). Aus dem Zustand vor der Montage kann die getrennte Lage so hergestellt werden, daß die Muttern auf dem Bolzen in die Entfernung $\overline{B_* B_0}$ voneinander gebracht werden, wie das bei der Montage tatsächlich gemacht wird. Jedoch nicht die gesamte Strecke $\overline{B_* B_0}$ entspricht einer Kraftwirkung und zwar die Projektion von $\overline{B_* A_0}$ ist diejenige Entfernung, die in waagerechter Richtung einer kräftefreien Verschiebung entspricht.

In Radialrichtung besteht zwischen den Punkten A_0 und B_0 der Maßunterschied Y_0, wo gemäß dem Bild

$$Y_0 = r_y - r_\eta \tag{17}$$

ist.

Entsprechend der Endlage des Zusammenbaus verschiebt sich der Ring um η_0, der Lamellenkörper um die Entfernung y_0, deren vorzeichenrichtige Summe offenbar Y_0 ist, d. h

$$Y_0 = \eta_0 - y_0. \tag{18}$$

Hinsichtlich der bei der Montage auftretenden Radialkraft Q_0 ist für den Lamellenkörper gemäß Zusammenhang (16)

$$y_0 = -\varepsilon_n Q_0. \tag{19}$$

Bild 3b entsprechend wurde nämlich die Federkonstante des gesamten Lamellenkörpers mit c_n bezeichnet, andererseits bedeutet Q_0 die Druckkraft auf den Lamellenkörper. Gleichzeitig ist für den Ring

$$\eta_0 = c_h Q_0 , \tag{20}$$

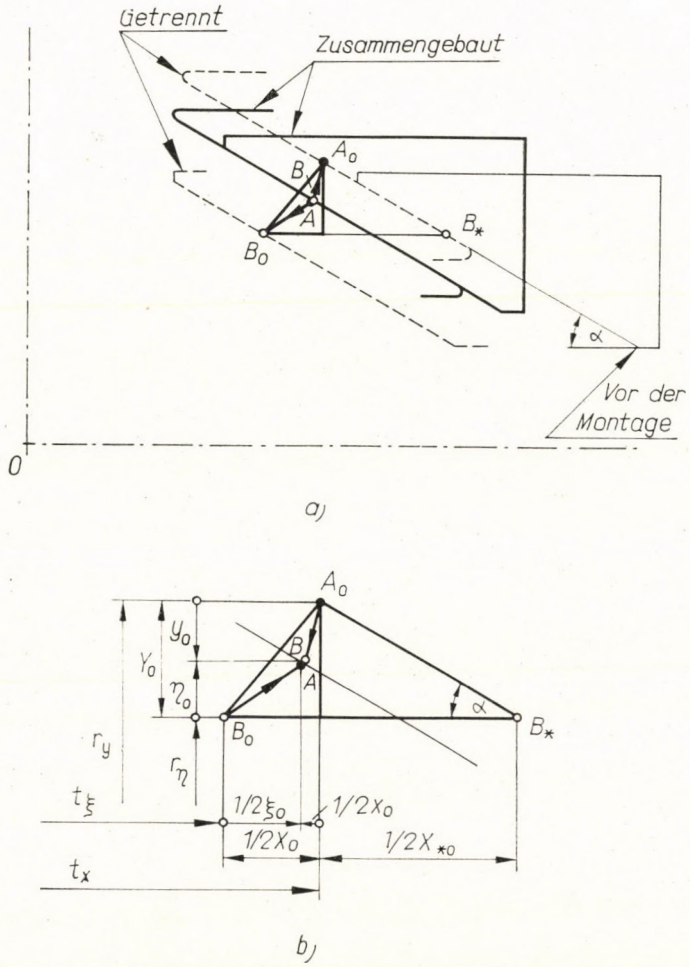


Bild 4. Verschiebungen bei der Montage

ebenfalls mit den Bezeichnungen von Bild 3b. Nachdem wir aber dort unter c_h die gemeinsame Federkonstante der beiden Ringe verstanden haben, muß natürlich die Summe Q_0 der auf beide Ringe separat wirkenden Kräfte $Q_0/2$ eingesetzt werden. Wenn schließlich für die resultierende radiale Federkonstante die Bezeichnung

$$c = c_h + c_n \tag{21}$$

eingeführt wird, erhalten wir auf Grund von (18), (19), (20) und (21) den Zusammenhang

$$Y_0 = cQ_0. \tag{22}$$

Dieser gibt in Kenntnis des geometrischen Maßunterschieds Y_0 die bei der Montage auftretende Radialkraft Q_0 an.

In Axialrichtung ist $1/2\, X_0$ der Maßunterschied zwischen den Punkten A_0 und B_0. Für beide Ringe zusammen also

$$X_0 = t_x - t_\xi. \tag{23}$$

Die Vorschiebung der Ringe von einander ist am Ende der Montage ξ_0, die Zusammendrückung des Lamellenkörpers ist x_0, d. h.

$$X_0 = \xi_0 - x_0. \tag{24}$$

Für die bei der Motage auftretende Axialkraft T_0 gilt unter Anwendung des Ausdruckes (16) für den Lamellenkörper

$$x_0 = c_{nx} T_0, \tag{25}$$

da gemäß Bild 3a die Federkonstante für den gesamten Lamellenkörper zu c_{nx} angenommen wurde, andererseits wirkt axial auf den Lamellenkörper beim Zusammenbau die Kraft T_0. Für die Ringe ist ähnlich

$$\xi_0 = c_{hx} T_0, \tag{26}$$

ebenfalls mit den Bezeichnungen von Bild 3b. Nach Einführung der resultierenden axialen Federkonstante

$$c_x = c_{nx} + c_{hx}, \tag{27}$$

ist mit Benützung der Beziehungen (24), (25), (26) und (27)

$$X_0 = c_x T_0, \tag{28}$$

was in Kenntnis des Maßunterschiedes X_0 die axiale Montagekraft T_0 ergibt.

Die Verbindung zwischen den radialen und axialen Federsystemen kann mit Hilfe der an den Berührungsstellen auftretenden Kräfte hergestellt werden. Die Kräfte sind in Bild 2 dargestellt. Da die Reibung zu Null angenommen wurde, muß nur mit den Normalkräften N_{s0} gerechnet werden, für deren beide Komponenten laut dem Bilde bzw. gemäß den Ausdrücken (3) und (4)

$$\frac{T_{s0}}{R_{s0}} = \tan\alpha = \frac{T_{s0}}{Q_{s0}/2},$$

oder unter Anwendung von (8) und (11)

$$\tan \alpha = \frac{2\, T_0}{Q_0}, \tag{29}$$

wobei die Indizes für die Montage angewandt worden sind.

Zufolge der Verbindung genügt es nur das eine Federsystem zu berechnen, in der anderen Richtung ergibt sich nämlich die Kraft aufgrund von (29) sofort.

Ebenso können nach Klärung des einen Systems auch die in der anderen Richtung vor sich gehenden Deformationen bestimmt werden; indem in die Zusammenhänge (29) die Ausdrücke (22) und (28) eingesetzt werden, erhalten wir den Zusammenhang:

$$\tan \alpha = 2 \frac{X_0}{Y_0} \frac{c}{c_x}, \tag{30}$$

der zwischen den beiden Deformationen eine tatsächlich einfache Verbindung schafft.

Schließlich kann aufgrund von Bild 4b die mit X_* bezeichnete Entfernung berechnet werden, die also der kräftefreien Näherung der beiden Ringe entspricht, wie aus obigem hervorgeht:

$$X_* = 2 \frac{Y_0}{\tan \alpha}. \tag{31}$$

Wenn also das Anziehen der Schrauben vorgeschrieben werden soll, gehört von dem zur vollständigen Verschiebung der Mutter nötigen Wert $X_* + X_0$ nur zu X_0 eine elastische Spannkraft. $\overline{A_0 B_*}$ ist die Verschiebung, das Gleiten des Ringes und des Lamellenkörpers im Verhältnis zu einander, wofür

$$\overline{A_0 B_*} = \frac{Y_0}{\sin \alpha} = \frac{X_*}{2 \cos \alpha} \tag{32}$$

gilt.

Die Ergebnisse zusammenfassend kann hinsichtlich der praktischen Ausführung der Berechnungen gesagt werden, daß es genügt, das eine Federsystem zu untersuchen und dann erhält man über die Umrechnungsgleichungen unmittelbar die für die andere Richtung charakteristischen Werte.

Hinsichtlich der Gesichtspunkte der praktischen Berechnungen muß in Betracht gezogen werden, daß jetzt nicht mehr von eindeutig bestimmbaren Maßen r_y, r_η die Rede sein kann, d. h. aufgrund der geometrischen Daten steht der Überdeckungswert Y_0 nicht zur Verfügung. Im allgemeinen wird die Spannkraft T_0 vorgeschrieben, die zur Gewährleistung einer stabilen Kom-

mutatorkonstruktion nötig ist. In deren Kenntnis gewinnen wir aus (29) unmittelbar die Radialkraft Q_0 und aus (22) die Überdeckung Y_0. Diese letztere kann nun folgendermaßen gedeutet werden: wählen wir auf dem Ring einen beliebigen Punkt B_0 (Bild 4), zu dem das Maß r_η gehört. Die Lage des entsprechenden Lamellenkörperberührungspunktes A_0 wird gemäß dem Zusammenhang (17) durch r_y bestimmt. Die im Verlaufe der Montage entstehenden Deformationen aber können aus den Ausdrücken (19), (20) berechnet werden. Im obigen Sinne kann das behandelte Federsystem im weiteren als ein in den beliebig gewählten Punkten A_0 und B_0 zusammenhängendes System angesehen werden.

Die Deformationen bzw. die Kräfte können auf die im Schrifttum [4] erörterte Weise auch in einem Diagram dargestellt werden.

VI. Erwärmung

Gehen wir davon aus, daß es anfangs keine Vorspannung gibt, d. h., die Ringe liegen auf dem Lamellenkörper kräftefrei auf. Die Ringe erwärmen sich zufolge der Einwirkung der Strombelastung und der OberflächenReibungskräfte wenig, der Lamellenkörper aber bedeutend. Dementsprechend tritt eine größere Dehnung des Lamellenkörpers auf, die durch den Ring begrenzt wird, es treten also zwischen diesen Kräftewirkungen auf.

Zwecks Verfolgung der Verhältnisse sind im Bild 5a die aneinander angeschlossenen Elemente dargestellt, u. zwar die der Erwärmung entsprechende Lage mit einer starken Linie. Wenn wir uns die Konstruktion in warmem Zustand getrennt vorstellen, nimmt sie die mit gestrichelter Linie bezeichnete Lage ein, die als kräftefreie Lage gilt. Hörte die Erwärmung auf, würden die Berührungspunkte in die Lagen A_0 und B_0 gelangen.

Aufgrund von Bild 5b können die Berechnungen ausgeführt werden. A_0 und B_0 bezeichnen den kalten Ausgangszustand, der kräftefrei ist. Während der Erwärmung dehnt sich der Lamellenkörper und auch der Ring in Radialrichtung aus. Dementsprechend ist die radiale Verschiebung von A_0 aus der im Umfang eingetretenen Längenänderung Δk:

$$y' = \frac{1}{2\pi} \Delta k = \frac{1}{2\pi} K (\alpha_s s_s \Delta t_s + a_l s_l \Delta t_l), \tag{33}$$

wo α_s bzw. α_l der Wärmeausdehnungskoeffizient der Lamellen bzw. der Segmentisolierung, Δt_s bzw. Δt_l die entsprechenden Temperaturänderungen sind. Von der Dehnung der Mika-Kappe nehmen wir wegen deren relativ geringen Wertes Abstand. Die radiale Verschiebung des Punktes B_0 ist aus der Wärmedehnung des Ringes berechnet

$$\eta' = \alpha_g r_{B_0} \Delta t_g , \tag{34}$$

wo α_g der Wärmeausdehnungskoeffizient des Ringes, Δt_g die Temperaturänderung ist.

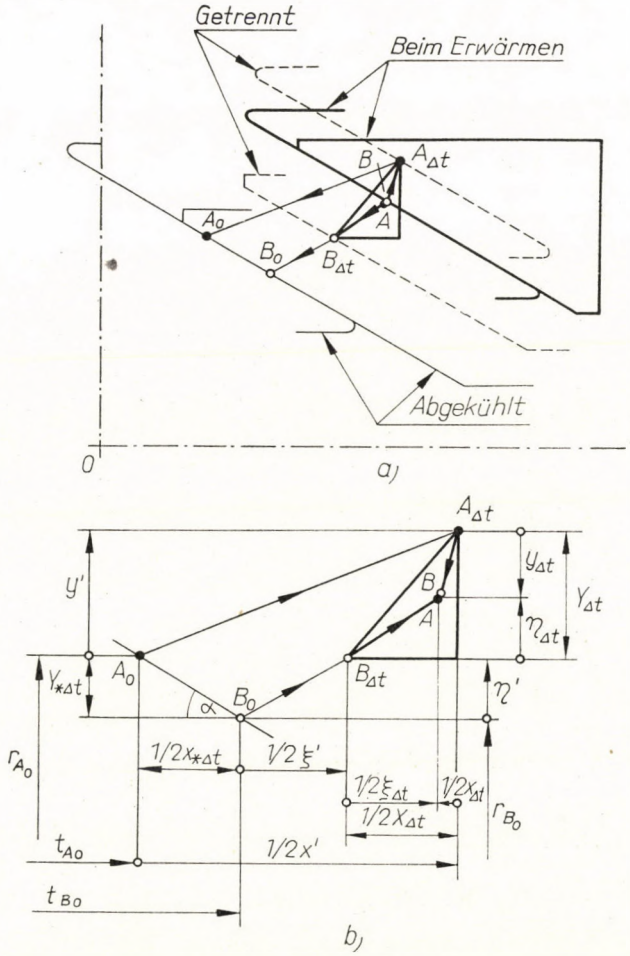


Bild 5. Verschiebungen bei der Erwärmung

Die Punkte $A_{\Delta t}$ und $B_{\Delta t}$ zeigen die Lage nach der Erwärmung, aber ohne Zusammenbau. Zwischen diesen besteht der Maßunterschied

$$Y_{\Delta t} = y' + Y_{*\Delta t} - \eta' . \tag{35}$$

Bei der Vereinigung verschiebt sich der Ring um den Wert $\eta_{\Delta t}$, der Lamellenkörper um $y_{\Delta t}$ in radialer Richtung. Hierfür sind A und B charakteristisch. Aufgrund des Bildes 5b kann geschrieben werden:

$$Y_{\Delta t} = \eta_{\Delta t} - y_{\Delta t}. \tag{36}$$

Mit der zur Erwärmung gehörenden Kraft $Q_{\Delta t}$ sind die einzelnen radialen Deformationen

$$y_{\Delta t} = -c_n Q_{\Delta t} \tag{37}$$

und

$$\eta_{\Delta t} = c_h Q_{\Delta t}. \tag{38}$$

Mit Verwendung des Ausdruckes (21), aufgrund von (36), (37) und (38) ist schließlich

$$Y_{\Delta t} = cQ_{\Delta t}. \tag{39}$$

Durch einen ähnlichen Gedankengang gelangen wir zur Bestimmung der axialen Kraft. Die axiale Dehnung des Lamellenkörpers ist

$$x' = \alpha_s t_{A_0} \Delta t_s. \tag{40}$$

Die Verschiebung der Punkte B_0 der Ringe wird durch die Ausdehnung der Schrauben bestimmt, für die

$$\xi' = \alpha_a t_{B_0} \Delta t_a \tag{41}$$

gilt, wo α_a der Wärmeausdehnungskoeffizient der Schrauben, Δt_a aber die Temperaturänderung der Schrauben ist.

Die Erwärmungslage wird jetzt durch

$$X_{\Delta t} = x' - (\xi' + X_{*\Delta t}) \tag{42}$$

gekennzeichnet.

Bei der Vereinigung werden die Verschiebungen $x_{\Delta t}$ und $\xi_{\Delta t}$, und durch dieselben ausgedrückt wird

$$X_{\Delta t} = \xi_{\Delta t} - x_{\Delta t}. \tag{43}$$

Die auf Einwirkung der Axialkraft $T_{\Delta t}$ entstehende Deformation ist hingegen

$$x_{\Delta t} = -c_{nx} T_{\Delta t} \tag{44}$$

und

$$\xi_{\Delta t} = c_{hx} T_{\Delta t}. \tag{45}$$

Unter Benützung der Gleichung (27), aufgrund der Zusammenhänge (43), (44) und (45) ist schließlich

$$X_{\Delta t} = c_* T_{\Delta t}. \tag{46}$$

Für die Beziehungen zwischen den radialen und axialen Kräften besteht auch jetzt ähnlich dem Zusammenhang (29)

$$\tan \alpha = \frac{2\, T_{\Delta t}}{Q_{\Delta t}} \tag{47}$$

bzw. ähnlich wie nach von (30)

$$\tan \alpha = 2\, \frac{X_{\Delta t}}{Y_{\Delta t}}\, \frac{c}{c_x}\,. \tag{48}$$

Aufgrund des Bildes ist ersichtlich, daß

$$Y_{*\,\Delta t} = 2\, \frac{Y_{*\Delta t}}{\tan \alpha} \tag{49}$$

ist. Gleichzeitig ist die relative Verschiebung beim Erwärmen:

$$\overline{A_0\, B_0} = \frac{Y_{*\Delta t}}{\sin \alpha} = \frac{X_{*\Delta t}}{2 \cos \alpha}\,. \tag{50}$$

Bei der praktischen Durchführung der Berechnung bleibt die Frage der Bestimmung von $Y_{*\Delta t}$ offen, nämlich nur in dessen Kenntnis kann $Y_{\Delta t}$ mit Hilfe des Zusammenhanges (35) berechnet werden.

Setzen wir in den Ausdruck (48) die Gleichungen (35) und (42) ein, ordnen wir ferner die Gleichung nach $Y_{*\Delta t}$:

$$Y_{*\Delta t} = 2\, \frac{c}{c_x}\, \frac{1}{\tan \alpha}\, (x' - \xi' - X_{*\Delta t}) - (y' - \eta').$$

Setzen wir aus Gleichung (49) $X_{*\Delta t}$ ein, so wird nach Ordnen

$$Y_{*\Delta t} = \frac{2\, c\, (x' - \xi') - c_x \tan \alpha\, (y' - \eta')}{4\, c + c_x \tan^2 \alpha}\, \tan \alpha\,, \tag{51}$$

womit die Aufgabe gelöst ist.

Die grafische Lösung ist nach dem Vorbild von [4] möglich.

VII. Rotation

Laut den Ausgangsbedingungen berühren sich der Lamellenkörper und die Preßringe vorspannungsfrei. Während der Rotation dehnen sich beide Teile entsprechend den Fliehkräften, wodurch der Lamellenkörper belastet wird.

Bild 6a zeigt das untersuchte System. Starke Linien bedeuten die der Rotation entsprechende Lage. Wenn der Lamellenkörper und der Ring von-

einander getrennt gedacht werden, so nehmen sie die gestrichelt gezeichnete Lage ein, die als kräftefreier Zustand gilt. Würde schließlich die durch die Rotation verursachte Deformation aufhören, so würden die Berührungspunkte in die Lage A_0 und B_0 gelangen.

Die Berechnungen können aufgrund von Bild 6b durchgeführt werden. A_0 und B_0 bezeichnen die ruhende Ausgangslage. Während der Rotation tritt eine Verschiebung sowohl am Lamellenkörper als auch am Ring ein. Die Verschiebung von A_0 kann laut (16) einfach aus der Federkonstante c_s des Lamellenkörpers berechnet werden:

$$y'' = c_s C. \tag{52}$$

Hier muß nämlich statt der Belastung Q sinngemäß die aufgrund von (12) definierte Fliehkraft des Lamellenkörpers eingesetzt werden. Die Berechnung derselben, ebenso wie die Berechnung der Konstante c_s, werden im Abschnitt Federkonstanten besprochen. Es muß jedoch bemerkt werden, daß in Wirklichkeit dieser Zustand nicht zustandekommen kann, die Ausdehnung des Lamellenkörpers würde nämlich Separation der Lamellen bedeuten, d. h. in Ermangelung einer zusammenhaltenden Kraft den Zerfall. Natürlich kann für die Zwecke der Berechnung unsere Vorstellung beibehalten werden, und der Lamellenkörper kann als zusammenhängender Ring betrachtet werden, unter Zulassung von Zugspannungen.

In ähnlicher Weise können auch die Verschiebungen der Ringe berechnet werden, also die radiale Verschiebung von B_0:

$$\eta'' = c'_g C_g, \tag{53}$$

wo die gemeinsame Federkonstante c'_g der beiden Ringe und ihre gemeinsame Fliehkraft C_g im Abschnitt über Federkonstanten bestimmt werden.

Die der Rotation entsprechende Lage, ohne Zusammenbau, charakterisiert die Punkte A_c und B_c. Die radiale Entfernung zwischen ihnen ist

$$Y_c = y'' - Y_{*c} - \eta''. \tag{54}$$

Bei der Vereinigung des Systems verschiebt sich der Ring als elastisches Element um η_c, der Lamellenkörper um y_c, und die Berührung wird jetzt durch das Punktpaar A und B gekennzeichnet.

Aufgrund vom Bild 6b ist jedoch

$$Y_c = \eta_c - y_c, \tag{55}$$

worin die einzelnen Deformationen, mittels der zwischen den beiden Systemen auftretenden Radialkraft ausgedrückt

$$y_c = -c_n Q_c \tag{56}$$

und

$$\eta_c = c_h Q_c \tag{57}$$

sind. Die Gleichungen (21), (55), (56) und (57) ergeben daher

$$Y_c = cQ_c. \tag{58}$$

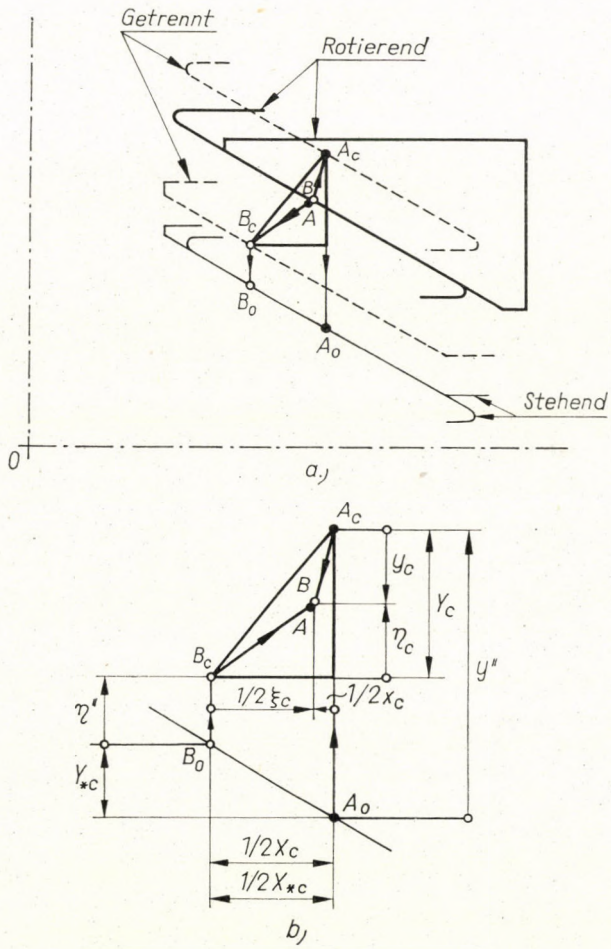


Bild 6. Verschiebungen bei der Rotation

Bei der Berechnung der Axialkräfte muß bloß darauf geachtet werden, daß bei den frei rotierenden Elementen keine Axialverschiebung auftritt. Übrigens tritt beim Übergang aus dem durch die Punkte A_c und B_c gekennzeichneten Zustand in den Zustand A und B natürlich auch eine axiale elastische Verschiebung auf, für welche

$$X_c = -\xi_c - x_c \tag{59}$$

ist. Die einzelnen Deformationen können durch die Axialkraft T_c ausgedrückt werden:

$$x_c = -c_{nx} T_c \tag{60}$$

und

$$\xi_c = c_{hx} T_c. \tag{61}$$

Aufgrund der Gleichungen (27), (59), (60) und (61) wird dann

$$X_c = c_x T_c. \tag{62}$$

Im übrigen gilt auch jetzt die Beziehung zwischen den Radial- und den Axialkräften:

$$\tan\alpha = 2\frac{T_c}{Q_c} \tag{63}$$

sowie für die Deformationen:

$$\tan\alpha = 2\frac{X_c}{Y_c}\frac{c}{c_x}. \tag{64}$$

Laut Bild 6b ist hingegen

$$X_{*c} = 2\frac{Y_{*c}}{\tan\alpha} \tag{65}$$

und die relative Verschiebung ist

$$\overline{A_0 B_0} = \frac{Y_{*c}}{\sin\alpha} = \frac{X_{*c}}{2\cos\alpha}. \tag{66}$$

Zur Durchführung der Berechnung muß der Wert von Y_{*c} bekannt sein, d.h. nur auf Grund desselben kann Y_c aus (54) berechnet werden.

Nach Einsetzen der Gleichung (54) in (64) und mit $X_{*c} = X_c$ gemäß Bild 6b wird

$$Y_{*c} = 2\frac{c}{c_x}\frac{1}{\tan\alpha}(-X_{*c}) + (y'' - \eta'');$$

und andererseits wird unter Benützung von (65)

$$Y_{*c} = c_x \frac{y'' - \eta''}{4c + c_x \tan^2\alpha}\tan^2\alpha. \tag{67}$$

Die Deformationen und die Kräftewirkungen können laut [4] auch grafisch dargestellt werden.

VIII. Betriebszustände

Die Deformationen und die Kräfte für einen beliebigen Betriebszustand können auf Grund der dargelegten Berechnungen einfach durch Superposition gewonnen werden. Ohne auf die Einzelheiten einzugehen, fassen wir hier unsere diesbezüglichen Ergebnisse zusammen.

Stillstand, kalt: Im Stillstand bei Umgebungstemperatur stimmt jede Charakteristik des Kommutatorsystems mit den für den Montagezustand eingehend besprochenen überein.

Stillstand, warm: Im Stillstand nach Unterbrechung des Betriebs der Maschine ist der Kommutator auf einer höheren Temperatur als seine Umgebung. Dann sprechen wir vom warmen Zustand.

Die Kräfte- und Deformationsverhältnisse können aufgrund der Tatsache bestimmt werden, daß der Kommutator aus dem Montagezustand in die erwärmte Endlage gekommen ist. Daher müssen die bei der Montage und bei der Erwärmung eingehend besprochenen Ergebnisse benützt werden.

Die für die Berechnung benötigten Gleichungen sind in Radialrichtung:

$$\begin{aligned} Y_\vartheta &= \eta_\vartheta - y_\vartheta = Y_c + Y_{\Delta t}\,, \\ Q_\vartheta &= Y_\vartheta / c = Q_0 + Q_{\Delta t} \end{aligned} \tag{68}$$

und in Axialrichtung:

$$\begin{aligned} X_\vartheta &= \xi_\vartheta - x_\vartheta = X_0 + X_{\Delta t}\,, \\ T_\vartheta &= X_\vartheta / c_x = T_0 + T_{\Delta t}\,. \end{aligned} \tag{69}$$

Da eine elektrische Maschine je nach ihrer Umgebung auch der Abkühlung ausgesetzt sein kann, wird fallweise auch die dementsprechende Kontrolle nötig. Die Berechnung wird im Sinne des vorangegangenen durchgeführt, aber die Temperaturänderung Δt ist negativ und die Vorzeichen der Deformationen sind umgekehrt.

Rotation, kalt: Beim Anfahren einer Maschine aus dem Dauerstillstand ist der Kommutator kalt. Eine gewisse Zeit lang kann er auch bei Betriebsdrehzahl als kalt betrachtet werden, d. h. zur Erwärmung wird eine gewisse Zeit benötigt. Beim charakteristischen kalten, rotierenden Zustand müssen außer der Montage auch die Fliehkräfte berücksichtigt werden. Daher müssen die für die Montage und die Rotation eingehend erörterten Ergebnisse summiert werden.

In *Radialrichtung:*

$$\begin{aligned} Y_\omega &= \eta_\omega - y_\omega = Y_0 + Y_c\,, \\ Q_\omega &= Y_\omega / c \quad\;\; = Q_0 + Q_c\,; \end{aligned} \tag{70}$$

und in *Axialrichtung:*

$$X_\omega = \xi_\omega - x_\omega = X_0 + X_c\,,$$
$$T_\omega = X/c_x = T_0 + T_c\,. \tag{71}$$

Bei der Werkserprobung der Maschinen erfolgt die Rotation mit größerer Drehzahl als im Betrieb. Bei diesem sogenannten Schleudern sind die Verhält-

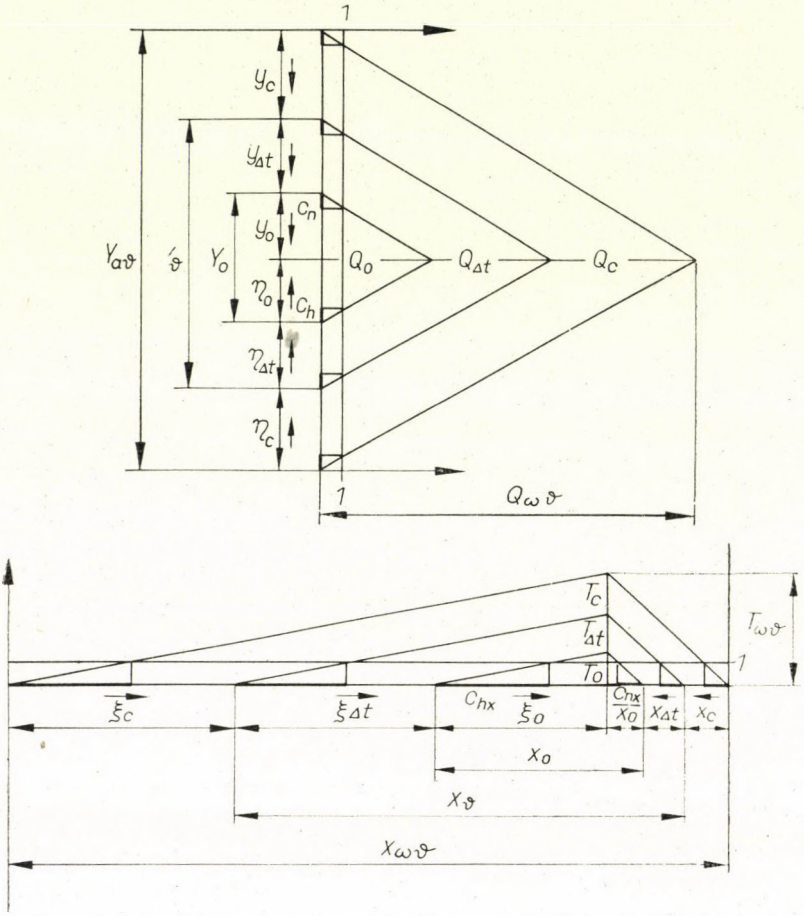


Bild 7. Darstellung der Deformation; $\tan\alpha = 2(T/Q) = 2 \cdot (X/Y)\,(c/c_x)$

nisse vollkommen gleich den vorangehend besprochenen, nur muß statt der Winkelgeschwindigkeit ω die Geschwindigkeit $\omega_{\max}$ in die Formeln eingesetzt werden.

Betriebszustand: Der Kommutator einer Maschine in normalem Betrieb arbeitet einerseits auf einer höheren Temperatur als die Umgebungstemperatur, andererseits rotiert er. In diesem Falle erscheint also gleichzeitig, was das Kräftespiel der Konstruktion anbetrifft, der Einfluß der Montage, der Erwärmung und der Rotation.

Die Daten des Betriebszustandes werden durch Summierung derjenigen für die Teilzustände gewonnen:

In *Radialrichtung*

$$\begin{aligned} Y_{\omega\vartheta} &= \eta_{\omega\vartheta} - y_{\omega\vartheta} = Y_0 + Y_{\Delta t} + Y_c, \\ Q_{\omega\vartheta} &= Y_{\omega\vartheta}/c \quad\;\; = Q_0 + Q_{\Delta t} + Q_c; \end{aligned} \tag{72}$$

und in *Axialrichtung*

$$\begin{aligned} X_{\omega\vartheta} &= \xi_{\omega\vartheta} - x_{\omega\vartheta} = X_0 + X_{\Delta t} + X_c, \\ T_{\omega\vartheta} &= X_{\omega\vartheta}/c_x \quad\;\; = T_0 \;\; + T_{\Delta t} \;\; + T_c. \end{aligned} \tag{73}$$

Der Vollständigkeit halber werden das Deformations- und Kräftediagramm gezeigt (Bild 7). Für die Radialkräfte wurde gemäß dem Schrifttum [4] vorgegangen, während für die Axialkräfte nach dessen Vorbild sinngemäß das Diagramm auch ohne besondere Erläuterung aufgezeichnet werden kann.

IX. Federkonstanten

Die Grundlage für die Berechnung der Federkonstanten ist Gleichung (16).

1. Radiale Federkonstanten

Bei der Berechnung von c_s wird berücksichtigt, daß einerseits unter Einwirkung des Gewölbedrucks p sich der ganze Lamellenkörper deformiert (c_s'), andererseits sich der Schwalbenschwanzteil unter der Oberflächenbelastung q durchbiegt (c_s''). c_s' wurde entsprechend [4] berechnet, deshalb genügt es, unter Weglassung der Einzelheiten das Ergebnis mitzuteilen:

$$c_s' = \frac{K}{4\,\pi^2\,A_s}\left(\frac{s_s + s_l}{E_s + E_l}\right), \tag{74}$$

wo E_s der Elastizitätsmodul der Lamellen, E_l derjenige der Segmentisolierung ist, und das Ausmaß von s_s und s_l durch Bild 2 gezeigt wird. c_s'' kann einfach mit den Bezeichnungen von Bild 8 annähernd aus einem Konsolträger berechnet werden, welchen die Kraft N_s belastet, unter Vernachlässigung der seitlichen Abstützung. Die Durchbiegung unter dem Einfluß der Komponente R_s ergibt sich einerseits aus der Biegung, andererseits aus dem Schub [5]:

$$y = \frac{l''^3\,R_s}{3\,E_s\,I''} + \frac{6\,l''\,R_s}{5\,G_s\,A''},$$

wo I und A das Trägheitsmoment bzw. die Fläche des Schwalbenschwanzteils sind. Nach Einführung der aus dem Bild ersichtlichen Abmessungen und des Wertes Q aus den Gleichungen (4) bzw. (11) anstelle von R_s, wird

$$y = \frac{1}{K}\left(\frac{2\,l''^2}{E_s\,h''^2} + \frac{3}{5\,G_s}\right)\frac{l''}{b''\,h''}\,Q\,.$$

So wird die gemeinsame Federkonstante für die beiden Schwalbenschwänze

$$c_s'' = \frac{1}{K}\left(\frac{2\,l''^2}{E_s\,h''^2} + \frac{3}{5\,G_s}\right)\frac{l''}{b''\,h''}\,. \tag{75}$$

Die resultierende Federkonstante des Lamellenkörpers ist:

$$c_s = c_s' + c_s''\,. \tag{76}$$

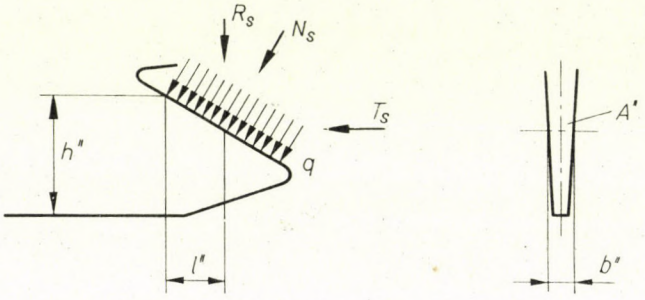


Bild 8. Teil des Schwalbenschwanzes

Die Berechnung von c_m erfolgt aus der durch die Belastung q verursachten Zusammendrückung (Bild 9). Das Maß h_m der Kappe verringert sich unter Einfluß des Druckes q um den Betrag

$$\Delta h_m = \frac{q}{E_m}\,h_m\,.$$

Unter Benützung des Zusammenhangs (1) und Berücksichtigung des Ausdrucks (10) sowie von

$$K\,(s_s' + s_l') = d_m\,\pi\,,$$

wird

$$\Delta h_m = \frac{h_m}{E_m}\,\frac{N_s}{l_m\,(s_s' + s_l')} = \frac{h_m}{E_m}\,\frac{N}{l_m\,d_m\,\pi}\,.$$

Das Bild zeigt die Längenänderung $\Delta\, h_m$ vergrößert.

Radial entspricht der Maßänderung $\Delta\, h_m$

$$y = \frac{\Delta h_m}{\cos\alpha} = \frac{h_m}{2\,\pi\,l_m\,d_m\,E_m\cos^2\alpha}\,Q\,,$$

wenn anstelle von N mit Hilfe der Beziehungen (10), (2), (4) und (11) der Ausdruck $Q/2 \cdot \cos \alpha$ geschrieben wird. Hieraus ist im Sinne von (16):

$$c_m = \frac{h_m}{2\,\pi \cos^2 \alpha\, E_m\, l_m\, d_m} \,. \tag{77}$$

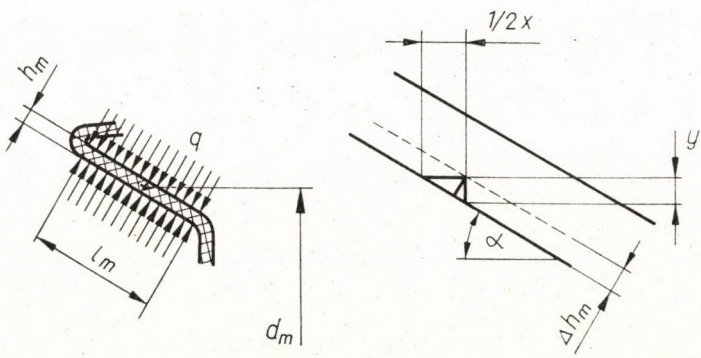


Bild 9. Teil der Kappe

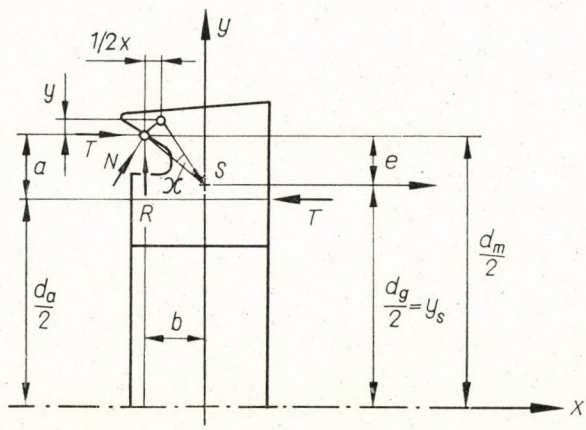


Bild 10. Teil des Ringes

Bei der Berechnung von c_g können wir uns auf die veröffentlichten Ergebnisse stützen [6] und die Bezeichnungen von Bild 10 verwenden. Die Verschiebung des Berührungspunktes mit der Kappe ist in y-Richtung

$$y = \varkappa b = \frac{M_*}{I_y}\,\frac{y_s}{E_g}\,b \,,$$

wo I_y das Trägheitsmoment des Querschnitts in Bezug auf die y-Achse ist, und andererseits

$$M_* = \frac{1}{2\pi}(aT + bR) = \frac{1}{4}(a \tan \alpha + b)\, Q$$

ist. Hiemit ist jedoch

$$y = \frac{a \tan \alpha + b}{4} \frac{y_s b}{I_y E_g} Q,$$

und auf Grund von (16) ist die Federkonstante

$$c_g = \frac{(a \tan \alpha + b)\, y_s b}{4\pi I_y E_g}. \tag{78}$$

Entsprechend Bild 3 ist die resultierende Federkonstante des gedrückten Teils

$$c_n = c_s + c_m. \tag{79}$$

Die resultierende Federkonstante des gezogenen Teils ist

$$c_h = c_g. \tag{80}$$

2. Axiale Federkonstanten

Bei der Berechnung von c_{sx} wird der Lamellenkörper als gedrückter Zylinder angesehen, dessen Deformation laut Bild 11

$$x = \frac{1}{K} \frac{l_s}{A' E_s + A_l E_l} T,$$

wo A' der Flächeninhalt des Querschnitts einer Lamelle und A_l derjenige einer Lamellenisolation ist. Auf Grund von (16) ist daher

$$c_{sx} = \frac{1}{K} \frac{l_s}{A' E_s + A_l E_l}. \tag{81}$$

Die Berechnung von c_{mx} aus der durch q verursachten Zusammendrückung kann auf Grund von Bild 9 durchgeführt werden. Die axiale Zusammendrückung ist nämlich

$$x = 2 \frac{\Delta h_m}{\sin \alpha} = \frac{2 h_m}{\pi l_m d_m E_m \sin^2 \alpha} T,$$

wenn anstelle von N mit Hilfe der Gleichungen (10), (3) und (8) der Wert $T/\sin\alpha$ eingesetzt wird. Schließlich wird im Sinne von (16)

$$c_{mx} = \frac{2\,h_m}{\pi\,l_m\,d_m\,E_m\sin^2\alpha} \tag{82}$$

und dies ist die gemeinsame Federkonstante der beiden Kappen.

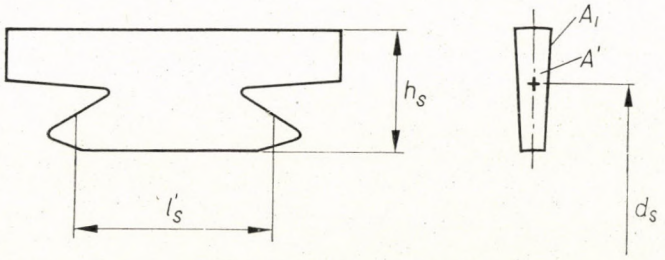


Bild 11. Kommutatorlamelle

c_{gx} wird gemäß dem Schrifttum [6] berechnet. Laut Bild 7 ist die Verschiebung des Berührungspunktes der Kappe in x-Richtung, für beide Ringe gemeinsam,

$$x = 2\,\varkappa e = 2\,\frac{M_*}{I_y}\,\frac{I_s}{E_y}\,e\,,$$

worin

$$M_* = \frac{1}{2\,\pi}\,(aT + bR) = \frac{1}{2\,\pi}\left(a + \frac{b}{\tan\alpha}\right)T,$$

und daher ist

$$x = \frac{a\tan\alpha + b}{\pi\tan\alpha}\,\frac{y_s\,e}{I_y E_g}\,T,$$

und die Federkonstante ist laut Gleichung (16)

$$c_{gx} = \frac{(a\tan\alpha + b)\,y_s\,e}{\pi\tan\alpha\,I_y\,E_g}\,. \tag{83}$$

Die Berechnung von c_{ax} kann einfach nach der Formel für gezogene Stäbe erfolgen. Die Dehnung der Schrauben ist

$$x = \frac{l_a\,T}{n A_a\,E_a}\,,$$

wo l_a die Länge des belasteten Abschnitts der Schrauben ist, A_c der Flächeninhalt des Querschnitts der Schrauben, E_a der Elastizitätsmodul des Schraubenwerkstoffs. Laut (16) ist die gemeinsame Federkonstante der Schrauben

$$c_{ax} = \frac{l_a}{n A_a E_a} . \tag{84}$$

Die resultierende Federkonstante des gedrückten Teils ist laut Bild 3

$$c_{nx} = c_{sx} + c_{mx} . \tag{85}$$

Die Konstante des gezogenen Teils ist

$$c_{hx} = c_{gx} + c_{ax} . \tag{86}$$

Die algebraische Summe der Fliehkräfte des Lamellenkörpers ist laut [4]

$$C = K \frac{A_s}{2g} (\gamma_s s_s + \gamma_l s_l) d_s \omega^2, \tag{87}$$

wo γ_s das spezifische Gewicht der Lamellen, γ_l dasjenige der Segmentisolierung ist; die weiteren Größen sind aus Bild 2 und Bild 11 ersichtlich.

Laut Bild 9 ist die algebraische Summe der Fliehkräfte der beiden Ringe (C'_g) und der beiden Schrauben (C''_g)

$$C_g = 2 \left(\frac{\gamma_g}{g} \pi A_g d_g \right) \frac{d_g}{2} \omega^2 + n \left(\frac{\gamma_a}{g} A_s L_a \right) \frac{d_a}{2} \omega^2 , \tag{88}$$

wo γ_g das spezifische Gewicht des Ringes, γ_a dasjenige der Schrauben, A_g der Flächeninhalt des Ringquerschnitts, L_a die gesamte Schraubenlänge ist. Die der eigenen Fliehkraft entsprechende Deformation ist

$$\eta''_g = \frac{\sigma}{E_g} \frac{d_g}{2} = \frac{\gamma_g}{g} \left(\frac{d_g}{2} \right)^2 \omega^2 \frac{1}{E_g} \frac{d_g}{2} = C'_g \frac{d_g}{8 \pi A_g E_g} = c'_g C'_g ,$$

d. h.

$$c'_g = \frac{d_g}{8 \pi A_g E_g} . \tag{89}$$

X. Bemerkungen zur Dimensionierung

Eine grundlegende Bedingung für die Stabilität von Kommutatorkonstruktionen ist, daß in keinem Zustand der Gewölbedruck zwischen den Lamellen auf Null sinken darf, da dann die Lamellen lose werden. Selbst-

verständlich muß aus Sicherheitsgründen ein größerer Wert als Null für den Gewölbedruck vorgeschrieben werden. Auf Grund einer eingehenden Untersuchung und von Erfahrungswerten kann der benötigte Gewölbedruck bestimmt werden [7].

Wenn die geometrischen Maße aufgrund der in der elektrischen Berechnung vorgeschriebenen Hauptabmessungen bestimmt worden sind, ist ein wichtiger Schritt in der Dimensionierung die Wahl des Montage-Gewölbedrucks, der größer als der minimale Wert sein muß. In Kenntnis von p_0 kann aus den Beziehungen (5), (6), (7), (4), (2), (3), (8) und (15) die notwendige Spannkraft T_c der Schrauben berechnet werden, mit welcher der Gewölbedruck hervorgebracht wird.

Dann muß das Kräftespiel für jeden Betriebszustand geklärt werden und der Zustand gesucht werden, welcher den kleinsten Gewölbedruck hervorruft. Es muß kontrolliert werden, ob der so berechnete Gewölbedruck nicht kleiner als der zulässige Minimalwert ist.

Die obere Grenze des Gewölbedrucks kann auch nicht irgend beliebig sein, wachsen doch mit dem Gewölbedruck sämtliche Kräfte, also auch die Spannungen. Eben deswegen ist es eine weitere wichtige Aufgabe der Festigkeitsberechnung die Berechnung der Spannungen. Unter Berücksichtigung derselben kann der maximale Gewölbedruck festgelegt werden.

Schließlich muß der Gewölbedruck so gewählt werden, daß derselbe zwischen den minimalen und den maximalen Wert fällt. Die Einzelheiten der Dimensionierung werden hier jedoch nicht besprochen, Zielstellung war nur die Klärung des der Bemessung zugrundeliegenden Kräftespiels.

SCHRIFTTUM

1. Liska: Villamosgépek, V. Szerkezettan (Elektrische Maschinen, V. Konstruktionslehre). Tankönyvkiadó, Budapest 1954.
2. Hütte: Elektrotechnik. A. Ernst & Sohn, Berlin 1957.
3. *Алексеев:* Конструкция электрических машин. Госэнергоиздат, Москва 1958.
4. Huszár: Berechnung des Kräftespiels von Kommutatoren mit einfachem Schrumpfring. *Acta Techn. Hung.* **55** (1966), 3—4.
5. ten Bosch: Berechnung der Maschinenelemente. Springer, Berlin 1953.
6. Huszár: Kommutátor szorítógyűrű feszültségvizsgálata (Spannungsprüfung von Kommutatorpreßringen). *Elektrotechnika* (1967).
7. Huszár: Bemerkungen zur Wahl des Gewölbedruckes von Kommutatoren. *Elektrische Bahnen*, 1965.

Calculation of the System of Forces of Arch-Bound Commutators. The mechanical model of the arch-bound commutator can be represented by two one-dimensional spring systems With the one the radial forces and deformations are considered, with the other the axial ones. The paper gives the relations by which the spring rates of the individual spring elements are determined. By clarifying the calculating method of the spring system the forces corresponding to the various states occurring in commutator service (assembly, rotation, warming, rotating warm state) may be determined. The stresses stated thus supply the data needed for dimensioning the commutator.

Calcul du jeu de forces des collecteurs à queue d'aronde. Le modèle mécanique du collecteur à queue d'aronde peut être représenté par deux systèmes de ressorts unidimensionnels, dont l'un tient compte des forces et déformations radiales, et l'autre de celles axiales. L'étude donne les relations servant à déterminer les constantes de rappel des éléments de ressort individuels. En mettant au point la méthode de calcul du système, il devient possible de déterminer les forces correspondant aux différents états survenant pendant le service du collecteur (montage, rotation, échauffement, rotation à chaud). Les sollicitations ainsi déterminées fournissent les données nécessaires pour le dimensionnement du collecteur.

Расчет работы коллекторов в виде ласточкина хвоста (*И. Хусар*). Механическую модель конструкции коллектора в виде ласточкина хвоста можно получить в качестве двух одноразмерных пружинных систем. Одна из них учитывает радиальные, а другая осевые усилия и деформации. В задаче приводятся зависимости, служащие для определения пружинных постоянных отдельных пружинных элементов. Выяснением методики расчета пружинной системы можно определить усилия, соответствующие различным состояниям (сборка, вращение, нагрев, вращающееся и горячее состояние), встречающимся при работе коллектора. Полученные таким образом нагрузки предоставляют необходимые данные для расчета конструкции коллектора.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 31—54 (1968)

ПРИМЕНЕНИЕ ЦИФРОВЫХ АНАЛОГОВ В УСТРОЙСТВАХ ПРОМЫШЛЕННОЙ АВТОМАТИКИ*

Й. ХАТВАНИ, П. КАРДОШ и Я. БАРАТ

ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ ПО АВТОМАТИЗАЦИИ
ВЕНГЕРСКОЙ АКАДЕМИИ НАУК, БУДАПЕШТ

[Поступило 24. марта 1967 г.]

После выяснения возникающих понятий статья занимается проблемами расчетов комбинированных систем. Авторы предлагают методику расчета, контролируемую измерениями, воспроизводимыми на опытной установке, вместо описанных ранее совершенно эмпирических методов. Для расчета и контроля систем большей сложности единственно подходящим путем является симуляция и оценка, выполняемые на цифровой вычислительной машине. Статья занимается вопросами симуляционной программы, общей действительности и основывающейся на применении теории конечных автоматов. Рассматривается теоретические проблемы применения выбранного принципа и говорится о доказательстве действительности применения этого принципа в случае данных типов систем. Экспериментальный контроль симуляционного метода с использованием вычислительной машины производился путем симуляции изложенной ранее измерительной схемы. Результаты, полученные вследствие использования симуляционной программы, дали значения, идентичные значениями непосредственных измерений.

I. Введение

Первым типом вычислительной машины, использованной для технических расчетов, является механический дифференциальный анализатор, построенный в 1931 году *Буш*ем [1] по ранее опубликованной идее (*Лорд Келвин* — 1876 г.; *Крылов* — 1904 г.). *Магиннис* [2] уже сообщает о машинном расчете таких проблем, как переходные процессы в индукционных электродвигателях, колебания самолетов при посадке, траектории электронов, переходные тепловые процессы в диэлектриках и т. д. Все эти многообразные задачи на дифференциальных анализаторах разрешаются по существу с помощью многократного применения комбинации блоков двух типов: интегратора и сумматора.

Шеннон в своей работе [3], имеющей коренное значение в теории дифференциальных анализаторов, доказал, что с помощью машины, построенной исключительно на интеграторах и сумматорах, можно генерировать любую рациональную и иррациональную функцию, а также все так называемые алгебраические трансцендентные функции (например, экспоненциальные, тригонометрические, гиперболические, вероятностные и Бесселевы функции), за исключением всего только гипертрансцендентных, как, например, Гамма-функция и Зета-функция Римана. Из этого следует, что можно показать относительно в самой общей форме дифференциальных уравнений (XI теорема

* Разделы I—IV написаны *Й. Хатвани* и *П. Кардош*; разделы V—VI *Я. Барат*.

цитированной работы *Шеннона*), что она разрешима с помощью конечного числа интеграторов и сумматоров, если она не содержит гипертрансцендентные функциональные связи.

Появление двух больших классов электронных вычислительных машин, работающих с аналоговой и кодированной цифровой формой информации, вытеснило из области практических расчетов медленно действующий и имеющий ограниченную точность механический дифференциальный анализатор. Принцип дифференциального анализатора вновь приобрел значение только после открытия в 1950 году методов цифрового интегрирования и суммирования, основанных на счете импульсов [4]. Вслед за этим были построены первые цифровые дифференциальные анализаторы и начались работы по их теории. Один из нашедших свое применение методов, который *Шпраг* [5] называет методом «*сита*», а другие «*последовательным*» или «*операционным*» (интегратор с последовательным переносом), формировался в ходе развития цифрового управления металлорежущими станками. Другой — «параллельный» метод (интегратор с параллельным переносом) применяется в скоростных вычислительных и симулирующих устройствах. Применение первого метода авторы предложили сравнительно давно [6], а затем построили измерительный преобразователь на интеграторе данного типа и предложили некоторые новые области его применения [7, 8]. Построенные блоки, которые в конечном счете можно рассматривать как интеграторы и сумматоры, сами по себе и в простых комбинациях со сравнительно медленным изменением входных параметров, могут быть проектированы и применены без учета некоторых таких свойств, которые в более сложных комбинациях препятствуют простой замене идеальных интеграторов и сумматоров цифровыми. Следовательно, предметом наших дальнейших исследований являлось установление возможностей, пределов и правил синтеза более сложных структур в пределах открытых *Шенноном* возможностей для идеальных интеграторов-сумматоров.

Необходимость проведения таких исследований выявилась в ходе проектирования устройств промышленной автоматики средней сложности. В соответствии с этим в каждом отдельном случае можно было определить значение, вес отдельных параметров. Для нахождения оптимального компромисса необходимо было выявить качественные характеристики возможных решений и характер взаимосвязи отдельных параметров.

На современном уровне науки и техники теоретически выполнимы любые задания. Важнейшим ограничением является экономичность. Значение оптимального проектирования вообще, а при проектировании промышленных устройств в особенности, весьма многообразное. Оно дает возможность снизить стоимость, количество элементов, повысить надежность работы устройства, а также выявить пути улучшения, важнейших для данного устройства параметров.

Оптимальное проектирование надо начать с оптимальной формулировки задания, а это нельзя свести к минимализации структуры. Очень важно найти не только оптимальный метод, но и оптимальные средства, причем следует учесть, что абсолютно оптимальных решений нет.

II. Проблема неравномерности

Большим преимуществом двоичного умножителя и построенного с его применением интегратора с последовательным переносом по сравнению с другими методами цифрового умножения и интегрирования является схемная простота и малое число деталей. Недостатками являются сравнительно небольшая скорость, а в случае более сложных структур в основном то, что ряд выходных импульсов даже при неизменности входных величин в большинстве случаев распределен по времени неравномерно. При построении первых таких устройств [9, 10] эта временная неравномерность не была учтена. Некоторые авторы описали несколько таких структур, которые на практике по предложенными ими методам не могли быть осуществлены (эти схемы позднее перешли в ряд обзорных работ). Влияние неравномерности могло оставаться незамеченным потому, что в более простых схемах только в исключительно редких случаях (а именно при некоторых сравнительно быстрых периодических изменениях входных величин) эта неравномерность приводит к накоплению ошибок, ошибка же мгновенного значения сравнительно мала и периодически становится равной нулю.

Ошибки вследствие неравномерности впервые выявил *Лунд* [11]. Он показал, что если стабильность какой-нибудь системы зависит от уравновешенного состояния реверсивного счетчика, при поступлении импульсов на суммирующий и вычитающий его входы с одинаковой средней частотой повторения, то в случае, когда на один из этих входов поступают импульсы, распределенные по времени неравномерно, возникает нестабильность.

Это воздействие может возрастать, особенно в том случае, когда упомянутый счетчик управляет одним из его входных сигналов.

Лунд в качестве примера приводит случай, показанный на рис. 1. Если в течении интервала времени t_1 управляющий умножителем код принимает значение 1/2, а в интервале t_2 3/8, тогда в течение $T = t_1 + t_2$ число выходных импульсов в зависимости от фазы может быть равно или 4 или 1. (*Джонсон* для одного особого случая показывает [12], что если управление четырехэлементного умножителя изменяется по временной диаграмме, показанной на рис. 2, между значениями 4/16 и 3/16, то интегрированное за один цикл значение вместо 3 1/2 станет равным нулю!)

Констатируя тот факт, что до сих пор не удалось установить математическую зависимость для расчета влияния неравномерности, *Лунд* предлагает средство коррекции: элемент «*холостого хода*». Схема этого «*холостого*

хооа», или по удачному названию *Вуда* [13] «*цифрового фильтра*», показана на рис. 3.

Предположим, что триггер находится в состоянии «сложения», и приходит импульс на вход «*вычитание*». Импульс не проходит через схему «*И*», потому что ее другой вход находится в состоянии отрицания, но через время τ он переводит триггер в состояние «*вычитания*». Таким образом, если следующий импульс тоже поступает на вход «*вычитание*», он уже проходит.

Однако, если импульсы на противоположные входы поступают чередующе, то на выходе импульс не появится.

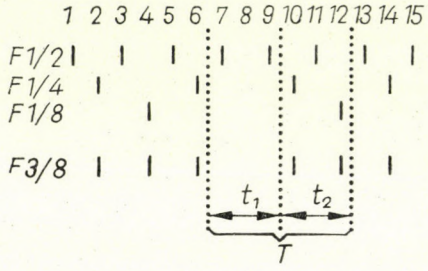


Рис. 1

Очевидно, что при последовательном включении таких «*фильтрующих*» звеньев, ценой некоторого увеличения времени вычисления, в любой мере можно уменьшить влияние неравномерности, так как эти звенья уменьшают амплитуду и частоту колебаний, но не приводят к ошибке в численности. Это открытие *Лунда* использовал *Вуд* [13] и ряд других авторов.

В случае сложения выходных импульсов уменьшение неравномерности возможно путем деления выходных импульсов. При этом длительность цикла соответственно удлиняется, что соответствует изменению масштабных коэффициентов. По эффективности одно звено фильтра примерно соответствует делению на два. Необходимое число элементов фильтра и делителя определили экспериментально. У *Вуда* при 6 двоичных элементах число фильтрующих звеньев равнялось 2, *Воронов* и *Соколов* [14] применяли в устройстве на 13 двоичных знаках 3 двоичных делителя.

По нашему мнению, недостающий по Лунду математический метод для определения необходимого числа фильтрующих или делительных звеньев найден в результате работ *Гарбузова* [15] и работавшего от него независимо *Джонсона*. Эти авторы опубликовали анализ точности двоичного умножителя в режиме интегрирования скачка. В результате получена формула для максимальной ошибки в зависимости от числа двоичных знаков

$$\eta_{k\,\max}^{\tau} = \frac{7}{18} + \frac{n}{6} + \frac{1}{9}(-1)^n 2^{-n}. \qquad (1)$$

Таким образом, максимальная ошибка неравномерности при 13 двоичных знаках равна примерно 2, 5, а в случае 6 двоичных знаков 1, 4.

При синтезе более сложных систем можно остановиться на подробном анализе структуры в целом, и расчет параметров отдельных блоков осуществить на основе этого анализа, находя для каждого отдельного блока допустимые значения ошибок неравномерности, или же создать такой комплект модульных блоков, входные и выходные параметры которых определены с известным ограничением ошибок и в дальнейшем рассматривать их с достаточной точностью приближения как идеальные. Во многих случаях второй

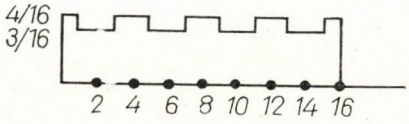


Рис. 2

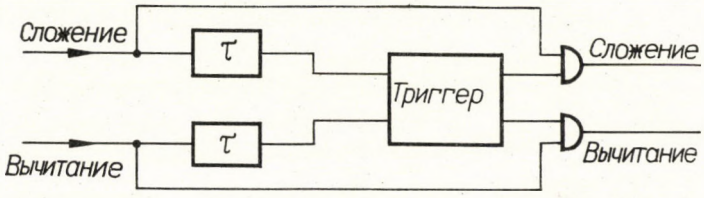


Рис. 3

метод может быть достаточно точным, а в других случаях может дать исходное первое приближение для дальнейших уточнений.

Попыткой нахождения такого метода можно считать работу *Шмид*а [16], который преобразовал входные информации отдельных блоков в длительность импульсов, измеряемую в квантах импульсов общего задающего генератора. Это преобразование требует дополнительных элементов, и при его применении теряется возможность одновременного использования информаций в виде частоты повторения и численности импульсов. Используя результаты работ *Гарбузов*а и *Лунд*а, подобное ограничение ошибок отдельных функциональных блоков стало возможным и без преобразования вида носителя информации. В этом случае, в известных пределах скорости расчета и масштабных коэффициентов, применимы широко освещенные во многих работах (напр. в обзорной работе [15] и в работе [17]), методы идеализированного синтеза. Точный анализ цифровых систем весьма затруднителен, а допущение упрощений может привести к ошибкам такого же порядка, как замена цифровой структуры идеализированной.

III. Динамические характеристики

При анализе свойств интегратора, по сказанным выше соображениям, цифровой интегратор с соответствующей фильтрацией или делением импульсов можно рассматривать в качестве интегратора непрерывного действия.

Передаточная функция интегратора имеет вид

$$\frac{x_{\text{вых}}(p)}{x_{\text{вх}}(p)} = \frac{1}{Tp}, \tag{2}$$

где значение интегральной постоянной равно:

$$T = \frac{2^{-N}}{f}, \tag{3}$$

где f частота повторения входных импульсов.

Так как в большинстве случаев частота повторения входных импульсов управляется каким-нибудь другим, связанным сумматором, постоянная времени интегрирования также может меняться во времени [18]. Для идеализированного случая это дает чаще всего экспоненциальную передаточную функцию, экспоненциальный отрицательный показатель степени которой в своем числителе содержит частоту повторения входных импульсов системы, а в своем знаменателе число двоичных знаков примененных счетчиков. Следовательно, скорость установления таких систем по экспоненциальному закону растет с частотой и убывает с увеличением точности (двоичных знаков). На основании вышесказанного можно сделать лишь качественные выводы, потому что не были учтены ни дискретный характер, ни действие неравномерности. *Лунд* [11], например, рассчитал время установления отдельных систем при предположении постоянства входных величин без учета других факторов. Расхождение между расчетными и измеренными значениями достигало 50%. К сожалению, найти отдельные составляющие ошибки приближения не удалось.

Классический анализ дает ценные данные, но практический расчет динамических свойств этим путем не выполним.

По нашим исследованиям основные шаги практического проектирования будут следующие:

а) Определить предельные параметры используемых элементов в наименее оптимальном случае, исключить возможность возникновения недозволенных режимов работы (частота, разрешающая способность и т. д.);

б) выбрать число двоичных знаков в зависимости от требуемой точности;

в) в соответствии с выбранным числом двоичных знаков определить необходимое число фильтрующих звеньев или делителей, исходя из формулы

(1), таким образом, чтобы ошибка неравномерности соответствующего выхода стала меньше единицы;

г) ограничить входные параметры по возможности также и снизу или применить переменное число двоичных знаков для сокращения времени установления и расчета;

д) с учетом полученной структуры определить продолжительность цикла каждого умножителя и позаботиться о том, чтобы минимальный период изменения управляющего сигнала во всех дозволенных случаях был больше этой продолжительности цикла.

Описанный метод нельзя считать расчетом в полном смысле слова. Он основывается на практике проектирования, экспериментального и теоретического исследования нескольких простых структур, содержащих двоичные умножители. К ним относится по порядку возрастающей сложности: измерительный преобразователь с одним умножителем для перемножения и интегрирования функции двух переменных, как, например, перенесенный ленточным конвейером вес как функция.

Скорости движения и удельной нагрузки (или протекшая через сечение трубопровода масса, как функция скорости и плотности) [19], управляющее устройство непрерывной пропорциональной дозировки нескольких компонентов [20], варианты частотомера непрерывного действия *Вуда*, варианты устройства управления движением по прямолинейной и круговой траектории [21]. Описание этих устройств выходит за рамки этой статьи за исключением одного экспериментального устройства.

IV. Экспериментальное устройство

Для проверки выводов, основанных на теоретических расчетах и на данных литературы, а также для экспериментального исследования было построено экспериментальное устройство. Блок-схема показана на рис. 4. Оно является вариантом частотомера Вуда для изучения роли отдельных

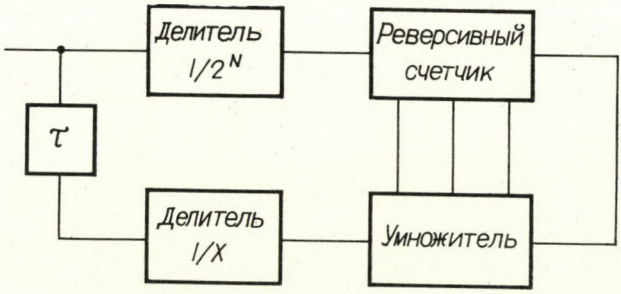


Рис. 4

факторов в «*чистом*» виде и без ошибки повторения. С этой целью «*опорная*» (f_A) и измеряемая (f_X) частоты образуются из сигналов одного генератора импульсов. Одновременное появление импульсов этих частот исключают временной задержкой одного из них. Кроме того, поскольку «*делитель*» $1/2^N$ и умножитель имеют одинаковое число двоичных знаков, то достигают, что отношение «*измеряемой*» частоты к единице измерения будет целым числом. (С точки зрения работы схемы не существенно, но следует учитывать, что изменяется опорная частота.)

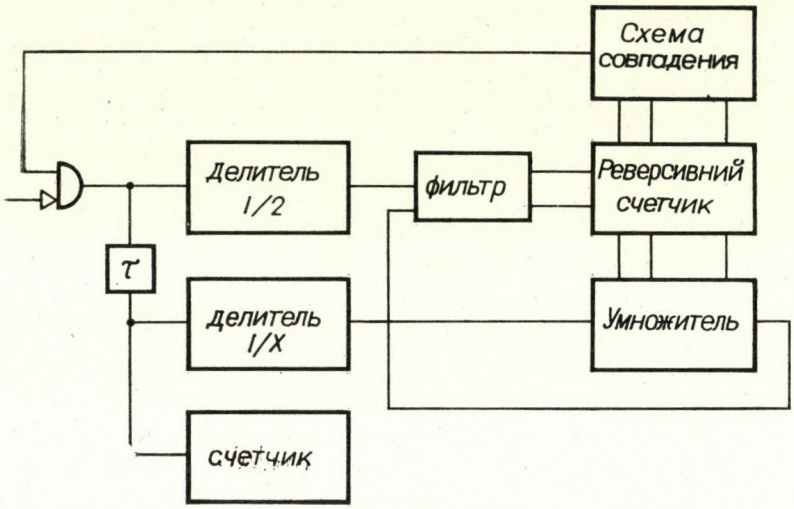


Рис. 5

В результате упомянутых выше мер исключается влияние колебания фазы опорных и измеряемых частот, а также колебания установившегося значения между соседними целочисленными значениями при дробных отношениях измеряемых и опорных частот.

При идеализированном анализе содержание реверсивного счетчика равно:

$$Z_n = \int_0^n \frac{dn}{2^N} - \int_0^n \frac{Z_n}{2^N X}\, dn\,, \tag{4}$$

откуда

$$\frac{dz_n}{dn} + \frac{Z_n}{2^N X} - \frac{1}{2^N} = 0\,, \tag{5}$$

следовательно

$$Z_n = X - Xe^{-\frac{n}{2^N X}}\,. \tag{6}$$

Обозначим число импульсов, необходимое для достижения значения $z_n = x - 1$, через n_b. В этом случае имеем

$$Xe^{-\frac{n_b}{2^N X}} = 1, \tag{7}$$

откуда

$$n_b = 2^N X \ln X. \tag{8}$$

Исследования, которые выполнимы с помощью этого устройства следующие:

Измерение фактического времени установления в зависимости от числа фильтрующих звеньев. Блок-схема измерения представлена на рис. 5.

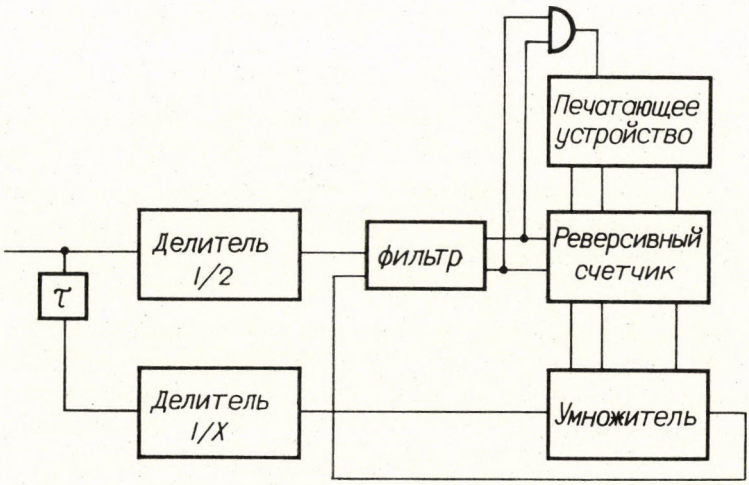


Рис. 6

Диодная схема при достижении установленного значения отключает входные импульсы и после этого время установления можно отсчитать с декадного счетчика.

Регистрация ступеней установления и исследования стабильности по блок-схеме рис. 6. Печатающее устройство срабатывает при каждом изменении с помощью диодной схемы. Если печатание производится при поступлении импульсов на входы фильтра с точки *B*, то можно исследовать влияние фильтра, а если печатанием управлять с точки *C*, то можно исследовать ход установления во времени. Во всех этих случаях вместо печатающего устройства можно применить инкрементальное чертежное устройство. В этом случае результаты получаются в наглядной, графической форме.

Результаты измерений подтвердили правильность предложенного метода проектирования. Экспериментальная проверка помогла выявить возможные источники ошибки измерений *Лунда* и других авторов:

a) Исследование таких структур, где ошибка неравномерности увеличивается (последовательное включение двоичных умножителей);

б) применение ограниченного числа двоичных знаков, большие ошибки округления;

в) неопределенность произвольных начальных условий;

г) неправильный выбор альгоритма осуществления математических операций, применение умножения вместо интегрирования;

д) измерение характеристик, расхождение которых с расчетными обусловлено неопределенностью, и не действительным отклонением работы схемы, время установления определяется точкой пересечения линий с малым углом.

Проведенные нами измерения показывают хорошее совпадение данных измерений с расчетными. Удалось выявить также некоторую цикличность изменения расхождения, хотя закон этого изменения сложный и нет оснований для обобщения. Измерения дали хорошие результаты для качественного анализа, однако количество цифровых данных даже при принятых ограничениях столь велико, что полная обработка их, проверка выводов о закономерностях возможна лишь с помощью цифровой вычислительной машины.

Проведенная работа, а также необходимость проектирования и проверки реальных структур, в том числе устройства для управления металлорежущими станками, заставили нас искать возможности моделирования, анализа и оптимизации структуры на ЦВМ. Метод симулирования основан на теории конечных автоматов.

V. Моделирование работы дискретных устройств на цифровых вычислительных машинах

1. Метод построения матриц состояний и реакций

Известно, что любая логическая сеть по *Берксу* и *Райту* может быть представлена в виде последовательностной машины [22] и поэтому к ней применимы все методы синтеза и анализа последовательностных машин. Авторам показалось удобнее всего применять матричный метод, предложенный *М. Л. Цетлиным*, исходя из соображений экономности хранения информации в памяти и малого количества вычислений. Ниже излагается новый метод построения матриц состояний и реакций, который удобнее, если схема задана в виде таблиц истинности.

Рассмотрим часто использованный метод представления последовательностных машин в виде $(n \times m)$ полюсника с обратными связями, куда вынесены все временные задержки, существующие в системе. Известно, что достаточно рассмотреть такие системы, где задержки составляет один такт по дискретной шкале времени, потому что все логические сети могут быть

сведены к такой последовательностной машине. Пусть нашу машину изображает рис. 7.

И пусть нам известны системы уравнений, описывающие эту машину:

$$\begin{aligned} y_i(t) &= f_i\left[x_1(t), \ldots, x_p(t), z_1(t-1), \ldots, z_{m-s}(t-1)\right]; \\ z_j(t) &= \varphi_j\left[x_1(t), \ldots, x_p(t), z_1(t-1), \ldots, z_{m-s}(t-1)\right], \end{aligned} \tag{9}$$

где

$$i = 1, \ldots, s\,; \quad j = 1, 2, \ldots, m - s\,,$$

и

$$z_j(t-1) = y_{s+j}(t-1).$$

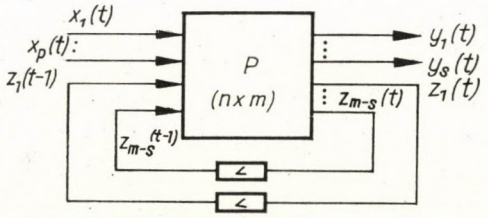


Рис. 7

Это суть две системы булевых уравнений, и если назвать набор переменных

$$\{x_1(t), \ldots, x_p(t)\}$$

входом, набор

$$\{y_1(t), \ldots, y_s(t)\}$$

выходом, а набор переменных

$$\{z_1(t), \ldots, z_{m-s}(t)\}$$

состоянием системы, то ясно, что уравнения (9) задают выход и состояние системы в момент времени t как функции входа в тот же самый t момент времени, и состояния в предыдущий такт.

Система уравнений (9) всегда может быть представлена в совершенной нормальной форме, а поскольку мы тривиальным случаем, когда уравнения являются постоянными (0 или 1) — заниматься не будем, можно считать, что они заданы в дизъюнктивной нормальной (совершенной) форме.

Вводим обозначения:

$$\begin{aligned} p_0 &= \overline{x_1(t)} \cdot \overline{x_2(t)} \ldots \overline{x_p(t)}\,, \\ p_1 &= \overline{x_1(t)} \cdot \overline{x_2(t)} \ldots \overline{x_{p-1}(t)} \cdot x_p(t)\,, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ p_{2p-1} &= x_1(t) \cdot x_2(t) \ldots x_p(t)\,; \end{aligned}$$

а также

$$r_0 = \overline{z_1(t-1)} \ldots \overline{z_{m-s-1}(t-1)} \cdot \overline{z_{m-s}}(t-1),$$
$$r_1 = \overline{z_1(t-1)} \ldots \overline{z_{m-s-1}(t-1)} \cdot z_{m-s}(t-1),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$r_{2^{m-s}-1} = z_1(t-1) \ldots z_{m-s-1}(t-1) \cdot z_{m-s}(t-1)$$

и будем называть эти конъюнкции *конституентами*. Ясно, что каждое уравнение (9) состоит из логических произведений таких конституент, соединенных дизъюнктивной связью, например:

$$p_{j_1} v_{i_1} \vee p_{j_2} r_{i_2} \vee \ldots \vee p_{j_k} r_{i_k}, \tag{10}$$

где j и i могут принимать значения из

$$(0, 1, 2, \ldots, 2^p - 1)$$

и

$$(0, 1, 2, \ldots, 2^{m-s} - 1)$$

соответственно.

Рассмотрим, например систему на рис. 8. Пусть уравнения, описывающие эту систему написаны уже в дизъюнктивной совершенной нормальной форме:

$$\left.\begin{aligned} y_1(t) &= \overline{x_1(t)} \cdot x_2(t)\, z_1(t-1) \cdot \overline{z_2(t-1)} \vee x_1(t) \cdot x_2(t) \cdot \overline{z_1(t-1)} \cdot z_2(t-1), \\ y_2(t) &= x_1(t) \cdot x_2(t)\, z_1(t-1) \cdot z_2(t-1); \end{aligned}\right\} \tag{11'}$$

$$\left.\begin{aligned} z_1(t) &= x_1(t) \cdot \overline{x_2(t)} \cdot \overline{z_1(t-1)} \cdot z_2(t-1), \\ z_2(t) &= \overline{x_1(t)} \cdot \overline{x_2(t)} \cdot z_1(t-1) \cdot \overline{z_2(t-1)}. \end{aligned}\right\} \tag{11''}$$

Тогда можно построить булеву матрицу для системы (11′) так, чтобы элемент b_{ij} матрицы $\boldsymbol{B}_1$ был

$$b_{ij} = \begin{cases} 0 \text{ если } e^i\text{-том уравнении системы не} \\ \quad \text{участвует } j\text{-тая конституента } p_j; \\ r_k \text{ если } e^i\text{-том уравнении } j\text{-тая} \\ \quad \text{конституента участвует,} \end{cases}$$

где

$$i = 1, 2, \ldots, s;$$
$$j = 0, 1, \ldots, 2^p - 1,$$

и r_k та конституента, которая как раз в данной конъюнкции участвует. И так, для (11′) мы имеем:

$$\boldsymbol{B}_1 = \begin{matrix} 1 \\ 2 \end{matrix} \begin{pmatrix} 0 & z_1(t-1) \cdot \overline{z_2(t-1)} & 0 & \overline{z_1(t-1)} \cdot z_2(t-1) \\ 0 & 0 & 0 & z_1(t-1) \cdot z_2(t-1) \end{pmatrix}.$$
$$\begin{matrix} 0 & 1 & 2 & 3 \end{matrix}$$

Как видно, у этой матрицы столько столбцов, сколько существует конституент входа. Элементы матрицы, вообще говоря, суть функции состояния. Покажем простой метод перехода от этой матрицы к матрице реакций *Цетлина*.

Правило перехода можно сформулировать следующим образом. Рассматривая один из столбцов матрицы B_1, могут быть два существенно разных случая: либо все элементы этого столбца равны нулю, либо же есть хотя бы один ненулевой элемент. В первом случае, если на вход поступает комбинация сигналов, которая сделает равной 1 именно *ту* конституенту, номер которой равен номеру столбца, то на выходе системы будет комбина-

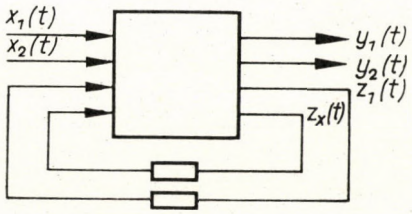


Рис. 8

ция сигналов (0, 0, . . . , 0), или короче 0. Если один или больше элементов данного столбца отличны от нуля, то есть они представляют собой конституенты состояния, то в зависимости от предыдущего состояния либо один, и только один из этих конституент примет значение 1, либо опять все элементы равны нулю. В первом случае на выходе появляется комбинация сигналов, равная двоичному числу, представленному этим столбцом если прочесть элементы сверху вниз, а во втором случае на выходе опять все сигналы нули.

Формально, если ввести обозначение e_β^* для s-мерного единичного вектора-строки, в котором единственный ненулевой элемент (равный 1) стоит на месте, определенным выражением $\log 2^\beta + 1$ (откуда видно, что такая вектор-строка составляется только для таких, β, которые суть целочисленные степени 2), то

$$l_{\alpha\beta} = \begin{cases} r_k & \text{если } [e_\beta^* \cdot b_\alpha^1] = r_k, \\ 1 & \text{если } b_\alpha^1 = 0 \text{ и } \beta = 0; \\ \overline{r_{k_1} \vee r_{k_2} \vee \ldots \vee r_{k_m}} & \text{при } b_\alpha^1 \neq 0 \text{ и } \beta = 0; \\ 0 & \text{если } b_\alpha^1 \neq 0 \text{ но } [e_\beta^* \cdot b_\alpha^1] = 0; \\ & \text{или} \\ & b_\alpha^1 = 0 \text{ и } \beta \neq 0, \end{cases}$$

где

$$\alpha = 0, 1, \ldots, 2^p - 1; \quad \beta = 0, 1, \ldots, 2^s - 1; \; b_\alpha^1 - \alpha\text{-тый}$$

столбец матрицы

$$B_1, [e_\beta^* \cdot b_\alpha^1]$$

скалярное произведение вектор-строки e_β^* и вектора-столбца

$$b_\alpha^1; \; a \; \overline{r_{k_1} \vee r_{k_2} \vee \ldots \vee r_{k_m}}$$

выражение, составленное из всех ненулевых элементов столбца b_α^1.

Доказательство такого метода перехода легко свести к теореме, доказанной в работе [23], и можно найти в работе [24].

Как известно, элемент $l_{\alpha\beta}$ матрицы реакций показывает, что если на входе системы появилась в данный момент комбинация сигналов, которую можно прочесть как двоичное представление десятичного числа α, то при каком состоянии в предыдущем такте дает система на выходе комбинацию сигналов, десятичным кодом которой служит β. Если этот элемент равен 0, то система никогда не может дать на данный вход рассмотренный выход. Если элемент равен 1, то в ответ на данный вход система на выходе выдает комбинацию сигналов, определяемой вторым индексом элемента, то-есть номером столбца матрицы реакций. Если же элемент есть функция состояния в предыдущий такт, то данный выход появляется при выполнении «*условия*», указанного логическим выражением составляющим эту функцию.

Например, рассмотрим вход

$$[x_1(t) = 0; \; x_2(t) = 1]$$

для системы (11′). Тогда из конституент входа лишь

$$\overline{x_1(t)} \cdot x_2(t) = p_1 = 1,$$

а все остальные равны 0. Определим, например, l_{12}. Берем $e_2^* = [1;0]$ и составим скалярное произведение

$$[e_2^* \cdot b_1^1] = [1; 0] \begin{bmatrix} z_1(t-1) \cdot \overline{z_2(t-1)} \\ 0 \end{bmatrix} = z_1(t-1) \cdot \overline{z_2(t-1)} = l_{12},$$

а для расчета, скажем l_{20}, надо взять вектор-столбец b_2^1, который равен 0. По правилу перехода $l_{20} = 0$, поскольку здесь $\beta = 0$. Продолжая определение всех элементов мы приходим к матрице реакций:

$$\boldsymbol{L} = \begin{array}{c|cccc} 0 & 1 & 0 & 0 & 0 \\ 1 & z_1(t-1) \vee z_2(t-1) & 0 & z_1(t-1) \cdot z_2(t-1) & \\ 2 & 1 & 0 & 0 & 0 \\ 3 & \overline{z_2(t-1)} & z_1(t-1) \cdot z_2(t-1) & \overline{z_1(t-1)} \cdot z_2(t-1) & 0 \\ & 0 & 1 & 2 & 3 \end{array}.$$

Аналогично составляется булева матрица B_2 для системы уравнений, определяющих следующее состояние системы. Элементы этой матрицы в общем случае будут логическими функциями входа. В нашем примере, для системы уравнений (11″)

$$\boldsymbol{B}_2 = \begin{array}{c|cccc} 1 & 0 & x_1(t) \cdot \overline{x_2(t)} & 0 & 0 \\ 2 & 0 & 0 & x_1(t) \cdot x_2(t) & 0 \\ & 0 & 1 & 2 & 3 \end{array}.$$

Эта матрица теперь однозначно определяет вид матрицы состояний по *Цетлину*. Правило перехода полностью аналогично правилу перехода для матрицы реакций:

$$a_{\alpha\beta} = \begin{cases} p_k \text{ если } [e^*_\beta \cdot b^2_\alpha] = p_k, \\ 1 \text{ если } b^2_\alpha = 0 \text{ и } \beta = 0; \\ \overline{p_{k_1} \vee p_{k_2} \vee \ldots \vee p_{k_n}} \text{ при } b^2_\alpha \neq 0 \text{ и } \beta = 0, \\ 0 \text{ если } b^2_\alpha \neq 0 \text{ но } [e^*_\beta \cdot b^2_\alpha] = 0, \\ \qquad\qquad \text{или} \\ \qquad b^2_\alpha = 0 \text{ и } \beta \neq 0. \end{cases}$$

Для системы уравнений (11″) это приводит к материце реакций:

$$\boldsymbol{A} = \begin{array}{c|cccc} & 0 & 1 & 2 & 3 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & \overline{x_1(t)} \vee x_2(t) & 0 & x_1(t) \cdot \overline{x_2(t)} & 0 \\ 2 & x_1(t) \vee x_2(t) & \overline{x_1(t)}\ \overline{x_2(t)} & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \end{array}.$$

На основе известных матриц реакций и состояния, а также и взаимных связей между отдельными частями системы можно рассчитать композицию этих систем, т. е. матрицы, описывающие работу составной системы. При этом можно пользоваться методом, данным *М. Л. Цетлин*ым в работе [25].

2. Естественный метод для представлении дискретных систем

Естественный метод, примененный авторами при представлении дискретных систем в вычислительной машине, состоит в следующем:

Система вводится в машину с помощью нескольких, по возможности

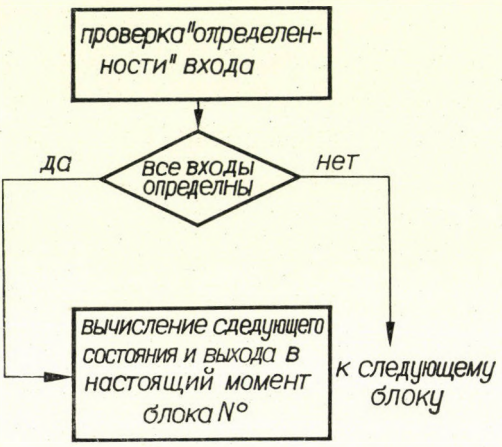


Рис. 9

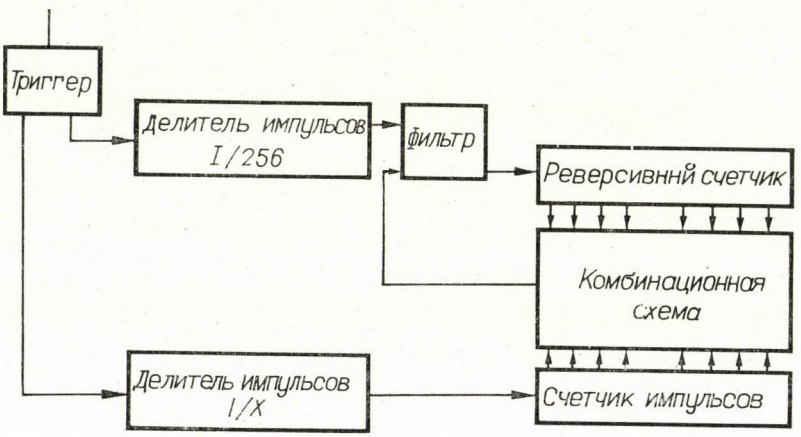


Рис. 10

немногочисленных данных, которые задают число основных блоков, их логическую структуру и число состояний, а также взаимные связи между блоками. В программе, разработанной авторами, возможно представление дискретной системы в виде таблиц, булевых матриц или же матриц реакций и состояний. Используя известные соотношения композиции схем, машина рассчитает полную матрицу реакций и состояний.

Основная схема работы алгоритма моделирования состоит из блоков для вычисления следующего состояния и настоящего выхода отдельных, главных частей системы (рис. 9). Чтобы сделать любые взаимосвязи между блоками допустимыми, все блоки состоят из двух главных частей, разделенных блоком условного перехода.

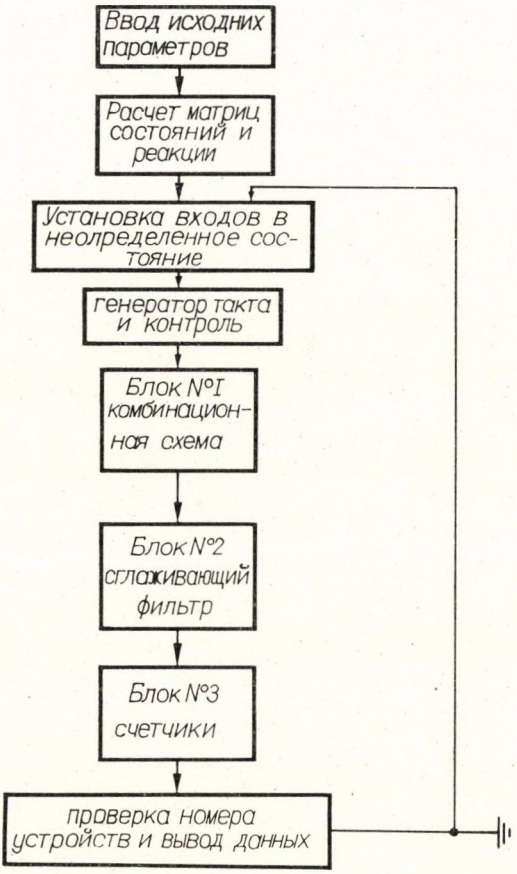


Рис. 11

«*Определенность*» входа означает следующее. Поскольку входные сигналы данного блока могут идти не только с входа всей системы, но и от другой части внутри системы, выходные сигналы которой в данном такте еще неизвестны, необходимо подождать пока эти сигналы вычисляются по естественной причинной цепи событий.

Был проведен эксперимент моделирования системы, показанной на рис. 10. Эксперимент был поставлен на вычислительной машине *Эллиотт 803* с программой, написанной на языке «*Автокод Марк – 2*». Небольшая ско-

рость вычислений и неприспособленность языка создали определенные неудобства, которые авторы намерены преодолеть переходом на более быстродействующие машины и к более совершенным языкам (*Алгол—60*). Система на рис. 10 состояла из 33-х триггеров и одной комбинационной схемы и 32-х

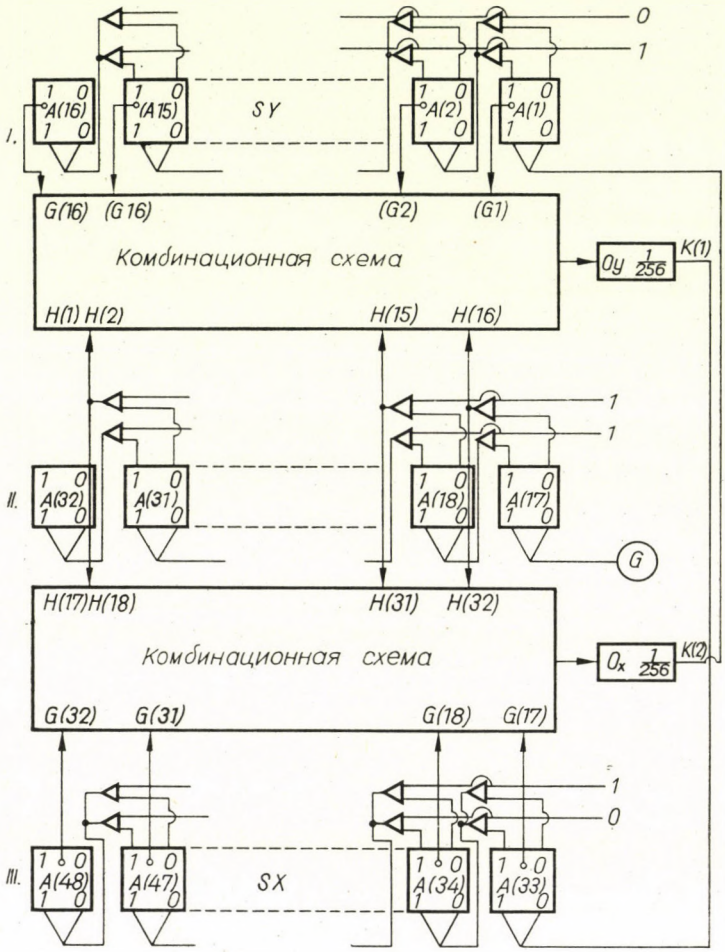


Рис. 12

схем совпадения. Один такт работы схемы занял на вычислительной машине около 4-х секунд. Принципиальная схема программы показана на рис. 11.

При разработке симуляционной программы для задачи, показанной на рис. 12, авторы перешли с универсального метода на специализированную. Это означает то, что программа организована уже в соответствии с потоком информации в данной схеме, и поэтому нет вынужденных остановок из-за нехватки информации на входе данного элемента или блока. В этом

случае отпадает необходимость и в введении лишнего состояния «определенность» в системе.

Эта программа была написана на языке *Алгол—60* для машины датской фирмы *GIER* [см. Приложение].

В программе имеются три процедуры:

TISZTITAS,

AZONOSITAS,

BISTABIL.

Процедура *TISZTITAS* после каждого цикла очищает входные и выходные полюсы комбинационных схем, а также внешние полюсы триггеров. Здесь установка состояния на Б не означало отдельное состояние, отличное от «0»-, и была введена лишь из-за соображений легкости контроля.

Процедура *AZONOSITAS* содержит фактически все параметры структуры и обеспечивает передачу информации внутри модели.

Процедура *BISTABIL* определяет выходные сигналы, а также следующее состояние триггеров.

Матрица $\boldsymbol{Y}[i, j]$ используется для определения выходов.

В программе имеются два вспомогательных такта $t = 1$ и $t = 2$ — за один реальный цикл. Эта задержка необходима для разъединения сигналов делителей частоты O_x и O_y от входного сигнала (см. рис. 12). Эта задержка конечно присутствует и в реальном устройстве.

За вспомогательный такт $t = 1$ вычисляются новые значения регистра *11*, а также выходы комбинационных схем. Во время же $t = 2$, имея выходы от O_x и O_y, вычисляются новые состояния и выходы регистров SX и SY.

Вспомогательная переменная h играет роль только в регулярном контроле программы и управляет выдачей группы из 4-х данных после каждого 128 цикла.

Программа приведена в приложении к данной статье.

ЛИТЕРАТУРА

1. Bush, V.: A Differential Analyzer. A New Machine for Solving Differential Equations. *Journal of the Franklin Institute* **212** (1931), 447.
2. Maginniss, F. J.: Differential Analyzer Applications. *General Electric Review* (1945), 54.
3. Shannon, D. E.: Mathematical Theory of the Differential Analyzer. *Journal of Mathematics and Physics* **18** (1939), 337.
4. Steele, F. G.—Sprague, R. E.—Wilson, B. T.: Digital Differential Analyzer. *U.S. Patent* No. 2.841,328 (заявлен).
5. Sprague, R. E.: Fundamental Concepts of the Digital Differential Analyzer Method of Computation. *Mathematical Tables and Other Aids to Computation* **6** (1952); No. 37, 41.
6. Hatvany, J.: Számjegyes szerszámgépvezérlés (Цифровое управление станками). MTA Kibernetikai Kutató Csoportja, Отчёт (рукопись).
7. Kardos, P.—Pödör, B.: Számlálórendszerű szorzómű (Умножитель частоты). *Mérés és Automatika* **12** (1964), 140.
8. Hatvany, J.: Számlálórendszerű műveleti elemek ipari felhasználása. *Mérés és Automatika* **12** (1964), 129.
9. Meyer, M. A.: Digital Techniques in Analog Systems. *IRE Transactions on Electronic Computers EC 3* (1954); No. 2, 23.

10. Gordon, B. M.: Adapting Digital Techniques for Automatic Controls, I. *Electr. Mfg.* (1954); Nov., 136.
11. Lundh, Y.: Sifferteknikk for enkle regneoperasjoner, II. *Teknisk Ukeblad* (1958); No 40, 105. Оттиск 7 стр. (на норвежском). Основные положения опубликованы на английском языке: Lundh, Y.: Digital Techniques for Small Computations. *J. Brit. I.R.E.* **19** (1959); No. 1, 37.
12. Johnson, E. C.: Digital Operational Techniques (Computer Handbook, Ed. Huskey, H. D.; Section 21.5.3.). McGraw-Hill, New York 1962, 21—94.
13. Wood, P.: A Frequency Meter with Continuous Digital Presentation. *J. Brit. I.R.E.* **26** (1963); No. 2, 109.
14. Воронов, А. А., Соколов, Г. Н.: Устройство для програмирования кривых второго порядка, основанное на цифровых интеграторах. *Автоматика и Телемеханика,* **20** (1959), № 2.
15. Гарбузов, А. Р.: Точность интегрирования ступенчатой функции (Гл. VIII. Цифровые аналоги для систем автоматического управления). Под ред. Воронова, А. А. и др. Изд. АН СССР, Москва—Ленинград. 1960, 163—176.
16. Schmid, H.: An Operational Hybrid Computing System Provides Analog-Type Computation with Digital Elements. *IEEE Trans. on El. Computers* (1963); Dec., 715.
17. Forbes, G. F.: Digital Differential Analyzers, 4 издание. Издание автора, Pacoima (Calif.) 1957.
18. Rechberger, H.: The Application of Digital Differential Analyzers in Control Loops. Доклад № 065/1 на II-ом конгрессе ИФАК, Basel 1963.
19. Kardos, P.: Szállítószalag mérleg digitális mérésadatfeldolgozó berendezése (Цифровое устройство для обработки данных конвеерных весов). *Mérés és Automatika* **13** (1965); No. 2—3, 89.
20. Simon, Z.-né: Impulzus digitális folyamatos adagoló berendezés (Цифровое устройство непрерывной пропорциональной дозировки). *Finommechanika* **4** (1965); No. 3, 81.
21. Hatvany, J.—Kardos, P.: Digitális interpolátorok számjegyes szerszámgépvezérléshez (Цифровые интерполяторы для управления металлорежущими станками). *Mérés és Automatika* **14** (1966), No. 1.
22. Кобринский Н. Е.—Трахтенброт Б. А.: Введение в теорию конечных автоматов. Физматгиз, Москва 1961.
23. Barát, J.: Digitális autonóm rendszerek, (Автономные цифровые системы). *Mérés és Automatika* (1962), 11—12.
24. Barát, J.: Szekvenciális áramkörök szintézise és analízise (Синтез и анализ секвенциальных схем).
25. Цетлин М. Л.: О примитивных схемах; Проблемы кибернетики. Вып. I. Москва 1958.

ПРИЛОЖЕНИЕ

PROGRAM PROTOKOLL

```
Szimulacio
begin comment Szimulacios program;
integer array A[1:48.1:7]. Y[0:31]. G[1:32]. H[1:32]. K[1:2].
      N[1:2]. P[:2]; integer SX.SY.h. t.i j.e.k.p.v;
real u; procedure TISZTITAS;
begin for i: = 1 step 1 until 32 do
      G[i]: = H[i]: = 5;
        for i: = 1 step 1 until 48 do
        for j: = 3 step 1 until 7 do
          A[i.j]: = 5;
end;
procedure AZONOSITAS; begin A[1.7]: = K/2/;
      A[33.7]: = K[1];
      for i: = 1 step 1 until 16 do
        begin A[i.4]: = A[i+32.3]: = 1;
          A[i.3]: = A[i+32.4]: = 0; A[i+16.3]: = A[i+16.4]: = 1;
          end;
          for i: = 2 step 1 until 16 do
          begin A[i.7]: = A[i—1.6]; A[i+16.7]: = A[i+15.6];
```

```
            A[i+32.7]: = A[i+31.5],
            end;
            for i: = 1 step 1 until 16 do
            H[i]: = H[i+16]: = A[33—i.5];
            if t = 2 then go to T2;
            for i: = 1 step 1 until 16 do
            begin G[i]: = A[i.2]; G[i+16]: = A[i+32.2];
            end;
T2: end;
procedure BISTABIL(k);
value k; integer k;
begin p; = A[k.7]×4+A[k.4]×2+A[k.3];
v: = 0; go to Ell; E8:v: = j+1;
        Ell: for j: = v step 1 until 3 do begin
        if Y[p×4+j] = 1 then goto El else if Y[p×4+j] = 5 then goto E2
        else if Y[p×4+j] = 6 then goto E3; end;
        El: if j = 0 then goto E4 else if j = 1 then goto E5 else if j = 2 then goto E6;
        E7:A[k.5]: = 1;
           A[k.6]: = 1;
        goto ALLAPOT;
        E4:A[k.5]: = 0;
           A[k.6]: = 0;
        goto ALLAPOT;
        E5:A[k.5]: = 0;
           A[k.6]: = 1;
        goto ALLAPOT;
        E6:A[k.5]: = 1;
           A[k.6]: = 0;
        goto ALLAPOT;
        E2:if t = 1 then goto E12;
if A[k.1] = 0 then goto E8 else goto E1; E12:if A[k.2] = 0 then goto E8 else goto E1;
        E3:if t = 1 then goto E13;
if A[k.1] = 1 then goto E8 else goto E1; E13:if A[k.2] = 1 then goto E8 else goto E1;
        ALLAPOT: if t = 2 then goto E14; if A[k.2] = 1 then goto E9 else goto E10;
        E9:if A[k.7] = 1 then A[k.1]: = 0 else A[k.1]: = 1;
        goto OUT;
        E10:if A[k.7] = 0 then A[k.1]: = 0 else A[k.1]: = 1;
E14:if A[k.1] = 1 then goto E15 else goto E16;
E15:if A[k.7] = 1 then A[k.2]: = 0 else A[k.2]: = 1;
goto OUT;
E16: if A[k.7] = 0 then A[k.2]: = 0 else A[k.2]: = 1;
OUT end outtext (≮<Szimulacios program 1966≯) i: = typein output (≮nd dd≯.i);
        for i: = 0 step 1 until 31 do
        input(Y[i]);
        TISZTITAS;
        begin comment KEZDETI FELTETELEK; h: = 0;
        t: = 1;
for i: = 1 step 1 until 48 do
for j: = 1 step 1 until 2 do
input(A[i.j]);
input(P);
K[1]: = K[2]: = 0;
        end;
L10 begin comment MUEKOEDTETES; h: = h+1;
        if t = 2 then goto T22;
        A[17.7]: = 1;
        goto L2;
T22: AZONOSITAS;
if K[1]+K[2] = 0 then goto L3;
        for i: = 1 step 1 until 16 do
        A[i+16.2]: = A[i+16.1];
        goto L4
```

```
L3: for i: = 1 step 1 until 16 do
        begin A[i.2]: = A[i.1];
        A[i+16.2]: = A[i+16.1];
        A[i+32.2]: = A[i+32.1];
        end;
        end;
        goto L9;
L2: begin comment t = 1 KEZDETE;
        for k: = 17 step 1 until 32 do
        begin AZONOSITAS;
        BISTABIL(k);
        end;
for i: = 1 step 1 until 16 do
begin  A[i.1]: = A[i.2];
       A[i+32.1]: = A[i+32.2];
end;
        begin comment LOGIKAAK;
        e: = 0;
        AZONOSITAS;
        for i: = 0 step 1 until 15 do
        e: = e+H[i+1]×G[i+1];
        if e≥1 then N[1]: = 1 else N[1]: = 0;
        e: = 0;
        for i: = 0 step 1 until 15 do
        e: = e+H[i+17]×G[i+17];
        if e>1 then N[2]: = 1 else N[2]: = 0;
        P[1]: = P[1]+N[1];
        P[2]: = P[2]+N[2];
if P[1] = 256 then K[1]: = 1 else goto L19;
P[1]: = 0;
L19: if P[2] = 256 then K[2]: = 1 else goto L20;
P[2]: = 0;
L20: end;
end t = 1;
        goto L9;
L4: begin comment t = 2 KEZDETE;
L12  if K[1] = 1 then goto L13 else goto L16;
L13  for k: = 33 step 1 until 48 do
        begin AZONOSITAS;
        BISTABIL(k);
        end; K[1]: = 0;
goto L14;
L16  for i: = 33 step 1 until 48 do
A[i.2]: = A[i.1];
L14  if K[2] = 1 then goto L15 else goto L17;
L15  for k: = 1 step 1 until 16 do
        begin AZONOSITAS;
        BISTABIL(k);
        end; K[2]: = 0;
goto L9;
L17: for i: = 1 step 1 until 16 do
A[i.2]: = A[i.1];
        L9: if t = 1 then goto L18;
SX: = 0;
SY: = 0;
for i: = 0 step 1 until 15 do begin
SX: = SX+A[i+33.2]×2 ↑ i;
SY: = SY+A[i+1.2]×2 ↑ i;
end; ifh>256 h<266 then output(≮nddddddddddd≯.SX.SY);
u: = abs(2 ↑ 16—1—sqrt(SX ↑ 2+SY ↑ 2)); if h<256 h<266 then
begin output(≮ndddd.ddddddddd≯.u); outcr; end; if h = 266 then h: = 0;
if u<1 then goto L21;
outcr;
```

```
output(≮n.dddd≯.u);
outcr; L21:
        TISZTITAS;
        t: = 1; if kb on then goto L23;
        goto L10; L18: t: = 2;
        TISZTITAS;
        goto L10;
        end; L23:outsp(30); for i: = 1 step 1 until 48 do
          for j: = 1 step 1 until 2 do
begin output(≮nd≯.A[i.j]);
outcr; end; output(≮nd≯.P[1] outcr, P[2]); end.
```

The Use of Counting Technique Instruments in Industrial Automation Equipments *(J. Hatvany, P. Kardos* and *J. Barát)*. After clarifying the concepts the paper deals with the dimensioning problems of the combined system. A counting procedure checked by reproducible measurements on an experimental installation is proposed instead of the completely empirical methods, which were proposed earlier. For dimensioning and checking more complicated systems the only practicable way is simulation and evaluation carried out on a digital computer. The paper deals with the problem of a generally valid simulation program based on the employment of the theory of finite automates. It discusses the theoretic problems of the application of the chosen principle and covers its proof for the case of systems of the given type. The experimental checking of the computer simulation procedure was carried out through the simulation of the measuring arrangement outlined earlier. The results delivered by the simulation program give the same values as the results gained by direct measurement.

Anwendung zähltechnischer Mittel bei industriellen Automatik-Anlagen *(J. Hatvani, P. Kardos* und *J. Barát)*. Nach Klärung der auftauchenden Begriffe befaßt sich der Aufsatz mit den Bemessungsproblemen zusammengesetzter Systeme. Ein auf einer Versuchseinrichtung durch reproduzierbare Messungen kontrolliertes Rechenverfahren wird anstatt der früher beschriebenen vollkommen empirischen Methoden vorgeschlagen. Zur Bemessung und Kontrolle von sehr komplizierten Systemen ist die mit Digitalrechnern ausgeführte Simulation und Auswertung der einzig gangbare Weg. Der Artikel behandelt Probleme eines auf der Anwendung der Theorie der endlichen Automaten beruhenden allgemeingültigen Simulationsprogrammes. Er bespricht die theoretischen Probleme der Anwendung des gewählten Prinzips und enthält den Beweis seiner Gültigkeit für den Fall von Systemen des gegebenen Typs. Die experimentelle Kontrolle des Simulationsverfahrens mit Rechenmaschinen erfolgt über die Simulation der früher demonstrierten Meßanordnung. Die mittels des Simulationsprogrammes gewonnenen Resultate ergaben mit den unmittelbaren Meßergebnissen identische Werte.

L'emploi des techniques opérationnelles dans les systèmes d'automatisation industrielle *(J. Hatvani, P. Kardos* et *J. Barát)*. Les unités et systèmes des techniques opérationnelles permettent d'établir, à partir d'un choix restreint d'unités opérationnelles, des simulateurs et des systèmes d'automation à l'échelle de temps réelle, dans les limites définies par SHANNON. Les notions utilisées une fois définies, l'article s'occupe des problèmes de dimensionnement des systèmes combinés. Un procédé de calcul contrôlé par des mesures reproduisibles sur un appareillage d'essai est proposé au lieu des méthodes intégralement empiriques décrites plus tôt. Pour le dimensionnement et le contrôle des systèmes plus compliqués, le seul chemin pratiquable est la simulation et l'évaluation effectuées sur une calculatrice digitale. L'article s'occupe des problèmes d'un programme de simulation de validité générale, fondé sur l'utilisation de la théorie des automates finis. Il discute des problèmes théoriques ayant trait à l'application du principe choisi et contient la démonstration de sa validité pour le cas des systèmes du type donné. La vérification expérimentale du procédé de simulation par calculatrice se faisait par simulation de l'unité opérationnelle déjà décrite. Les résultats fournis par le programme de simulation présentaient des valeurs identiques aux résultats de mesures directes.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 55—66 (1968)

SOME CONCEPTS ON THE PROGRAMMING OF DIRECT DIGITAL PROCESS CONTROL SYSTEMS

J. GERTLER

AUTOMATION RESEARCH INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST

[Manuscript received April 10, 1967]

The paper deals with some programming problems of the centralized direct digital control of continuous industrial processes. It shows that the program language ALGOL 60 can be extended for process control purposes and introduces the concept of classes and forms of variables. As standard procedure of the extended program language it summarizes the fundamental operations of direct digital process control and it contains considerations regarding the optimum form of the standard procedures. In order to simplify programming it introduces as a further extension of ALGOL 60 the concept of latent and substitution parameters and demonstrates the use of the extended language on a sample program.

I. Introduction

The centralized direct digital control (DDC) of continuous industrial processes has an increasing importance nowadays. The heart of most DDC systems is a universal process control computer. The programming of such computers, as was pointed out by several authors, is an extremely time-consuming and expensive task. The lack of a generally accepted high-level DDC programming language was specified as one of the main sources of the difficulties.

An attempt for solving some problems of DDC programming is presented in this paper. Two extensions of ALGOL 60 for process control purposes are proposed, the standard procedures of DDC are collected and some new considerations and results are summarized regarding the optimal form of these procedures.

II. The first extension of ALGOL 60

The algorithmic language ALGOL 60 [2], although it was constructed primarily for the programming of arithmetic and logic problems, is potentially suitable for DDC programming, too. Its exactly defined structure, high flexibility, and a set of powerful programming facilities (such as logical variables, conditional expressions and statements, arrays, procedures, etc.) make ALGOL 60 a possible and practical basis for a DDC programming language. This choice is supported also by the general use of ALGOL 60. A DDC language, based on the

ALGOL 60, may be utilized not only for the programming of a process control computer, but as a publication language of DDC, too.

The DDC programming, on the other hand, requires some special features of the language. The necessary extensions will be treated in the following chapter.

1. The classes of variables

An arithmetic computer has a relatively restricted connection with its environment, through the alphanumeric peripheral units only. The input and output informations, in this case, are connected with variables within the machine only. If the computer controls a process, however, its connection with the outside world (especially with the controlled process) broadens considerably, and the input and output informations (the informations to and from the process) belong to variables out of the machine, too. In the latter case, therefore, besides the usual internal variables there also exist external variables. The internal variables may be divided into operative and fixed ones (depending on the sort of memory which they are stored in) and the external variables may be input or output ones. By introducing the new concept of the variable-classes, every DDC variable may be arranged into one of the four classes: operative, fixed, input and output.

The possible effect of the program, regarding the variables, depends upon the class of the given variables. While there is no restriction concerning the operative variables, the program can only be passive, if the variable belongs to the fixed or input class, and only active, if it is an output one. This means, that fixed and input variables are not allowed to appear on the left side of assignment statements, and output ones are not allowed to stand in expressions.

Because of the practical significance of the variable-classes, not only the type but also the class of the variables must be declared and specified. For this purpose the class declarators — **fixed, input** and **output** — have been introduced. The operative variables, for convenience sake, are declared and specified by empty class-declarator.

Regarding the non-operative variables some additional restrictions should be considered:

a complete identifying list must be supplied on the compilation of the program;

a mutually unique correspondence is required between the several variables and their identifiers; the *own* concept is inapplicable.

Some declaration examples are given at the end of the chapter.

2. The form of the variables

If our language serves for the programming of a real process control system (not only for publication purposes), a further concept has to be introduced regarding the external variables: the form of the variables. The form means the real technical form of the variables at stepping in or out of the range of programm (i.e. the part of the control system, working under the direct guidance of the programm). The knowledge of the form of the external variables is essential for the proper performance of the program, as this information determines the sort of the peripheral units to be employed.

The proposed form assortment and the corresponding form-declarators are given below:

Forms	Declarators
Two-state variables	
Digital variables	
binary digital variables	
parallel b.d.v.	**(bp)**
series b.d.v.	**(bs)**
coded digital variables	
parallel c.d.v.	**(cp)**
series c.d.v.	**(cs)**
direct digital variables	
parallel d.d.v. (state of a selector)	**(dp)**
series d.d.v. (number of pulses)	**(ds)**
Analogue variables	
continuous analogue variables	
low-level c.a.v. (millivolts)	**(al)**
high-level c.a.v. (volts)	**(ah)**
discontinuous analogue variables	
parallel d.a.v. (amplitude-coded pulse)	**(ep)**
series d.a.v. (duration-coded pulse)	**(es)**
frequency	**(f)**

3. Declaration examples

Fixed real *c1, c2, c3;*
input boolean *a1, a2;*
input integer (bp) *b1, b2;* **(dp)** *b3, b4, b5,* **(ds)** *b6;*
input real (al) *c4, c5* **(ah)** *c6, c7, c8;*
output integer (cp) *b7, b8,* **(ds)** *b9;*
output real (ep) *c9,* **(es)** *c10, c11.*

III. The collection of the standard functions

A survey of the usual basic DDC operations is given in this section, in accordance with the literature [3, 4]. The basic operations have been written as standard functions in the extended ALGOL language. The DDC standard functions show a hierarchy. Sampling and time-manipulations form the lowest level, differentiation and integration the middle, and the various monitoring and controlling operations the highest one.

1. The lowest level

a) Sampling

The function

$$samp\ (y, a, b) \tag{A1}$$

produces the value $a \cdot y + b$, when y is usually an input variable, a and b are constants. This operation is not indispensable for activating an input variable, but may be economic, if the same variable is used for a few operations.

b) Time-operations

The time data have a one week period, and are stored in homogeneous form (in pure second units). There is a pair of functions for the conversions of these data and another for their adding and substracting.

2. Differentiation and integration

The differentiation and integration algorithms are intended for cyclic action: they use some results from the last cycle and store new ones for the next. The differentiating procedure

$$der(y, yl, tl) \tag{A2}$$

realizes the formula

$$D_n = \left[\frac{dy}{dt}\right]_n \approx \frac{y_n - y_{-1}}{t_n - t_{n-1}} \tag{F1}$$

where the parameters mean y_n, y_{n-1} and t_{n-1}, respectively.

The integrating algorithms

$$int0(y, int, tl) \text{ and } int1(y, yl, int, tl) \tag{A3}$$

differ in the method of approximation: they are based, respectively, no the relationships

$$I_n = \int_{t_0}^{t_n} y dt \approx \sum_{j=1}^{n} y_j \cdot (t_j - t_{j-1}) = I_{n-1} + y_n \cdot (t_n - t_{n-1}), \tag{F2}$$

$$I_n = \int_{t_0}^{t_n} y dt = \sum_{j=1}^{n} \frac{y_j + y_{j-1}}{2} \cdot (t_j - t_{j-1}) = \\ = I_{n-1} + 0{,}5 \cdot (y_n + y_{n-1}) \cdot (t_n - t_{n-1}). \tag{F3}$$

The parameter *int* refers to I_{n-1} while the meaning of the others is the same as in (A2).

3. *The highest level*

a) Monitoring

The limit-monitoring procedure

$$limr(y, ub, lb) \tag{A4}$$

generates a three-state result depending on the position of the variable y as compared to the upper and lower bounds (*ub* and *lb*) and produces a printed announcement, if the variable is out of limits.

The trend-monitoring function

$$trend(y, yl, trub, trlb, tl) \tag{A5}$$

is a *limr*-like procedure for the first derivate of the variable y. Here *trub* and *trlb* are the limits, *yl* and *tl* are y_{n-1} and t_{n-1}, respectively.

The last pair of the monitoring algorithms

$$av0(y, int, tint, tl) \text{ and } avl(y, yl, int, tint, tl) \tag{A6}$$

compute the time-average of variable y, based on the integrating formulas (F2) and (F3), respectively. The parameter *tint* is the integrated value of time.

b) Control

Two groups of the control procedures have been developed: the direct and the recoursive ones.

The direct procedures are simply based on the general control algorithm:

$$v(t) = K_0 + K_P \cdot e(t) + K_I \cdot \int_{t_0}^{t} e(\tau)\, d\tau + K_D \cdot \frac{de(t)}{dt}, \tag{F4}$$

when

$$e(t) = y_0 - y(t) \tag{F5}$$

and y_0 is the set-point for y.

Function

$$p(y, sp, k0, kp) \tag{A7}$$

involves only the constant and proportional terms.

Functions

$$pi0(y, sp, kp, ki, erint, tl)$$

and

$$pil(y, sp, kp, ki, erl, erint, tl) \tag{A8}$$

consist of the proportional and integral term: the evaluation of the latter is carried out in accordance with the rules (F2) and (F3), respectively.

Finally the procedure

$$pid(y, sp, kp, ki, kd, erl, erint, tl) \tag{A9}$$

implies the proportional, integral [evaluated by (F3)] and differential [evaluated by (F1)] terms. In these algorithms *sp* is the setpoint, *k*0, *kp*, *ki* and *kd* are the constants, *erl* is the last and *erint* is the integrated value of the error *e*. If integration is involved, the constant term is unnecessary.

The recoursive (velocity) algorithms generate, instead of the total value of the variable v, only the increment Δv:

$$\begin{aligned}\Delta v(t_n) = K_P \cdot [e(t_n) - e(t_{n-1})] + K_I \cdot \int_{t_n-1}^{t_n} e(\tau)\, d\tau + \\ + K_D \cdot \left[\frac{de(t_n)}{dt} - \frac{de(t_{n-1})}{dt}\right] \approx \\ \approx K_P(e_n - e_{n-1}) + K_I \cdot 0{,}5 \cdot (e_n + e_{n-1}) \cdot (t_n - t_{n-1}) + \\ + K_D \left[\frac{e_n - e_{n-1}}{t_n - t_{n-1}} - \left(\frac{\Delta e}{\Delta t}\right)_{n-1}\right].\end{aligned} \tag{F6}$$

Procedure

$$rpi(y, sp, kp, ki, erl, tl) \tag{A10}$$

implies the terms, corresponding to proportional and integral actions, and procedure

$$rpid(y, sp, kp, ki, kd, erl, erdt, tl) \tag{A11}$$

performs the entire algorithm (F6). The parameter *erdt* means the last value of the error-derivative. A recoursive algorithm, implying only the first term of (F6), would be unusable.

IV. The optimal structure of the standard functions

All the standard functions, which involve differentiation or integration, require the calculation of a time-difference and the storage of a time- and, occasionally, of a parameter-value. These additional operations are rather time

and memory space consuming. Considerable savings can be achieved, if the additional operations are, whenever it is possible, carried out in common for a few procedures. These considerations raise the demand, that the structure of the standard functions should permit the programmer to freely decide concerning the application or omission of every single additional operation.

The detailed studies, taking into account the view-points of the required running-time and memory-space as well as those of the programming ease, have led to the result, that the optimal solution of the problem may be obtained, if the procedures concerned are developed in several variants, which differ from each other only in the additional operations. The variants are denoted by the characters *p* and *t*, denoting the parameter and time manipulations, respectively. The *der* procedure, for example, was developed in four variants as follows:

$$der(y, yl);\ derp(y, yl);\ dert(y, yl, tl);\ derpt(y, yl, tl). \qquad \text{(A12)}$$

The *der* variant implies no additional operations, *derp* implies the parameter-one and *dert* the time-one only, and *derpt* implies both.

The performed estimations have shown, that the probable savings, due to the introduction of the procedure-variants, amount to about 30 per cent in running-time and 45 per cent in operative memory space (regarding the standard functions only).

The application of the standard functions will be illustrated in section VI. 1.

V. The second extension of the ALGOL 60

The standard functions of DDC are multi-parameter procedures. This situation causes some inconveniences in writing and reviewing of the programs. In order to overcome these difficulties a second extension of the ALGOL 60 has been accomplished.

1. Latent parameters

The parameters of a standard function are not of the same rank from the point of view: of process control one of the parameters (the first in the list) is particularly characteristic of the given application of the procedure. This parameter, which is usually a direct or deduced process-variable in the actual parameter list, may be regarded as the basic parameter of the procedure. If all the other parameters are assigned to the corresponding basic parameters in advance, it is sufficient to denote the actual basic parameter, when the procedure is called. In this case all the parameters except the basic one are latent parameters of the procedure.

The notation of the latent parameters is implemented by standardized composite identifiers. These identifiers consist of two parts, separated by a dot. The first part is the identifier of the basic parameter, which the latent one belongs to, and the second one specifies the role of the given parameter within the procedure, in accordance with the designations in the formal parameter lists of the functions (A2)...(A11). The identifier *z.erint*, for example, denotes the actual parameter, which is to replace *erint* if the function *pil*(*z*) or *pid*(*z*) is called.

The latent parameters are declared and treated (outside of the procedures) as ordinary variables.

If a set of basic parameters form an array, the corresponding latent parameters are also declared as arrays, but the bound pair lists must be supplied only with the basic parameters. The subscripts of elements of latent parameter arrays take place in the first part of the composite identifiers:

$$z[i].erint,$$

for example.

The introduction of the latent parameters raises some difficulties as the actual parameters of a procedure must not be selected freely within the program and a basic parameter must not be but a variable (single or subscripted). These difficulties however, if they arise, may be easily eluded.

The application of the latent parameters is illustrated by a simple example:

```
begin integer i;
      input array y[1:10];
      fixed array y.a, y.b;
      array z[1:10];
      for i := 1 step 1 until 10 do;
          z[i] := samp(y[i]);                    (A13)
end.
```

2. *Substituting variables*

If the value of a latent parameter is the same for every element of a basic parameter array, it is uneconomical to treat this latent parameter as an array. The situation is similar, if a procedure is called in a cycle with the consecutive elements of an array as basic parameters, and the actual value of these elements is generated immediately before the procedure-call. The idea of the substituting parameters was introduced for these cases.

The substituting variable is a simple variable, which replaces all the elements of an array (basic or latent) whenever such an element is applied as an actual parameter (basic or latent) of a procedure. The identifier of a

substituting parameter is identical to that of the substituted array. The substituting parameters are declared as simple variables, marked one by one by the declarator **subst.** They are treated (outside of the procedures) as ordinary variables.

If a procedure is applied to elements of arrays, the subscripts of the basic parameter, in order to assign the adequate basic parameters, must be indicated even if the array of the basic parameters is substituted. If a basic parameter is substituted, the bound pair list has to be supplied with the first non-substituted latent parameter array.

The substitution of a latent parameter array may be illustrated by the modification of the program (A13) in such a manner, that $y[i] \cdot a = y \cdot a$ for every value of i:

```
begin integer i; fixed real subst y·a;
      input array y[1:10];
      fixed array y·b;
      array z[1:10];
      for i :=1; step 1 until 10 do;
          z[i] := samp(y[i]);                    (A14)
end.
```

To illustrate the substitution of the basic parameters our program has been somewhat extended:

```
begin integer i;
      fixed real subst z · k0; real subst z;
      input array y[1:10]; array v[1:10];
      fixed array y · a, y · b, z[1:10] · sp, z · kp;
      for i := 1 step 1 until 10 do;
      begin;
      z := samp(y[i]);
      v[i] := p(z[i]);
      end;                                       (A15)
end.
```

VI. Illustrative program

An illustrative program will be presented in two versions. The work is the cyclic watching and control of 100 process-variables: limit and trend monitoring, calculation of the parameter and error-square averages and generating control values by PID algorithm. While the first program-version is based on the first extension of ALGOL 60 and the collection of the multivariant standard functions only, also the second extension is applied to the second version.

1. The first program-version

```
begin
    real z, h; integer i;
    real array zl, tl, int, tint, ersqint, ersqtint,
              erl, erint [1:100];
    fixed real array a, b, ub, lb, trub, trlb, sp, kp,
                     ki, kd[1:100];
    input real array (al) y[1:100];
    output integer array (bp) v, (dp) rep, trep [1:100];
    ⋮
    for i :=1 step 1 until 100 do;
    begin;
      z := samp(y[i], a[i], b[i]);
      rep[i] := limr(z, ub[i], lb[i]);
      trep[i] := trendt(z, zl[i], trub[i], trlb[i], tl[i]);
      h := avlp(z, zl[i], int[i], tint[i]);
      h := avl((sp[i]-z) ↑ 2, erl[i] ↑ 2, ersqint[i], ersqtint[i]);
      v[i] := pidp(z, sp[i], kp[i], ki[i], kd[i], erl[i], erint[i])
    end;
    ⋮
    end.                                                    (A16)
```

2. The second program-version

```
begin;
    real h, subst z, subst ersq, subst ersq · l;
    integer i;
    input real array (al) y[1:100];
    real array z[1:100] · l, z · tl, z · int, z · tint,
               ersq[1:100] · int, ersq · tint, z · erl, z · erint;
    fixed real array y · a, y · b, z · ub, z · lb, z · trub;
                     z · trlb, z · sp, z · kp, z · ki, z · kd;
    output integer array (bp) v, (dp) rep, trep[1:100];
    ⋮
    for i := 1 step 1 until 100 do;
  begin;
    z := samp(y[i]);
    rep[i] := limr(z[i]);
    trep[i] := trendt(z[i]);
```

```
h := avlp(z[i]);
ersq := (z[i] · sp-z) ↑ 2;
ersq · l := z · erl ↑ 2;
h := avl(ersq[i]);
v[i] := pidp(z[i])
end;
⋮
end.                                                    (A17)
```

REFERENCES

1. GERTLER: Közvetlen digitális folyamatirányító rendszerek programozása és organizációja (Programming and Organization of Direct Digital Process Control Systems; in Hungarian). Dissertation, Budapest 1966.
2. NAUR: Revised Report on the Algorithmic Language ALGOL 60. *The Computer Journal* (1963), 1.
3. MORLEY—CUNDALL: The Ferranti System and Experience with Direct Digital Control. *Process Control and Automation* (1965), 7.
4. *IEEE* International Convention Record. Symposium on the Direct Digital Control, 1965

Einige Probleme der Programmierung unmittelbarer Digital-Prozeß-Steuersysteme. Der Aufsatz behandelt einige Probleme der Programmierung des zentralisierten unmittelbaren Digitalsteuerns von kontinuierlichen Industrieprozessen. Er weist nach, daß die Programmsprache ALGOL 60 für Zwecke der Prozeßsteuerung erweitert werden kann und führt den Begriff der Klassen und Formen der Veränderlichen ein. Als Standardverfahren der erweiterten Programmsprache faßt er die grundlegenden Verfahren des unmittelbaren Digital-Prozeßsteuerns zusammen und enthält Überlegungen bezüglich der optimalen Form der Standard-Verfahren. Zur Vereinfachung der Programmierung führt er als weitere Ausdehnung der ALGOL 60 den Begriff der latentenund Substitutionsparameter ein und veranschaulicht die Anwendung der erweiterten Sprache an einem Musterprogramm.

Quelques problèmes de la programmation des systèmes de commande digitale directe des processus. L'article traite de quelques problèmes de programmation de la commande digitale directe centralisée des processus industriels continus. Il démontre que le langage algorithmique ALGOL 60 peut être développé pour les besoins de la commande des processus et introduit les notions de classe et de forme des variables. Comme procédé standard du langage algorithmique développé, l'article résume les opérations fondamentales de la commande digitale directe des processus et contient des considérations relatives à la forme optimum des procédés standard. Pour simplifier la programmation, il introduit comme nouveau développement de l'ALGOL 60 la notion des paramètres latents et de substitution, puis illustre l'emploidu langage développé sur un programme modèle.

Некоторые вопросы программирования систем непосредственного цифрового управления процессами *(Я. Гертлер)*. Статья занимается несколькими вопросами программирования централизованного непосредственного цифрового управления непрерывными промышленными процессами. Показано, что программный язык ALGOL 60 можно так же использовать для целей управления процессами, и при этом вводится понятие классов и форм переменных. В качестве стандартного метода распространенного программного языка обобщаются основные операции непосредственного цифрового управления процесса и приводятся соображения по оптимальной форме стандартных методов. В интересах упрощения программирования в качестве дальнейшего распространения языка ALGOL 60 вводится понятие латентных и субституционных параметров, а после этого иллюстрируется применение распространенного языка на образцовой программе.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 67—74 (1968)

СРЕДНЯЯ ИНТЕНСИВНОСТЬ ОТКАЗОВ НА ЗАДАННОМ ИНТЕРВАЛЕ ВРЕМЕНИ РАБОТЫ АППАРАТУРЫ

Ж. Б. ЛИНКОВСКИЙ

В работе рассматриваются определение данной характеристики надежности, доверительный интервал для истинного значения, а также средняя надежность на заданном интервале времени и вопросы ее оценки по выборке.

I. Определение данной характеристики надежности*

Если $P(t)$ - надежность аппаратуры на отрезке времени $[0; t]$, и при $t = 0$ она была исправной, $P(0) = 1$, то

$$P(t) = e^{-\int_0^t \lambda(t)dt}; \quad t \geqslant 0, \tag{1}$$

где $\lambda(t)$ - интенсивность отказов. В ряде случаев интересуются средним значением $\lambda_{cp,T}$ интенсивности отказов на заданном интервале $[0; T]$ времени:

$$\lambda_{cp,T} = \frac{1}{T}\int_0^T \lambda(t)\,dt. \tag{2}$$

Из (1) следует, что

$$\lambda_{cp,T} = -\frac{1}{T}\ln P(T), \tag{3}$$

т. е. эта характеристика надежности определяется лишь величиной надежности $P(T)$ в конце заданного интервала времени ($t = T$). Например, для усеченного экспоненциального закона надежности:

$$P(t) = \frac{e^{-\Theta t} - e^{-\Theta T_\Phi}}{1 - e^{-\Theta T_\Phi}}; \quad 0 \leq t \leq T_\Phi; \quad 0 < \Theta = \text{const}, \tag{4}$$

имеем на заданном интервале времени $[0; T]$, $T < T_\Phi$:

$$\lambda_{cp.T} = -\frac{1}{T}\ln\frac{e^{-\Theta T} - e^{-\Theta T_\Phi}}{1 - e^{-\Theta T_\Phi}}, \quad t < T_\Phi. \tag{5}$$

* Написанию данной работы автор обязан ознакомлению с монографией *П. М. Рабиновича* [3], где, в частности, обсуждаются показатели надежности (средний показатель интенсивности отказов и т. д.).

II. Доверительный интервал для истинного значения

Из опытных данных известна эмпирическая (выборочная) надежность $P^*(t)$, определяемая либо по параметрическим методам статистики (при известности закона надежности, и неизвестности его параметров), либо по непараметрическим методам — в случае неизвестности этого закона. Известна также дисперсия $DP^*(t)$, причем эти данные обычно получаются при бóльшем объеме выборки n, поэтому случайная величина $P^*(t)$ асимптотически нормальна со средним $P(t)$ и дисперсией $DP^*(t)$. Соответственно получаем эмпирическое значение средней интенсивности отказов:

$$\lambda^*_{cp,T} = -\frac{1}{T}\ln P^*(T). \tag{6}$$

Случайная величина $\lambda^*_{cp,T}$, как функция асимптотически нормально распределенной случайной величины $P^*(T)$, по теореме Крамера [1] распределена также асимптотически нормально со средним и дисперсией:

$$M\lambda^*_{cp,T} \approx \lambda_{cp,T} \approx \lambda^*_{cp,T}; \tag{7}$$

$$D\lambda^*_{cp,T} \approx \frac{DP^*(T)}{T^2 P^2(T)} \approx \frac{DP^*(T)}{T^3 P^{*2}(T)}, \tag{8}$$

где знаки приближенных равенств соответствуют учету больших n.

Вследствие асимптотической нормальности $\lambda^*_{cp,T}$ формула доверительного интервала для истинного значения $\lambda_{cp,T}$ будет (при больших n):

$$P = P\left\{\lambda^*_{cp,T} - u\sqrt{D\lambda^*_{cp,T}} < \lambda_{cp,T} < \lambda^*_{cp,T} + u\sqrt{D\lambda^*_{cp,T}}\right\} = \frac{1}{\sqrt{2\pi}}\int_{-U}^{U} e^{-z^2/2}\,dz, \tag{9}$$

где $U > 0$ (обычно $U = 3$ и доверительная вероятность

$$P = 0{,}9973;\ u = 1{,}96 : P = 0{,}95),$$

и в правой части содержится табулированная функция [2].

В частности, по непараметрическим методам статистики

$$P^*(T) = 1 - F^*(T),$$

где эмпирическая функция распределения равна:

$$F^*(T) = n_T/n,$$

n_T — число отказов, расположенных левее T на оси t, n — общее число наблюденных отказов. Тогда имеем [2]:

$$DP^*(T) = DF^*(T) = \frac{F(T)\,[1 - F(T)]}{n} = \frac{P(T)[1 - P(T)]}{n}\,; \tag{10}$$

$$D\lambda^*_{cp,T} \approx \frac{1 - P(T)}{nT^2\,P(T)} \approx \frac{F^*(T)}{nT^2\,P^*(T)}\,. \tag{11}$$

К рассмотренным выше вопросам примыкает понятие средней надежности.

III. Средняя надежность на заданном интервале времени

Она равна:

$$P_{cp,t} = \frac{1}{t}\int_0^t P(\tau)\,d\tau, \tag{12}$$

и указывает среднее значение надежности $P(t)$ аппаратуры от момента включения $t = 0$ до заданного момента времени t. Из (3) следует, что

$$P(t) = \exp\,[-t\lambda_{cp,t}]\,, \tag{13}$$

где

$$\lambda_{cp,t} = \frac{1}{t}\int_0^t \lambda(\tau)\,d\tau\,.$$

Тогда (12) примет вид:

$$P_{cp,t} = \frac{1}{t}\int_0^t e^{-t\lambda_{ap,t}}\,dt\,. \tag{14}$$

Таким образом, средняя надежность на заданном интервале времени выражается через среднюю интенсивность отказов на том же интервале времени.

Например, для неусеченного экспоненциального закона надежности

$$P(t) = e^{-\lambda t},\ 0 < \lambda = \text{const}\,. \tag{15}$$

имеем согласно (12):

$$P_{cp,t} = \frac{1}{t}\int_0^t e^{-\lambda t}\,d\tau = \frac{1}{\lambda t}(1 - e^{-\lambda t})\,. \tag{16}$$

Для линейного закона надежности:

$$P(t) = 1 - t/T_\Phi\,,\ 0 \le t \le T_\Phi \tag{17}$$

получим:

$$P_{cp,t} = 1 - \frac{t}{2\,T_\Phi} \,. \tag{18}$$

Важен частный случай, когда заданный интервал, на котором оценивается средняя надежность $P_{cp,T}$, равен всему возможному периоду функционирования аппаратуры T_Φ, т. е. $t = T_\Phi$, в том числе и при $T_\Phi \to \infty$:

$$P_{cp,T_\Phi} = \frac{1}{T_\Phi} \int_0^{T_\Phi} P(\tau)\, d\tau \,. \tag{19}$$

Легко показать, что

$$\int_0^{T_\Phi} P(t)\, dt = T_{cp}\,, \tag{20}$$

где T_{cp} — среднее время безотказной работы аппаратуры. В самом деле, пусть $\xi > 0$ — случайное время первого отказа аппаратуры, включенной при $t = 0$. Оно имеет функцию распределения $q(t)$ — ненадежность аппаратуры:

$$q(t) \equiv F_\xi(t) = P\left\{\xi < t\right\};\ 0 \leq t \leq T_\Phi\,, \tag{21}$$

с плотностью вероятности $P_\xi(t)$, равной:

$$P_\xi(t) = -\frac{dq(t)}{dt} = -\frac{dP(t)}{dt};\ 0 \leq t \leq T_\Phi\,, \tag{22}$$

так как

$$P(t) = P\left\{\xi > t\right\} = 1 - P\left\{\xi < t\right\} = 1 - q(t)\,. \tag{23}$$

Но среднее время T_{cp} безотказной работы аппаратуры является математическим ожиданием (средним) случайной величины ξ:

$$T_{cp} = M\xi = \int_0^{T_\Phi} t p_\xi(t)\, dt = -\int_0^{T_\Phi} t\, \frac{dP(t)}{dt} \,. \tag{24}$$

Интегрируя по частям, из (24) имеем:

$$T_{cp} = -\left[tP(t)\right]_0^{T_\Phi} + \int_0^{T_\Phi} P(t)\, dt\,. \tag{25}$$

Так как

$$P(0) = 1 \quad \text{и} \quad P(T_\Phi) = 0\,,$$

то отсюда следует формула (20) [для конечных T_Φ]. Если $T_\Phi \to \infty$, то для неусеченных справа распределений, встречающихся в теории надежности, всегда имеет место:

$$\lim_{t \to \infty} tP(t) = 0\,.$$

Итак, из (19) и (20) получаем:

$$P_{cp,T_\Phi} = \frac{T_{cp}}{T_\Phi}, \tag{26}$$

т. е. средняя надежность на всем возможном интервале времени работы аппаратуры равна отношению среднего времени T_{cp} безотказной работы к значению максимального времени T_Φ.

Например, для усеченного экспоненциального закона надежности (4):

$$T_{cp} = \frac{1}{1 - e^{-\Theta T_\Phi}} \int_0^{T_\Phi} (e^{-\Theta t} - e^{-\Theta T_\Phi})\, dt = \frac{\frac{1}{\Theta}(1 - e^{-\Theta T_\Phi}) - T_\Phi e^{-\Theta T_\Phi}}{1 - e^{-\Theta T_\Phi}}; \tag{27}$$

или

$$T_{cp} = \frac{1}{\Theta} - \frac{T_\Phi}{e^{\Theta T_\Phi} - 1}; \tag{28}$$

т. е.

$$P_{cp,T_\Phi} = \frac{1}{\Theta T_\Phi} - \frac{1}{e^{\Theta T_\Phi} - 1}. \tag{29}$$

Во многих случаях заранее неизвестен закон надежности $P(t)$ аппаратуры, но на основании опытных данных можно считать известным T_Φ. Кроме того, по большой выборке n наблюденных значений времени первого отказа (случайной величины ξ) известны эти времена $t_1, t_2, \ldots, t_n$. Тогда непараметрической оценкой среднего времени T_{cp} безотказной работы (т. е. не зависящей от вида распределения вероятностей случайной величины ξ) может служить среднее выборки:

$$T^*_{cp} = \frac{t_1 + t_2 + \ldots t_n}{n} \equiv \frac{1}{n} \sum_{i=1}^{n} t_i, \tag{30}$$

которое, как известно [1], распределено асимптотически нормально (с ростом n) со средним

$$MT^*_{cp} = T_{cp}$$

и дисперсией

$$DT^*_{cp} = \frac{D\xi}{n}.$$

Ввиду больших n ($n \geq 100$) можно приближенно положить в вычислениях:

$$MT^*_{cp} = T_{cp} \approx T^*_{cp}, \tag{31}$$

$$DT^*_{cp} = \frac{D\xi}{n} \approx \frac{1}{n}\left[\frac{1}{n-1}\sum_{i=1}^{n}(t_i - T^*_{cp})^2\right]. \tag{32}$$

Учитывая асимптотическую нормальность случайной величины T^*_{cp}, после ряда несложных преобразований под знаком доверительной вероятности P, легко построить доверительный интервал для истинного значения T_{cp} (при больших n):

$$P = P\left(T^*_{cp} - u\sqrt{DT^*_{cp}} < T_{cp} < T^*_{cp} + u\sqrt{DT^*_{cp}}\right) = \\ = \frac{1}{\sqrt{2\pi}}\int_{-u}^{u} e^{-z^2/2}\,dz, \tag{33}$$

где любое $u > 0$, а в правой части по-прежнему содержится табулированная функция [2].

Эмпирическое (выборочное) значение P^*_{cp}, T_Φ средней надежности P_{cp,T_Φ} будет равно:

$$P^*_{cp,T_\Phi} = \frac{T^*_{cp}}{T_\Phi}, \tag{34}$$

и оно также распределено асимптотически нормально со средним (математическим ожиданием) и дисперсией, равными соответственно:

$$MP^*_{cp,T_\Phi} = P_{cp,T_\Phi};\ DP^*_{cp,T_\Phi} = \frac{DT^*_{cp}}{T^2_\Phi}. \tag{35}$$

Поэтому, аналогичным путем строится доверительный интервал для истинной средней надежности (при больших n):

$$P = P\left(P^*_{cp,T_\Phi} - \frac{u}{T_\Phi}\sqrt{DT^*_{cp}} < P_{cp,T_\Phi} < \right. \\ \left. < P^*_{cp,T_\Phi} + \frac{u}{T_\Phi}\sqrt{DT^*_{cp}}\right) = \frac{1}{\sqrt{2\pi}}\int_{-u}^{u} e^{-z^2/2}dz. \tag{36}$$

Таким образом, оценивается средняя надежность при $t = T_\Phi$ по бóльшей выборке, при неизвестности закона надежности. Если $T_\Phi \to \infty$, то средняя надежность на таком бесконечном интервале времени обычно равна нулю ввиду обычной конечности математического ожидания случайной величины ξ времени первого отказа аппаратуры.

ЛИТЕРАТУРА

1. *Крамер, Г.:* Математические методы статистики. *ИИЛ, Москва* 1948, 401.
2. *Дунин-Барковский, И. В.—Смирнов, Н. В.:* Теория вероятностей и математическая статистика в технике (общая часть). *ГИТТЛ, Москва* 1955; 501, 180, 289.
3. *Рабинович, П. М.:* Статистические методы исследования производственных резервов промышленного предприятия. *Госстатиздат, Москва* 1963, 183.

The Mean Intensity of Instrument Breakdown in a Given Time Interval (*J. Linkovski-Conde*) The author introduces a new function for the reliability of an instrument which gives the mean intensity of breakdown for a given time interval. He deduces its relation with the other reliability functions and determines the confidence interval of the real values of the function. He examines the mean reliability for a given time interval and deals with its evaluation on the base of sampling.

Durchschnittliche Intensität des Schadhaftwerdens von Instrumenten in einem gegebenen Zeintintervall (*J. Linkowski-Conde*). Der Verfasser führt eine neue Funktion für die Verläßlichkeit eines Instruments ein, welche für ein gegebenes Zeitintervall die durchschnittliche Intensität des Schadhaftwerdens angibt. Er zeigt den Zusammenhang dieser Funktion mit sonstigen Verläßlichkeitsfunktionen und berechnet den Konfidenzbereich der realen Werte der Funktion. Er untersucht weiterhin die durchschnittliche Verläßlichkeit für ein gegebenes Zeitintervall und faßt sich mit deren Auswertung auf Grund der Probenahme.

L'intensité moyenne des pannes des instruments dans un intervalle de temps donné (*J. Linkovski-Conde*). L'auteur introduit une fonction nouvelle relative à la sécurité de fonctionnement des instruments, qui donne l'intensité moyenne des pannes pour un intervalle de temps donné. Il montre la relation entre cette fonction et d'autres fonctions de sécurité et trace l'intervalle de confidence des valeurs réelles de la fonction. Il examine ensuite la sécurité de fonctionnement moyenne pour l'intervalle donné et s'occupe de son interprétation en se basant sur l'échantillonage.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 75—100 (1968)

ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ ЦИФРОВОГО РЕГУЛЯТОРА

Я. ШОМЛО

ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ ПО АВТОМАТИЗАЦИИ
ВЕНГЕРСКОЙ АКАДЕМИИ НАУК, БУДАПЕШТ

[Поступило 26. апреля 1967 г.]

В настоящей работе проводится исследование устойчивости цифрового регулятора нелинейного действия методом гармонической линеаризации. Путем упрощений становится возможным определение аналитических соотношений, служащих для выбора параметров, обеспечивающих устойчивость системы. Правильность выбора может быть проверена при помощи точных коэффициентов гармонической линеаризации. В работе нашли применение результаты, которые ранее получены автором. Исследования связаны с одной конкретной прикладной задачей.

I. Введение

В работах [2, 3] описывается обобщенный метод гармонической линеаризации кусочно-линейных характеристик. Этот метод дает возможность решить такие проблемы, которые из-за трудностей вычислительного характера раньше не представлялось возможным решать.

Ниже, на конкретном примере, покажем использование данного метода. Так как этот пример связан с осуществлением одного конкретного устройства, постановка задачи имеет специальный характер, но несмотря на это, полученные результаты имеют общий характер. Эти результаты сами по себе интересны, а демонстрированный метод решения общеприменим для решения подобных задач (например, в случае другого объекта регулирования).

В настоящей работе рассматривается исследование устойчивости цифрового регулятора, работающего по нелинейному принципу, в регуляторном режиме.

Одна регулирующая цепь исследуемого цифрового регулятора показана на рис. 1.

Значение регулируемой величины измеряется периодически, через каждый T_m промежуток времени. Значения, соответствующие этим моментам времени, поступают на вход цифрового регулятора. Предположим, что изменение значения регулируемой величины между двумя квантованиями не изменяется больше, чем на величину, соответствующую одному двоичному значению. В зависимости от значения регулируемой величины в момент квантования на выходе цифрового регулятора появляется положительный или отрицательный импульс, или не появляется импульс. Когда на

выходе регулятора появляется импульс, шаговый двигатель, установленный на его выходе, повернется на определенную (небольшую) величину. Угол поворота оси шагового двигателя представляет собой регулирующий сигнал.

Угол повтора оси шагового двигателя «считает» импульсы с соответствующим знаком. Если за входной сигнал шагового двигателя принимаем среднюю частоту импульсов (число импульсов за единицу времени), шаговый двигатель может рассматриваться как интегрирующий член.

От входного сигнала исполнительного устройства до управляемой величины система работает непрерывными сигналами.

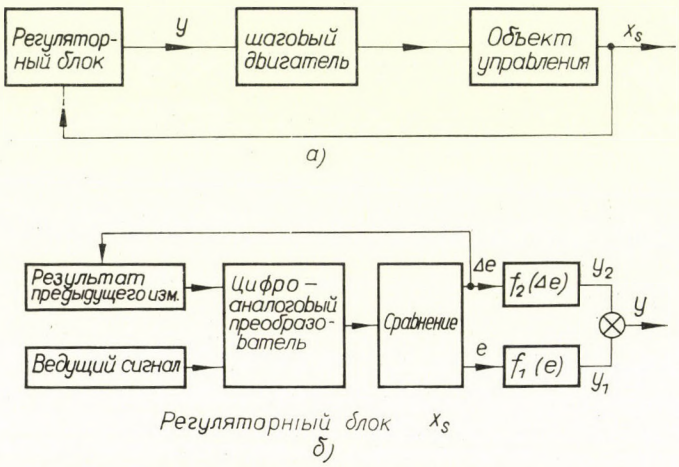


Рис. 1

Принцип действия регулятора следующий:

Когда сигнал рассогласования «больше», чем некоторое значение ($s\pm$), регулятор при каждом A_0-вом квантовании выдает такой импульс, который уменьшает это рассогласование. Если сигнал рассогласования превышает одно бóльшее значение ($s\pm\pm$), тогда через каждое C_0-вое квантование выдается такой импульс. Очевидно, $C_0 \leq A_0$, $C_{0\,\min} = 1$ (при каждом квантовании выдается импульс, $C_0 = 1, 2, 3, 4, \ldots$).

Так как шаговый двигатель работает как интегратор, в практических случаях, если предполагается обеспечить хорошие динамические свойства и точность, возникают проблемы в связи с обеспечением устойчивости системы.

Эта проблема часто может быть решена при дополнении регулятора такой частью, которая уменьшает интенсивность воздействия при уменьшении рассогласования. В то же самое время параллельно с этим улучшаются динамические свойства системы, если интенсивность воздействия регулятора возрастет с ростом сигнала рассогласования. Это действие аналогично дифференциальным воздействиям в линейных системах.

Рассмотрим работу регулирующего блока, изображенного на рис. 1б. Регулируемая величина x_s является непрерывной, а сигнал рассогласования цифровым. Цифро-аналоговый преобразователь определяет те границы, в пределах которых цифровой регулятор воспринимает аналоговые сигналы как данные цифровые (двоичные) значения. На рис. 2 цифро-аналоговый преобразователь изображается этими границами или «уровнями».

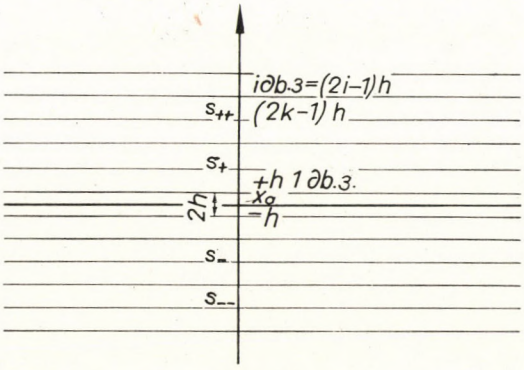


Рис. 2

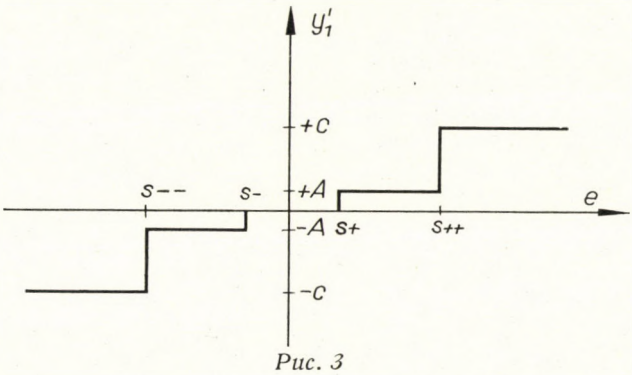


Рис. 3

Аналоговое значение $2h$ соответствует диапазону в 1 бит, т. е. одному двоичному значению.

Воздействие по рассогласованию следующее (рис. 3). До тех пор, пока сигнал рассогласования меньше чем s_+ и больше чем s_-, выходного сигнала нет. Когда абсолютное значение сигнала станет больше этих предельных величин, на входе появляются импульсы частотой $+A$ или $-A$. Когда значение сигнала рассогласования станет больше значения s_{++} (меньше s_{--}), частота выходных импульсов возрастает до $+C$ (или $-C$).

Частоты импульсов A и C определяются соотношениями

$$A = \frac{1}{T_m A_0}; \qquad C = \frac{1}{T_m C_0} \left[\frac{имп}{ед.\ времени} \right].$$

Одна из возможностей образования воздействия по изменению сигнала рассогласования следующая. Когда аналоговый сигнал рассогласования изменяется с меньшего двоичного значения на бóльшее (с возрастанием пересекается двоичная граница — 1 бит), регулятор выдает за время τ при каждом квантовании положительные импульсы. При противоположном изменении выходным сигналом регулятора являются отрицательные импульсы за время τ. Если за время τ происходит новый переход через двоичную границу, тогда влияние этого «идет дальше» за время τ, считая от момента этого нового перехода.

Части регулятора, которые воздействуют по рассогласованию и по изменению рассогласования, на рис. 1б обозначены через $f_1(e)$ и $f_2(\Delta e)$ соответственно. Так как на выходе регулятора может появиться только положительный или отрицательный импульс или же его нет, между этими двумя частями регулятора необходимо было создать логическую связь. Она состоит в том, что часть $f_2(\Delta e)$ преобладает над частью $f_1(e)$, то есть после каждого перехода через каждую двоичную границу за время τ часть $f_1(e)$ не работает.

Когда по сравнению с изменениями регулируемой величины время квантования мало, то данный цифровой регулятор может рассматриваться как сложный нелинейный член, входной сигнал которого является аналоговым сигналом рассогласования $(g - x_s)$, а выходным сигналом — частота или число импульсов.

II. Исследование устойчивости

Рассмотрим систему, где передаточная функция объекта регулирования имеет вид:

$$W_1(p) = \frac{K}{(1 + T_1 p)(1 + T_2 p)} . \tag{1}$$

Предположим, что постоянные времени намного превышают период квантования.

Так как можно предполагать, что система обладает обобщенным свойством фильтра [1], то при анализе применим метод гармонической линеаризации.

Объект регулирования вместе с шаговым двигателем составляет линейную часть системы. Коэффициенты гармонической линеаризации зависят не только от амплитуды, но также и от частоты

$$W_H = b(B, \omega) + ja(B, \omega) , \tag{2}$$

где B и ω — амплитуда и частота синусоидального сигнала, предположенного на входе нелинейного регулятора

$$(\tilde{e} = B \sin \omega t);$$

замкнутая цепь регулирования показана на рис. 4.

Без ограничения общности можно принять, что $g(t) = 0$.

При существовании автоколебаний:

$$\frac{x_s(j\omega)}{e(j\omega)} = \frac{K(b + ja)}{j\omega(1 + T_1 j\omega)(1 + T_2 j\omega)} = -1 \tag{3}$$

устойчивость системы можно исследовать, например, в аналитической форме. По формуле (3) имеем

$$j\omega(1 + T_1 j\omega)(1 + T_2 j\omega) + K(b + ja) = 0.$$

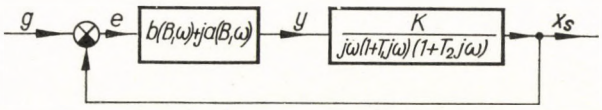


Рис. 4

Разделив действительную и мнимую части, получим:

$$\begin{aligned} (1 - \omega^2 T_1 T_2) + Ka(B, \omega) &= 0, \\ -\omega^2(T_1 + T_2) + Kb(B, \omega) &= 0. \end{aligned} \tag{4}$$

Если уравнение (4) имеет решения, то они дают параметры возможных автоколебаний (B и ω). Отметим, что возможен, в данном случае, и другой подход к решению задачи. Известно, по частотным методам исследования систем, что автоколебания не могут возникать, если выполняется условие:

$$\frac{a}{b} \geq \omega T_1,$$

$$(T_2 > T_1).$$

Однако, в дальнейшем будем применять аналитический подход к решению задачи, так как это имеет более общий характер и, хотя рассмотрим конкретный пример, стараемся дать общий способ для решения подобных задач.

III. Коэффициенты гармонической линеаризации

Их можно определить следующим образом (рис. 5). Рассмотрим какие зависимо определяют выходной сигнал нелинейности, когда на входе его действует сигнал

$$\tilde{e} = B \sin \omega t.$$

Сперва рассмотрим случай малых амплитуд. Значения s_+ и s_- целесообразно выбирать как можно меньшими. Пусть:

$$s \pm = \pm h \quad \text{и} \quad B \leq 3h.$$

Графическая зависимость между y и $\tilde{e}$ определяется следующим образом.

Когда величина выходного сигнала превышает значение

$$\tilde{e} = B \sin \alpha = h,$$

пересекается двоичная граница, и выходной сигнал, в виде частот импульсов, за время τ, то есть за фазовый угол $\Delta\varphi = \omega\tau$ будет иметь максимальную величину (H).

На рис. 5а отношения амплитуды, частоты (B и ω) и времени τ таковы, что:

$$\omega\tau \leq \alpha$$

и, кроме того

$$\alpha \leq \frac{\pi}{6},$$

где

$$\alpha = \operatorname{arc\,sin} \frac{h}{B}.$$

По истечении времени τ, так как сигнал ошибки больше чем s_+, частота импульсов принимает значение $+ A$. Это имеет место до тех пор, пока сигнал ошибки, уменьшаясь, опять не пересекает двоичной границы. После этого отрицательные импульсы будут иметь за время τ максимальную частоту, а затем частота их станет равной нулю. В последующем процесс симметрично повторяется.

На основе вышесказанного можно построить зависимость

$$y = f(\tilde{e}).$$

Полученная кусочно-линейная характеристика нечетно симметрична и имеет точки излома в точках (рис. 5б):

$$F_1^{(1)} = h\,; \; F_2^{(1)} = B \sin(\alpha + \omega\tau), \tag{5}$$

$$F_1^{(2)} = B \sin(\pi - \alpha + \omega\tau) = B \sin(\alpha - \omega\tau)\,; \; F_2^{(2)} = h. \tag{6}$$

Значения коэффициентов гармонической линеаризации при определенных амплитудах и частотах можно определить при помощи обобщенных формул работы [2, 3].

Увеличением частоты ω форма этой зависимости изменяется (рис. 5в, 5г, 5д и 5е). Эту форму при каждой амплитуде и частоте можно определить. Таким образом, при каждой частоте, изменяя значения амплитуд, можно определить коэффициенты гармонической линеаризации. Этот процесс совершенно такой же и в случае больших амплитуд, но характеристики очевидно будут более сложными, состоящими из нескольких участков.

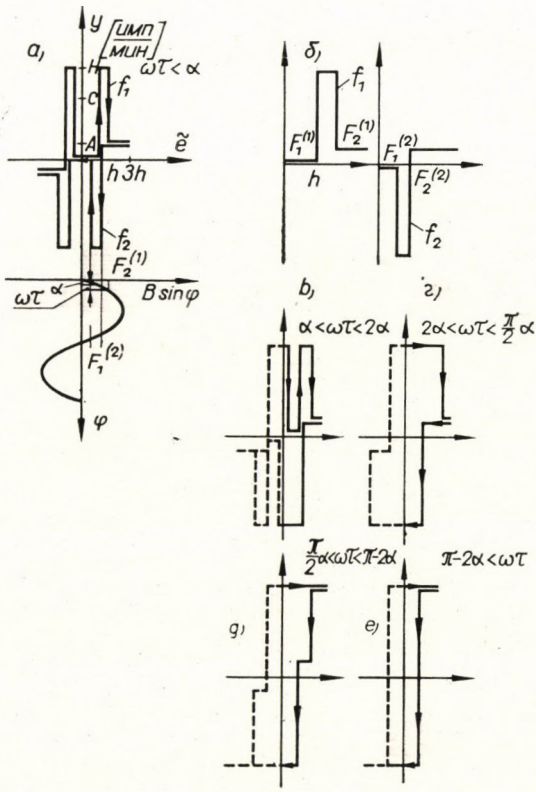


Рис. 5

Таким образом, исследование вопроса устойчивости систем может быть произведено, но даже в случае использования обобщенного метода определения коэффициентов, расчетные трудности являются неизменно бóльшими. В практических случаях дается возможность для упрощения проблем.

Рассмотрим процесс регулирования температуры в консервном автоклаве, где имеются следующие параметры:

$$K = 0{,}8^\circ\, C/имп; \quad T_1 = 1\, мин,\ T_2 = 1440\, мин;$$

$$h = 0{,}2\ ^\circ C;\ s \pm = \pm 0{,}2\ ^\circ C;\ T_m = 0{,}01\, s = \frac{1}{6} 10^{-3}\, мин.$$

В результате данного алгоритма регулятора коэффициенты гармонической линеаризации оказываются всегда положительными, итак цифровой регулятор всегда дает положительный фазовый сдвиг

$$\left(\varphi = \operatorname{arc\,tan} \frac{a}{b}\right).$$

Рассматривая частотные характеристики системы, видно, что те частоты, которые оказываются ниже действительной оси, можно не рассматривать, так как положительный фазовый сдвиг исключает возможность автоколебаний (фазовый сдвиг $-180°$).

Амплитудно-фазовая диаграмма системы пересекает действительную ось при частоте:

$$\omega = \frac{1}{\sqrt{T_1 T_2}} = 0{,}0264 \frac{1}{мин}\ .$$

Эта частота является одновременно частотой автоколебаний, когда регулятор воздействует только по рассогласование (рис. 3), то есть

$$f_2(\Delta e) = 0.$$

В этом случае

$$a(B, \omega) = 0\,,$$

и коэффициент гармонической линеаризации $b = b(B)$ лежит на действительной оси. [По уравнению (4) при значении $a = 0$ получим частоту $\omega^2 = 1/T_1 T_2$.]

Время одного периода колебаний:

$$T_n = \frac{2\pi}{0{,}0264} = 238\,мин\,.$$

При такой небольшой частоте, — так как по отношению к T_n время τ очевидно очень мало, — можно принять, что выдача импульсов τ^i происходит мгновенно и, следовательно, ось шагового двигателя мгновенно повернется на соответствующий угол. Это показано на рис. 6. Согласно этому каждый переход через двоичную границу дает пропорциональное положение оси шагового двигателя. Весь регулятор — в этой частотной области — работает подобно линейному регулятору *ПИ*.

Рассмотрим коэффициенты гармонической линеаризации характеристик на рис. 6б и 6в. Если введем обозначения $b_1 = b_1\,(B)$ и $b_2 = b_2\,(B)$ на коэффициенты, определенные по характеристикам рис. 6б и 6в, соответственно, тогда на основе рис. 6а коэффициент гармонической линеаризации получим в виде:

$$W_{H'}(B, \omega) = \frac{b_1(B)}{j\omega} + b_2(B) = (b_1 + j\omega b_2)\frac{1}{j\omega} = (b + ja)\frac{1}{j\omega},$$

где

$$b = b_1(B),$$
$$a = b_2(B)\omega.$$

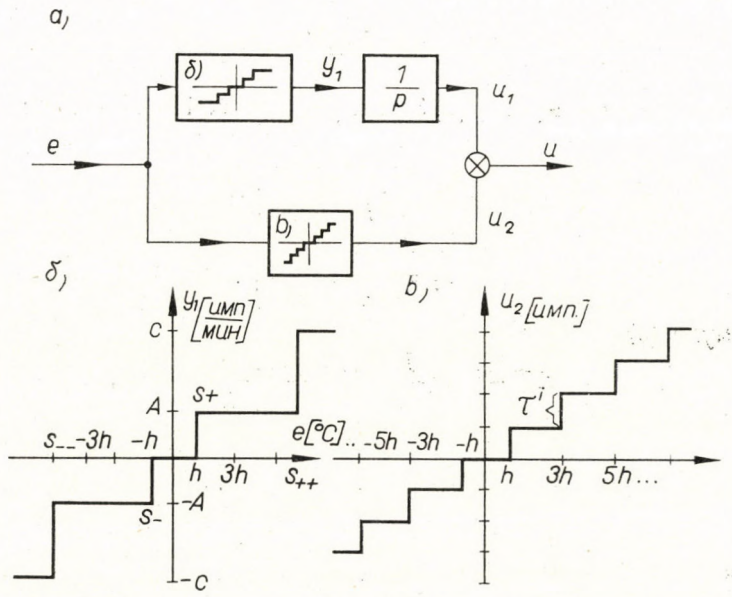


Рис. 6

Подобно сказанному ранее (подключая член $1/j\omega$ к линейной части), нелинейную часть системы характеризуем в виде

$$W_H(B, \omega) = b + ja.$$

По уравнениям (4) получим

$$1 + Kb_2 - \omega^2 T_1 T_2 = 0,$$

отсюда

$$\omega^2 = \frac{1 + Kb_2}{T_1 T_2}. \tag{7}$$

Подстановкой (7) во второе уравнение (4) имеем

$$Kb_1(B) = \frac{1 + Kb_2(B)}{T_1 T_2}(T_1 + T_2) = \frac{1 + Kb_2(B)}{T}, \tag{8}$$

где

$$T = \frac{T_1 T_2}{T_1 + T_2} \approx T_1.$$

Автоколебания в системе не могут существовать, если уравнение (8) не имеет действительного, положительного решения для амплитуды B. Если:

$$h \leq B \leq 3h,$$

по обобщенным формулам [2, 3] получаем (см. Приложение):

$$b_1 = \frac{A}{h} l\left(\frac{B}{h}\right),$$

$$b_2 = \frac{\tau^i}{h} l\left(\frac{B}{h}\right).$$

По формуле (8)

$$KT \frac{A}{h} l\left(\frac{B}{h}\right) = 1 + K \frac{\tau^i}{h} l\left(\frac{B}{h}\right),$$

откуда

$$TA - \tau^i = \frac{h}{Kl\left(\frac{B}{h}\right)}.$$

Так как

$$l_{\text{макс}}\left(\frac{B}{h}\right) = l(\sqrt{2}) \cong 0{,}637$$

и в данной системе $h = 0{,}2$; если

$$TA - \tau^i < \frac{0{,}2}{K \cdot 0{,}637} = \frac{0{,}314}{K}, \tag{9}$$

тогда в системе автоколебания с амплитудой

$$h \leqslant B \leqslant 3h$$

невозможны.

Пусть будет

$$\tau^i = vTA, \tag{10}$$

где v какое-нибудь положительное число. На основе (9) условием того, чтобы автоколебания не возникали, будет

$$1 - v < \frac{0{,}314}{KTA}. \tag{11}$$

В виду того, что в практических случаях произведение KTA имеет большое значение, вместо (11) можно потребовать выполнения условия

$$v \geq 1, \tag{12}$$

учитывая (10), получим

$$\tau^i \geq TA.$$

Пусть

$$3h \leq B \leq 5h.$$

Выражение для определения коэффициента b_1 (B) остается неизменным до значения $B < s_{++}$

$$b_2(B) = \tau^i \left[\frac{1}{h} l\left(\frac{B}{h}\right) + \frac{1}{3h} l\left(\frac{B}{3h}\right) \right].$$

Если в соответствии с условием

$$v = l + n > 1$$

(где n какое-нибудь положительное число), то по (8), учитывая (10), получим

$$KTA \frac{1}{h} l\left(\frac{B}{h}\right) = 1 + KTA \left[\frac{1}{h} l\left(\frac{B}{h}\right) + \frac{n}{h} l\left(\frac{B}{h}\right) + \frac{v}{3h} l\left(\frac{B}{3h}\right) \right].$$

В правой части первая слагаемая выражения в квадратной скобке всегда дает значение, одинаковое левой части уравнения. Имея в виду, что все остальные члены правой части являются положительными величинами, равенство никогда не может быть получено.

При увеличении амплитуды до $B < s_{++}$:

$$b_2 = \tau^i \left[\frac{1}{h} l\left(\frac{B}{h}\right) + \frac{1}{3h} l\left(\frac{B}{3h}\right) + \frac{1}{5h} l\left(\frac{B}{5h}\right) + \dots \right]$$

первый член выражения всегда дает (подобно вышеуказанному), такое слагаемое в уравнении «автоколебаний», которое исключает возникновение их. Наоборот, как это видно, при увеличении амплитуды, увеличивается запас устойчивости.

Рассмотрим условия того, чтобы автоколебания при $B > s_{++}$ не возникали. В этом случае

$$b_1 = \frac{A}{h} l\left(\frac{B}{h}\right) + \frac{C}{s_{++}} l\left(\frac{B}{s_{++}}\right)$$

$$b_2 = \tau^i \left[\frac{1}{h} l\left(\frac{B}{h}\right) + \frac{1}{3h} l\left(\frac{B}{3h}\right) + \frac{1}{5h} l\left(\frac{B}{5h}\right) + \right.$$

$$\left. + \frac{1}{7h} l\left(\frac{B}{7h}\right) + \dots + \frac{1}{s_{++}} l\left(\frac{B}{s_{++}}\right) \right].$$

[Предположим, что при $B > s_{++}$ регулятор не воздействует по изменению рассогласования, то есть $f_2(\Delta e) = 0$.] По уравнению (8)

$$KTA\,\frac{1}{h}\,l\left(\frac{B}{h}\right) + KTC\,\frac{1}{s_{++}}\,l\left(\frac{B}{s_{++}}\right) = 1 + K\tau^i\left[\frac{1}{h}\,l\left(\frac{B}{h}\right) + \frac{1}{3h}\,l\left(\frac{B}{3h}\right) + \ldots + \frac{1}{s_{++}}\,l\left(\frac{B}{s_{++}}\right)\right].$$

Отсюда, учитывая (10), получим

$$KTC\,\frac{1}{s_{++}}\,l\left(\frac{B}{s_{++}}\right) = 1 + K\tau^i\left[\left(1-\frac{1}{v}\right)\frac{1}{h}\,l\left(\frac{B}{h}\right) + \frac{1}{3h}\,l\left(\frac{B}{3h}\right) + \ldots + \frac{1}{s_{++}}\,l\left(\frac{B}{s_{++}}\right)\right] = 1 + K\tau^i\left[\frac{v-1}{vh}\,l\left(\frac{B}{h}\right) + \frac{1}{3h}\,l\left(\frac{B}{3h}\right) + \ldots + \frac{1}{s_{++}}\,l\left(\frac{B}{s_{++}}\right)\right].$$

После деления на КТС получим

$$\frac{\tau^i}{TC} = \frac{\dfrac{1}{s_{++}}\,l\left(\dfrac{B}{s_{++}}\right) - \dfrac{1}{KTC}}{J(v, B)}, \tag{13}$$

где

$$J(v, B) = \frac{v-1}{v}\,\frac{1}{h}\,l\left(\frac{B}{h}\right) + \frac{1}{3h}\,l\left(\frac{B}{3h}\right) + \frac{1}{5h}\,l\left(\frac{B}{5h}\right) + \ldots + \frac{1}{s_{++}}\,l\left(\frac{B}{s_{++}}\right). \tag{14}$$

В числителе правой части уравнения (13) вторая слагаемая мала и поэтому ею можно пренебречь. Тогда получим

$$\frac{\tau^i}{TC} = \frac{l\left(\dfrac{B}{s_{++}}\right)}{s_{++}\,J(v, B)}. \tag{15}$$

Когда амплитуда стремится «сверху» к значению

$$s_{++}(B = s_{++} + 0), \quad \text{то} \quad l(B/s_{++})$$

стремится к нулю. Так как $J(v; B)$ имеет конечное значение, правая часть уравнения (15) также стремится к нулю. При увеличении амплитуды правая часть будет представлять дробное соотношение двух положительных чисел.

По работе [2]:

$$l\left(\frac{B}{F_i}\right) = \frac{4}{\pi}\,\frac{F_i}{B}\sqrt{1-\left(\frac{F_i}{B}\right)^2},$$

$$\lim_{B\to\infty} l\left(\frac{B}{F_i}\right) = \lim_{B\to\infty} \frac{4}{\pi}\,\frac{F_i}{B},$$

По уравнению (15), учитывая (14), имеем

$$\lim_{B\to\infty} \frac{l\left(\frac{B}{s_{++}}\right)}{s_{++}\,J(v,B)} = \frac{1}{\frac{v-1}{v}+k-1}, \tag{16}$$

где k равняется числу двоичных значений величины

$$s_{++}\left(s_{++} = (2k-1)\,h;\ \text{см. рис. 2}\right).$$

Рассмотрим выражение (15). Дробь правой части выражения при увеличении амплитуды B увеличивается от нуля до некоторого экстремального значения, и далее стремится к значению (16).

Если параметры системы выберем с таким расчетом, что

$$\frac{\tau^i}{TC} > \left.\frac{l\left(\frac{B}{s_{++}}\right)}{s_{++}\,J(v;\,B)}\right|_{\text{макс}} = s_{\text{макс}}(v,B), \tag{17}$$

тогда в системе колебаний не будет.

Таблица I

		$1{,}1s_{++}$	$1{,}4s_{++}$	$2s_{++}$	$3s_{++}$	$4s_{++}$	B
$s_{++} = 7h$	$v = 1$	0,178	0,280	0,312	0,324	0,328	0,333
$(k = 4)$	$v \to \infty$	0,120	0,201	0,230	0,240	0,246	0,250
$s_{++} = 11h$	$v = 1$	0,110	0,136	0,185	0,193	0,138	0,200
$(k = 6)$	$v \to \infty$	0,087	0,114	0,153	0,159	0,164	0,167

Рассмотрим, как влияет выбор значения v на вид функции $S_{\text{макс}}(v, B)$. По выражению (14) эта функция будет по v ($v \geq 1$) максимальной тогда, когда $v = 1$, а минимальной при $v \to \infty$.

Из-за трудности формального определения экстремального значения, рассмотрим вид функции $S(v; B)$ численно. Вычислим значения функций $S(1\ B)$ и $S(\infty; B)$ для нескольких значений амплитуды. Результаты сведем в табл. I.

По таблице I можно предположить, что в практических случаях функция $S(v; B)$ принимает свое наибольшее значение при $B \to \infty$, и это значение определяется выражением (16). Итак, критерием работы системы без автоколебаний, согласно (16) и (12), будет:

$$\frac{\tau^i}{TC} > \frac{1}{k-1}, \tag{18}$$

$$\frac{\tau^i}{TA} > 1. \tag{19}$$

Условия (18) и (19) позволяют подбирать параметры системы — на основе низкочастотного поведения — так, чтобы в системе не возникали автоколебания с такими частотами. Легко доказать (или графическим путем проверить), что система при таком выборе параметров является устойчивой.

Однако, такое упрощенное описание работы регулятора возможно только в определенном частотном и амплитудном диапазоне. При каждой частоте, выше определенной амплитуды и, соответственно, выше определенной частоты, при каждой амплитуде воздействие по переходу двоичных границ полностью «заполняет» (с импульсами максимальной частоты) время возможных периодов колебаний.

Рассмотрим рис. 7. Если

$$\varphi_2 - \varphi_1 \leq \omega\tau$$

(τ время подачи импульсов τ^i), тогда импульсы «по переходам» полностью «заполняют» время периода. Так как

$$\varphi_2 - \varphi_1 = \pi - 2\varphi_1 = \pi - 2 \arcsin \frac{2n-3}{2n-1},$$

вышеуказанное условие можно записать в форме

$$\arcsin \frac{2n-3}{2n-1} \geq \frac{\pi - \omega\tau}{2}. \tag{20}$$

Следовательно, для тех амплитуд

$$B = (2n - 1)h,$$

для которых выполняется условие (20), получим полное «заполнение» импульсами.

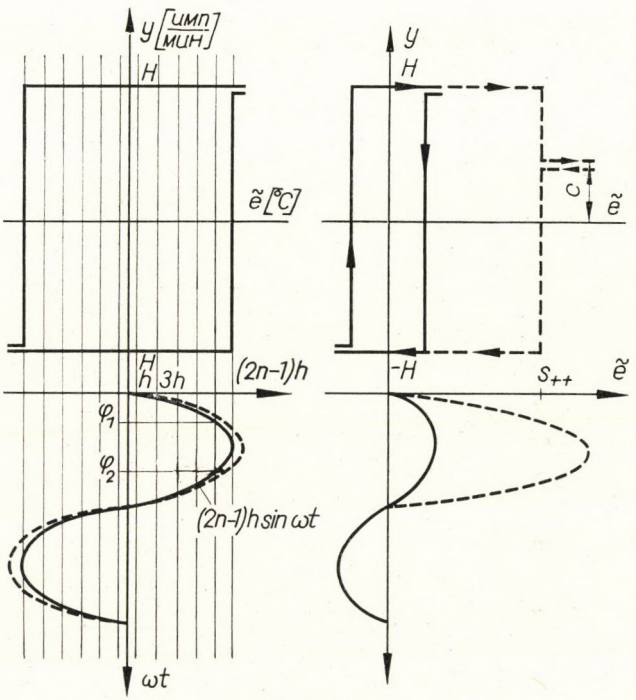


Рис. 7

Принимая значение $n = 2$, получим то значение частоты $\omega = \omega_{kp}$, выше которого «заполнение» всегда полное:

$$\arcsin \frac{1}{3} = \frac{\pi - \omega_{kp}\tau}{2}.$$

Отсюда имеем

$$\omega_{kp} = \frac{2{,}461}{\tau}. \tag{21}$$

В этом случае регулятор можно характеризовать характеристикой, приведенной на рис. 7. Соответствующие выражения для определения коэффициентов гармонической линеаризации зависят только от амплитуды колебаний.

Рассмотрим, по управлению (4), условие, когда автоколебания в системе не могут существовать.

Если

$$B = (2n-1)h + \alpha h,$$

где

$$0 < \alpha < 2$$

$$n = 1, 2, 3, \ldots,$$

тогда по обобщенным формулам работ [2, 3] имеем

$$a(B) = \frac{4H(2n-1)h}{\pi B^2}, \tag{22}$$

$$b(B) = \frac{H}{(2n-1)h} l\left(\frac{B}{(2n-1)h}\right). \tag{23}$$

Если

$$\alpha \to 0,$$

$$b(B) \to 0.$$

В этом случае по уравнению (4) автоколебания не существуют ($\omega \to 0$).

Когда $0 < \alpha < 2$, подстановкой (23) в (4) получим

$$\omega^2 = \frac{KH}{(2n-1)h(T_1+T_2)} l\left(\frac{B}{(2n-1)h}\right). \tag{24}$$

Автоколебания в системе не существуют, если ω_{kp} столь велико, что при любом значении амплитуды B, уравнение (24) не может дать равенства. По выражению (21), учитывая, что $\tau^i = \tau H$ и что максимальное значение правой части получим при значениях

$$n = 1 \quad \text{и} \quad l(\sqrt{2}) = 0{,}637;$$

условие устойчивости будет

$$\left(\frac{2{,}461}{\tau^i} H\right)^2 > \frac{HK}{h(T_1+T_2)} 0{,}637$$

или

$$\frac{Hh(T_1+T_2)}{K(\tau^i)^2} > 0{,}105. \tag{25}$$

Таким же образом можно проверить, что для частот, меньших ω_{kp} — для амплитуд выше граничных, определенных по (20) — на основе выражения (24) автоколебания могут ли существовать.

Границы применимости упрощенной схемы на рис. 6 тоже могут быть установлены.

Рассмотрим рис. 8. Принимаем какое-нибудь достаточно малое отклонение βh от принятого мгновенного действия. В этом случае — при данной частоте ω — можно определить то значение амплитуды, ниже которого отклонение от идеального случая остается ниже βh. — По рис. 9 имеем

$$(2n-1)\sin(\alpha_1+\omega\tau)\leq(1+\beta)h\,, \tag{26}$$

где

$$\alpha_1=\arcsin\frac{1}{2n-1}\,.$$

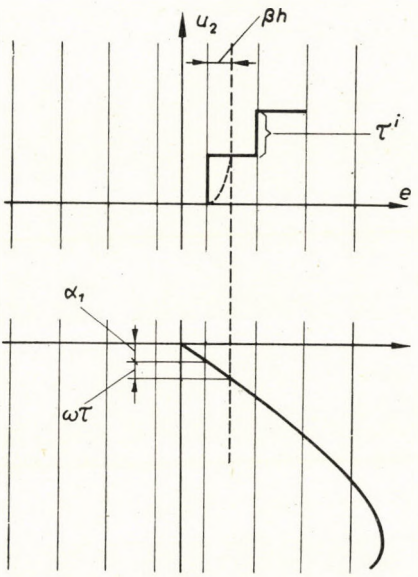


Рис. 8

При расчете регулятора можно использовать соотношения (18), (19) и (25). Для тех частот и амплитуд, для которых упрощенная схема работы неприменима, целесообразным является сделать проверочные расчеты, используя при них точные коэффициенты гармонической линеаризации.

Ниже приводится пример, который является как бы продолжением ранее анализированного случая и который хорошо демонстрирует описанный выше метод исследования.

Пример. Рассмотрим регулятор, который, кроме ранее приведенных имеет еще следующие параметры

$$s_{++}=3\,[^\circ\mathrm{C}],\ [k=8]\ ;$$

$$H = \frac{1}{T_m} = 6000 \left[\frac{имп}{мин}\right];$$

$$C = 2000 \left[\frac{имп}{мин}\right];$$

$$A = 300 \left[\frac{имп}{мин}\right];$$

$$\tau^i = 800 \ [имп], \ \tau = 0{,}133 \left[\frac{имп}{мин}\right].$$

По формуле (21)

$$\omega_{kp} = \frac{2{,}461}{0{,}133} = 18{,}5 \left[\frac{1}{мин}\right].$$

По интуитивным соображениям принимаем значение $\beta = 0{,}4$. (В этом случае максимальное отклонение от «мгновенного» действия будет 20%.)

В этом случае по выражению (26) имеем

$$(2n - 1)\, 0{,}2 \sin(\alpha_1 + 0{,}133\, \omega) \leq 0{,}28$$

$$\alpha_1 = \text{arc sin} \frac{1}{2n - 1}.$$

Принимая различные значения n, определяем соответствующие частоты ω (табл. II).

Таблица II

n	2	3	4	5
ω [1/*мин*]	1,1	0,5	0,435	0,370

По таблице II, например, при частотах меньших, чем

$$\omega = 0{,}5 \left[\frac{1}{мин}\right];$$

для амплитуд, меньших

$$B = (2n - 1)\, h = 5 \cdot 0{,}2 = I \ [°C];$$

отклонение от идеального случая составляет меньше 20%.

Из изложенного выше видно, что запас устойчивости увеличивается при увеличении амплитуды; итак упрощенная схема работы регулятора применима в крайнем случае до

$$\omega = 0{,}5 \left[\frac{1}{мин}\right].$$

По условию (18) получаем

$$\frac{\tau^i}{TC} > \frac{I}{K - 1},$$

так как

$$T \approx T_z = 1,$$

$$\frac{800}{2000} > \frac{1}{7} = 0{,}143,$$

$$0{,}4 > 0{,}143.$$

Условие (18) выполняется с достаточным запасом. По условию (19) имеем

$$\frac{\tau^i}{TA} > 1,$$

$$\frac{800}{300} = 2{,}66 > 1.$$

Условие (19) выполняется тоже с достаточным запасом. По условию (25) имеем

$$\frac{Hh(T_1 + T_2)}{K(\tau^i)^2} = \frac{6000 \cdot 0{,}2 \cdot 1441}{0{,}8 \cdot 800^2} = 3{,}38 > 0{,}105.$$

Это условие выполняется с очень большим запасом, что позволяет уменьшить верхнюю частотную границу на основе следующих рассуждений.

По рис. 7 определяем — выше частоты $\omega_{кр}$ коэффициенты гармонической линеаризации.

Если

$$B < s_{++} \quad \text{и} \quad (2n-1)h \leq B < (2n+1)h,$$

$$b(B) = \frac{2H}{(2n-1)h}\, l\left[\frac{B}{(2n-1)h}\right] = \frac{60\,000}{2n-1}\, l\left[\frac{B}{(2n-1)\,0{,}2}\right],$$

$$a = \frac{4H(2n-1)h}{\pi B^2} = 7650\,\frac{(2n-1)\,0{,}2}{B^2}.$$

Если

$$B > s_{++},$$

$$b(B) = \frac{1}{2}\,\frac{C-H}{s_{++}}\, l\left(\frac{B}{s_{++}}\right) + \frac{1}{2}\,\frac{H+C}{s_{++}}\, l\left(\frac{B}{s_{++}}\right) =$$

$$= \frac{C}{s_{++}}\, l\left(\frac{B}{s_{++}}\right) = 666{,}7\, l\left(\frac{B}{3}\right),$$

a

$$(B) = \frac{4Hs_{++}}{\pi B^2} = \frac{23\,000}{B^2}.$$

Для нескольких значений амплитуды вычисляем коэффициенты (табл. III).

Таблица III

B	0,2(+0)	$0{,}2\sqrt{2}$	0,6(+0)	$0{,}6\sqrt{2}$	1(+0)	1,4(+0)	1,8(+0)	3(+0)	$3\sqrt{2}$
a	38 200	19 100	12 780	6 390	7 650	5 460	4 250	2 560	1 280
b	0	38 200	0	12 780	0	0	0	0	425

Максимум абсолютного значения комплексного коэффициента усиления нелинейной части

$$|W_H|_{\text{макс}} = |\sqrt{a^2 + b^2}|_{\text{макс}} = \sqrt{19100^2 + 38\,200^2} = 42\,700 = 92{,}6\ дб.$$

Абсолютное значение комплексного коэффициента усиления, соответствующего воздействию с максимальной интенсивностью (рис. 7), больше, чем значение его в том случае, когда частота импульсов является максимальной не за весь период.

Поэтому, на основе рассчитанного выше максимального абсолютного значения можно определять ту частоту, выше которой суммарный коэффициент усиления всей цепи будет меньше единицы.

Построим логарифмические амплитудные и фазовые характеристики линейной части (рис. 9). Амплитуда при частоте

$$\omega = 3{,}5\left[\frac{1}{мин}\right];\ |A| = -94\ дб.$$

Поэтому при частотах выше этого значения суммарный коэффициент усиления будет меньше 1.

Остается все еще невыясненным вопрос поведения системы в области частот

$$\omega = 0{,}5 \div 3{,}5\left[\frac{1}{мин}\right].$$

Внести ясность в этот вопрос возможно, используя точные коэффициенты гармонической линеаризации.

Рассмотрим частоту

$$\omega = 1\left[\frac{1}{мин}\right].$$

При значении: $B = 1{,}2 \cdot 0{,}2 = 0{,}24°$ С в формулах (5) и (6):

$$\alpha = \arcsin\frac{0{,}2}{1{,}2 \cdot 0{,}2} = 56°\ 30';\ \omega\tau = 0{,}133\ [\mathrm{rad}] = 7°\ 40';$$

$$F_2^{(1)} = 0{,}24 \sin 64°\ 10' = 0{,}216;$$

$$F_1^{(2)} = 0{,}24 \sin 48°\ 50' = 0{,}181.$$

По рис. 5б имеем

$$b(B) = \frac{1}{2}\left[\frac{H}{0{,}2}\,l(1{,}2) - \frac{H-A}{0{,}216}\,l\left(\frac{0{,}24}{0{,}216}\right) - \frac{H}{0{,}181}\,l\left(\frac{0{,}24}{0{,}181}\right) + \frac{H+A}{0{,}2}\,l(1{,}2)\right],$$

$$b(B) = \frac{1}{2}\left[\frac{2}{0{,}2}\,l(1{,}2) - \frac{1}{0{,}216}\,l\left(\frac{0{,}24}{0{,}216}\right) - \frac{1}{0{,}181}\,l\left(\frac{0{,}24}{0{,}181}\right)\right]H +$$

$$+ \frac{1}{2}\left[\frac{1}{0{,}216}\,l\left(\frac{0{,}24}{0{,}216}\right) + \frac{1}{0{,}2}\,l(1{,}2)\right]A =$$

$$= 0{,}07 \cdot 6000 + 2{,}585 \cdot 300 = 1195;$$

$$a = 2\,\frac{0{,}016\,(H-A) + 0{,}019\,H}{\pi \cdot 0{,}24^2} = \frac{2}{\pi \cdot 0{,}24^2}\,(0{,}035\,H - 0{,}016\,A) = 2280.$$

Абсолютное значение комплексного коэффициента усиления

$$|W_H| = \sqrt{a^2 + b^2} = 2574 = 68{,}3\ дб.$$

Фазовый сдвиг нелинейной части будет

$$\varphi_H = \arctan\frac{2280}{1195} = 62°\ 20'.$$

Так как амплитудно-фазовые характеристики линейной части при

$$\omega = 1\left[\frac{1}{мин}\right]$$

будут следующими

$$|W_л| = -63\ \partial\delta\ ;\ \varphi_л = -225° = -180° - 45°,$$

то для всей цепи получаем амплитудные и фазовые характеристики:

$$|W| = |W_H|\,|W_л| = 5{,}3\ \partial\delta\ ;\ \varphi = \varphi_H + \varphi_л = -180° + 17°\,20'.$$

Подобным же образом для ряда амплитуд и частот определили характеристики цепи. Результаты приведены в табл. IV.

Рассмотрим одну бóльшую амплитуду при частоте

$$\omega = 1\left[\frac{1}{мин}\right].$$

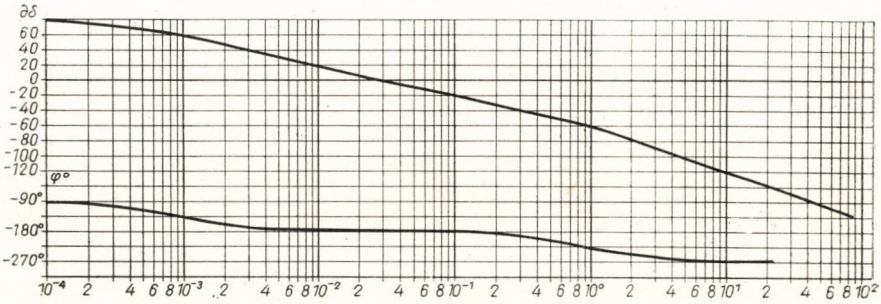


Рис. 9

Пусть будет

$$B = 3{,}6\ [°C].$$

На основе рис. 10

$$\alpha_1 = \arcsin\frac{0{,}2}{3{,}6} = 3°\ 10',$$

так как

$$\omega\tau = 7°\ 40',\ F_1^{(1)} = 2{,}2\ ;$$

$$\alpha_2 = 9°\ 40',\ \alpha_3 = 16°\ 10',\ \alpha_4 = 23°,\ \alpha_5 = 30°,\ \alpha_6 = 37°\ 50'\ ;$$

$$F_2^{(1)} = 2{,}2,\ F_3^{(1)} = 2{,}57\ ,\ F_4^{(1)} = 2{,}6\ ,\ F_5^{(1)} = 2{,}91\ ,\ F_6^{(1)} = 3{,}24\ ;$$

$$\alpha_7 = 46°\ 20',\ \alpha_8 = 56°\ 40'\ ;$$

$$F_1^{(2)} = 1{,}8\ ,\ F_2^{(2)} = 1{,}8\ ,$$

$$F_3^{(2)} = 2{,}2\ ,\ F_4^{(2)} = 2{,}245\ ,\ F_5^{(2)} = 2{,}6\ ,\ F_6^{(2)} = 2{,}71\ ,\ F_7^{(2)} = 3.$$

По обобщенным формулам работ [2, 3]:

$$b(B) = -\frac{1}{2}\left\{(H-A)\left[\frac{1}{2{,}6}\,l\left(\frac{3{,}6}{2{,}6}\right) + \frac{1}{3}\,l\left(\frac{3{,}6}{3}\right) - \frac{1}{2{,}57}\,l\left(\frac{3{,}6}{2{,}57}\right) - \right.\right.$$

$$\left. - \frac{1}{2{,}91}\,l\left(\frac{B}{2{,}91}\right)\right] - (H-C)\,\frac{1}{3{,}24}\,l\left(\frac{3{,}6}{3{,}24}\right) +$$

$$+ (H+A)\left[\frac{1}{2{,}2}\,l\left(\frac{3{,}6}{2{,}2}\right) + \frac{1}{2{,}6}\,l\left(\frac{3{,}6}{2{,}6}\right)\right] + (H+C)\,\frac{1}{3}\,l\left(\frac{3{,}6}{3}\right) -$$

$$-(H+A)\left[\frac{1}{2{,}245}\,l\left(\frac{3{,}6}{2{,}245}\right)+\frac{1}{2{,}71}\,l\left(\frac{3{,}6}{2{,}71}\right)\right]\Bigg\}=$$

$$=0{,}0245\,H+0{,}1725\,C+0{,}013\,A=495\,,$$

$$a(B)=\frac{4\,H\,2{,}2+2(H-A)\,(0{,}37+0{,}31)+2(H-C)\,0{,}24+2(H+A)\,(0{,}355+0{,}29)}{\pi\,3{,}6^2}=$$

$$=0{,}2935\,H-0{,}0118\,C-0{,}0017\,A=1730\,;$$

$$|w_H|=\sqrt{495^2+1730^2}=1800=65{,}1\,\partial\delta\,;$$

$$\varphi_H=74^\circ\,30'\,;$$

$$|w|=2{,}1\,\partial\delta\,:$$

$$\varphi=-180^\circ+29^\circ\,30'.$$

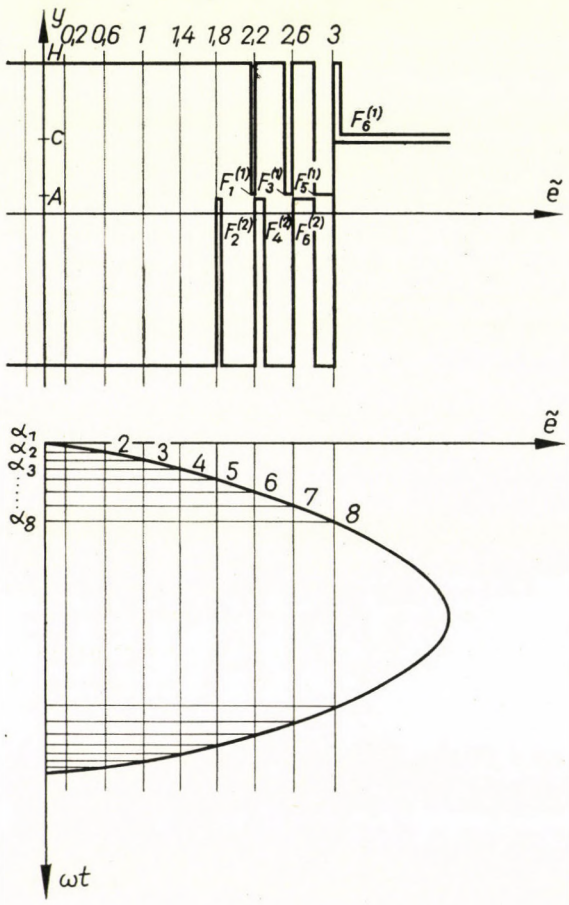


Рис. 10

Таблица IV

B		ω			
		0,5	1	2	3
	$\lvert W_л\rvert\ \partial\delta$	—49	—63	—79	—90
	$\varphi_л\sim$	—180°—26°36′	—180—45°	—180°—63°24′	—180°—74°30′
		1240	1200	1520	
	b	1240	1200	1520	1755
	a	1130	2280	4440	6750
	$\lvert W_H\rvert$	1680	2575	4700	6970
0,2 ·1,2=0,24	$\lvert W_H\rvert\ \partial\delta$	64,5	68,3	73,4	76,9
	φ_H	42°10′	62°20′	71°10′	75°30′
	$\lvert W\rvert\partial\delta$	15,5	5,3	—5,6	—13,1
	φ	—180°+15°34′	—180°+17°20′	—180°+7°46′	—180°+1°
	b	1175	1170	1740	2240
	a	1120	2450	4950	7300
	$\lvert W_H\rvert$	1625	2715	5340	7640
0,2 · 14=0,28	$\lvert W_H\rvert\ \partial\delta$	64,2	68,7	74,6	77,7
	φ_H	43°40′	64°30′	70°40′	72°50′
	$\lvert W\rvert\ \partial\delta$	15,2	5,7	—4,4	—12,3
	φ	—180°+17°04′	—180°+19°30′	—180°+7°16′	—180°—1°40′
	b	660	710	1015	1470
	a	785	1560	3075	4585
	$\lvert W_H\rvert$	1025	1715	3240	4815
0,2 · 3=0,6	$\lvert W_H\rvert\ \partial\delta$	60,2	64,7	70,3	73,6
	φ_H	50°	65°30′	71°50′	72°20′
	$\lvert W\rvert\ \partial\delta$	11,2	1,7	—8,7	—16,4
	φ	—180°+23°30′	—180°+20°30′	—180°+8°26′	—180°—2°10′
	b	410	470	905	1515
	a	1205	2170	4145	6090
	$\lvert W_H\rvert$	1270	2220	4510	6230
0,2 · 4=0,8	$\lvert W_H\rvert\ \partial\delta$	61	66,9	73,1	75,9
	φ_H	71°20′	77°50′	77°40′	77°50′
	$\lvert W\rvert\ \partial\delta$	12	3,9	—5,9	—14,1
	φ	—180°+44°50′	—180°+32°50′	—180°+14°16′	—180°+3°20′

Оценивая результаты приведенных выше расчетов, заметим следующее. Данная система при заданных выше параметрах является устойчивой. С ростом амплитуды запас устойчивости возрастает. Запас устойчивости имеет наименьшее значение в области

$$\omega = 1 \div 2\left[\frac{1}{мин}\right]; \; B = 0{,}2 \div 0{,}3\ ^\circ\mathrm{C}.$$

Запас устойчивости можно увеличить при заданном значении τ^i, уменьшив значение A. (Хотя в случае $B > s_{++}$ запас устойчивости, по-видимому, достаточен, можно увеличить его уменьшением значения C.)

Рассмотрим, например, как изменяются характеристики системы около критических значений амплитуды и частоты при изменении значения A на

$$150\left[\frac{имп}{мин}\right]$$

табл. V.

Таблица V

B		ω	
		1	2
0,24	b	810	930
	a	2300	4460
	$\lvert W_H\rvert$	2440	4555
	$\lvert W_H\rvert\ \partial\delta$	67,7	73,2
	φ_H	70°40′	78°10′
	$\lvert W\rvert\ \ \partial\delta$	4,7	—5,8
	φ	—180°+25°40′	—180°+14°46′
0,28	b	725	1340
	a	2480	5000
	$\lvert W_H\rvert$	2580	5180
	$\lvert W_H\rvert\ \partial\delta$	68,2	74,3
	φ_H	73°40′	75°
	$\lvert W\rvert\ \ \partial\delta$	5,2	—4,7
	φ	—180°+28°40′	—180°+11°36′

Таблица V показывает значительное увеличение запаса устойчивости.

Наши расчеты показывают, что для получения достаточного запаса устойчивости правую часть выражений (18) и (19) следует умножить на значение $(3 \div 5)$. Этот вопрос, конечно, надо рассматривать осторожно, так как этот вывод получен только на основе конкретного числового примера в конкретных условиях.

Наконец, отметим, что в результате уменьшения τ^i (и следовательно A и C) можно достичь того, что частотные границы двух упрощенных схем (рис. 6 и 7) соприкасаются и, таким образом, применение точных коэффициентов гармонической линеаризации станет излишним. Однако в данном случае это очевидно ведет к снижению динамических качеств системы.

IV. Заключение

Устойчивость данного нелинейного цифрового регулятора можно хорошо исследовать методом гармонической линеаризации. В настоящей работе дается способ для определения точных, зависящих от частот, коэффициентов гармонической линеаризации регулятора. Однако использование его для исследования устойчивости невозможно. В практических случаях появляется возможность для упрощений, которые позволяют получить соотношения, хорошо применимые при проектировании систем. (Выражения (18), (19) и (25). В тех частотных диапазонах, где эти упрощения не могут применяться, результаты необходимо проверить, используя точные коэффициенты гармонической линеаризации.

Исследования, проведенные аналитическим путем, могут служить исходным этапом для моделирования систем на аналоговых или цифровых вычислительных машинах для определения оптимальных параметров. — О результатах моделирования конкретного примера, рассмотренного в настоящей работе, можно получить сведения из работы [4].

ПРИЛОЖЕНИЕ

Автор настоящей работы разработал обобщенный способ гармонической линеаризации [2, 3] кусочно-линейных характеристик. При этом были получены обобщенные формулы для определения коэффициентов гармонической линеаризации. В настоящей работе фигурируют только характеристики конечными разрывами (без линейных участков). Для таких характеристик в обобщенной формуле определения „вещественного" коэффициента гармонической линеаризации фигурируют только величины:

$$\frac{D_i}{F_i}\ l\left(\frac{B}{F_i}\right),$$

где D_i величина разрыва;

F_i абсцисса точки разрыва;

$l\left(\frac{B}{F_i}\right) = \frac{4}{\pi}\frac{F_i}{B}\sqrt{1-\frac{F_i^2}{B^2}}$ специальная функция, введенная для определения коэффициента.

ЛИТЕРАТУРА

1. *Попов, Е. П.—Пальтов, И. П.:* Приближенные методы исследования нелинейных автоматических систем. Физматгиз, Москва 1960.
2. **Somló J.: Egyenes szakaszokkal közelített hiszterézis hurok típusú nemlinearitások leíró függvényei** *(Шомло, Я.:* Описывющие функции нелинейностей типа петли гистерезиса, кусочно-линейной аппроксимацией). **II. Automatizálási Kollokvium, Budapest 1963.**
3. *Шомло, Я.:* Обобщенный метод для исследования систем с кусочно-линейными характеристиками I. *Автоматика и Телемеханика* (1966), № 7.
4. **Singer D.: Vizsgálatok a közvetlen digitális szabályozásról** *(Зингер, Д.:* Исследования в области непосредственного цифрового управления). **MTA Automatizálási Kutató Intézet zárójelentése, Budapest 1967.**
5. *Федоров, С. М.—Литвинов, А. П.:* Автоматические системы с цифровыми управляющими машинами. *Энергия* (1965).
6. *Круг, Е. К.—Александриди, Т. М.—Дилигенский, С. Н.:* Цифровые регуляторы. *Энергия* (1966).

Investigation of the Stability of a Digital Controller *(J. Somló)*. The paper deals with the investigation of the stability of a non-linear digital controller by the describing function method. By simplifications it is possible to determine the analytical relations which serve to select the parameters guaranteeing stable working. Whether the selected parameters are suitable or not can be checked with the aid of the accurate describing functions. In the paper the earlier results of the author are used. The investigation is connected with a definite problem of application.

Untersuchung der Stabilität eines digitalen Reglers *(J. Somló)*. Die Arbeit behandelt die Stabilitätsuntersuchung eines nichtlinear funktionierenden Reglers mit Hilfe von Funktionen für die Beschreibung der Nichtlinearität. Vereinfachungen ermöglichen, diejenigen analytischen Zusammenhänge zu bestimmen, welche zur Auswahl der stabiles Funktionieren sichernden Parameter dienen. Ob die ausgewählten Parameter entsprechen, kann mit Hilfe der genauen Beschreibungsfunktionen kontrolliert werden. In der Arbeit werden die früheren Ergebnisse des Verfassers mit verwendet. Die Untersuchungen knüpfen sich an einen konkreten Anwendungsfall an.

Examen de la stabilité d'un régulateur digital *(J. Somló)*. L'étude traite de l'examen de la stabilité d'un régulateur digital de fonctionnement non-linéaire par la méthode des fonctions décrivant la non-linéarité. Des simplifications permettent de déterminer les relations analytiques appropriées pour le choix des paramètres assurant un fonctionnement stable. Pour vérifier si les paramètres choisis sont convenables, on a recours aux fonctions descriptives exactes. L'étude utilise les résultats antérieurs de l'auteur; l'examen décrit est lié à un cas d'application concret.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 101—136 (1968)

EFFECT OF STRUCTURAL IRREGULARITIES ON AVALANCHE BREAKDOWN IN $p-n$ STEP JUNCTIONS

A. P. KEMÉNY

INDUSTRIAL RESEARCH INSTITUTE FOR ELECTRONICS "HIKI", BUDAPEST

[Manuscript received June 13, 1967]

Besides surface phenomena the avalanche breakdown voltage of transistors and diodes is reduced by geometric or resistivity irregularities along the junction. Recent theories based on the one-dimensional solution of the Poisson-equation do not hold for the practical case when a hemispherical or hemicylindrical curvature is "embedded" in the planar section of the junction. Though the problem is two-dimensional and seems not accessible to analytical solution, an approximation based on the known interrelation between the volumetric breakdown voltages BV_{CBO} and BV_{CEO} of transistors has been stated. High-current pulsed measurements eliminate disturbances due to surface avalanche. It was found that firstly, the higher the resistivity of the light doped side of the junction is (i.e. by high voltage devices), the larger is the breakdown-voltage reducing effect of volumetric irregularities; secondly, the breakdown-voltage decreases most if the radius of the embedded curvature (or irregularity) amounts to appr. one fifth of the depletion layer width. Irregularities very small in comparison with the depletion layer width do not cause appreciable degradation.

I. Introduction

The breakdown voltage of an $n - p^+$ or a $p - n^+$ germanium step junction is given by S. L. MILLER and others [1—3] in the semi-empirical form

$$BV_{CBO} = 85\,\varrho^{0,65} = 85\,(\mu q N_I)^{-0,65}, \tag{1}$$

where ϱ is the resistivity of the lower doped side [ohmcm], μ the majority carrier mobility [cm²/Vs], q the electron charge [As] and N_I the net impurity concentration [cm^{-3}]. It is a well-known fact that in practice this expression is only valid in the rather low resistivity range of 0,1 ohmcm. As the resistivity increases, with an ever-growing spread the measured breakdown voltage BV_{CBOD} falls short of the theoretical breakdown voltage, determined by equation (1) [D stands for "degraded"]. The deviation from this value is considerable, since for example at 15 ohmcm in an unfavourable case the measured breakdown voltage BV_{CBOD} may be as low as one fourth of the theoretical amount. Owing to this phenomenon no germanium diode and transistor with an operating voltage higher than 150 ÷ 200 volts is available.

It is a well-known fact that especially at low current levels, when the carrier multiplication does not exceed the order of magnitude of 10^3, the decrease in surface breakdown stems from the adsorbed and partly ionized

impurities i.e. charges, near the external boundary of the junction causing a field which penetrates the transistor body to a depth of about one Debye-length (some 1000 Ångstroms). This breakdown is sometimes multiplicative [1, 4, 5, 11, 13] and reduces the overall breakdown of the device, when fixed surface charges have a sign opposite to that of the fixed body impurities in the lower doped side of the junction. In germanium $n - p^+$ junctions the positive surface charge, due e.g. to adsorbed water, diminishes the breakdown voltage narrowing the depletion layer at the very surroundings of the surface, thus raising the field strength [4]. The amount of degradation depends mainly on the fixed net charge concentration in the body of the lower doped side. At heavy doping, i.e. low resistivity and small Debye-length, as well as at not too low multiplication factors, the effect of a relatively small charge concentration on the surface can be neglected [4, 10, 11]. So the significance of this problem grows with increasing resistivity.

It must be emphasized, however, that this behaviour is characteristic only for low multiplication factors, e.g. relatively low avalanche current, when the breakdown current flows mainly toward the surface. At higher current levels, when multiplication factors approaching 10^5 occur, a high voltage drop will be developed across the high dynamic resistance of $10^3 \div 10^4$ ohms of the small cross-sectioned surface structure [2]. Thus the avalanche current is forced to flow mainly in the body, and the voltage drop across the junction may approach the breakdown voltage of the transistor body (the latter having the much lower dynamic resistance of $10 \div 10^2$ ohms, due to the body spreading resistances and the space charge pinch-in effect of the carrier flow pattern [1, 2]).

The importance of body breakdown rises when current multiplication is high or surface breakdown voltage exceeds the body breakdown. The role of body irregularities or junction geometry is widely discussed [1—4, 6—9].

According to our investigations the decrease of BV_{CBO} results, besides surface breakdown, from body causes, due to deviations of the junction geometry from the ideal planar homogeneous one. This is especially important when the device is working in the high-current avalanche mode or some measures are taken to raise the surface avalanche voltage considerably above the body breakdown, e.g. with the aid of the surface contouring technique, described by R. L. Davies and F. E. Gentry [6].

The direct causes of reduced BV_{CBO}, i.e. the increase in the field strength E_{MB} or the narrowing of the depletion layer width X_B can be attributed, either to a geometric curvature or a low resistivity spot inside the depletion layer. Irrespective of the resistivity range, in case of ideal planar homogeneous junctions the body breakdown occurs exactly as prescribed by equation (1), down to the Zener region and up to resistivities approaching the intrinsic value [2].

A. P. SHOTOV [2] generated very high multiplications (about $10^4 \div 10^5$) in abrupt germanium $p - n$ junctions of $\varrho \leqq 27$ ohmcm with high current pulses and found, in accordance with our results, a high spread at high resistivities, due obviously to body irregularities, though the highest values measured were in agreement with equation (1). In this work we demonstrated the validity of expression (1) at least up to $\varrho = 10$ ohmcm.

Our investigations are concerned mainly with the determination of the highest possible degradation caused by either a geometric curvature embedded in the surrounding planar junction or a spot of elevated doping inside the depletion layer. It was concluded that both irregularities equally decrease the body breakdown voltage by raising the electric field strength along the irregularity. If a very low resistivity inhomogeneity spot within the lower doped material lies directly on the boundary of the abrupt junction, where maximum electric field strength occurs, the question shrinks to the two-dimensional field problem of a geometric junction curvature embedded in the planar section and may be treated similarly.

The chief task is to determine the effect of curvature radius of the imperfection. Earlier works deal with solitary curved junctions [7—9], e.g. a hemisphere, involving a one-dimensional field problem and pure analytical solutions could easily be found on grounds of the Poisson equation. In this case by decreasing r_0 (radius of curvature) the field is forced to concentrate at a point and the maximum field strength along the junction tends to infinity: with diminishing radius the breakdown voltage drops monotoneously (cf. Fig. 1d).

In our case, when the curvature is surrounded by the planar junction as shown in Figs 1b and 1c, the radius of curvature does not define the situation unambiguously: with vanishing radius r_0 (Fig. 1c) the range of elevated field strength becomes negligible compared to the depletion layer width, though the gradient of field strength increases considerably only at the very edge of the curvature. Under such circumstances the avalanche multiplication cannot be determined solely by the maximum field strength E_{MB}, as in planar step junction (Fig. 1a) characterized by a linear increase of the field strength along the depletion layer. Severe degradation takes place if the radius of imperfection approaches the depletion layer width (Fig. 1b), when the field pattern is the most distorted and the range of elevated field strength covers the whole width X_{BD} of the depletion layer under the curvature. When the radius r_0 and with it the region of raised field strength along the imperfection becomes comparable to, or falls short of, the magnitude of the mean free path Π_f for interactions between minority carriers and phonons (about 140 Ångstroms according to S. L. MILLER [1]), it is obvious that the multiplication does not further grow considerably and BV_{CBO} cannot approach zero. On the other hand, with vanishing r_0 the field pattern will be more and more similar to that of the planar junction; the field strength ascends almost linearly along

the depletion layer width except in the immediate vicinity of the imperfection, without any appreciable growth in avalanche multiplication. *Because of the planar surroundings the concentration of field does not take place* as it does in the case of Fig. 1d, so E_{MBD} tends to E_{MB} and consequently the breakdown voltage approximates the non-degraded theoretical value of the planar homogeneous case. It follows that the degradation function $BV_{CBOD}(r_0)$ must have

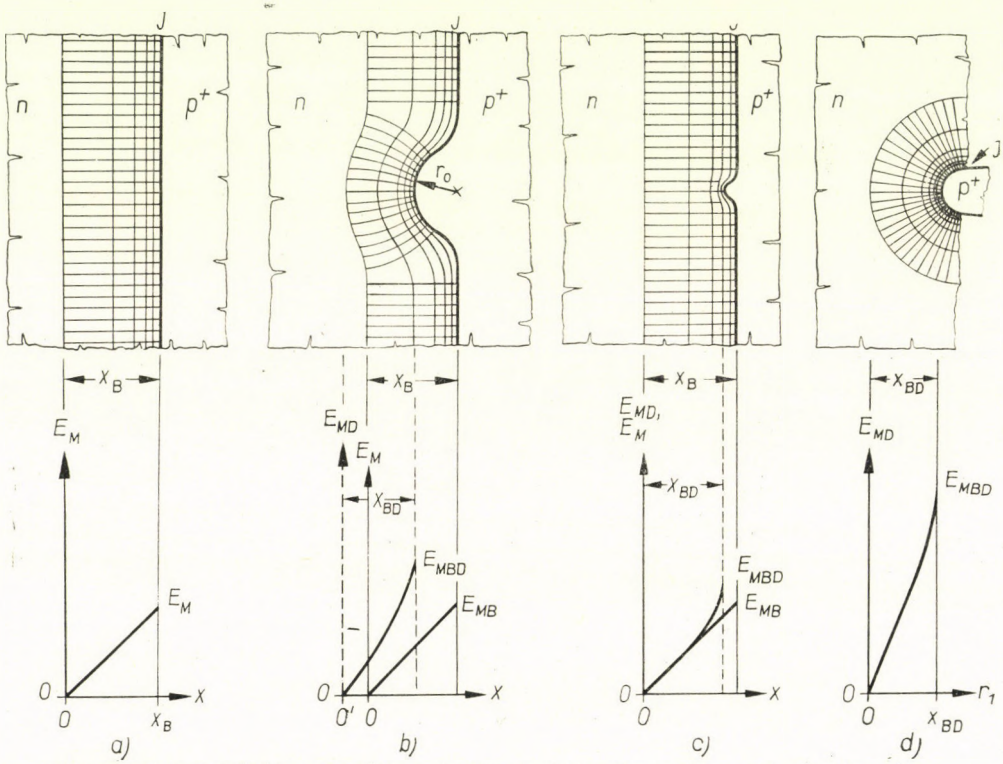


Fig. 1. Electrical field at breakdown in the depletion layer of $n - p^+$ step junctions: Planar junction or $r_0 \gg X_B$ (*a*); curvatures embedded in a planar section, $r_0 \cong X_B$ (*b*); same as (*b*) but $r_0 \ll X_B$ (*c*); solitary hemispherical junction (*d*)

a minimum i.e. there has to be a radius of imperfection r_{0p} at which the highest degradation of breakdown voltage occurs. This most dangerous radius must be near to the depletion layer width. At $r_0 = \infty$ and $r_0 = 0$ there is no degradation at all in breakdown voltage.

It is almost hopeless to solve the two-dimensional field problem with pure analytical means based on the Poisson equation and the complicated boundary conditions. An approximative solution could be found if it would be possible to determine the ratio of the degraded breakdown voltage BV_{CBOD} to the non-degraded breakdown voltage BV_{CBO} in individual $p - n$ junctions.

In transistors the known relation between collector-base and collector-emitter breakdown voltages is made use of. It will be shown in the following how to determine the highest degradation with a sufficient number of such measurements.

In the present work only germanium $p - n - p$ alloyed devices will be considered.

II. Recent theories and their limitations

The theories concerning geometric curvatures have been developed by H. L. ARMSTRONG et al. [7, 8] and J. SHIELDS [9] for solitary hemispherical and cylindrical junctions (see Fig. 1d), the practical case of gold-bonded diodes. In junction diodes and transistors, when the curvatures of the junction are embedded in the surrounding planar region, the validity of this theory is restricted.

Expressing the results obtained from the Poisson equation in normalized terms of BV_{CBOD}/BV_{CBO}, and taking r_0/X_B as independent variable (X_B denotes the width of the depletion layer in the planar system in case of breakdown, at a given ϱ), for cylindrical junctions we obtain

$$\frac{BV_{CBOD}}{BV_{CBO}} = \frac{1}{2}\left(\frac{r_0}{X_B}\right)^2 \cdot \left[\left(1+\frac{2X_B}{r_0}\right)\cdot \ln\left(1+\frac{2X_B}{r_0}\right) - \frac{2X_B}{r_0}\right] \tag{2}$$

and for hemispherical junctions

$$\frac{BV_{CBOD}}{BV_{CBO}} = \left\{\frac{2r_0/X_B}{1+\left(\frac{r_0}{X_B}\right)^2 \cdot \left[\left(1+3\frac{X_B}{r_0}\right)^{2/3}-1\right]}\right\}^2 \tag{3}$$

as shown in Fig. 10.

At $r_0/X_B \geq 10$ both equations yield approximately unity, the characteristic value for planar junctions. Here the depletion layer width of the planar case at breakdown amounts to

$$X_B = (2\,\varepsilon BV_{CBO}/qN_I)^{1/2} \tag{3a}$$

with ε the dielectric permittivity of germanium $= \varepsilon_r \times \varepsilon_0 = 16{,}5 \times 8{,}86 \times 10^{14}$ farads/cm.

One of the effects so far neglected in the above expressions is that of the "resolving" effect of Π_f. When r_0 reaches or falls short of the mean free path Π_f of the carrier-phonon collision, it is obvious that BV_{CBOD} cannot approach zero. Thus irregularities comparable to Π_f, e.g. dislocations, do not

affect the breakdown voltage [3]. As to the second effect, if the irregularities of hemispherical or cylindrical shape are embedded in the planar junction (Figs 1b and 1c) we have found that the validity of equations (2) and (3) was limited by the following two criteria:

$$r_0/X_B \geqslant 0{,}4 \tag{4}$$

and

$$X_B \rangle\rangle\rangle \Pi_f. \tag{5}$$

Equation (4) will be explained later; equation (5), expressing the "resolving" effect of the mean free path, holds only when the resistivity is higher than some 10 ohmcm. First of all this "resolving" effect of the mean free path can be made responsible for the fact that at low resistivities the inhomogeneity of the junction causes but a negligible drop in BV_{CBO}.

III. Determination of n in the multiplication function by breakdown measurements

According to S. L. MILLER [1] in case of step junction transistors the ratios of breakdown voltages can be expressed in the form

$$BV_{CBO}/BV_{CEO} = (1 + \beta_0)^{1/n} = (I_{CEO}/I_{CBO})^{1/n} \tag{6}$$

where BV_{CEO} is the avalanche breakdown voltage between collector and emitter, β_0 the small-signal common-emitter current amplification factor at the reverse current I_{CEO} and n is the exponent in the multiplication function, determined on germanium $n - p^+$ junctions by S. L. MILLER in the range $0{,}1 \div 2$ ohmcm and with pulsed breakdown measurements up to 20 ohmcm by the author, is as follows.

If the lower doped side of the junction has high resistivity, there are two factors rendering the determination of the function $n(BV_{CBO})$ or $n(\varrho)$ very difficult. First of all the decrease in BV_{CBO} is subject to surface phenomena, e.g. surface avalanche. This disturbing effect can be avoided if the measurement of BV_{CBO} is performed at a very high current, since for surface currents the small crosssection of surface breakdown represents a high dynamic resistance [1], i.e. a shunt, negligible compared to the low dynamic resistance (appr. $20 \div 100$ ohms), exhibited at the high current of body breakdown [1, 2], as will be pointed out in Appendix 1. To minimize heat generation, a train of current pulses with a low duty cycle should be applied [1, 2].

Besides surface phenomena, the body effects presently discussed can also reduce the measured value of BV_{CBO}. To remove this difficulty the following measurement has been carried out.

R. A. SCHMELTZER [12] determined the collector-emitter breakdown voltage (BV_{CE}) of transistors with forward biased emitter-base diode as a function of the stability factor S_i:

$$BV_{CE} = BV_{CBO} \cdot S_i^{-1/n} \tag{6a}$$

where

$$S_i = \frac{dI_C}{dI_{CBO}} = \frac{(1+\beta)(R_B + R_E)}{R_B + R_E(1+\beta)} \tag{6b}$$

and β denotes the low voltage common-emitter current amplification factor at the given d.c. collector current I_C; R_E stands for the emitter resistor and R_B for the total resistance of the base circuit between the internal base point and the "cold" terminal of R_E (for example R_B is the sum of the series base resistor and the spreading resistance $r_{BB'}$).

Plotting BV_{CE} against S_i one may use the circuit depicted in Fig. 2. Suitable choice of the external resistances R_E and $R_B - r_{BB'}$, allows the adjustment of S_i between unity ($R_B = 0$ and $R_E \gg r_{BB'}$) and $1 + \beta$ (when R_B is high and $R_E = 0$). At zero output of the pulse generator ($V_C = 0$) the required emitter current can be set by the voltage V_E with the aid of ammeter I_{EO}. At the low voltage V_{CO} a current $I_C \cong I_E$ flows without multiplication. Gradually increasing the voltage V_C of the pulse generator, the multiplicated current can be adjusted by means of the peak meter $M' \times I_C$, measuring the peak voltage across R_m at a chosen multiplication factor M' ranging from 20 to 50. BV_{CE} is indicated by a peak-reading VTVM. When the punch-through voltage V_{pt} falls short of the avalanche voltage BV_{CE}, falsifying the measurement, the pointer of the instrument V_{BE} deflects considerably.

On determining the function $BV_{CE}(S_i)$ in the pulsed test arrangement of Fig. 2, we must find that the log-log plot of the function BV_{CE} versus S_i results in a straight line of slope $-1/n$ decreasing towards growing S_i-values (see Fig. 3a) in accordance with Schmeltzer's results. At $S_i = 1$, (thus $R_E \gg R_B$), equation (6a) gives $BV_{CE} = BV_{CBO}$, the same which can be calculated from ϱ with formula (1). If log BV_{CE} is plotted against log S_i, for low voltage transistors made with a base material of low resistivity (lower than 1 ohmcm), really straight lines are obtained, defining at $S_i = 1$ the breakdown voltage $BV_{CE} = BV_{CBO}$ and the slope which is in good agreement with the value n given by S. L. MILLER [1], corresponding to a resistivity evaluable by putting the measured breakdown voltage $BV_{CE} = BV_{CBO}$ into equation (1) (see Fig. 3a).

Irrespective of ϱ, at high S_i values a straight line has always been observed and at $S_{i\,\max} = 1 + \beta$, according to equation (6a), the curve terminates in the breakdown voltage BV_{CEO}, which is in good agreement with the results of a simple measurement to be described in the next chapter.

In transistors designed for higher operating voltages, i.e. with $\varrho > 2$ ohmcm, at low S_i-values the function $BV_{CE}(S_i)$ will fall short of the theoretical straight line as shown in Fig. 3b. The reduction of BV_{CE} with reference to the theoretical breakdown voltage is the greater, the lower is S_i. The degraded breakdown voltage BV_{CBOD}, determined by means of measurement at $S_i = 1$, does not reach the avalanche voltage calculated from ϱ by formula (1). This phenomenon is presumably the result of body irregularities in the collector-base junction, as will be shown in the following, and perhaps partly of remaining surface breakdown effects.

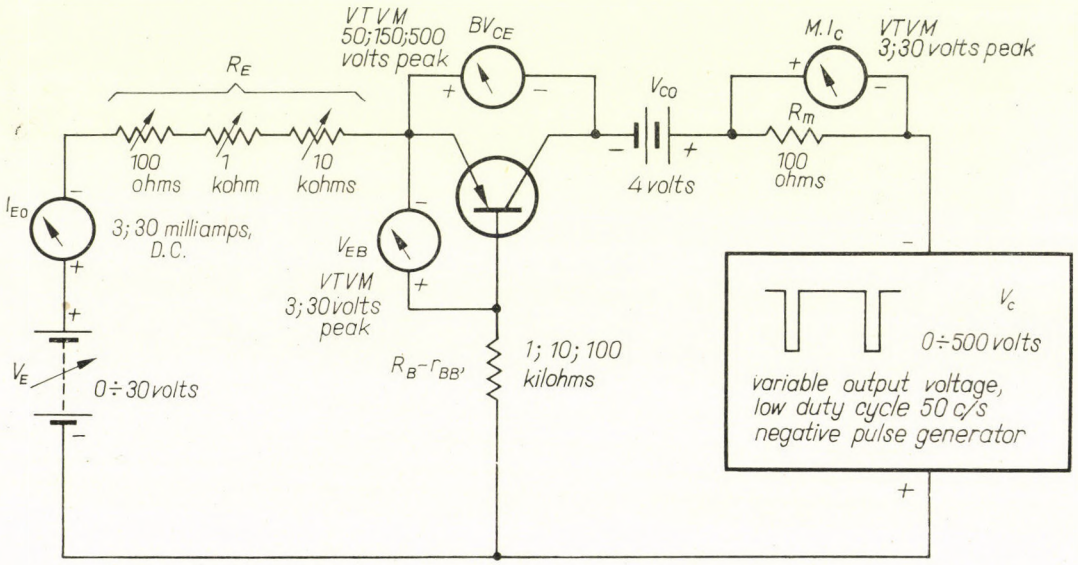


Fig. 2. Block diagram for pulsed measurements of the relationship $BV_{CE}(S_i)$

According to the measurements undertaken on a great number of alloyed germanium transistors, the degradation of $BV_{CE}(S_i)$ increases at low S_i values with the base resistivity ϱ, i.e. with the maximum operating voltage of the transistor.

When BV_{CE} is measured at a sufficiently high number of S_i-values, the breakdown voltage BV_{CBO} underlying equation (1) and the value of n subject to formulae (6) and (6a) can be determined even in case of severe degradation if the straight line obtained at high S_i-values is extrapolated to $S_i = 1$ as shown in Fig. 3b. This method presupposes that at high S_i-values the slope $-1/n$ corresponds with the required value of n. This assumption is supported by two facts. First, the n-values, computed at high S_i-values from the straight section of the $BV_{CBO}(S_i)$ curve in transistors with a relatively low base resistivity ($\varrho < 2$ ohmcm) by the author, on one hand, and from the extrapolated breakdown voltage BV_{CBO} with aid of equation (1) and the

function $n(\varrho)$ given by S. L. MILLER [1], on the other, are exactly equal. This obviously holds for high resistivities, too, so it can be assumed that the slope of the straight section gives the required n-value in all circumstances. Secondly, at $S_{i\,\max} = 1 + \beta$, in the multiplication function

$$M^{-1} = 1 - (V_{CB}/BV_{CBO})^n \tag{7}$$

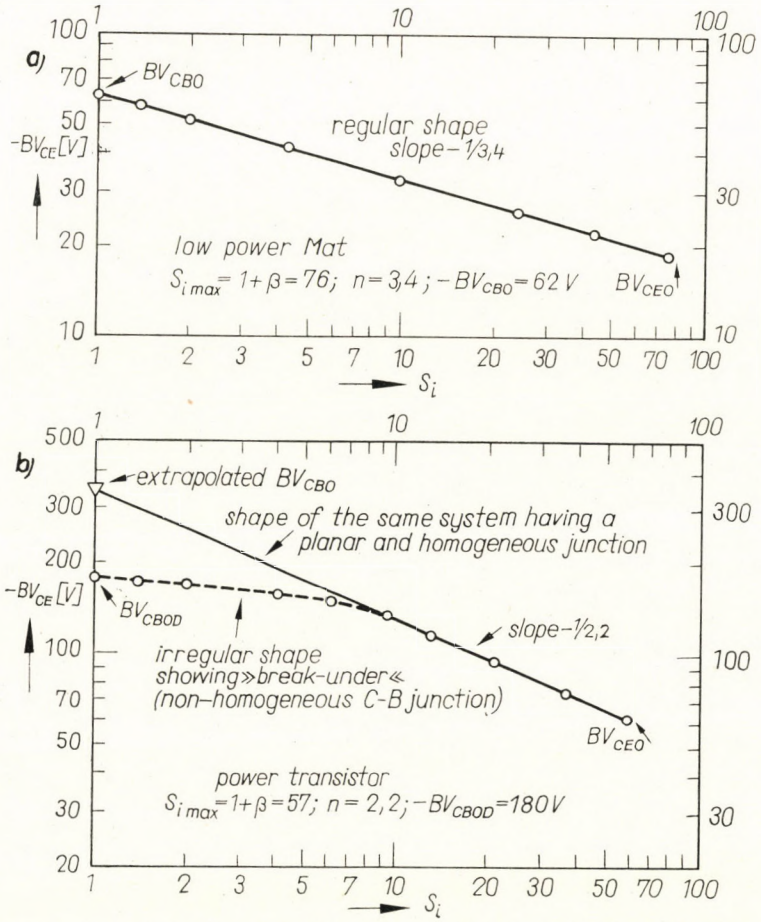


Fig. 3. Log-log plots of the function $BV_{CE}(S_i)$: regular shape (*a*); irregular shape exhibiting heavy degradation of BV_{CBO} at low S_i (*b*)

of S. L. MILLER [1] M hardly increases above unity, since at the corresponding $V_{CB} = BV_{CEO}$ voltage

$$M = \alpha^{-1} = 1 + \beta^{-1}$$

and even at high S_i-values the carrier multiplication remains close to unity. Thus, at high S_i values the increase of multiplication due to elevated field strength at the spot of irregularity will be negligible, but the breakdown

voltage BV_{CEO} is unambiguously determined by BV_{CBO} and β of the planar part of the junction, that is to say, by the theoretical expression (6) and is not affected by carrier multiplication. On the contrary, at low S_i-values the multiplication becomes very high and the increased field strength along the irregularity could thoroughly degrade the value of the breakdown voltage.

On measuring the breakdown voltage and by extrapolation of the results obtained at high S_i-values, it was possible on the one hand, to prove with a

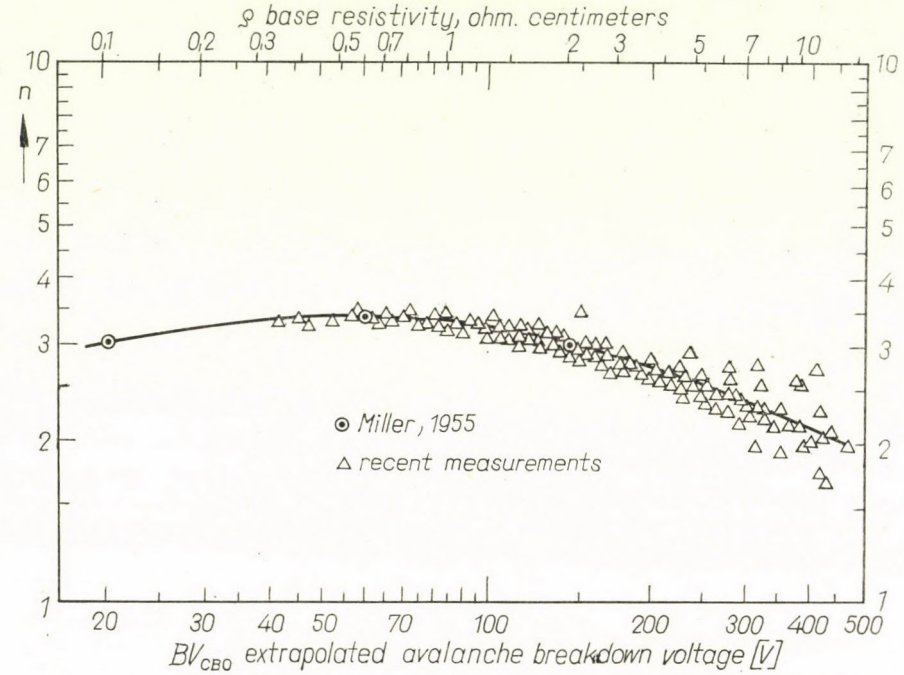


Fig. 4. Function $n(BV_{CBO})$ or $n(\varrho)$ resp. measured on a great number of alloyed germanium *pnp* transistors in the circuit shown in Fig. 2

different method, on different germanium *pnp* alloy-junction transistors the accuracy of S. L. MILLER's investigations in the resistivity range $0{,}3 \div 2$ ohmcm. On the other hand, the theoretical value of the parameter n could in this way be determined as a function of resistivity in the range $2 \div 15$ ohmcm, not investigated by S. L. MILLER supposedly on account of the disturbing effect of the body phenomenon described previously (see Table I). The

Table I

ϱ [ohmcm]	0,1	0,25	0,6	1	2	3	5	7	10	15	20
n (approx.)	3	3,3	3,4	3,3	3	2,85	2,55	2,35	2,17	2	1,9
BV_{CBO}[volts]	20	35	60	85	140	175	240	300	380	480	600

values of $n(\varrho)$, found at high resistivities, are to be seen as the natural continuation of S. L. MILLER's results, as is shown in Fig. 4.

IV. Measurement of the breakdown voltages BV_{CBOD} and BV_{CEO} — the measure of irregularity

It is rather elaborate to determine the degradation of the breakdown voltage with the method described in the previous chapter, as for proper statistical confidence a great number of measurements would be required. In possession of the "theoretical" function $n(\varrho)$ given in Fig. 4 and Table I it is enough if BV_{CBOD} and BV_{CEO} are measured in the two extreme positions corresponding to $S_i = 1$ and $S_i = 1 + \beta_0$. According to formula (6), $1 + \beta_0$ can be obtained with the required accuracy by observing the ratio of the reverse currents I_{CEO} and I_{CBO} ensuring that both reverse current measurements are performed at the same temperature and at low voltages where multiplication does not take place. This measurement is simple and suitable for mass investigation.

Naturally, the reverse currents and the breakdown voltages BV_{CBO} and BV_{CEO} must be determined at the same temperature — most conveniently at room temperature — otherwise equation (6) does not hold owing to the temperature dependence of the current amplification factor β_0. It is known that the reverse current I_{CBO} has a significant surface component that possibly disturbs the measurement. At high temperatures, however, the reverse current consists only of a body component, so the author performed the reverse current measurements at an elevated temperature (e.g. 75 °C) on alloyed transistors, extrapolating the obtained I_{CBO} to 25 °C on the basis of the known relationship

$$I_{CBO(T_2)} = I_{CBO(T_1)} \cdot \exp\left[\frac{qV_f}{kT_1^2}(T_2 - T_1)\right] =$$

$$= I_{CBO(T_1)} \cdot \exp\left[\Phi\,(T_2 - T_1)\right],$$

where qV_f is the forbidden band gap energy in electronvolts (V_f 720 mV for germanium), k the Boltzmann's constant, so $\Phi = 8 \times 10^{-2}/°C$ in this temperature range. Since the reverse currents determined at high temperature are void of surface components, the extrapolated reverse currents will be lower than those, measured at room temperature. Nevertheless the *ratio* of the extrapolated reverse currents agrees fairly well with that of the reverse currents at room temperature. Thus, $1 + \beta_0$ may be determined by means of measuring

$$I_{CEO} \text{ and } I_{CBO} \text{ at } T_a = 25\,°C\,.$$

Due to the irregularity of the collector-base junction as well as to surface effects, the measured value of BV_{CBOD} is in most cases smaller than BV_{CBO}. Since the breakdown voltage BV_{CEO} can be precisely measured at small multiplication, the „degraded" value n_D, computed in accordance with equation (6) as it is given later by equation (8a), will *exceed* the n-value of equation (6), belonging to the "non-degraded" BV_{CBO} value of the planar-homogeneous junction. To eliminate surface avalanche and thermal breakdown effects, the avalanche voltages BV_{CBOD} and BV_{CEO} are to be determined by means of high current pulses, that is $30 \div 100$ mA in low power and at $200 \div 500$ mA in power transistors by a method treated in the following chapter. The value of n_D could thus be obtained with the required accuracy.

Measuring a large number of transistors of different base resistivity ($\varrho = 0{,}4 \div 15$ ohmcm), Fig. 5 gives n_D as a function of the measured breakdown voltage BV_{CBOD}. It can be seen that the majority of the transistors with $\varrho < 0{,}6$ ohmcm or $BV_{CBOD} < 60$ V have a parameter $n_D \simeq n$ in accordance with the theoretical value of Table I, showing that these devices are not subject to degradation. With increasing ϱ, i.e. in parallel to the rise of BV_{CBOD}, the median value and the spread of n_D rapidly grow, especially in power transistors (see Fig. 5b), since there is a higher probability of finding irregularities of critical curvature on the extended alloyed areas.

In knowledge of the degraded breakdown voltage BV_{CBOD} we can determine a corresponding value n either from Table I, or from Fig. 4, substituting BV_{CBOD} with BV_{CBO} in $n(BV_{CBO})$. Naturally n_F, a parameter obtained by this arbitrary substitution* will be equal to the calculated value n_D only if the measured collector-base breakdown voltage is not degraded ($BV_{CBOD} = BV_{CBO}$), i.e. in case of a planar-homogeneous junction and in the absence of surface effects.

As $n_D = n_F = n$ means that the breakdown voltage is not degraded and the resistivity calculated from equation (1) with the assumption $BV_{CBO} = BV_{CBOD}$ yields the mean value of ϱ in the base wafer of the transistor in question. On the contrary, when the ratio n_D/n_F is substantially higher than unity, there is a severe degradation of BV_{CBO}. Thus

$$\lambda = n_D/n_F$$

can be taken as a rough measure of inhomogeneity. It can easily be read from Fig. 6, where the median BV_{CBOD} of power transistors is plotted against the median λ, that, as stated before, the higher the BV_{CBOD}, the higher is the possible degradation.

* F stands for fictive, since n_F has no physical content and is only an auxiliary parameter for further calculation.

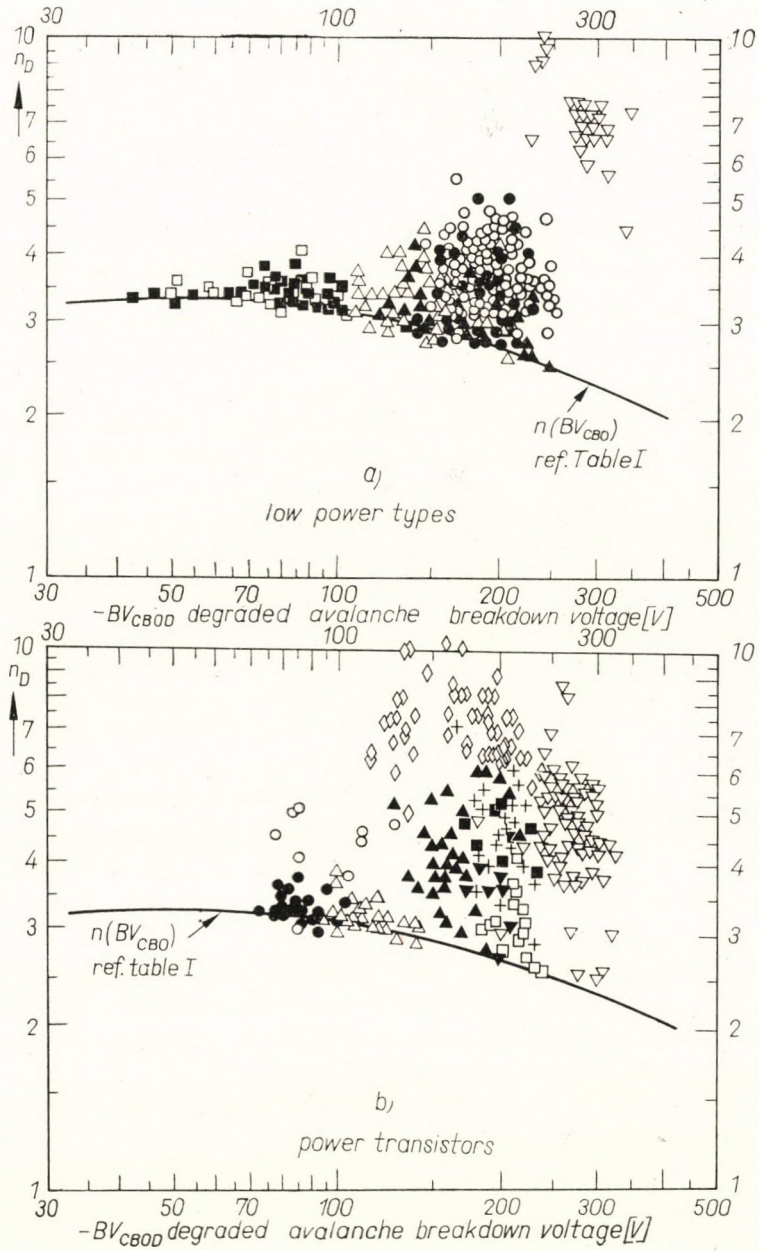


Fig. 5. n_D versus BV_{CBOD} plots, obtained from measurement and calculation for a great number of alloyed germanium *pnp* transistors

Low power types (*a*): △, ▲ *OC 70, OC 71, OC 72*; ○, ● *OC 76, OC 77*; □, ■ *OC 44, OC 45, 2N218, 2N219*; ▽ *2SB150*

Power types (*b*): ○, ● *2N301, 2N301A*; △ *2SB236*; ▲ *AD1202, AD 1203*; ▽ *ADZ 12*; ▼*2N1046*; □, ■ *AD 103, AD 105*; + *ASZ 15, ASZ 18*; ◇ early experimental type

The number of transistors investigated amounts to 15 ÷ 370, depending on the type. The transistors 2*N 1046* and *AD 103* lie under the straight line indicating a better than average homogeneity.

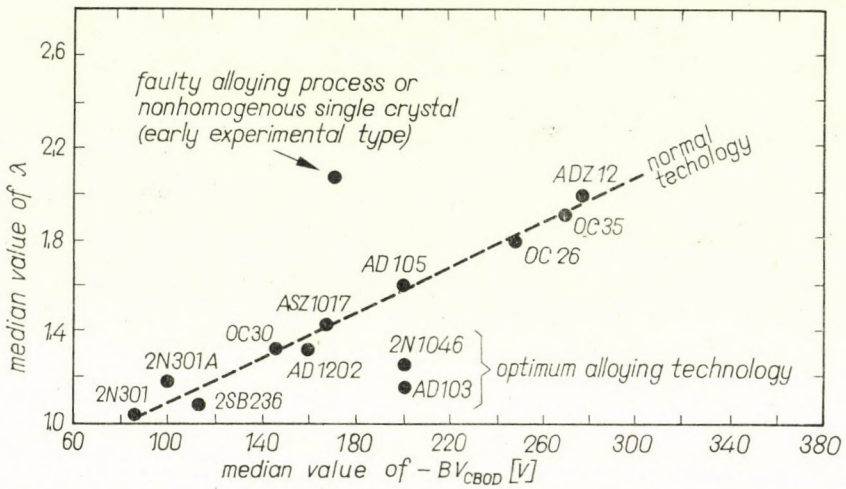


Fig. 6. Median λ plotted against median BV_{CBOD} for alloyed germanium *pnp* power transistors

V. Exclusion of surface effects on breakdown measurement

In another paper [13] the author describes how to determine the true body breakdown voltage along the $p - n$ junction itself by measurement, excluding the following disturbing effects:

a) Surface breakdown due to surface charges;

b) excess voltage drop across body resistance on both sides of the junction, especially due to the spreading base resistance $r_{BB'}$ in the measurement of BV_{CBOD};

c) effect of the space charge of carriers in the base region of raised field strength, with increasing current (pinch-in effect); this results in a considerable voltage drop.

Ad a) As discussed in Section I, fixed surface charges (e.g. adsorbed water on the surface oxide structure in the form of dipole multilayers [14] representing a positive charge with respect to the oxide coated body) can affect the body breakdown seriously only in case of low avalanche multiplication ($M < 10^3$), high surface charge density and high body resistivity on the lower doped side. The field of these fixed surface charges narrows the depletion layer of the *n*-type side in the very vicinity of the surface, raising the field

strength [4] and causing a surface breakdown which occurs at lower voltages than body breakdown.

At moderate charge densities the field perturbation due to such surface states does not penetrate, according to literature [4, 10, 11], deeper than one Debye-length ranging from some hundred Ångstroms to about 6000 Ångstroms, as a function of the resistivity of the lower doped side (see Appendix 1). The active cross section of the surface breakdown is, therefore, very small in spite of the large junction periphery (in practical cases $2\times10^{-4} \div 8\times10^{-3}$ mm^2), compared to the large junction area of the body. It is pointed out in the Appendix 1 with the aid of an approximative calculation that the voltage-current characteristic of the small cross-sectioned surface breakdown exhibits a differential resistance of $3\times10^2 \div 10^4$ ohms depending on geometry and resistivity: types of higher base resistivity and lower power yield higher differential resistances.

This high differential resistance stems primarily from the narrow current path in the base region, as it can be seen in Fig. 12. At low multiplication ($M < 10^3$) or junction current densities smaller than $1 \div 3$ mA/mm^2 there is no obstacle for the avalanche current in flowing mainly toward the surface owing to the fact that its breakdown voltage falls short of the body breakdown voltage and the relatively low voltage drop across the (series) differential resistance. This breakdown current range characterized by the high differential resistance may be attributed to pure surface effects.

Contrary to this, when multiplication or avalanche current density becomes high ($M > 10^4$ or $i > 30$ mA/mm^2) the voltage drop across the differential resistance of the surface rises enormously, forcing the avalanche current to flow in the body structure where the differential resistance (due mainly to $r_{BB'}$) is about two orders of magnitude lower than in the surface structure, allowing to reach the body breakdown voltage. In this region of avalanche characteristics the breakdown can be considered as pure body breakdown.

Though the dry germanium oxide is electronegative with respect to the n-type body [14] and should exhibit as a result of the decreased field strength near the surface theoretically a surface breakdown voltage higher than the body breakdown voltage [4], in practice an opposite behaviour is observed. This fact may be regarded partly as a consequence of a very small quantity of water adsorbed so strongly to the porous oxide film that below 130 °C no heat treatment performed for some hours before encapsulation [15] can dehydrate perfectly the surface of the alloyed junction. This small amount of strongly adsorbed water forming a dipole layer alters the electrical properties of the oxide structure, which attains a positive charge as a net result, in spite of the electronegative behaviour of the oxide. On the other hand C. G. B. Garrett and W. H. Brattain [4] have pointed out that this behaviour of the "dry" oxidized surface may be caused by microscopic irregularity spots

initiating avalanches at a rather low voltage. In some way or other the surface breakdown voltage is always somewhat lower than the body breakdown voltage irrespective of the sign of surface charges.

The situation discussed above may be modified in case of excess surface charges, ionic or channel conductance, due to considerable surface contamination [4, 14], e.g. owing to an ambient of high humidity (e. g. as a result of leaky encapsulation).

Much water generates a (with respect to the n-type body) positive surface potential and seriously reduces BV_{CBO} at low multiplication factors, owing to the accumulation layer. Apart from this, excess reverse current can flow due to elevated ion migration. On the other hand, a large density of the adsorbed negative (e.g. oxygen or chlorine) ions on the oxidized n-type material may give rise to an inversion layer or p-type channel beneath the surface [4, 14]. Both channel and excess ionic reverse current lead to puzzling phenomena in the low current range of the reverse characteristics: the rise in reverse current seems to stem from carrier multiplication. This difficulty in measuring the surface breakdown at low multiplication ($M < 10^2$), may be overcome by considering the power n in the multiplication function (7). If n varies strongly with current, either channels occur or an excess ionic shunt current flows.

High ambient humidity resulting in elevated surface charge density, by raising the apparent differential resistance of the breakdown voltage-current characteristics in the range $10^2 < M < 10^4$ via increased surface breakdown, alters the breakdown voltage but insignificantly, when multiplication is sufficiently high ($M > 3 \times 10^4$) and humidity remains considerably below saturation. This statement is in agreement with A. P. Shotov [2], who investigated high resistivity junctions in different ambients at very high multiplication ($M \simeq 10^5$). Apart from this, measuring at different humidity levels high voltage *pnp* transistor junctions subjected to various electrolytic and chemical etching-oxidizing treatments, it could be undoubtfully proved that the breakdown voltage determined above an avalanche current density of $40 \div 50$ mA/mm^2 or at multiplication factors over $M = 10^4$ really constitutes the body breakdown voltage, if relative humidity does not exceed 60 per cent, even at base resistivities as high as $10 \div 12$ ohmcm [13] (cf. Appendix 2).

Although surface breakdown has little influence on high current density breakdown, it is worth noting that at very high relative humidity levels approaching or reaching the dew point, junctions without encapsulation show an abrupt and large decrease in breakdown voltage. This phenomenon is due in all probability to a thick water film on the (cooler) surface, providing a shunt for the pulsed measuring current. Sudden variations in ambient temperature must, therefore, be avoided during test.

Ad b) and *c)* Both the still remaining influence of the surface breakdown

and the body voltage drops due to lead-in resistances and pinch-in effect can be avoided if the avalanche current is plotted against avalanche voltage at adequately high current values, say in the range 20 ÷ 200 mA in case of medium power types and at 20 ÷ 500 mA in testing high power ones.

Fig. 7 shows the breakdown characteristics of 3 transistor groups of typical behaviour. Curves *a* hold for transistors of regular demeanor, which exhibit an approximately constant differential resistance $r_d = dBV_{CBO}/dMI_{CBO}$

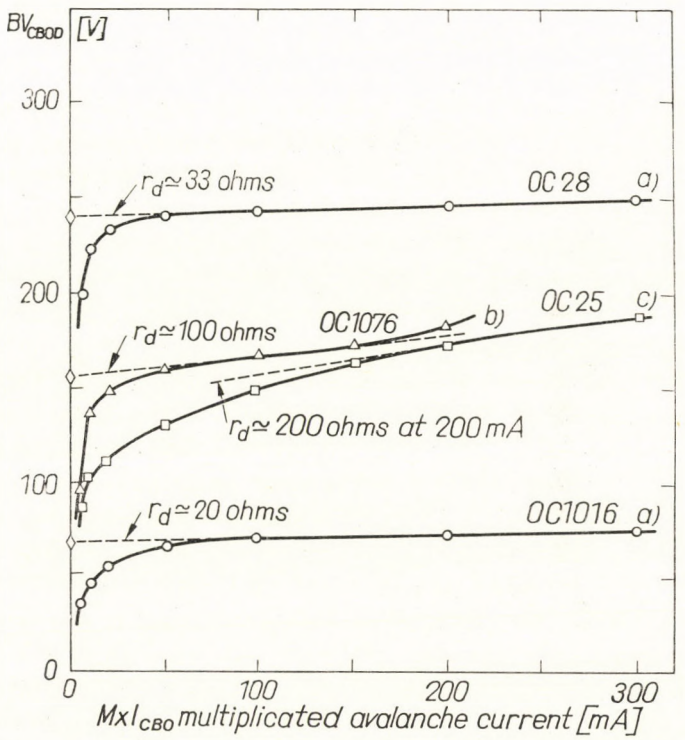


Fig. 7. Three typical avalanche breakdown voltage-avalanche current characteristics: Normal shape with low differential resistance (*a*); inflexion type due to pinch-in effect (*b*); high differential resistance due to enhanced humidity (*c*)

above a high enough current (or multiplication factor). As can be seen from Fig. 7, the differential resistance is in the order of the base spreading resistance $r_{BB'}$ ($r_d \cong 23$ ohms for the low voltage and 33 ohms for the high voltage power types). Curve *b* furnishes an example for "pinch-in" of the carrier flow pattern in the base region of increased field strength occuring suddenly at a high current, marked by an inflexion on the characteristics, due to the gradual increase in effective base spreading resistance. Curve *c* presents a typical "bad" surface, generated by removing the cap and the molecular sieve

in a power transistor and enabling room humidity to penetrate the surface oxide film. Differential resistance soared to 200 ohms (about one order of magnitude above the value at dry ambient) even at 200 mA measuring current. All curves feature a differential resistance of about 3000 ohms at $MI_C = 20$ mA, corresponding to a multiplication factor of appr. 500, characteristic for pure surface breakdown.

Body voltage drops can simply be eliminated in curves like *a* and *b*. In curve *a* the extension of the straight part gives the body breakdown voltage along the junction void of all lead-in voltage drops $r_d MI_{CBO}$, while in curve *b* the tangent drawn through the inflexion point locates it at zero current. In curve *c* the exact value of the body breakdown voltage cannot be determined because of the high and ever-decreasing differential resistance shown at the highest measured current. The tangent to this point gives, however, a fairly good approximation of the true body breakdown voltage.

The determination of the accurate body breakdown voltage, essential for the further evaluation of the measurements, has been performed with the technique described above. To avoid all possible disturbances (due to surface and body effects other than described) no samples with a high current differential resistance (exceeding the characteristic value of $r_{BB'}$ by a factor higher than 3), have been considered. Finally, if the measured punch-through voltage was falling short of the measured collector-emitter breakdown voltage BV_{CEO} plus 10 V, the transistor was excluded, since the low punch-through could eventually influence BV_{CEO}.

The collector-emitter breakdown voltage BV_{CEO} was determined in the same way as the collector-base breakdown voltage BV_{CBOD}. Though multiplication ($M'I_{CEO}$) is much smaller here than in the measurement of BV_{CBOD} (by a factor of about $1 + \beta_0$), the curves did not betray any appreciable disturbance at the high measuring currents used for the statement of BV_{CBOD}. The differential resistance is, of course, much smaller due to the lack of $r_{BB'}$.

VI. Estimates of maximum degradation due to body causes, on grounds of measurements

If equation (6) is applied for the theoretical planar homogeneous case and the inhomogeneous system referred to with subscript D, we have

$$BV_{CBO} = BV_{CEO}\left(\frac{BV_{CBOD}}{BV_{CEO}}\right)^{n_D/n} = \\ = BV_{CEO} \cdot \operatorname{num\,log}\left(\frac{n_D}{n} \cdot \log \frac{BV_{CBOD}}{BV_{CEO}}\right), \tag{8}$$

or solving equation (6) for n_D

$$n_D = \frac{\log (I_{CEO}/I_{CBO})}{\log (BV_{CBOD}/BV_{CEO})} . \tag{8a}$$

With the equations (8) and (8a) the "non-degraded" collector-base body breakdown voltage BV_{CBO} of the planar-homogeneous section can be determined by successive approximation and with the aid of relationship $n(BV_{CBO})$, depicted in Fig. 4, as follows:

On grounds of the measured values I_{CEO}, I_{CBO}, BV_{CBOD} and BV_{CEO} equation (8a) yields n_D, while (under the arbitrary substitution $BV_{CBO} = BV_{CBOD}$) we obtain n_F from Fig. 4. Insertion of n_D and $n = n_F$ into equation (8) results in a collector-base breakdown voltage BV'_{CBO}, higher than BV_{CBOD}. Substituting $BV_{CBOD} = BV'_{CBO}$ into equation (8a) we obtain another factor n'_D, smaller than n_D. With $BV_{CBO} = BV'_{CBO}$ Fig. 4 furnishes n'_F, so that we have a value-triplet consisting of BV'_{CBO}, n'_D and n'_F to characterize a transistor with slightly degraded breakdown voltage. Repeating this cycle n''_D (or n'''_D etc.) will approach more and more n''_F (or n'''_F etc.) until to the equality $n'''_D = n'''_F = n$ holds with the desired accuracy. (The number of primes designates the approximation steps.)

The collector-base breakdown voltage BV''_{CBO} (or BV'''_{CBO} etc.) obtained in the last step of the approximation process will be nearly equal to the breakdown voltage BV_{CBO} of the planar-homogeneous transistor section. The number of necessary approximation steps depends on the degree of degradation; 2 or 3 steps are usually enough (cf. Appendix 3). In possession of BV_{CBO} Equ. (1) yields the corresponding mean resistivity ϱ for the planar homogeneous section.

The described approximation method is made possible by the fact that the function $n(BV_{CBO})$ decreases only sligtly with growing BV_{CBO} in the whole practical range of BV_{CBO} (or ϱ) for alloyed transistors, because the maximum value of the curve lies in the neighbourhood of 0,4 ohmcm [1] as it is to be seen in Fig. 4. The quickest convergency near the peak of the curve $n(BV_{CBO})$ should be attributed to the same fact.

The function $BV_{CBOD}(\varrho)$ determined by the described method is shown in Fig. 8 for the measured transistors having maximum degradation (i.e. highest λ), so that the line indicating the supposedly largest degradation can be drawn. The considerable deviation from the curve representing equation (1) can readily be observed above 0,5 ohmcm, especially at high resistivities. The results corresponding to the maximum degradation are summarized in Table II.

Transistors of the smallest inhomogeneity, falling in Fig. 5 on or close to the theoretical curve $n(BV_{CBO})$, show a good agreement with equation (1) at $\varrho < 8$ ohmcm (see Fig. 9), while above this resistivity value it was impossible

Table II

N_I [atoms/cm^3]	$3{,}4 \cdot 10^{13}$	10^{14}	$2 \cdot 10^{14}$	$5 \cdot 10^{14}$	10^{15}	$2 \cdot 10^{15}$	$5 \cdot 10^{15}$
ϱ [ohmcm]	46	18	9	3,6	1,8	0,9	0,36
BV_{CBOD} min [V]	180	128	107	85	69	57	41
$\frac{BV_{CBOD}\ \text{min}}{BV_{CBO}}$ (approx.)	0,15	0,215	0,29	0,415	0,55	0,73	1,0
r_{op} (approx.), microns	46	21	15	11	9	7,5	5,8

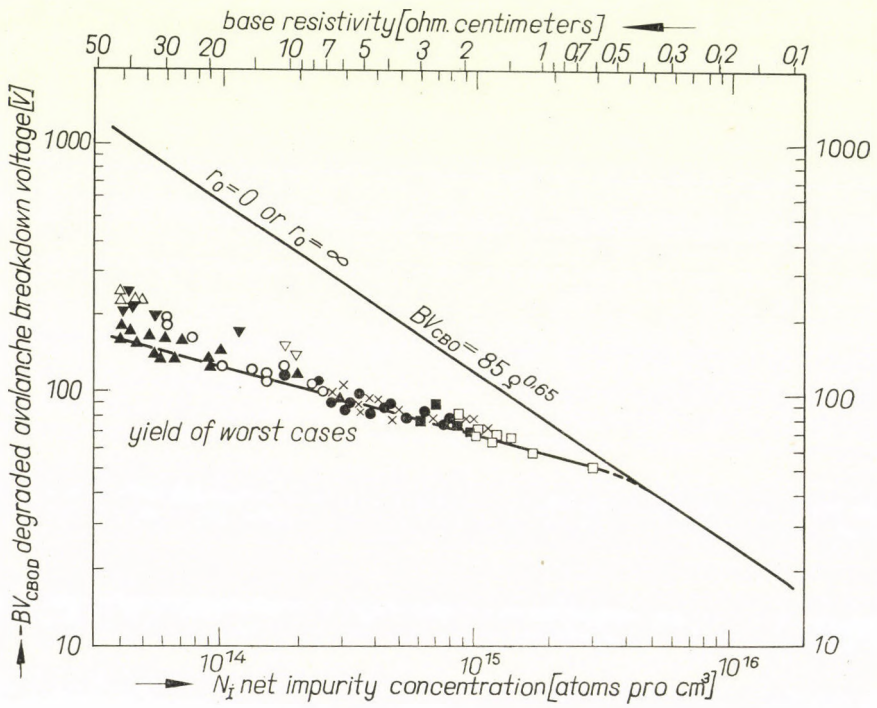


Fig. 8. Estimates of maximum degradation versus ϱ (or N_I) plotted on basis of cases exhibiting most severe degradation:

○ *AD 1202, AD 1203*; ● *OC 76, OC 77*; ▽ *AD 105*; ▼ *ADZ 12*; △ *2SB150*; ▲ *ASZ 15, ASZ 18, OC 28, OC 35*; □ *2N218, 2N219, OC 44, OC 45*; ■ *2N301, 2N301 A*; × *OC 1016*

to find any transistor with a breakdown voltage higher than $BV_{CBOD} = 360$ V. The possible cause of this phenomenon is the cylindrical curvature at the edge of the collector-base junction, which can hardly be avoided even in ideal alloying processes.

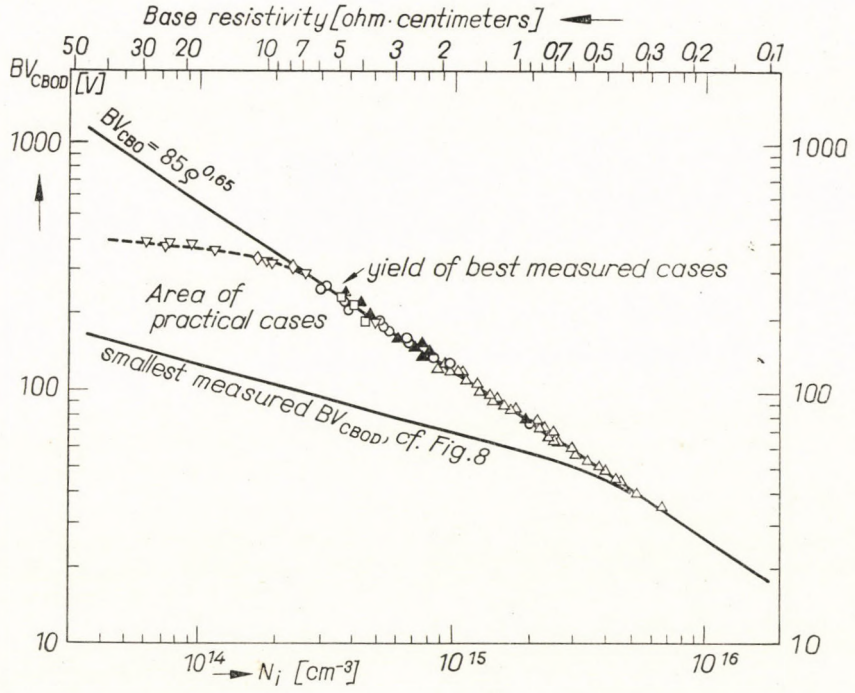


Fig. 9. Highest measured breakdown voltages versus ϱ (or N_I); below 8 ohmcm the results are in good agreement with equation (1)

VII. Critical size of irregularity

In the previous chapters the maximum possible degradation of BV_{CBO} was estimated, caused undoubtedly by body imperfections in or near the junction, as shown in Fig. 8. The estimate is based on measurements performed on more than 900 alloyed *pnp* germanium transistors, so that transistors featuring the "most dangerous" imperfection geometry and, as a consequence, maximum degradation, occur supposedly with satisfactory statistical confidence.

The next problem to be attacked is the determination of the normalized degradation BV_{CBOD}/BV_{CBO} as a function of the irregularity radius r_0 or of its normalized value r_0/X_B, when the imperfection is embedded in the planar-homogeneous surrounding. This function obviously conforms with that of H. L. ARMSTRONG et al. [7, 9] and equation (3), respectively, when the size of irregularity is not too small compared to the depletion layer width X_B and at high base resistivity the resolving effect of the mean free path for carrier-phonon collisions (Π_f) has no significant influence. The condition $X_B \ggg \Pi_f$ holds primarily if the lighter doped junction side exhibits intrinsic resistivity

(appr. 46 ohmcm at room temperature). In such a theoretical $i - p^+$ junction X_B equals 230 microns, which, in fact, considerably exceeds $\Pi_f = 140$ Å.

Extrapolation of the curve denoting the maximum possible degradation to 46 ohmcm yields $BV_{CBOD}/BV_{CBO} = 0{,}15$ in Fig. 8. The abscissa marking the minimum ordinate or "valley" of the normalized degradation function can be estimated by the assumption that (i) the function has no sharp edge and (ii) irregularities with $2\,r_0/X_B \leq 10^{-2}$ do not cause any degradation of BV_{CBO}, so that BV_{CBOD}/BV_{CBO} approaches unity again as shown in Fig. 10. On grounds of the above assumptions one can construct a function shape

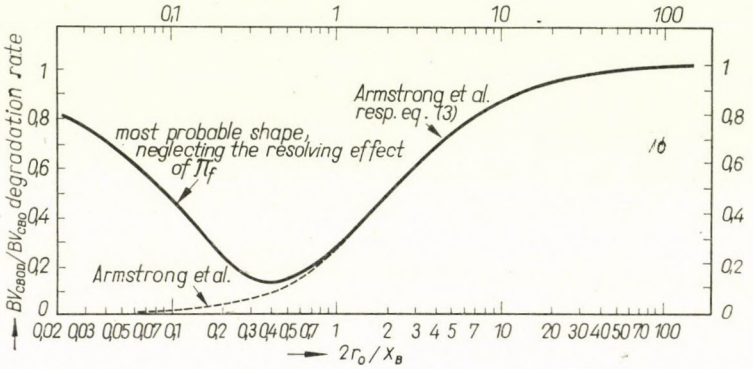


Fig. 10. An approximation of the normalized degradation versus r_0/X_B in case of embedded irregularity and at very high resistivity (solid line; the dashed line represents the results of H. L. ARMSTRONG et al. [7])

having a minimum of about $r_0/X_B = 0{,}2$, in all probability symmetrical with respect to the valley (cf. Fig. 10; solid line). Compared to the curve of H. L. ARMSTRONG a deviation can be detected only beginning within the immediate vicinity of the minimum locus if the valley is approached from high r_0/X_B values. The fairly small deviation in this region of the curves suggests a relatively high accuracy in the determination of the minimum abscissa.

The curve in Fig. 10 thus presents the correction of H. L. ARMSTRONG's degradation function, see equation (3), for cases when the junction curvature is surrounded by a planar junction. It can be seen that the results of H. L. ARMSTRONG (dashed line) hold also at

$$r_0/X_B \geq 0{,}4\,, \tag{4}$$

naturally as long as the base resistivity remains very high ($\varrho > 20$ ohmcm) and criterion (5) is fulfilled.

The degradation function obtained does not consider the resolving effect of Π_f. The shape of the degradation function is equally important in the range

of moderate and low resistivity values, when Π_f exerts a growing influence. With decreasing resistivity (i.e. increasing impurity concentration N_I) the depletion layer shrinks according to the expression

$$X_B = \left(\frac{2\,\varepsilon}{qN_I}\, BV_{CBO}\right)^{1/2}, \tag{3a}$$

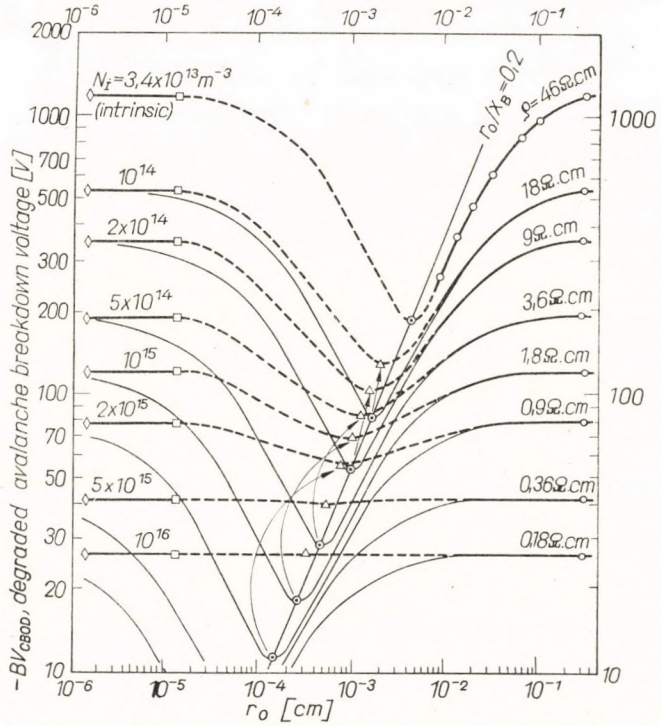


Fig. 11. Estimate of the function BV_{CBOD} (r_0) with N_I as parameter

deduced for the planar-homogeneous case. Thus, with rising N_I the depletion layer width X_B becomes more and more comparable with Π_f and consequently the resolving effect will increasingly counteract the degradation due to tiny imperfections. Since according to S. L. MILLER [1] a potential gradient acting on a path shorter than about 10 Π_f has no significant influence on avalanche carrier multiplication, one can assume that imperfections of $r_0 \leq 10\,\Pi_f$ cause no degradation in BV_{CBO} despite of the elevated field strength. On the other hand, the estimated curve in Fig. 8, representing the maximum possible degradation plotted against the impurity concentration, yields the minimum ordinates of the functions $BV_{CBOD}(r_0)$, with N_I as parameter.

The above premise makes it possible to plot the degraded breakdown voltage versus irregularity radius r_0 as shown in Fig. 11. At high r_0 the curves closely follow the degradation function of Fig. 10 (——⊙——), and at $r_0 \leq 10\,\Pi_f$ no degradation occurs (◇————◇). As to the region between the above extremities only the minimum ordinate belonging to the different N_I values (— — — △ — — —) is known. Due to the resolving effect the curve shapes will be more and more shallow near the valleys with dropping resistivity and the minimum abscissa shifts towards higher ratios r_{0p}/X_B in comparison with the minimum locus of the $i - p^+$ junction (indicated with arrows). The results of the above estimating method are summed up in Table II, where r_{0p} represents the *most dangerous radius of curvature,* at which the degradation of BV_{CBO} supposedly peaks.

It should be emphasized that the valley of the normalized degradation function, the minimum locus $r_0/X_B = 0{,}2$ as well as the results of Table II are but estimates on grounds of measurements and they cannot be regarded as infallible investigations supported by exact analysis. Besides a qualitative description of the phenomenon involved only an approximate quantitative accuracy can be expected in this method.

VIII. Inhomogeneity in resistivity

The possible maximum degradation owing to the inhomogeneities of the resistivity in the depletion layer are given in Table II, too, where in this case r_{0p} stands for the radius of inhomogeneity. The degradation at a given r_0 depends, irrespective of the impurity type, on the ratios ϱ_D/ϱ (ϱ_D denoting here the resistivity of the irregularity spot). There is no degradation when $\varrho_D \geqslant \varrho$. At $\varrho_D < \varrho$ the degradation gradually approaches with decreasing ϱ_D the values given in Table II.

The presence of resistivity inhomogeneities in high-resistivity single crystals gives a plausible explanation of the described degradation, especially in high-voltage power diodes or transistors with extended junction area, cf. the investigations of H. UEDA [16] and G. MELTZER [17], devoted to microstructural resistivity inhomogeneities in germanium single crystals. According to the above papers, the resistivity variation with periodicities of $50 \div 1000$ microns grows with dropping impurity concentration, in particular in the direction of crystal pulling. The fluctuation of ϱ is especially big in *n*-type germanium (i.e. in *pnp* transistors) produced by the zone-levelling techniques, generally used in the fabrication of alloyed germanium devices. The resistivity fluctuation within an *n*-type germanium single crystal of $8 \div 10$ ohmcm manufactured with the zone-levelling method can reach or even exceed 1 : 3 [17], and it may be considerable even in the direction 111 at the outer periphery

of the single crystal ingot. Therefore, it is advisable to discard the peripheral parts of the 111 germanium slice in manufacturing high-voltage devices. — Thus in most high-voltage power transistors of considerable inhomogeneity the resistivity of the planar-homogeneous base section, calculated from equations (8) and (8a), approaches the intrinsic value, as shown in Fig. 8. This fact stands to reason since the level of base impurity is determined primarily by n-type regions of lower resistivity in the almost intrinsic base wafer. Obviously a melt much less pure can be chosen if "homogeneous" single crystals of approximately 0,5 ohmcm are to be drawn, while at resistivities of $5 \div 10$ ohmcm the purity of the material and skill in the crystal drawing technology become the most critical. Nowadays, a high voltage alloyed germanium transistor below the average quality features $p^+ - n - p^+$ islands embedded in a nearly $p^+ - i - p^+$ substrate.

Although the method given in this article cannot distinguish a geometrical curvature from resistivity inhomogeneity as a cause of degradation, the above train of ideas suggests that at high ϱ (i.e. at high-voltage devices) the resistivity inhomogeneity dominates, while at lower ϱ the geometrical irregularities of the junction take the upper hand, because the average size of the alloy faults approaches the depletion layer width of medium-voltage products. Apart from these, the production of high-voltage germanium devices is primarily encumbered by the rim of the junction: a cylindrical one limits the breakdown voltage at about 400 V, while sharply edged junctions due to indium buttons containing aluminium show a better behaviour. In two identical high-voltage transistor groups, produced with the same technology except for the collector button material and checked by ground and etched cuts, the sharply edged junctions resulted in a 6 per cent rise of the median value of BV_{CBOD} (individual breakdown voltages differed as much as 18 per cent). In accordance with former statements the sharp edge with a very small radius of curvature does not decrease the avalanche breakdown voltage.

IX. Conclusions

The theoretical avalanche breakdown voltage in agreement with the function $BV_{CBO} = 85\ \varrho^{0,65}$ cannot be usually achieved due to the increased field strength caused by surface charges as well as by geometrical irregularities of the junction and the inhomogeneous resistivity within the depletion layer. With increasing ϱ this effect becomes more and more significant. Applying the described pulsed method of breakdown measurement it was possible to exclude surface effects and to recognize the structural irregularities in the collector-base junction of germanium *pnp* transistors and to calculate the base resistivity corresponding to the planar section enclosing the irregularity. Examining a

great number of alloyed *pnp* transistors with a wide range of base resistivity the maximum degradation of BV_{CBO} could be estimated as a function of the resistivity of the lower doped side. A two-dimensional exact analysis based on the Poisson equation appears to be very difficult, if possible at all, due to the complicated boundary conditions and the resolving effect of the mean free path for carrier-phonon interactions, which counteracts the degradation stemming from tiny imperfections.

The maximum degradation of BV_{CBOD} seems to be reliably determinable by means of combined measurement and estimation at least above resistivities of $1 \div 2$ ohmcm, since it is very improbable that in a lot of more than 900 pieces no transistors with critical r_{0p} would occur (see Fig. 8). Below 1 ohmcm the minimum value of BV_{CBOD} (Table II) could not be properly investigated owing to the lack of power transistors having an extended alloyed area in this resistivity range. It is possible that in case of a large junction surface where due to alloying faults irregularities exist with greater probability, the discussed effect may give rise to a significant degradation of BV_{CBO} even at around 0,2 ohmcm, but this resistivity range has no importance in practice or in the qualitative analysis of the irregularity effect.

It should be pointed out that the practice, generally adopted in manufacturing high-voltage germanium transistors with base wafers having a resistivity of $6 \div 8$ ohmcm, is not advisable since the inherent increase in $r_{BB'}$ spoils the high frequency and switching properties of the transistor. It seems to be much more useful to insure the required high operating voltage with base wafers of relatively low resistivity and with homogeneous as well as planar junctions. An improved technology (e.g. careful crystal drawing with the Czochralsky-method and "slow" alloying process) will easily supply germanium transistors with an operating voltage of 200 V or higher.

The above investigations confirm the results obtained by S. L. MILLER in the resistivity range of $0{,}3 \div 2$ ohmcm. In addition, it was possible to plot the function $n(\varrho)$ up to resistivities of $10 \div 20$ ohmcm. It must be emphasized, however, that the phenomenon discussed in this article has nothing to do with the secondary breakdown described by C. G. THORNTON and C. D. SIMMONS [18], H. A. SCHAFFT and J. C. F. FRENCH [19] and F. WEITZSCH [20], where collector-emitter breakdown occurs — mostly at reverse base current — possibly as a result of thermal generation of carriers. A measurement of BV_{CEO} incorporating secondary breakdown would yield an *unusually low* value of n_D, contrary to the empirical $n_D > n$.

The object of this article can easily be adopted to extend itself to diodes as well as pn^+-type and np^+-type alloyed or diffused junctions.

ACKNOWLEDGEMENT

The author is indebted to Professor I. P. VALKÓ, leader of our laboratory, for the patronage of this work and to M. KOCSIS for reading and commenting on the manuscript.

REFERENCES

1. MILLER, S. L.: Avalanche Breakdown in Germanium. *Phys. Rev.* **99** (1955), 1234.
2. SHOTOV, A. P.: *J. Techn. Phys.* (USSR) **26** (1956), 1634.
3. TOKUYAMA, T.: Zener Breakdown in Alloyed Germanium $p - n$ Junctions. *Solid-State Electronics* **5** (1962), 161.
4. GARRETT, C. G. B.—BRATTAIN, W. H.: Some Experiments on, and a Theory of, Surface Breakdown. *J. Appl. Phys.* **27** (1956), 292.
5. VUL, B. M.—SHOTOV, A. P.: *J. Techn. Phys.* (USSR) **27** (1957), 2189.
6. DAVIES, R. L.—GENTRY, F. E.: Control of Electric Field at the Surface of $p - n$ Junctions. *IEEE Trans. on Electron Devices ED* **11** (1964), 313.
7. ARMSTRONG, H. L.—METZ, E. D.—WEIMANN, J.: Design Theory and Experiments for Abrupt Hemishperical $p - n$ Junction Diodes. *IRE Trans. on Electron Devices ED* **3** (1956), 86.
8. ARMSTRONG, H. L.: A Theory of Voltage Breakdown of Cylindrical $p - n$ Junctions, with Applications. *IRE Trans. on Electron Devices ED***4** (1957), 15.
9. SHIELDS, J.: Avalanche Breakdown Voltage in Hemispherical ($p - n$) Junctions. *J. of Electronics and Control* **IV** (1958), 58.
10. SHOCKLEY, W.: The Theory of $p - n$ Junctions in Semiconductors and $p - n$ Junction Transistors. *Bell Syst. Techn. J.* **28** (1949), 435.
11. GARRETT, C. G. B.—BRATTAIN, W. H.: Physical Theory of Semiconductor Surfaces. *Phys. Rev.* **99** (1955), 376.
12. SCHMELTZER, R. A.: Maximum Stable Collector Voltage for Junction Transistors. *Proc. IRE* **48** (1960), 332.
13. KEMÉNY, A.: Untersuchung und Messung des Lawinendurchbruchs schroffer $p - n$ Übergänge unter Berücksichtigung von Oberflächeneffekten. *Tungsram Technische Mitteilungen* (to be published).
14. KINGSTON, R. H.: Review of Germanium Surface Phenomena. *J. Appl. Phys.* **27** (1956), 101.
15. FURUSHO, K.—ONO, K.: Effects of Vacuum Drying on the Surface Properties of Germanium Transistors. *Japan J. Appl. Phys.* **1** (1962); No. 3, 148.
16. UEDA, H.: Resistivity Striations in Germanium Single Crystals. *J. Phys. Soc. Japan* **16** (1961), 61.
17. MELTZER, G.: Minute Resistivity Variations in Germanium Crystals and their Effect on Devices. *J. Elchem. Soc.* **109** (1962), 947.
18. THORNTON, C. G.—SIMMONS, C. D.: A New High Current Mode of Transistor Operation. *IRE Trans. on Electron Devices ED* **5** (1958), 6.
19. SCHAFFT, H. A.—FRENCH, J. C.: Second Breakdown in Transistors. *IRE Trans. on Electron Devices ED* **9** (1962), 129.
20. WEITZSCH, F.: Zum Einschnüreffekt bei Transistoren, die im Durchbruchsgebiet betrieben werden. *AEÜ* **16** (1962), 1.

APPENDIX 1

Approximative Determination of the Differential Resistance at Surface Breakdown

The avalanche current flow pattern for moderately low multiplication with dominating surface breakdown is shown schematically in Figs 12a and 12b. It can be seen that the current density, concentrated under the surface, decreases gradually towards the base bulk material, while it peaks at the edge of the depletion layer and drops considerably towards the ohmic base contact. Naturally, the high differential resistance stems primarily from the small volume section in the very vicinity of the surface and the boundary of the depletion layer where current density culminates.

In the following the three-dimensional problem will be approximated with a two-dimensional treatment.

The assumption that no breakdown current flows towards the junction in the bulk structure holds at relatively low surface breakdown voltage, small carrier multiplication and avalanche current considerably higher than the low-voltage reverse currents (i.e. $10^2 < M < 10^3$).

The body section marked in Fig. 12a by $X_B \times L_D \times 2\,\pi\, r_1$ may be left unconsidered since it represents a depletion layer (X_B stands for the width of the depletion layer, $2\,\pi\, r_1$ denotes the perimeter of the collector junction). In addition, one can neglect the lead-in resistance of the p^+-zone.

The section which essentially determines the diferential resistance has a nearly triangle-shaped cross section in Figs 12a and 12b, where the current flows in the high resistivity base material. Let us assume as a rough approximation that the current density does not vary but with the distance from the barrier of the depletion layer, and in the direction perpendicular to the base surface it remains constant up to $y(r)$, sinking hereafter to zero. Thus, a semi-conical ring will be obtained having an inner radius r_1, composed of the radius of the collector junction and of X_B (the latter can be neglected); the outer radius r_2 of the ring presents the mean value of the outside radius of the base wafer and the inside radius of the base contact ring.

As to the inner width of the above conical ring, it can be seen from Fig. 4 that the current density decreases considerably even along the Debye-length L_D. The field strength sinks exponentially towards the bulk and at a depth of one Debye-length it is $1/e = 0{,}368$ down on its surface value. While the current density may be regarded as proportional to the field strength, and such an exponentially decreasing current density pattern may be substituted by a homogeneous one with constant current density, ranging to a depth of one Debye-length, the inner width of the conical ring will be assumed to be L_D.

The three-dimensional problem, in this way reduced to a two-dimensional one, can easily be solved by simple integration in the polar coordinate system according to Fig. 12b.

Neglecting variations in resistivity the resistance of the conical ring amounts to

$$dR = \frac{\varrho}{2\,\pi r \cdot y(r)}\, dr\,,$$

where r is the distance of the ring from the origo, dr thickness of the ring and

$$y(r) = L_D + \frac{B - L_D}{r_2 - r_1} \cdot (r - r_1)$$

where B denotes the width of the base wafer.

Introducing

$$S = \frac{B - L_D}{r_2 - r_1}$$

the resistance of the ring becomes

$$dR = \frac{\varrho}{2\,\pi r \cdot [L_D + S\,(r - r_1)]}\, dr\,,$$

which can be transformed into

$$dR = \frac{\varrho}{2\,\pi S} \cdot \frac{dr}{r\left(r - r_1 + \frac{L_D}{S}\right)}\,.$$

The substitution

$$\frac{L_D}{S} - r_1 = C$$

leads to

$$dR = \frac{\varrho}{2\,\pi SC}\left(\frac{1}{r} - \frac{1}{r + C}\right) = \frac{\varrho}{2\,\pi\,(L_D - Sr_1)}\left(\frac{1}{r} - \frac{1}{r + C}\right),$$

so that the resistance of the ring can be expressed in the form

$$R = \frac{\varrho}{2\pi(L_D - Sr_1)} \cdot \int_{r_1}^{r_2} \left(\frac{1}{r} - \frac{1}{r+C}\right) dr =$$

$$= \frac{\varrho}{2\pi(L_D - Sr_1)} \left| \ln \frac{r}{r+C} \right|_{r_1}^{r_2}.$$

The above function integrates immediately to give the resistance between the inner ring surface $2\pi r_1 L_D$ and the outer ring surface $2\pi r_2 B$:

$$R = \frac{\varrho}{2\pi(L_D - Sr_1)} \cdot \ln\left(\frac{r_2}{r_1} \cdot \frac{r_1 + C}{r_2 + C}\right).$$

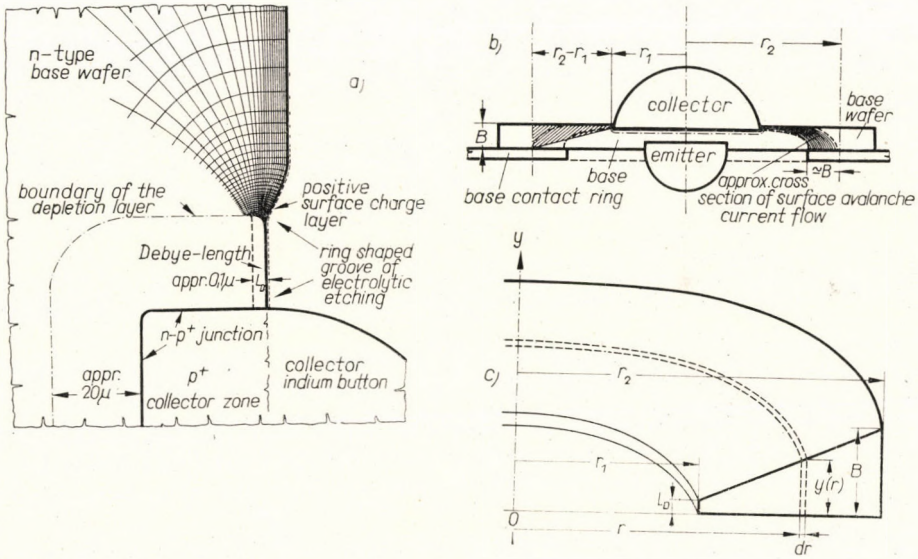


Fig. 12. Surface avalanche current flow pattern of a conventional *pnp* alloyed transistor
In the vicinity of the junction periphery, the scale of the Debye-length L_D is magnified tenfold (*a*); in the whole base cross-section (*b*); in the section of current-flow, considered for the numerical estimation of the differential resistance (*c*)

Combining this result with the expressions for S and C yields

$$R = \frac{\varrho}{2\pi L_D \left[1 + \frac{r_1}{r_2 - r_1}\left(1 - \frac{B}{L_D}\right)\right]} \cdot \ln\left(\frac{r_2}{r_1} \cdot \frac{L_D}{B}\right). \tag{9}$$

The above equation can be considerably simplified, as

$$L_D \simeq 10^{-6} \div 10^{-5}\ \text{cm},\ B \simeq 10^{-2}\ \text{cm}\ \text{and}\ r_1/(r_2 - r_1) \simeq 1,$$

so that one may neglect both unities in the square brackets. Making use of the interrelation $-\ln x = \ln 1/x$, formula (9) is found to be

$$R = \frac{\varrho\,(r_2 - r_1)}{2\pi r_1 \cdot B} \cdot \ln\left(\frac{r_1}{r_2} \cdot \frac{B}{L_D}\right). \tag{9a}$$

In power transistors, where the inner rim of the base contact lies extremely near to the outside edge of the collector junction, i.e. $r_2 - r_1 \ll r_1$, the quantity r_1/r_2 equals approximately 1 and can remain unconsidered. The resistance of the conical ring is given in this case by

$$R = \frac{\varrho\,(r_2 - r_1)}{2\,\pi\,r_1 \cdot B} \cdot \ln \frac{B}{L_D} \,. \tag{9b}$$

On grounds of the approximations (9a) and (9b) let us first of all compute the Debye-engths, determined by W. SHOCKLEY [10] as well as by C. G. B. GARRETT and W. H. BRATTAIN [9, 11] in analogy with the Debye—Hückel theorem for strong electrolytes:

$$L_D = \left(\frac{\varepsilon_r\,\varepsilon_0\,kT}{2\,\pi q^2 \cdot N_I}\right)^{1/2} = 1{,}94 \times 10^2\, N_I^{-1/2} \ \text{[cm]}, \tag{10}$$

where $\varepsilon_r = 16{,}5$ denotes the dielectric constant of germanium, $\varepsilon_0 = 8{,}86 \times 10^{-14}$ A/Vcm the dielectric permittivity of vacuum, $kT/q = 26$ mV, $q = 1{,}6 \times 10^{-19}$ As the electron charge, and

$$N_I = |N_D - N_A| = 1/\varrho\mu q$$

s the net impurity concentration in the base, in cm^{-3}.

We shall now consider two specific examples:

a) The investigated power transistor in Appendix 2, made of an *n*-type base material with $\varrho = 10$ ohmcm, thus

$$N_I = 1/\varrho\mu q \simeq 1{,}75 \times 10^{14} \ \text{cm}^{-3}.$$

Using equation (10) the Debye-length amounts to

$$L_D = 1{,}94 \times 10^2\,(1{,}75 \times 10^{14})^{-1/2}\,1/2 =$$
$$= 1{,}94 \times 10^2\,(1{,}33 \times 10^7) = 1{,}45 \times 10^{-5}\ \text{cm}.$$

If $r_1 = 0{,}13$ cm, $r_2 = 0{,}23$ cm and $B = 2{,}5 \times 10^{-2}$ cm; equation (9a) yields $R = 350$ ohms.

b) In a low power alloyed *pnp* germanium transistor manufactured of an *n*-type base material with $\varrho = 10$ ohmcm with $r_1 = 1{,}6 \times 10^{-2}$ cm, $r_2 = 0{,}12$ cm and $B = 1{,}6 \times 10^{-2}$ cm we have $R = 3140$ ohms.

The difference in the resistances amounting to one order of magnitude truly represents the divergence in the geometry of low power and high power types, though the resistances fall by half to one an order of magnitude short of the expected and measured values. This deviation stems primarily from the assumed linearity of the function $y(r)$, a rough approximation contradictory to Fig. 4, where a considerably higher resistance of the body next the junction edge is suggested by the strongly curved contour of the current flow.

As another rough approximation let us assume that the surface avalanche current does not flow trough a conical but a cylindrical ring having a constant thickness of L_D. Thus, from equation (9), with $B = L_D$ the resistance of this volume section amounts to

$$R = \frac{\varrho}{2\,\pi\,L_D} \ln \frac{r_2}{r_1} \,. \tag{11}$$

Using (11), the differential resistance equals about 52 000 ohms in case *a)* and 182 000 ohms in case *b)*. These values exceed the expected ones by one to one and half orders of magnitude showing that the geometric mean value of the solutions of equations (9) and (11) presents a good approximation of the measured differential resistance, reminding one that the real current flow pattern has a strongly curved contour in Fig. 12a. Omitting π from the denominator or perhaps introducing it into the numerator of equations (9), (9a) or (9b), one obtains fairly good results compared to the measured ones in most cases of practical alloyed transistors.

Nevertheless, the above approximative computation proved suitable to show the prospective surface breakdown, to clarify the dominant role of the body resistance and contributed to the correct interpretation of the experimental data.

APPENDIX 2

Breakdown Measurements at Various Surface Treatments and Ambient Humidity

Summing up the results of another paper [13] it can easily be proved that humidity (at least up to 60 per cent relative value) and the various etching and oxidizing treatments do not exert any significant influence on the body breakdown voltage measured at a sufficiently high multiplication ($M \geq 10^4$), even if the resistivity of the lower doped side reaches 10 ohmcm.

The experiments were performed on 48 *pnp* alloyed germanium power transistor junctions with a nominal base resistivity 6 ÷ 12 ohmcm. The dimensions of the junction coincide with those in example *a)* of Appendix 1. The *n*-type single crystals doped with antimony were produced in a zone-leveling process; the resistivity of the wafers was determined by the four-probe method. After slicing and polishing, the pellets were alloyed with pure indium. After

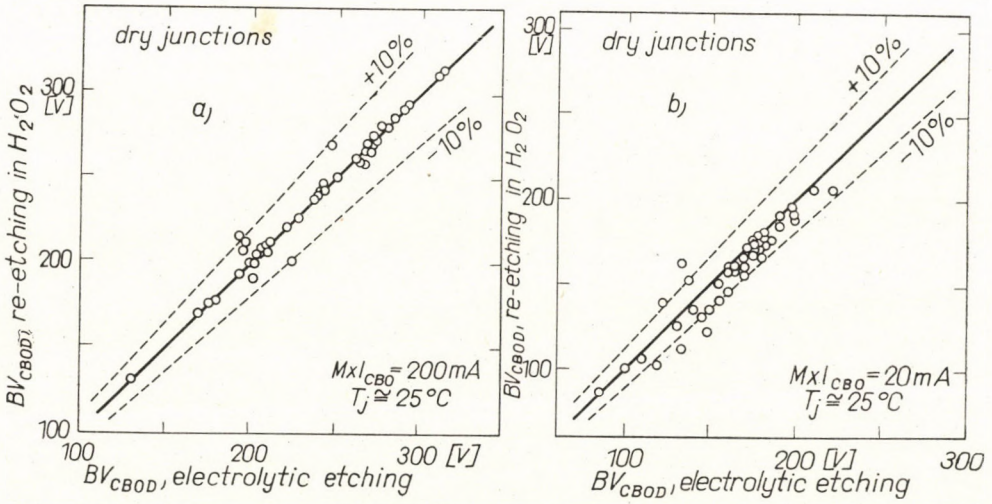


Fig. 13. Breakdown voltage of high-voltage power transistors as a function of the surface-etching or oxidizing treatment, in dry environment:

Horizontal axis: electrolytic etching; *vertical axis:* chemical re-etching in H_2O_2 at a measuring current of 200 mA (*a*); at a measuring current of 20 mA (*b*)

electrolytical etching in 15 per cent KOH, the junctions were rinsed for 3 hours in deionized water, subsequently dried, heat-treated for 6 hours at 120 °C and temporarily encapsulated (without cold-welding, but employing silicon grease on the the cap rim, pressed firmly against the base plate) in dry air (the dew point of which did not exceed —40 °C). To assure the required dry environment, pieces of molecular sieve were introduced into the cap to sustain a dew point lower than —60 °C.

To compute the degradation ratios according to equations (8) and (8a) with the aid of Fig. 4, first the breakdown voltages BV_{CBOD} and BV_{CEO} were determined in the current range 20 ÷ 300 mA followed by the measurement of $1 + \beta_0 = I_{CEO}/I_{CBO}$. After the removal of the cap the junctions were subjected to chemical re-etching in a bath containing 30 per cent H_2O_2 until the appearance of dark-yellow interference. Before permanent encapsulation of junction and molecular sieve in dry air the previous rinse and heat-treatment had been duplicated. Subsequent to a second measurement performed exactly in the aforementioned manner, the caps and the molecular sieves were removed and the junctions were exposed to room temperature for at least 48 hours, putting them in chambers with controllable relative humidity of 35 ÷ 60 per cent. The most important results of the repeated measurements are given in the form of correlation diagrams or histograms, the horizontal scale of which always refers to electrolytical etching or dry ambient.

Figs 13a and 13b demonstrate the interrelation between various dry surface oxide structures and the breakdown voltage BV_{CBOD}. At high multiplication (at 200 mA

corresponding to $M \cong 5 \cdot 10^3$) electrolytic and chemical surface oxide films yield the same breakdown voltage. As can be seen from Fig. 13a, the two differently treated transistor lots have the same average breakdown voltage, i.e. the oxide film structures do not substantially alter the avalanche breakdown voltage. At low-multiplication (at 20 mA, corresponding to $M \simeq 5 \times 10^2$) the average breakdown voltage underwent an insignificant variation only, despite the increased spread (cf. Fig. 13b). As the calculated degradation ratios BV_{CBOD}/BV_{CBO} (Fig. 14) are not subject to surface treatment either, the discussed degradation evidently stems from body imperfections.

Though it is a well-known fact that surface-dependent parameters such as reverse currents, current amplification factor and low-multiplication avalanche breakdown voltage are strongly affected by ambient humidity, the dependence of high-multiplication breakdown voltage on humidity has not been clearly recognized in literature. Figs 15a and 15b show that at high multiplication (i.e. at 200 mA) there is no significant difference between the

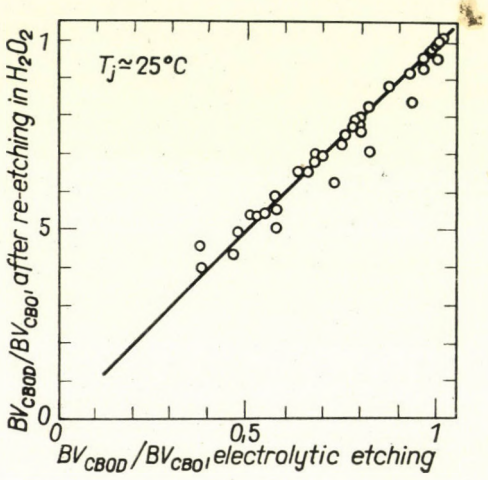


Fig. 14. Calculated normalized degradation as a function of the surface-etching treatment in power transistors referred to in Fig. 13:
Horizontal axis: electrolytic etching; *vertical axis:* chemical re-etching in H_2O_2

BV_{CBOD}-values measured in dry ambience and at 60 per cent relative humidity, apart from the range of high breakdown voltages, where the curve representing the average correlation deviates up to —5 per cent from the 45° line. This is obviously due to the unremovable surface breakdown, causing a small decrease in breakdown voltage, comparable with the accuracy of the measurement.

By contrast, at low multiplication ($M \cong 5 \times 10^2$, i.e. at 20 mA) due to the high humidity BV_{CBOD} drops sharply, as can be seen in Fig. 15b: in dry ambience the average breakdown voltage is 33 per cent down on its initial value and the spread of the individual values increases considerably. Fig. 16 gives the voltage-current characteristics of a typical transistor, measured in dry ambience (solid line) and at 60 per cent relative humidity (dashed lines). It can clearly be seen from the above experiments that in germanium devices with a base resistivity of $10 \div 12$ ohmcm the ambient humidity (even 60 per cent) has but little effect on the breakdown voltage if multiplication exceeds 10^4 (or the avalanche current density surpasses $40 \div 50$ mA/mm^2).

Thus one can conclude that the breakdown voltage measured at such a high multiplication may be regarded as a pure body parameter.

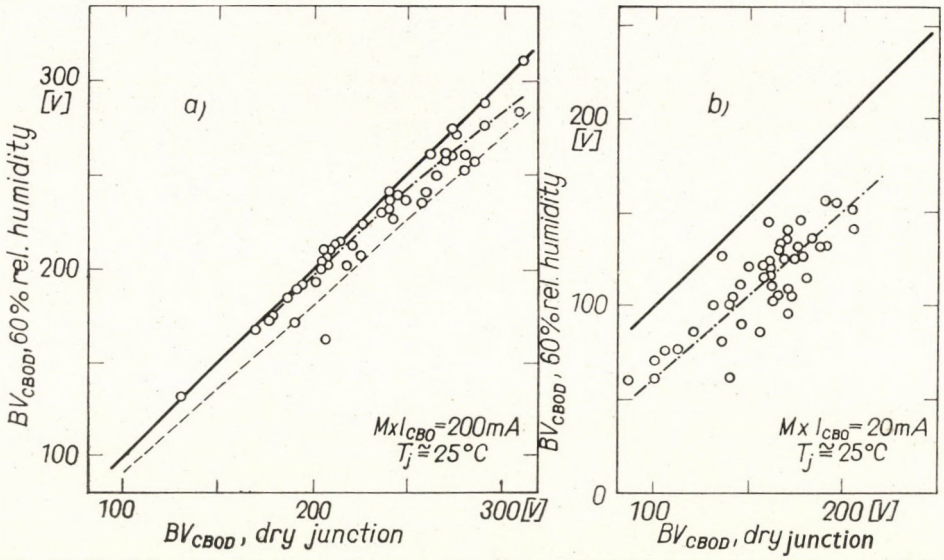


Fig. 15. Breakdown voltage as a function of ambient humidity in power transistors referred to in Figs 13 and 14:

horizontal axis: dry environment; *vertical axis:* 60 per cent relative humidity — at a measuring current of 200 mA (*a*); at a measuring current of 20 mA (*b*)

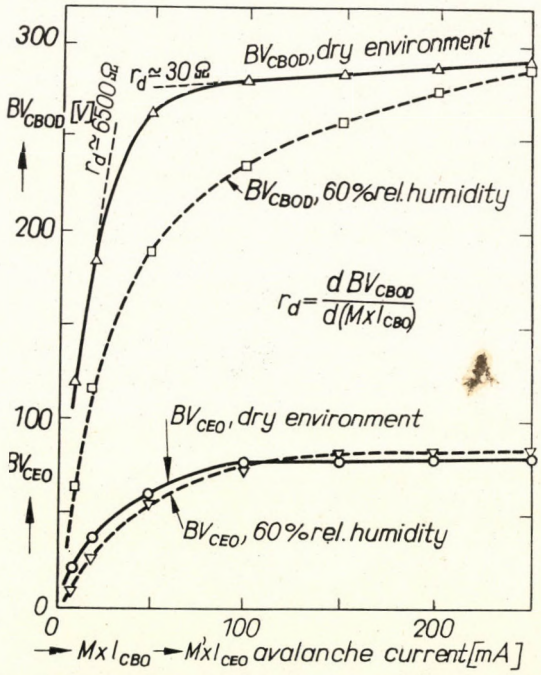


Fig. 16. Avalanche breakdown voltage-avalanche current characteristics of a typical high-voltage power transistor, measured in dry environment (solid lines) and in 60 per cent relative humidity (dashed lines) — the differential resistance in the regions of typical surface breakdown and typical body breakdown are indicated on the curve

APPENDIX 3

Numerical Example for the Calculation of the Degradation Ratio

To demonstrate the application of the described approximation method let us consider a power transistor with the following characteristics:

$$BV_{CBOD} = 265\ V,$$
$$BV_{CEO} = 98\ V,$$
$$1 + \beta_0 = I_{CEO}/I_{CBO} = 29.$$

According to equation (8a)

$$n_D = \frac{\log 29}{\log (265/95)} = 3{,}48.$$

Substituting $BV_{CBO} = BV_{CBOD} = 265$ V into Fig. 4 we obtain $n_F = 2{,}48$. Thus,

$$\lambda = \frac{3{,}48}{2{,}48} = 1{,}4.$$

With the above values equation (8) yields

$$BV'_{CBO} = 98 \text{ num log} \left(\frac{3{,}48}{2{,}48} \log \frac{265}{98} \right) = 388\ V.$$

Repeating the cycle, first we introduce BV'_{CBO} into equation (8a):

$$n'_D = \frac{1{,}463}{\log (388/98)} = 2{,}45.$$

From Fig. 4 with BV'_{CBO} we have $n'_F = 2{,}14$ and

$$BV''_{CBO} = 98 \text{ num log} \left(\frac{2{,}45}{2{,}14} \log 3{,}96 \right) = 480\ V.$$

Duplicating the cycle with BV'_{CBO} we obtain

$$n''_D = \frac{1{,}463}{\log (480/98)} = 2{,}12.$$

in Fig. 4 at $BV''_{CBO} = 480$ V we get $n''_F = 2{,}0$, thus

$$BV'''_{CBO} = 98 \text{ num log} \left(\frac{2{,}12}{2} \log 4{,}9 \right) = 535\ V.$$

It can be seen that the iteration leads to fast convergence. Since in Fig. 4 $BV'''_{CBO} = 535$ V yields $n''_F = 1{,}96$, being in a fairly good agreement with

$$n'''_F = \frac{1{,}463}{\log (535/98)} = 1{,}97 \simeq n,$$

$BV_{CBO} = BV'''_{CBO} = 535$ V represents the real breakdown voltage of the planar-homogeneous section of the junction. With the above breakdown voltage we get a degradation ratio

$BV_{CBOD}/BV_{CBO} = 265/535 = 0{,}496$, suggesting heavy degradation due to body imperfections, and according to equation (1) the resistivity of the planar-homogeneous transistor base section averages 16 ohmcm.

Der Einfluß der strukturellen Irregularitäten von p—n Übergängen auf die volumetrische Lawinendurchbruchspannung. Außer Oberflächeneinflüssen kann die Lawinendurchbruchspannung von Transistoren und Dioden auch durch die geometrischen Fehler des »ebenen« Übergangs und die Inhomogenitäten des spezifischen Widerstands (innerhalb der Verarmungsschicht stark herabgesetzt werden. Die bekannten Theorien lösen das Problem nur für Halbkugel- und Zylindermantelflächen mittels der eindimensionalen Poisson-Gleichung und passen nicht für den praktischen Fall, wo diese Verzerrungen in den ebenen Übergang »eingebettet« sind. Das Problem ist zweidimensional und daher der Rechnung nicht zugänglich. Der Verfasser beschreibt eine Näherung mittels Messung, welche den Zusammenhang zwischen den BV_{CBO} und BV_{CEO} Durchbruchspannungen bei Ge-Transistoren ausnützt. Die genannten volumetrischen Durchbruchspannungen wurden bei großer Stromdichte mittels einer Impulsmethode gemessen. Es wurde festgestellt daß je größer der spezifische Widerstand der leichter dotierten Seite ist (je größere Spannung das Halbleiter-Element aufweist), desto stärker ist die verringernde Wirkung der volumetrischen Irregularitäten auf die Durchbruchspannung; die Herabsetzung der volumetrischen Durchbruchspannung ist dann am größten, wenn der Krümmungsradius der geometrischen Unregelmässigkeiten ungefähr einem Fünftel der Verarmungsschicht gleich ist.Die Rolle von Irregularitäten, die gegenüber der Verarmungsschicht sehr klein sind, kann vernachlässigt werden.

L'influence des irrégularités structurales des jonctions p—n abruptes sur la tension de rupture en avalanche spatiale. Outre les phénomènes de surface, les défauts géométriques de la jonction «plane» et les inhomogénéités de la résistivité à l'intérieur de la couche vide peuvent réduire considérablement la tension de rupture en avalanche des transistors et des diodes. Les théories ne résolvent le problème que pour des surfaces demi-sphériques et cylindriques, en utilisant la forme unidimensionnelle de l'équation de Poisson, et ne conviennent pas pour le cas pratique où la jonction plane subit ces déformations. Le problème étant bidimensionnel, n'est pas accessible au calcul. L'auteur décrit une méthode de mesure utilisant la relation entre les tensions de rupture BV_{CBO} et BV_{CEO} pour les transistors au Ge. Ces tensions de rupture spatiales ont été mesurées aux densités de courant élevées par une méthode d'impulsions. Les résultats ont été les suivants: Plus la résistivité du côté contenant moins d'impuretés (soit la tension de l'élément) est élevée, plus l'effet des irrégularités spatiales diminuant la tension de rupture sera important; la réduction de la tension de rupture spatiale arrive au maximum quand le rayon de courbure des irrégularités géométriques correspond à peu près au cinquième de l'épaisseur de la couche vide. Les irrégularités très petites par rapport à la couche vide peuvent être négligées.

Воздействие нерегулярности структуры абруптных p—n переходов на объемное напряжение лавинного пробоя *(А. П. Кемень)*. Напряжение лавинного пробоя у транзисторов и полупроводниковых диодов, наряду с поверхностными воздействиями, сильно снижается также геометрическими дефектами «плоского» перехода и неоднородностями резистивности (в пределах эвакуированного слоя). Известные теории решают данную проблему только для случая поверхностей оболочки в виде полусферы и цилиндра посредством уравнения Пуассона, для практических же случаев они не подходят, ибо такие искажения расположены на плоском переходе. Проблема является двухмерной и, таким образом, не может быть решена при помощи вычислений. Описывается приближение методом измерения, при котором используется связь между напряжениями пробоя BV_{CBO} и BV_{CEO} в случае германиевых транзисторов. Упомянутые выше объемные напряжения пробоя измеряются при больших плотностях тока импульсированным напряжением. Установлено, что чем более высокую резистивность имеет сторона перехода с меньшим количеством примеси (чем более высокое напряжение имеет средство), тем более сильное снижающие действие будет иметь объемные нерегулярности на напряжение пробоя; снижение объемного напряжения пробоя будет наибольшим по масштабу, когда радиус кривизны геометрической нерегулярности составляет приблизительно одну пятую эвакуированного слоя. Ролью нерегулярностей, являющихся по отношению к эвакуированному слою очень небольшими, можно пренебречь.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 137—154 (1968)

UN EFFET THERMIQUE DES FOURS-TUNNEL

M. KORACH

MEMBRE DE L'ACADÉMIE DES SCIENCES DE HONGRIE

et

J. FÜLÖP

[Manuscrit présenté le 21. juin, 1967]

Les auteurs établissent les équations différentielles de la répartition des températures des fours-tunnels. En concevant ces fours comme des ensembles de 3 appareils de chauffage disposés en série, ils obtiennent un système de 3×2 équations différentielles, liées par des conditions aux limites particulières. Ces équations ont été résolues par la méthode approximative de la relaxation et les résultats ont été résumés par une série de diagrammes de température. Ces diagrammes montrent un bon accord avec les résultats de l'expérience industrielle aussi bien que de laboratoire et présentent un effet de retard de la température du noyau sur celle de la surface, qui croît avec la vitesse d'avancement de la charge. La méthode permet le calcul approché des températures dans les fours-tunnels.

Notations

c chaleur spécifique de la charge [kcal/kp°C];
c' chaleur spécifique du gaz [kcal/kp°C];
h épaisseur de la charge [m];
h' hauteur de la section des gaz [m];
H_1 longueur du préchauffage du four [m];
H_2 longueur de la section de chauffage [m];
H_3 longueur de la section de refroidissement [m];
$H = H_1 + H_2 + H_3$ longueur totale du four [m];
H^v longueur virtuelle du four [m];
k coefficient de transmission des parois du four [kcal/m²°Ch];
q quantité de chaleur transmise au gaz en H_2 [kcal/m²h];
q_b quantité de chaleur transmise dans un volume élémentaire (fig. 2) [kcal];
q_k quantité de chaleur sortant du volume élémentaire [kcal];
$T = T(x; y)$ température de la charge en direction x dans les points y [°C];
$T' = T'(x)$ température moyenne du gaz en direction x [°C];
$T_0 = T(0; y)$ température initiale de la charge [°C];
T'_2 température initiale du gaz [°C];
T_k température de l'air extérieur [°C];
v vitesse d'avancement de la charge [m/h];
v' vitesse du courant de gaz [m/h];
α coefficient de transmission de chaleur par convection [kcal/m²°Ch];
γ poids de l'unité de volume de la charge [kp/m³];
γ' poids de volume du gaz [kp/m³];
λ coefficient de conductibilité de la charge [kcal/mh°C];
$a = \lambda/c\,\gamma\,v$ coefficient de conductibilité de la température [m];
$b = \alpha/v'c'\,\gamma'\,h'$, $h^* = \alpha/\lambda$, $Q = q/v'c'\,\gamma'$, $p = v\,\gamma\,c\,h/2\,v'\,\gamma'\,c'\,h'$.

I

A *l'Institut de Génie Chimique de l'Académie des Sciences de Hongrie* nous avons exécuté une série d'expériences [1], pour voir si nos résultats théoriques et industriels obtenus par la méthode de cuisson que nous avons

nommée «cuisson sandwich» et publiée en plusieurs essais, peuvent être reproduits dans des conditions de laboratoire plus rigoureuses et sur une échelle réduite [2]. Ces expériences ont confirmé, qualitativement aussi bien que quantitativement, la validité de nos conclusions précédentes. En plus, nous avons constaté un phénomène nommé par nous «hysthérésis» de la température de noyau de la charge, c-à-d. le retard sensible du maximum de cette température par rapport à la température maximum de la surface de la charge, en cas d'accélération de la vitesse d'avancement de cette dernière.

Comme cet effet est minime dans les fours-tunnel traditionnels à grande section de charge, la littérature technique ne l'a pas traité jusqu'ici. Mais l'effet est devenu intéressant aussitôt que dans les fours tunnel «sandwich» nous avons pu obtenir jusqu'à des vitesses d'avancement dix fois plus grandes que celle des fours traditionnels. La longeur de ces derniers peut varier entre 50 et 150 m environ et leur hauteur de charge est considérable. Par conséquent les durées de cuisson varient, selon les dimensions et la qualité de la charge, entre 2 1/2 et 10 jours. Dans ces conditions, puisque la vitesse d'avancement de la charge est de 1 m/heure au maximum, les courbes des températures de la surface et du noyau sont très différentes, mais les maxima des deux courbes se situent pratiquement dans la même section transversale du four et se rapprochent entre eux d'une façon technologiquement acceptable (voir fig. 4 et 8). Mais si la vitesse augmente, les deux maxima s'éloignent dans le sens de la longeur du four et précisément d'une façon telle que la température du noyau croît encore, quand celle de la surface va déjà descendant. La fig. 1 représente les courbes de température obtenues dans les expériences publiées en [1]. On y voit que la courbe a de la température des gaz (qui est proche de celle de la surface de la charge) est arrivée au maximum quand celle du noyau est encore de 25 °C environ au-dessous de cette température. La courbe des températures du noyau rejoint à son tour son maximum, après 0,4 m de parcours, quand c'est la température du gaz qui devient inférieure de 30 °C environ. L'hystérésis du maximum de température du noyau est donc de 18′ environ, temps de parcours de la distance $e = 0{,}4$ m. En déplaçant la courbe b dans la position c, c'est-à-dire en faisant correspondre les deux maxima, on constate la coïncidence, approximative, des deux courbes, avec 15 °C de différence de température environ.

II

1. L'hystérésis apparentée déjà à une vitesse relativement modérée de 1,3 m/heure qui a été maintenue dans l'expérience selon la fig. 1, nous a suggéré l'étude théorique de la question, c'est-à-dire le calcul des courbes de température du noyau et de la surface de la charge dans un système donné, en prenant la vitesse d'avancement de la charge comme paramètre. Les vitesses choisies par nous étaient de 1, 2, 4, et 8 m/heure. Les courbes ainsi obtenues permettent

d'établir les valeurs de l'hystérésis mentionnée et celles de la différence des températures maxima des deux courbes. Nous verrons que toutes deux — l'hystérésis du maximum de température du noyau et la différence des maxima des deux courbes de température — vont croissant avec la vitesse; et cela pouvait être prévu car à la limite, c'est-à-dire en cas d'une vitesse infinie, la température du noyau doit rester évidemment invariable.

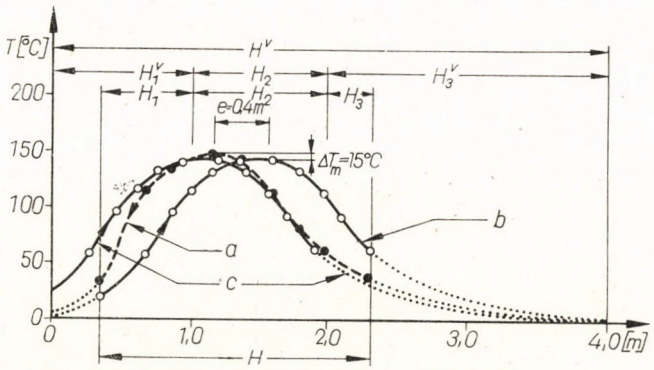


Fig. 1. Courbes des températures du modèle expérimental en fonction de la longueur du four (-- ● -- ● -- ● -- température du gaz; — ○ — ○ — ○ — température du noyau; vitesse d'avancement $v = 1{,}3$ m/h)

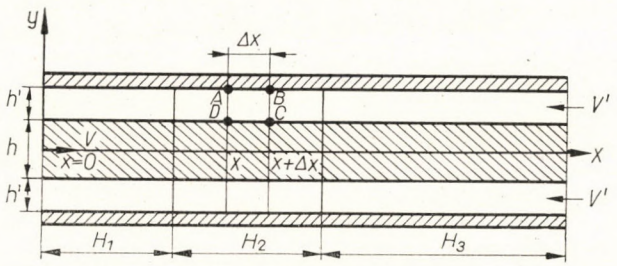


Fig. 2. Section longitudinale du modèle de four théorique

Le modèle théorique du four (fig. 2) pris pour nos calculs est un four-tunnel «sandwich» de 40 m de longueur, pour la cuisson d'une charge de 15 cm d'épaisseur, et d'une largeur beaucoup plus grande, de 1 m. La comparaison des données *expérimentales* du four dix fois plus petit de la fig. 1, avec les données *calculées* du modèle ci-dessus, nous a permis d'arriver, comme on verra au point 7 du chapitre 3—2 ci-dessous, à certaines conclusions relatives à l'effet d'échelle et à ce que nous avons nommé «similitude technologique» [3].

La charge et les gaz se déplacent en contre-courant, la charge allant de gauche à droite. Le modèle de la fig. 2 est le schéma de tous les fours-tunnel: H_1 est la longueur de la zone de préchauffage, où les gaz de combustion (ou en cas de chauffage électrique, l'air chaud) préchauffent la charge qui avance.

Dans la zone de chauffage H_2 la charge reçoit la chaleur de combustion ou celle des résistances électriques, et dans la zone de refroidissement H_3 c'est l'air entrant dans le four qui se préchauffe et la charge qui se refroidit. De cette façon, comme il est notoire, on réalise dans tous les fours-tunnel une récupération de chaleur.

Pour la simplicité du calcul, le modèle est conçu avec les suppositions suivantes:

a) dans la charge, il n'y a pas de propagation de chaleur longitudinale, c'est-à-dire on suppose la charge divisée en tranches minces isolées entre elles;

b) la transmission de chaleur latérale est supposée nulle, la largeur de la charge étant prise, comme on l'a vu, très grande en comparaison de son épaisseur;

c) la température T' du courant des gaz, au-dessus et au-dessous de la charge, est constante dans chaque section transversale, c.-à-d. le long de y;

d) la distribution de température est symétrique par rapport à l'axe x du four.

2. En partant du bilan thermique d'un élément de l'ensemble, nous pouvons écrire une équation différentielle pour la charge aussi bien que pour le gaz. A cet effet, considérons en H_2 l'élément $ABCD$ de l'espace de gaz, de longueur x et d'unité de largeur (fig. 2). Dans la zone de chauffage H_2 la quantité de chaleur que l'élément $ABCD$ *contient*, respectivement *reçoit* est:

$$\Delta q_b = 2\, v'\, h'\, c'\, \gamma'\, (T' + \Delta T') + 2\, q h'\, \Delta x\,, \tag{1}$$

où

v'	la vitesse du gaz;
h'	la hauteur de l'espace libre au-dessus de la charge;
c'	la chaleur spécifique du gaz;
γ'	le poids spécifique du gaz;
T'	la température en AD;
$T' + \Delta T'$	la température en BC;
q	la chaleur reçue dans l'espace H_2 par unité de volume dans l'unité de temps.

La chaleur qui *sort* est:

$$\Delta q_k = 2\, v'\, h'\, c'\, \gamma'\, T' + 2\, \alpha\, \Delta x\, \left[T' - T(x; h/2)\right] + \\ + 2\, k \Delta x (T' - T_k)\,, \tag{2}$$

où

α	le coefficient de transmission convectif;
$T(x_1 h/2)$	la température dans le point des coordonnées $(x; h/2)$, c'est-à-dire à la surface;
k	le coefficient de transmission de la chaleur à travers la paroi du four;
T_k	la température de l'air extérieur.

Comme

$$\Delta q_b = \Delta q_k$$

nous avons

$$\frac{\Delta T'}{\Delta x} = \frac{\alpha}{v' h' c' \gamma'} [T' - T(x;\ h/2)] +$$

$$+ \frac{k}{v' c' h' \gamma'} (T - T_k) - \frac{q}{v' c' \gamma'},$$

et passant à la limite:

$$\frac{dT'}{dx} = b[T' - T(x, h/2)] + \frac{bk}{\alpha}(T' - T_k) - Q. \tag{3}$$

où

$$b = \frac{\alpha}{v' h' c' \gamma'};$$

$$Q = \frac{c}{v' c' \gamma'}.$$

A la surface de la charge, selon NEWTON:

$$\lambda \frac{\partial T(x;\ h/2)}{\partial y} = \alpha [T' - T(x;\ h/2)], \tag{4}$$

où λ est le coefficient de conductivité de la charge.

Dans la charge on a, pour la distribution de température selon FOURIER:

$$\frac{\partial T}{\partial x} = a \frac{\partial^2 T}{\partial y^2}, \tag{5}$$

où

$$a = \frac{\lambda}{c\gamma v}$$

est le coefficient de conductivité de la température.

Selon le principe de la cuisson sandwich, la distribution de la température en direction y étant supposée symétrique, le long de x nous avons:

$$\frac{\partial T(x;\ 0)}{\partial y} = 0. \tag{6}$$

Les mêmes équations sont valables aussi pour les zones H_1 et H_3, mais avec $q = 0$ (et donc $Q = 0$) en (3).

Quant aux surfaces d'entrée, de sortie et de contact entre les trois zones, on a:

Pour la *surface d'entrée de la charge* ($x = 0$) la condition initiale:

$$T(0\,;\ y) = T_0. \tag{7}$$

En se servant des index 1 pour la zone H_1, 2 pour H_2 et 3 pour H_3 et en supposant la continuité dans les courbes de température, on a pour les *zones de contact*, dans la direction x:

$$T_1'(H_1) = T_2'(H_1)\,, \tag{8}$$

$$T_1(H_1\,;\ y) = T_2(H_1\,;\ y)\,, \tag{9}$$

$$T_2'(H_1 + H_2) = T_3'(H_1 + H_2)\,, \tag{10}$$

$$T_2(H_1 + H_2; y) = T_3(H_1 + H_2; y)\,, \tag{11}$$

$$T_3'(H) = T_0'. \tag{12}$$

Pour la *surface d'entrée du gaz* (et donc de la sortie de la charge) on a enfin la condition initiale:

$$T'(H) = T_0' = T_0. \tag{13}$$

En résumé:

Les équations (3) et (4) expriment, avec l'indice 1 des températures pour la zone H_1, avec l'indice 2 pour la zone H_2 et avec l'indice 3 pour la zone H_3, la variation des températures le long de x à la surface de la charge (4) et dans la charge (5), en posant $q = 0$ dans les zones H_1 et H_3.

Les équations (7) et (8) expriment les conditions périphériques dans la direction x, zone H_1, les équations (9) et (10) les conditions périphériques dans la direction x, zone H_2, et les équations (11) et (12) celles dans la zone H_3. Les équations (7), (12) et (13) expriment les conditions initiales.

3. Le système d'équations différentielles partielles (3), (4), (5) et (6) donne des formules trop compliquées pour être facilement résolues par voie analytique. Nous avons donc choisi la solution approchée, en transformant les équations différentielles en équations aux différences finies.

Comme le système est stationnaire, les températures sont constantes dans chaque section transversale. Nous pouvons couvrir tout le modèle de la fig. 2 par un réseau rectangulaire et déterminer les températures de la charge et du gaz dans les points d'intersection, discrets, du réseau. Nous renonçons donc à la détermination de $T = T(x; y)$ et de $T' = T'(x)$ en tant que fonctions *continues* de $(x; y)$, respectivement de (x), et cherchons des formules qui permettent le calcul des températures dans la section correspondante à $x_i + x$, en connaissance de celles de la section x_i.

Les températures des points d'intersection du réseau nécessitent évidemment un calcul qui, pour le transport de chaleur dans le voisinage de la surface de la charge, tient compte aussi de la transmission superficielle (voir fig. 3a),

mais qui, à l'intérieur de la charge, ne considère que la conduction (fig. 3b). Les deux cas nécessitent donc aussi des indices différant des points d'intersection des réseaux et notamment:

$$\rightarrow \eta, (\eta, F), (\eta, F-1) \text{ et } \eta+1, (\eta+1, F), (\eta+1), F-1$$

pour la surface,

$$(\eta, n+1), (\eta, n), (\eta+1, n+1)\,(\eta+1, n), (\eta\, n-1), (\eta+1, n-1)$$

pour la charge, indices utilisés sur les figures 3a et 3b.

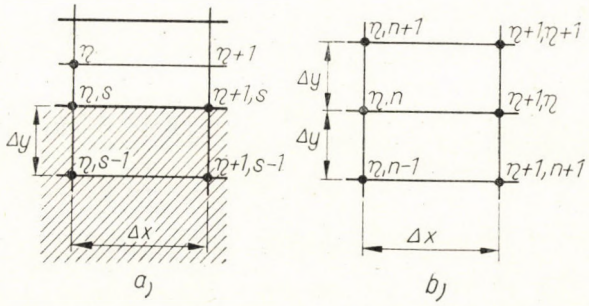


Fig. 3. Réseaux superposés au modèle de la fig. 2; le réseau (*a*) correspond au voisinage de la surface, le réseau (*b*) au noyau de la charge

Nous allons transformer maintenant les équations différentielles (3), (4), (5) et (6) avec la méthode des différences finies, en équations de différences, en donnant des formules approximatives pour les dérivées:

$$\left(\frac{dT}{dx}\right), \left(\frac{\partial T}{\partial x}\right), \left(\frac{\partial T}{\partial y}\right), \left(\frac{\partial^2 T}{\partial y^2}\right).$$

Nous écrivons donc approximativement, pour le *gaz*:

$$\left(\frac{dT'}{dx}\right)_\eta = \frac{T'_{\eta+1} - T'_\eta}{\Delta x}. \tag{14}$$

Alors l'équation (3) donne:

$$T'_{\eta+1} = T'_\eta \left(1 + b\,\Delta x + \frac{bk}{\alpha}\,\Delta x\right) - b\Delta x\, T_{\eta,F} - \left(Q\Delta x + \frac{bk}{\alpha}\, T_k\right). \tag{15}$$

Pour le *voisinage de la surface* nous écrivons, approximativement:

$$\left(\frac{\partial T}{\partial y}\right)_F = \frac{T_{\eta,F} - T_{\eta,F-1}}{\Delta y} \quad . \tag{16}$$

Pour la surface, où $y = h/2$, on peut écrire:

$$T(x\,;\, h/2) = T_{\eta,F}$$

$$T' = T'_\eta$$

et

$$\frac{\partial T_{y,F}}{\partial y} = \left(\frac{\partial T}{\partial y}\right)_F .$$

Posons

$$\frac{\alpha}{\lambda} = h^* .$$

Alors (4) devient:

$$\left(\frac{\partial T}{\partial y}\right)_F = h^*(T'_\eta - T_{\eta,F}) .$$

En substituant de (16), on a

$$T_{\eta,F} = \frac{T_{\eta,F-1} + h^* \Delta y T'}{1 + h^* \Delta y} \tag{17}$$

ce qui est valable pour tous les autres points de la surface, si on remplace

$$T_{\eta,F} \text{ par } T_{\eta+1,F},\ T_{\eta,F-1} \text{ par } T_{\eta+1,F-1}, T'_\eta \text{ par } T'_{\eta+1},$$

et ainsi de suite.

Substituons maintenant (17) en (15), et nous obtenons:

$$T'_{\eta+1} = T'_\eta \left(1 + \frac{b\Delta x}{1 + h^* \Delta h^-} + \frac{bk}{\alpha} \Delta x\right) - \\ - \frac{b\Delta x}{1 + h^* \Delta y} T_{\eta,F-1} - \left(Q\Delta x + \frac{bk}{\alpha} T_k\right) . \tag{18}$$

En connaissance de $T_{\eta,F-1}$, (18) permet donc de calculer la température de gaz, c'est-à-dire la température au *voisinage de la surface*, au lieu de (14) qui permet ce même calcul en connaissance de $T_{\eta,F}$, c'est-à-dire de la température *à la surface*.

Pour *l'intérieur de la charge* nous écrivons approximativement

$$\left(\frac{\partial T}{\partial x}\right)_{\eta,n} = \frac{T_{\eta+1,n} - T_{\eta,n}}{\Delta x} \quad . \tag{19}$$

Substituons en (5):

$$a\frac{\partial^2 T}{\partial y^2} = \frac{T_{\eta+1,n} - T_{\eta,n}}{\Delta x} \,. \tag{20}$$

Pour avoir la valeur approximative de $\partial^2 T/\partial y^2$, développons

$$T_{\eta,n+1} \quad \text{et} \quad T_{\eta,n-1}$$

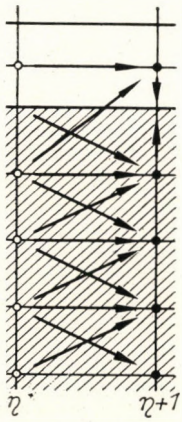


Fig. 4. Réseau pour indiquer les directions des calculs des températures

en série Taylor, en conservant seulement les trois premiers termes:

$$T_{\eta,n+1} = T_{\eta,n} + \left(\frac{\partial T}{\partial y}\right)_{\eta,n} \Delta y + \frac{1}{2}\left(\frac{\partial^2 T}{\partial y^2}\right)_{\eta,n} \Delta y^2 , \tag{21}$$

$$T_{\eta,n-1} = T_{\eta,n} - \left(\frac{\partial T}{\partial y}\right)_{\eta,n} \Delta y + \frac{1}{2}\left(\frac{\partial^2 T}{\partial y^2}\right)_{\eta,n} y^2 \tag{22}$$

et en faisant la somme de (21) et de (22) on a

$$\left(\frac{\partial T^2}{\partial y^2}\right)_{\eta,n} = \frac{1}{\Delta y^2}(T_{\eta,n+1} + T_{\eta,n-1} - 2\,T_{\eta,n}). \tag{23}$$

Substituons (23) en (20) et nous avons

$$T_{\eta+1,n} = \frac{a\,\Delta x}{\Delta y^2}(T_{\eta,n+1} + T_{\eta,n-1}) + T_{\eta,n}\left[1 - 2\,\frac{a\Delta x}{\Delta y^2}\right]. \tag{24}$$

Les équations (17), (18) et (24) permettent le calcul des températures dans chaque section $\eta + 1$, si on connaît celles de la section η, selon le schéma suivant (fig. 4). Dans la figure, des flèches partent des températures connues et se dirigent vers les températures qui en doivent être calculées.

III

1. A l'aide de ces équations récursives nous pouvons déterminer maintenant, par le calculateur électronique, les valeurs des températures pour 7 programmes différents, c'est-à-dire:

pour $\alpha = 50$ kcal/m² h °C, aux vitesses d'avancement $v = 1, 2, 4$ et 8 m/h;

pour $\alpha = 100$ kcal/m² h °C, aux vitesses d'avancement $v = 1, 2$ et 4 m/h.

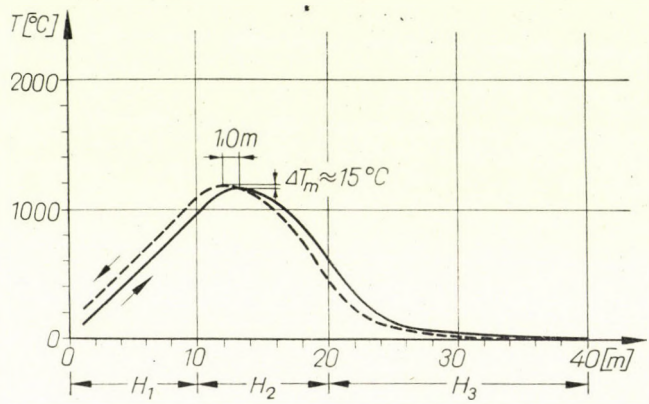


Fig. 5. Diagramme des températures dans le modèle fig. 2 (- - - - - - - - - - - - température du gaz; ———————— température du noyau; vitesse d'avancement $v = 1$ m/h; coefficient de transmission $\alpha = 50$ kcal/m²h°C)

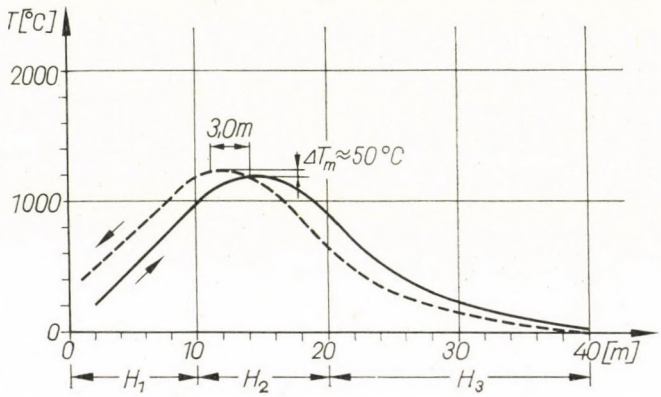


Fig. 6. Diagramme des températures dans le modèle fig. 2 (- - - - - - - - - - - - température du gaz; température du noyau ————————; vitesse d'avancement $v = 2$ m/h; coefficient de transmission $\alpha = 50$ kcal/m²h°C)

On a supposé que les constantes qui figurent dans les équations ont des valeurs moyennes fréquentes dans l'industrie céramique, proches de celles figurant dans les expériences publiées en [1] (voir tableau I).

Tableau I

Constantes et paramètres du modèle fig. 2

a) Dimensions

$H_1 = 10$ m, $H = H_1 + H_2 + H_3 = 40$ m;
$H_2 = 10$ m, $h = 0{,}10$ m,
$H_3 = 20$ m; $h' = 0{,}15$ m.

b) Intervalles du réseau (fig. 3a et 3b)

$\Delta x = 0{,}02$ 0,05 0,10 m,
si $v = 1$ 2 4 et 8 m/h;
$\Delta y = 0{,}015$ m.

c) Constantes de la charge

$\alpha = 50$ et 100 kcal/m² °C h, $\gamma = 1500$ kp/m³,
$\lambda = 1$ kcal/m h °C, $v = 1, 2, 4$ et 8 m/h,
$c + 0{,}2$ kcal/kg °C.

d) Constantes du gaz

$c' = 0{,}24$ kcal/kp °C, $\gamma' = 1{,}3$ kp/m³.

e) Constantes des parois

$k = 0{,}71$ kcal/m² °C h.

f) Valeurs initiales des températures

$$T_0 = T_0' = T_k = 0\ °C.$$

g) Paramètres calculés

En supposant

$$p = \frac{v\gamma ch}{2v'\gamma' c' h'} = 0{,}9$$

on a

$$v' = 801{,}28\ v \text{ m/h}.$$

En écrivant

$$b = \frac{\alpha}{v' h' c' \gamma'}$$

on a

v	1		2		4		8	
α	50	100	50	100	50	100	50	100
b	2	4	1	2	0,5	1	0,25	5

Pour le coefficient de conductivité de température $a = \lambda / c\,\gamma\,v$ on a

v	1	2	4	8
$a \cdot 10$ [m³]	3,333	1,667	0,833	0,4167

En prenant

$$Q = \frac{q}{c'\gamma' v'} = 60\ °C/m\,,$$

$$q = 15{,}6\ v'\,.$$

La symétrie thermique supposée ne permet de faire le calcul que pour la moitié supérieure de hauteur $h' + h/2$ de la fig. 2, comme cela résulte des équations (2), (3) et (4), où la variable $y = h/2$ figure partout.

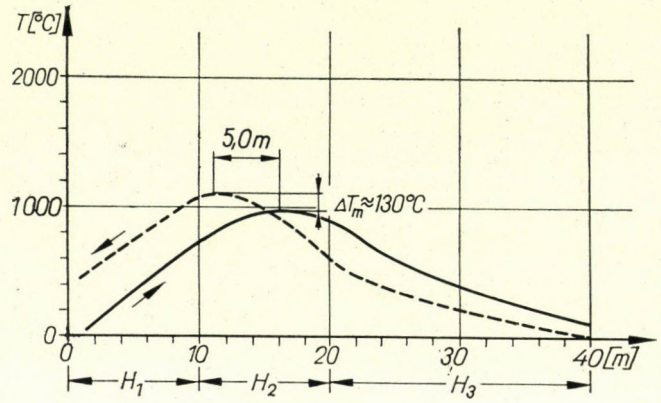


Fig. 7. Diagramme des températures dans le modèle fig. 2 (- - - - - - - - - température du gaz; ———— température du noyau; vitesse d'avancement $v = 4$ m/h; coefficient de transmission $\alpha = 50$ kcal/m²h°C)

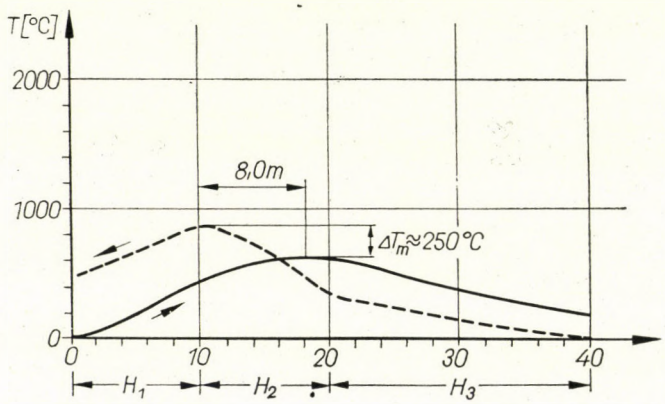


Fig. 8. Diagramme des températures dans le modèle fig. 2 (- - - - - - - - - - température du gaz; ———— température du noyau; vitesse d'avancement $v = 8$ m/h; coefficient de transmission $\alpha = 50$ kcal/m²h°C)

A l'aide des températures, calculées pour les points d'intersection du réseau des fig. 3a et 3b, nous avons construit les figures 5 à 11 en reliant les valeurs discrètes par des lignes continues. Les figures montrent d'une façon évidente les caractéristiques suivantes:

La position du maximum des températures du noyau de la charge se déplace d'une longueur e croissant avec la vitesse par rapport au maximum des températures du gaz (tableau II et fig. 12).

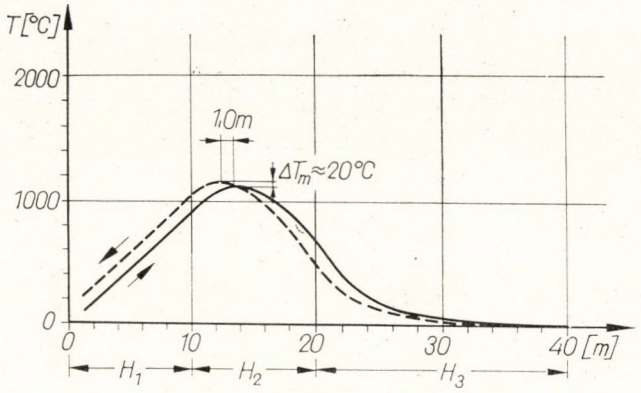


Fig. 9. Diagramme des températures dans le modèle fig. 2 (- - - - - - - - - température du gaz; ———— température du noyau; vitesse d'avancement $v = 1$ m/h; coefficient de transmission $\alpha = 100$ kcal/m²h°C)

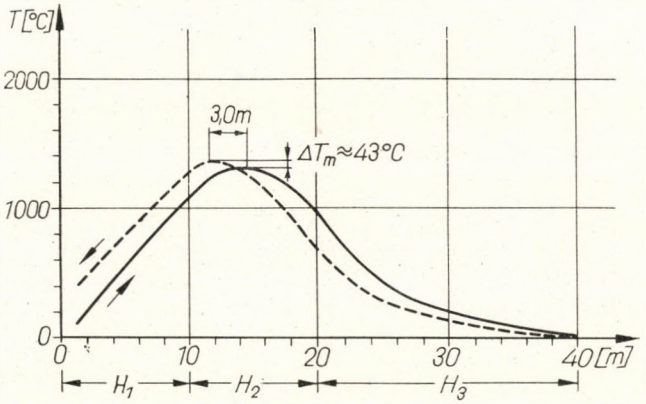


Fig. 10. Diagramme des températures dans le modèle fig. 2 (- - - - - - - - - température du gaz; ———— température du noyau; vitesse d'avancement $v = 2$ m/h; coefficient de transmission $\alpha = 100$ kcal/m²h°C)

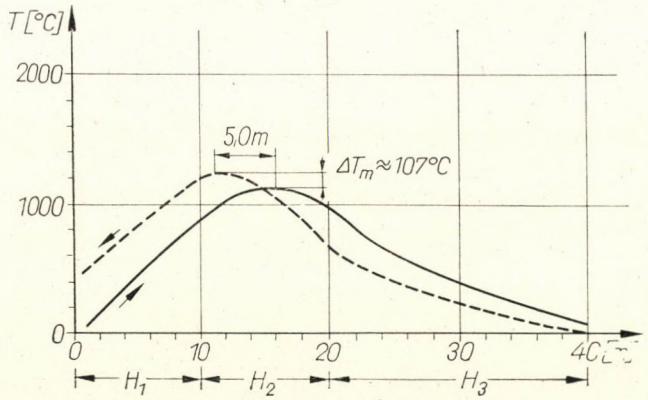


Fig. 11. Diagramme des températures dans le modèle fig. 2 (- - - - - - - - - température du gaz; ———— température du noyau; vitesse d'avancement $v = 4$ m/h; coefficient de transmission $\alpha = 100$ kcal/m²h°C)

Tableau II

Déplacements du maximum de température du noyau de charge en fonction de la vitesse

a)

v [m/h]	e [m]
1	1
2	3
4	5
8	8

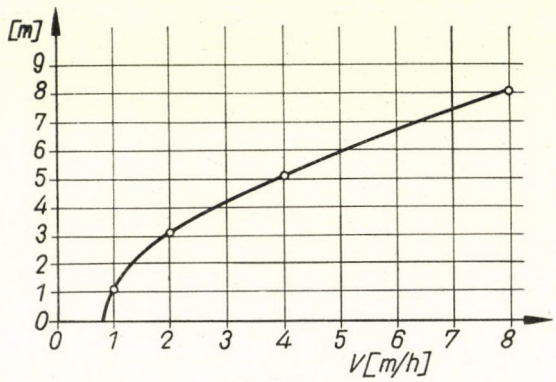


Fig. 12. Hystérésis du maximum de température du noyau de la charge en fonction de la vitesse d'avancement dans le four Fig. 2:

$$e = \frac{v - 0{,}85}{0{,}12} .$$

b) Les températures ont un maximum à la vitesse de 2 m/h, après quoi elles décroissent avec l'augmentation de la vitesse, tandis que les différences entre les deux maxima, du gaz et du noyau, vont croissant (fig. 13).

c) L'augmentation du coefficient de transmission de chaleur α détermine une augmentation des températures correspondantes pour les mêmes vitesses, et les différences entre les deux maxima diminuent.

d) Le calcul permet la détermination du maximum admissible de l'hystérésis, et donc le maximum de vitesse d'avancement de la charge. Si, dans le modèle calculé, on a p. ex. $\alpha = 100$ kcal/m² h °C, et l'on admet une différence de température de 100 °C entre les deux maxima (fig. 11), on ne peut pas dépasser une vitesse de 4 m/h. Mais nous avons supposé une charge compacte. En cas de vaisselle légère, le poids de l'unité de volume qui doit comprendre les parties vides représente seulement 1/5 à peu près de celui de notre modèle, c'est-à-dire $\gamma = 300$ kp/m³. Dans l'équation (5) le coefficient $a = \lambda / c\,\gamma\,v$ reste inchangé si la vitesse v augmente proportionnellement avec γ. Dans notre exemple, on peut alors arriver à 20 m/h d'avancement, ce qui explique l'équilibration presque totale des températures entre gaz et noyau de charge avec 10 m/h d'avancement.

2. Si on superpose les courbes mesurées en [1] ($v = 1{,}3$ m/h) et celles obtenues par interpolation des courbes publiées ici, pour la même vitesse de

1,3 m/h, en choisissant l'échelle de 1 : 200 pour les premières, et 1 : 20 pour les secondes (fig. 14a et 14b), on obtient une comparaison caractéristique des courbes de température. A cet effet nous avons dessiné dans la fig. 14a ce que nous avons nommé «longueur virtuelle» H^v du four d'essai [1], celle qui correspondrait à un four industriel d'échelle 20 fois plus grande, et géométri-

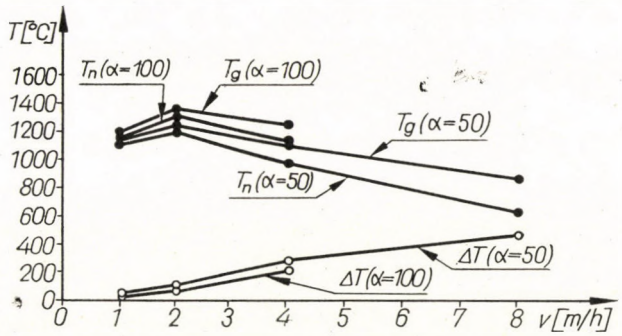


Fig. 13. Allure des maxima des températures dans le four fig. 2 (T_g température du gaz, T_n température du noyau)

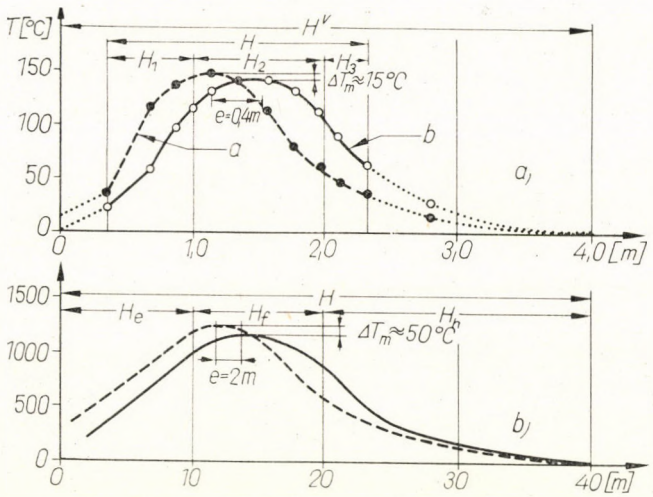


Fig. 14. Confrontation des courbes de température des fours correspondant à la fig. 1 (ici fig. 14a) et à la fig. 2 (ici fig. 14b), avec la même vitesse d'avancement de 1,3 m/h (- - ● - - ● - - ● - - température du gaz; — ○ — ○ — ○ — température du noyau; vitesse d'avancement v = = 1,3 m/h)

quement semblable. La prolongation des courbes 1 (voir fig. 1) n'est évidemment pas problématique. La fig. 14 permet de constater que

a) les courbes des température du gaz et de la charge présentent un point d'intersection dans les deux fours, de sorte qu'en H_1 c'est la courbe des gaz qui est plus haute, tandis qu'en H_3 on a le contraire;

b) les maxima des deux courbes, du gaz et du noyau, montrent un déplacement en 14a aussi bien qu'en 14b, qui est précisément, de 0,4 m pour les deux premières et de 2 m, soit 5 fois plus grand pour les deux secondes, ce qui correspond à l'épaisseur de la charge 5 fois plus grande en 14b qu'en 14a;

c) les différences de température entre les deux courbes, du gaz et du noyau, sont plus grandes avant et après les maxima, c'est-à-dire en H_1 et en H_3, que dans leur intervalle (15 °C en 14a, 30 °C en 14b), c'est-à-dire en H_2. Ces différences sont *relativement* plus considérables en 14a, où la température de cuisson (150 °C) était environ 8 fois plus basse qu'en 14b (1250 °C). Cela correspond au fait, connu dans la pratique, que les températures du noyau et du gaz se rapprochent d'autant plus que les températures de la zone de cuisson H_2 sont plus élevées, évidemment par la prévalence de l'irradiation sur la convection et la conduction. La vitesse d'avancement peut donc être accélérée avec l'élévation de la température de cuisson;

d) en exprimant les températures en °C, on voit que leur échelle, pour le gaz, présente le même rapport, dans les deux fours, que l'échelle des dimensions géométriques. En effet si, en 14a, la température de cuisson avait été de 125° et non de 150 °C, c'est-à-dire $1/10^e$ et non $1/8^e$ de celle du four théorique 14b, les deux courbes du gaz seraient pratiquement coïncidentes. Dans les fours-tunnel, la similitude géométrique des courbes des températures du gaz semble donc être parallèle à celle des dimensions des fours. C'est le contraire de ce qu'on trouve pour les courbes du noyau, lesquelles, comme nous l'avons vu (au point 3), sont d'autant moins semblables que la différence des températures de cuisson est plus grande.

e) De tout cela on doit conclure que *plus grande est la ressemblance technologique, c'est-à-dire l'identité relative des courbes de traitement thermique, moins accentuées devront être les ressemblances géométriques.*

Nous remarquons enfin que les types de courbes de température de fours-tunnel obtenus dans l'industrie ont toujours la même allure que nous avons calculée. La méthode développée *rend possible, en plus, le calcul des courbes de cuisson dans les fours-tunnel,* si on dispose des caractéristiques de la charge, du four et des températures maxima de cuisson.

Les expériences rapportées dans cette publication et en [1] ont été exécutées par Mme P. GÁCSI-BOLDOG avec la collaboration de Mme S. NAGY. Le four expérimental a été construit par M. G. STAHL. Dans l'exécution des expériences nous avons reçu l'aide du M. J. NÉMETH, chef de la Section Mécanique, et dans la réalisation des calculs celui de M. G. SASVÁRI, chef de la section théorique de l'Institut de Génie Chimique de l'Académie des Sciences de Hongrie. Qu'ils veuillent trouver ici l'expression de nos remerciements.

BIBLIOGRAPHIE

1. KORACH, M.: Résultats d'expériences Aérodynamiques et Thermiques sur Modèles de Fours-Tunnel. World Power Conference, Tokyo 1966.
2. KORACH, M.: «Théorie du four-tunnel et cuisson rapide «sandwich» (Essais sur modéles de fours-tunnel théoriques, I). *Acta Techn. Hung.* **11** (1955), 161—184.
   KORACH, M.: *Acta Techn. Hung.* **25** (1959), 25—62.
   KORACH, M.: *Acta Techn. Hung.* **33** (1961), 327—350.
3. KORACH, M.: Léptékhatás a kémiai technológiában (L'effet d'échelle dans la technologie chimique). *MTA Kémiai Tud. Oszt. Közl.* **11** (1959), 205.

A Heat Effect in Continuous Furnaces. The authors establish the differential equations for the temperature distribution in continuous furnaces. The furnace is considered as a series of three units and thus a system of three times two differential equations is obtained, which are connected by suitable conditions. The differential equations are approximately solved by the relaxation method and the results are presented in the form of diagrams. The diagrams agree well with the results of experiences carried out in the laboratory and in operation and with increasing pushing velocity the core temperature increasingly lags behind the surface temperature. The method permits the approximate calculation of the temperature curves in continuous furnaces.

Ein wärmetechnischer Effekt der Tunnelöfen. Die Verfasser stellen die Differentialgleichung der Temperaturverteilung des Tunnelofens auf. Der Ofen wird als eine Reihenschaltung von drei Apparaten aufgefaßt und so werden dreimal zwei Differentialgleichungen erhalten, die durch entsprechende Schaltbedingungen verknüpft sind. Die Differentialgleichungen werden nach der Relaxationsmethode näherungsweise gelöst, und das Ergebnis der Rechnung wird in Form von Diagrammen mitgeteilt. Aus den Diagrammen kann festgestellt werden, daß diese mit den Laboratoriums- und den Betriebserfahrungen gut übereinstimmen, und daß mit steigender Durchschiebegeschwindigkeit die Kerntemperatur gegenüber der Oberflächentemperatur mehr und mehr zurückbleibt. Die Methode ermöglicht die näherungsweise Berechnung der Temperaturkurven in den Tunnelöfen.

Один теплотехнический эффекттуннельных печей *(М. Корах и Я. Фюлёп).* Авторы вывели дифференциальные уравнения распределения температуры туннельной печи. Туннель принимается в качестве трех последовательно соединенных аппаратов и в конечном счете получается система, состоящая из трижды двух дифференциальных уравнений, которые связаны между собой определенными условиями связи. Эти дифференциальные уравнения решаются приближенным методом релаксации и расчетные результаты даются на диаграммах. В отношении диаграмм можно определить, что они удовлетворительно совпадают с производственными и лабораторными опытными данными, и с ростом скорости проталкивания показывают возрастающее отставание температуры, сравнивая с температурой поверхности. Метод позволяет осуществить приближенное вычисление температурных кривых для туннельных печей.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 155—176 (1968)

CIRCUITS WITH IMPEDANCES HAVING Z-CHARACTERISTICS

L. GÁDOR
CAND. OF TECHN. SC.
RESEARCH INSTITUTE OF THE ELECTRICAL INDUSTRY, BUDAPEST

[Manuscript received July 3, 1967]

A group of non-linear elements with extreme properties are those passive circuit elements between whose characteristic data the relation may be expressed by Z-type diagrams. Their examination shown in the paper brings into evidence qualitative relations which facilitate the handling of non-linear circuits.

I. Introduction

Our classic text-books write about the laws of electricity in such a manner that the material and system constants are constants indeed and that the relations are linear. Those among others are Ohm's law, the laws of induction, the Maxwell equations. The so described system in the relations are independent of the absolute values of the variables and from the effects of the surrounding, as temperature, mechanical stress etc., and the principle of superposition is valid in these systems. In the equations the system constants such as R, L, C and the constants of the media and materials being the place of electrical phenomena as σ, μ, ε, figure without restrictions as constants.

Though it soon became known that the validity of the laws stated for the systems, idealized in this way is limited, these limits for a long time stayed outside the range of practical application and linear electrotechnics remained the ruling science in the practice of electrotechnical machines, equipments and apparatuses.

Technical development extended, however, not only to the choice of electrotechnical products, but also set foot on new paths regarding the employed materials, the applied phenomena and the range of quantitative characteristics increased by orders of magnitude.

In this ever more complicated world the facts of non-linearity could no longer be ignored and the methods for handling non-linear phenomena and the system of relations between them had to be built up, just as the relativistic physics have advanced from their sphere of theoretical and cosmological interest to the routine work of the experimental physicists, since they carry out their nucleonic experiments with particles at a speed approaching the velocity of light.

The present paper wishes to contribute to the science of non-linear electrotechnics by examining a special group of passive circuit elements, the behaviour of the conductive, inductive and capacitive elements with Z-characteristics in the a.c. circuit.

The Z-characteristic is from a certain point of view an extreme manifestation of non-linearity, which just due to its extremeness may lead to statements characteristic of this type of circuits. The non-linearities of intermediate type occurring in practice can generally be produced as a combination of Z-characteristics and of classic linear elements. The examination of the behaviour of elements with Z-characteristic might also be useful for this reason.

II. Definitions

Before treating the meritory part, some definitions, concepts, denominations and symbols must be agreed upon, using and generalizing such as was given previously in a paper dealing with a similar subject [2].

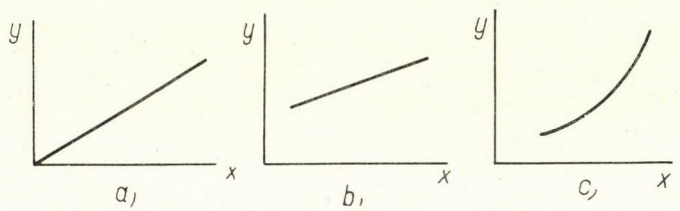


Fig. 1

Proportionate element is that circuit element, where the relation between the standard characteristics is represented by a straight through the origin (Fig. 1a). Such is the relation between the voltage across the ohmic resistance and the current flowing through it, or the relation between the current of the air-gap choke and the flux.

Linear element is that circuit element, where the relation between the standard characteristics is linear, but the straight does not contain the origin (Fig. 1b). Such is e.g. the relation between the current and the voltage on the section beyond the bend of the selenium rectifier or the relation current — voltage, or flux — current, respectively, of the saturating choke on the saturated section.

Non-linear is called the circuit element, where there are non-linear relations between its standard characteristics (Fig. 1c). Such is the relation between current and voltage of the incandescent lamp, this is the characteristic of the iron-core choke apart from the initial section, which can be regarded as proportionate and from the saturation section, which can be regarded as linear.

Resistor is called that circuit element, on whose terminals current and voltage can always act only in the same direction and which dissipates the electrical power taken up in the form of some kind of energy. Its behaviour in the circuit is determined by its resistance and its conductivity.

The static resistance of the resistor is

$$R = \frac{u}{i}, \tag{1}$$

its dynamic resistance

$$R_d = \frac{du}{di}, \tag{2}$$

its static conductivity

$$G = \frac{i}{u}, \tag{3}$$

its dynamic conductivity

$$G_d = \frac{di}{du} \tag{4}$$

where u is the instantaneous value of the voltage and i that of the current.

Choke is that passive circuit element, where on the effect of the current flowing through it, a magnetic field develops. Across its terminals a voltage may appear whilst no current is flowing through it and conversely. If current and voltage appear at the same time, their direction may be identical, or opposite. In the case of identical directions the consumed power is completely transformed into magnetic energy. The behaviour of the choke in the circuit is determined by its inductance.

The static inductance of the choke is

$$L = \frac{i}{\psi}, \tag{5}$$

its dynamic inductance

$$L_d = \frac{di}{d\psi} \tag{6}$$

where ψ is the total flux linking with the exciting circuit (magnetization).

Condenser is that passive circuit element, where the voltage appearing across its terminals is incident to accumulation of charge. Across its terminals voltage may appear while no current is flowing through it and conversely, the direction of current and voltage may be identical or opposite, as with the choke. In case of identical directions the consumed power totally transforms into electrical energy.

The behaviour of the condenser in the circuit is determined by its capacity. The static capacity of the condenser is

$$C = \frac{q}{u}, \tag{7}$$

its dynamic capacity

$$C_d = \frac{dq}{du} \tag{8}$$

where q is the instantaneous value of the charge accumulated on the armature of the condenser.

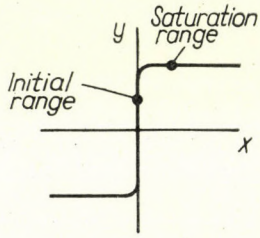


Fig. 2

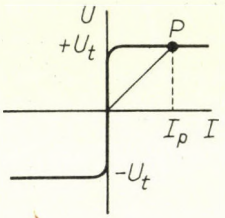


Fig. 3

Up till now in our definitions we avoided the expression inductive and capacitive *resistance,* connected in traditional electrotechnics with the physical devices representing the circuit elements, which, owing to their inductive or capacitive properties, let the current pass at a given alternating voltage to a lesser or higher degree. The Z-elements may open in a part of the period completely and may close in another part completely, so the traditional expression is misleading. We could accept at most a statistic interpretation of some kind, if we examine the mean value of current and voltage for the totality of the circuit also containing Z-elements. Fig. 2 shows the typical relation between the variables x and y of a circuit element with a Z-characteristic. For the *initial range* of the characteristic $x = 0$, in the sense of the above definitions this section of the characteristic is proportionate. Consequently in this range the values of the static and of the dynamic parameters are identical, namely, depending on the definition of the parameter they are zero or infinite. For instance, both the static and dynamic resistances of the voltage limiting resistor (Fig. 3) are infinite, both its static and dynamic conductivities are zero.

The *saturation range* is linear, its function can be given in the form of $y = (\mathrm{sg}\, x) \cdot \mathrm{const}$. The saturation is a typical non-linear property. In this range the value of the static parameter varies from point to point, the value of the dynamic parameter is constant, namely, zero or infinite depending on its definition. E.g. adhering to the voltage limiting resistor mentioned above,

its static resistance and conductivity, changing from point to point, are $R = U_t/I$, and $G = I/U_t$, respectively, while its dynamic resistance is zero everywhere and its dynamic conductivity is infinite everywhere.

Thus the Z-characteristic is not continuous, but it consists of sections which are simple and easy to handle mathematically.

We limit our examinations to the elements of instantaneous response and free of hysteresis. Further we assume the non-linearity as being non temperature dependent, but its origin to be some other physical phenomenon, as for instance, a blocking layer, or the field strength dependent generation of the charge carriers.

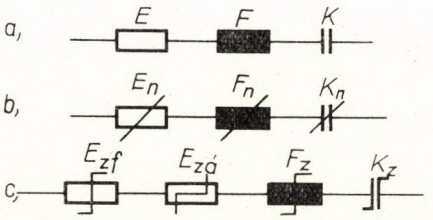


Fig. 4

While the first limit excludes frequency dependence, the latter supposition involves the requirement that the characteristic of the examined element changes neither in consequence of effects of the surrounding, nor under service conditions. Finally we assume the symmetry of the characteristic, i.e. that the behaviour of the examined circuit element is independent of the direction of the current and of the voltage, no rectifying effect arises.

On the circuit diagrams the traditional symbols according to Fig. 4a show the proportionate resistor (E), the choke (F) and the condenser (K), respectively, the crossed symbols of Fig. 4b mark the non-linear elements, while the crossing-out by broken lines of Fig. 4c mark the elements with Z-characteristic. We maintain the traditional letters R, L, C, to mark the parameters themselves, the resistance, the inductance and the capacitance. Those, who deal with linear electrotechnics have an easier job: it does not matter, when the choke and its inductance are both marked by L, the inductance of the choke being a constant value. For impedances with Z-characteristics, on the contrary, there is a separate static and a separate dynamic inductance and moreover both may change from point to point.

III. Voltage limiting Z-resistor

We get a voltage limiting Z-resistor with parameters well approaching the ideal characteristic, for instance, from two opposite connected Zener-diodes (Fig. 5). The opposite connection guarantees the symmetry, the current

starts in whatever direction only when exceeding the saturation voltage U_t, in our case when exceeding the Zener-voltage.

The limiting resistor E_a and the limiting resistance R_a, respectively — in general the useful consumer — ensure that the current should not reach too high values, which could endanger the circuit.

The supply is an alternating voltage

$$u = U \sin \omega t . \tag{9}$$

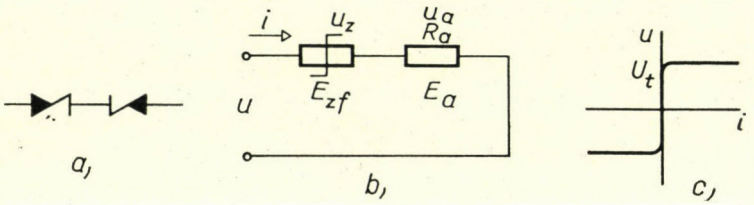


Fig. 5

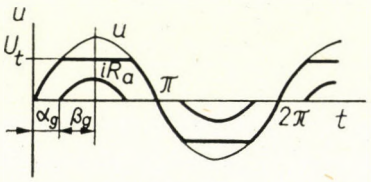


Fig. 6

As long as the voltage u does not reach the value of saturation voltage U_t, no current flows in the circuit, the whole exterior voltage is taken up by E_z (Fig. 6), $u_z = u, u_a = 0$.

In the section

$$\alpha_g < \omega t < (\pi - \alpha_g)$$

the dynamic resistance of E_z is practically zero:

$$R_{zd} = \frac{du_z}{di} = 0 , \tag{10}$$

the voltage above U_t is taken up by E_a and the current is correspondingly

$$i = \frac{u - U_t}{R_a} . \tag{11}$$

As long as $U \leqq U_t$, E_z comports itself like an open switch, it takes up all the voltage and no current flows in the circuit.

Is $U > U_t$, then also for a time no current flows, it starts, however, in that phase of the period, when u attains the value of U_t, i.e. in the moment corresponding to the relation $\omega t = \alpha_g$, but stops at the moment $\omega t = \pi - \alpha_g$. Hence

$$U \sin \alpha_g = U_t \,, \tag{12}$$

$$\alpha_g = \text{arc sin} \frac{U_t}{U} \,. \tag{13}$$

Thus summing up

$$u_a = 0, \; u_z = u, \; i_a = i_z = 0 \,,$$

if $0 < \omega t < \alpha_g$ and $(\pi - \alpha_g) < \omega t < \pi$;

$$u_a = u - U_t; \; u_z = U_t, i_a = i_z = \frac{u - U_t}{R_a} \,, \tag{14}$$

if $\alpha_g < \omega t < (\pi - \alpha_g)$.

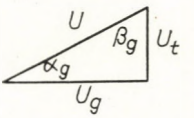


Fig. 7

Let us calculate the linear mean value of the voltage drop across the load resistance R_a:

$$U_{ak} = R_a I_k = \frac{2}{\pi} \int_{\alpha_g}^{\pi/2} (u - U_t) \, d\omega t \,, \tag{15}$$

$$U_{ak} = \frac{2}{\pi} \int_{\alpha_g}^{\pi/2} U \sin \alpha \, d\omega t - \frac{2}{\pi} \int_{\alpha_g}^{\pi/2} U_t \, d\omega t \,, \tag{16}$$

$$U_{ak} = \frac{2}{\pi} \left[U \cos \alpha_g - U_t \left(\frac{\pi}{2} - \alpha_g \right) \right], \tag{17}$$

$$U_{ak} = \frac{2}{\pi} (U_g - U_t \beta_g) \,. \tag{18}$$

where $U_g = U \cos \alpha_g$ and $\beta_g = \pi/2 - \alpha_g$ (Fig. 7).

Both the current and voltage of the voltage limiting resistor and those of the charging resistance are odd functions, i.e. their fundamental harmonic is in phase with the exterior voltage u. This has to be, because the circuit contains only conductive elements, the closing and opening of the non-linear element depends on the momentary value of the voltage, therefore, the curves showing the state of the voltage and of the current are arranged symmetrically within the period.

We obtain the power consumed in the circuit by multiplying the mains voltage with the fundamental harmonic of the flowing current, being in phase with the mains voltage. The amplitude of the fundamental harmonic of the current is

$$I_1 = \frac{4}{\pi} \int_{\alpha_g}^{\pi/2} i(t) \sin \omega t \, d\omega t \,, \tag{19}$$

$$I_1 = \frac{4}{R_a \pi} \int_{\alpha_g}^{\pi/2} u_a \sin \omega t \, d\omega t \,, \tag{20}$$

$$I_1 = \frac{4}{R_a \pi} \int_{0}^{\pi/2} (U \sin \omega t - U_t) \sin \omega t \, d\omega t, \tag{21}$$

$$I_1 = \frac{2}{R_a \pi} (U \beta_g - U_t \cos \alpha_g) \,. \tag{22}$$

The total power taken up by the circuit is

$$N_s = \frac{U I_1}{2} = \frac{U}{2} \frac{2}{R_a \pi} (U \beta_g - U_t \cos \alpha_g) \,, \tag{23}$$

$$N_s = \frac{U^2}{2 R_a} \frac{2 \beta_g}{\pi} - \frac{U_g U_t}{R_a \pi} \,. \tag{24}$$

From this, the resistance E_z dissipates the power

$$N_z = \frac{2}{R_a \pi} \int_{\alpha_g}^{\pi/2} U_t (U \sin \omega t - U_t) \, d\omega t = \frac{U_t U_{ak}}{R_a} = U_t I_{zk} \tag{25}$$

the rest is consumed on the load resistance R_a. Here $I_{zk} = U_{ak}/R_a$ is the mean value of the current flowing through the Z-resistor.

IV. Current limiting Z-resistor

Current limiting Z-resistors for alternating current can be produced by a simulating network. Their idealized characteristic is conform with (Fig. 8b). Till reaching the saturation current I_t, the current is let through without resistance, its value is determined in this range by the load resistance R_a. The supply voltage is, as it was also in the previous paragraph, $u = U \sin \omega t$.

As long as the current flowing in the circuit does not reach the saturation current I_t, E_z comports itself as if it would not even be present in the circuit, the total voltage is taken up by R_a (Fig. 9),

$$u_a = u, \; u_z = 0 \,.$$

In the section

$$\alpha_g < \omega t < (\pi - \alpha_g)$$

the dynamic conductivity E_z is practically zero:

$$G_{zd} = \frac{di}{du_z} = 0\,, \tag{26}$$

the voltage above $I_t R_a$ is taken up by E_z, the current at the same time being $i = I_t = \text{const.}$

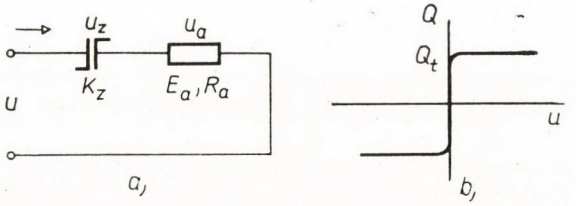


Fig. 8

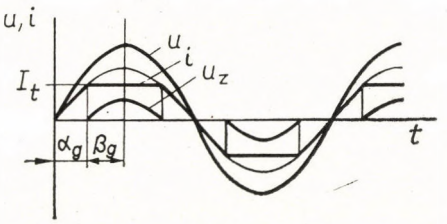


Fig. 9

As long as $U/R_a \leqq I_t$, E_z behaves as a closed switch, the consumer takes up the total voltage, in the circuit flows the current determined by R_a.

If $U/R_a > I_t$ the current also flows freely in a part of the period, but when it reaches the value I_t in the moment corresponding to the relation $\omega t = \alpha_g$, it does not increase any longer, but stabilizes itself at this value. In the same way the voltage $I_t R_a$ becomes stable on the consumer, the exceeding voltage $u - I_t R_a$ is taken up in the range

$$\alpha_g < \omega t < (\pi - \alpha_g)$$

by E_z.

Thus summing up

$$u_a = u,\ u_z = 0,\ i_a = i_z = \frac{u}{R_a}\,,$$

if $0 < \omega t < \alpha_g$ and $(\pi - \alpha_g) < \omega t < \pi$;

$$u_a = I_t R_a, u_z = u - I_t R_a, \; i_a = i_z = I_t, \tag{27}$$

if $\alpha_g < \omega t < (\pi - \alpha_g)$.

The mean values and the outputs may be calculated in a similar way as in the case of the voltage limiting resistor.

Mean voltage on the Z-resistor

$$U_{zk} = \frac{2}{\pi}(U_g - I_t R_a \beta_g). \tag{28}$$

Mean voltage on the load resistance

$$U_{ak} = U_k - U_{zk}, \tag{29}$$

where $U_k = (2/\pi)\, U$ is the mean value of the mains voltage.

The total power taken up by the circuit is

$$N_s = \frac{U^2}{2 R_a} \frac{2 \alpha_g}{\pi} + \frac{U_g I_t}{\pi}. \tag{30}$$

From this the Z-resistor dissipates the output

$$N_z = I_t U_{zk}, \tag{31}$$

the rest is consumed on the load resistance R_a.

Let us summarize the results obtained for the voltage limiting and for the current limiting resistors in Table I.

Table I

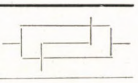
		
U_{ak}	$\frac{2}{\pi}(U_g - U_t \beta_g)$	$U_k - U_{zk}$
U_{zk}	$U_k - U_{ak}$	$\frac{2}{\pi}(U_g - I_t R_a \beta_g)$
N_a	$N_s - N_z$	$N_s - N_z$
N_z	$U_t I_{zk}$	$I_t U_{zk}$
N_s	$\frac{U^2}{2 R_a} \frac{2 \beta_g}{\pi} - \frac{U_g U_t}{R_a \pi}$	$\frac{U^2}{2 R_a} \frac{2 \alpha_g}{\pi} + \frac{U_g I_t}{\pi}$

These same values are given in Table II, if

$$U_t \geq U \text{ and } I_t \geq U/R_a, \quad \alpha_g = \pi/2, \quad \beta_g = 0, \quad U_g = 0,$$

respectively.

Table II

		
U_{ak}	0	U_k
U_{zk}	U_k	0
N_a	0	N_s
N_z	0	0
N_s	0	$U^2/(2R_a)$

The two resistors are each other's duals. During the saturation range E_{zf} operates as an open switch, it takes up all voltage, E_{za} comports itself as a closed switch, all voltage is taken up by R_a. Hence, in the first case there is no current and no power at all, in the second case the circuit behaves as if R_a would have simply been connected to the mains.

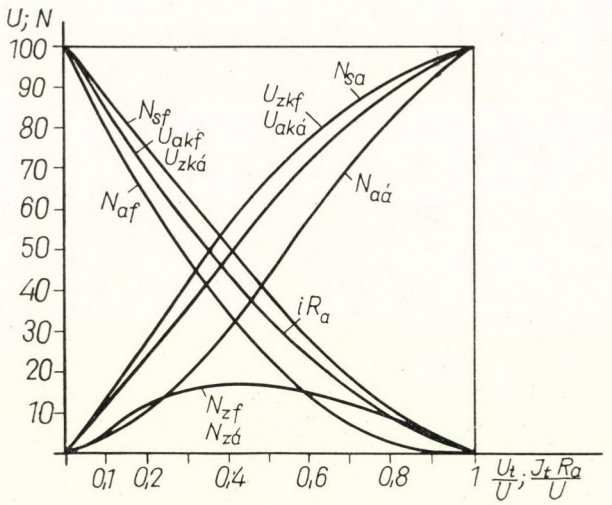


Fig. 10

Fig. 10 also gives a view on the intermediate values of the voltages and powers. The abscissae are increasing values of U_t and $I_t R_a$, respectively, for a fixed U, the ordinates are of values U and N, respectively, where 100 is the

mean value, U_k of the constant mains voltage and the maximum dissipated power, respectively.

In an extreme case there is either no current, or no voltage on the *Z*-element, thus the consumed power N_z is zero, it reaches its maximum about at the central point.

The *ATSZ—3* type voltage regulating device developed in the *Research Institute of the Electrical Industry* serves for automatic control of step voltage regulating transformers. The error signal actuates an integrating organ with a Ferraris disc. The stable adjustment of the steps requires the use of a dead

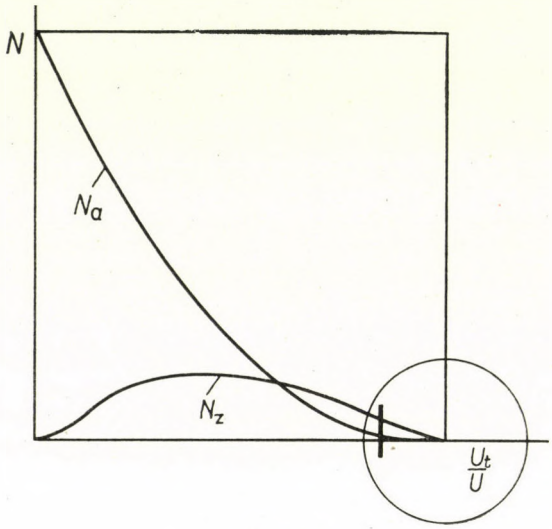


Fig. 11

zone, and this was solved by a *Z*-resistor containing a pair of Zener-diodes. The encircled area on Fig. 11 is the range, where the control voltage starts emerging from the dead zone limited by U_t. As is to be seen the N_a employed for actuation starts extraordinarily flat and in these surroundings, the power N_z dissipated in the limiting resistor is much greater than the useful N_a. This makes the start of the integrating organ uncertain, therefore in practice we moved the place of the start somewhat off by means of an independent magnetic moment limiter, to a steeper section of N_a and so the definite starting was ensured.

One more lesson: the output levels strongly impair the efficiency of the limitation, therefore we get a reasonable solution only by amplifying the useful output. I.e. if amplification is necessary anyway, it is advisable to apply the limitation in the low-level stage.

V. Z-choke

The choke of great initial permeability made of special magnetic alloys approach the Z characteristic of the Fig. 12b well enough. We neglected the hysteresis in our discussion. Thus, the iron core of the choke can be magnetized up to an induction B_t, or a flux Φ_t, or a magnetization ψ_t, respectively, by a practically negligible excitation current; above this, however, the increasing of the excitation current does not increase the magnetization, the choke is saturated.

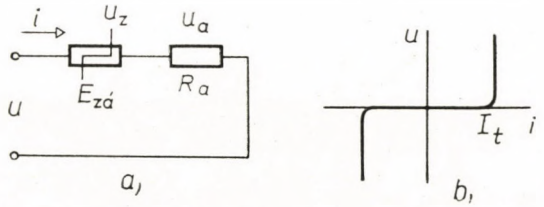


Fig. 12

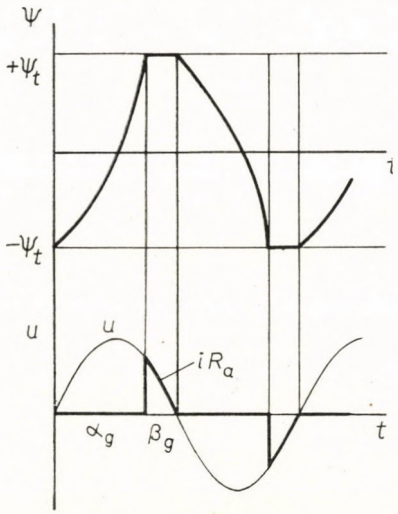


Fig. 13

Hence, the dynamic permeability is infinite till it reaches ψ_t, beyond this it is zero. The static and the dynamic inductance are infinite in the initial range, the dynamic inductance is zero in the saturation range.

While the magnetization in the circuit changes according to Fig. 12a from $-\psi_t$ to $+\psi_t$, the induced voltage across the terminals of the choke is just great enough to balance the total exterior voltage $u = U \sin \omega t$. Now thus $u = d\psi/dt$ across the terminals of the choke, and in the circuit $i = 0$. The conditions for the whole cycle are shown on Fig. 13. The reversal of the

magnetization is accomplished at the time α_g, from this time on the current flows freely in the circuit, its value is limited only by the load resistance R_a. During the period of the magnetization the voltage area is

$$\int_0^{\alpha_g/\omega} u\,dt = 2\,\psi_t. \tag{32}$$

Summing up

$$\begin{aligned} &\text{while} \quad 0 < \omega t < \alpha_g\,, \;\; u_a = 0,\;\; u_z = u = d\psi/dt\,,\;\; i = 0\,; \\ &\text{while} \quad \alpha_g < \omega t < \pi\,, \;\; u_a = u,\;\; u_z = 0,\;\; i = u/R_a\,. \end{aligned} \tag{33}$$

Let us calculate in a similar way as above the mean values appearing across the terminals of the individual elements.

Up to now we could integrate according to ωt, because the phenomena are frequency independent.

The saturation phenomenon occurs on each occasion, when the voltage or the current reaches the saturation value of the corresponding parameter of the Z-element. In the case of the choke the saturation is defined by the integral relation (32), the voltage area appearing on the left side of the relation determines the starting angle α_g not only in dependence on U, but also on the magnitude of the period, i.e. on the frequency. Hence, in our calculation we have to proceed correspondingly.

While $0 < \omega\, t < \alpha_g$, the total voltage is taken up by the choke, no current is flowing in the circuit. The mean value of the voltage consumed by the choke is

$$U_{zk} = \frac{\omega}{\pi} \int_0^{\alpha_g/\omega} u\,dt. \tag{34}$$

As in the sense of relation (32) the value of the integral is $2\,\psi_t$

$$U_{zk} = \frac{2}{\pi}\,\omega\psi_t\,, \tag{35}$$

$$U_{ak} = U_k - U_{zk} = \frac{2}{\pi}\,(U - \omega\psi_t)\,. \tag{36}$$

It is to be noted that the current and voltage shapes, produced by the Z-choke and represented in Fig. 13 are well known to electricians in our days, the same are produced by the thyristor started at phase α_g.

Hence as long as $0 < \omega t < \alpha_g$, the choke is blocked, at the moment α_g it opens, in the range $\alpha_g < \omega t < \pi$ the current flows freely with the intensity determined by the consumer R_a. In this stage $u_z = 0$ on the choke, thus

it does not consume power, the total power used in the circuit appears on the consumer R_a,

$$N_z = 0, \quad N_s = N_a.$$

Let us calculate the amount of this power:

$$N_a = \frac{\omega}{\pi} \int_{\alpha_g/\omega}^{\pi/\omega} \frac{u^2}{R_a} dt = \frac{\omega U^2}{R_a \pi} \int_{\alpha_g/\omega}^{\pi/\omega} \sin^2 \omega t \, dt, \tag{37}$$

$$N_a = \frac{U^2}{2 R_a} \frac{\beta_g}{\pi} - \frac{U^2}{2 R_a \pi} \sin \alpha_g \cos \alpha_g ; \tag{38}$$

$$\cos \alpha_g = \frac{U - 2 \omega \psi_t}{U} = \frac{\pi}{2} \frac{U_{ak} - U_{zk}}{U}, \tag{39}$$

$$\sin \alpha_g = \frac{2 \sqrt{\omega \psi_t (U - \omega \psi_t)}}{U} = \frac{\pi}{2} \frac{2 \sqrt{U_{ak} U_{zk}}}{U}, \tag{40}$$

substituting the values of $\cos \alpha_g$ and $\sin \alpha_g$ into Equ. (38),

$$N_a = \frac{U^2}{2 R_a} \frac{\beta_g}{\pi} - \frac{1}{R_a \pi} (U - 2 \omega \psi_t) \sqrt{\omega \psi_t (U - \omega \psi_t)} \tag{41}$$

or

$$N_a = \frac{U^2}{2 R_a} \frac{\beta_g}{\pi} - \frac{\pi}{4 R_a} (U_{ak} - U_{zk}) \sqrt{U_{ak} U_{zk}}. \tag{42}$$

VI. Z-condenser

Let us examine for the sake of completeness also the behaviour of the condenser having a Z-characteristic, though no such is found nowadays among the ordinarily employed means. There are, however, also to-day non-linear condensers: the blocking layer of some semi-conductors behave in the blocking direction like condensers with changing capacity and the antiparallel pair formed of them approaches the Z-condenser. The ferro-electric dielectrics result as well in non-linear condensers but these show an important hysteresis.

We carry out our investigations on the ideal Z-condenser corresponding to Fig. 14b. This element can accumulate only a determined charge Q_t.

Until the saturation charge is not accumulated in the condenser, this comports itself like a closed switch and allows the current to pass unhindered, whose value is determined by the voltage u and the load resistance R_a. When saturation is reached the condenser interrupts the circuit and behaves like an open switch (Fig. 15).

Hence

$$\begin{aligned} &\text{while} \quad 0 < \omega t < \alpha_g, \ u_a = u, \ u_z = 0, \ i = u/R_a, \\ &\text{while} \quad \alpha_g < \omega t < \pi, \ u_a = 0, \ u_z = u, \ i = 0. \end{aligned} \tag{43}$$

While current flows, $i = dQ/dt$, the saturation charge is given by

$$\int_0^{\alpha_g/\omega} i \, dt = 2 Q_t. \tag{44}$$

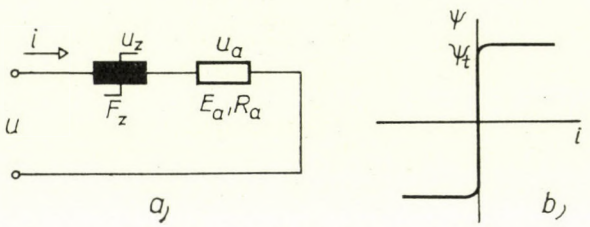


Fig. 14

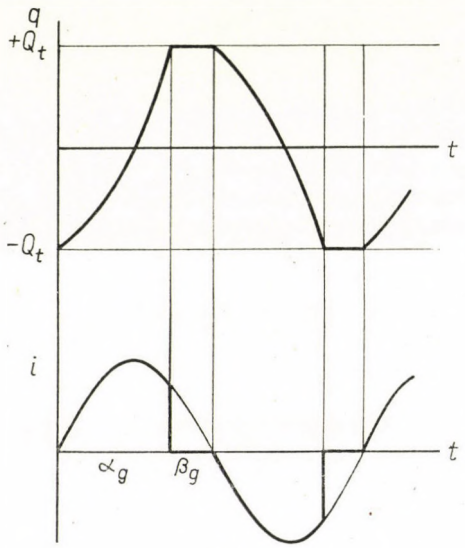


Fig. 15

Substituting the value of the current into (44)

$$\frac{U}{R_a} \int_0^{\alpha_g/\omega} \sin \omega t \, dt = 2 Q_t, \tag{45}$$

$$\frac{U}{R_a} \int (1 - \cos \alpha_g) = 2 \omega Q_t. \tag{46}$$

Let us calculate the mean values of the voltages appearing across the individual elements

$$U_{ak} = \frac{\omega}{\pi} \int_0^{\alpha_g/\omega} u\, dt = \frac{\omega R_a}{\pi} \int_0^{\alpha_g/\omega} i\, dt\,, \tag{47}$$

$$U_{ak} = \frac{2}{\pi} R_a \omega Q_t\,; \tag{48}$$

$$U_{zk} = \frac{2}{\pi} (U - R_a \omega Q_t)\,. \tag{49}$$

When the voltage appears across the condenser, the current is interrupted, thus, the condenser does not consume any power, the load Q_t accumulated at zero voltage does not represent any energy. Thus, $N_z = 0$, $N_s = N_a$. Let us calculate this power:

$$N_a = \frac{\omega}{\pi} \int_0^{\alpha_g/\omega} \frac{u^2}{R_a}\, dt\,, \tag{50}$$

$$N_a = \frac{U^2}{2 R_a} \frac{\alpha_g}{\pi} - \frac{U^2}{2 R_a \pi} \sin \alpha_g \cos \alpha_g\,; \tag{51}$$

$$\cos \alpha_g = \frac{U - 2 R_a \omega Q_t}{U}\,, \tag{52}$$

$$\sin \alpha_g = \frac{2 \sqrt{R_a \omega Q_t (U - \omega Q_t)}}{U}\,; \tag{53}$$

substituting the values $\cos \alpha_g$ and $\sin \alpha_g$ into Equ. (51)

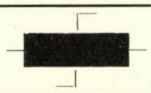
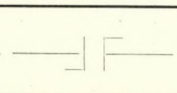
$$N_a = \frac{U^2}{2 R_a} \frac{\alpha_g}{\pi} - \frac{1}{R_a \pi} (U - 2 R_a \omega Q_t) \sqrt{R_a \omega Q_t (U - R_a \omega Q_t)} \tag{54}$$

or employing the mean voltages

$$N_a = \frac{U^2}{2 R_a} \frac{\alpha_g}{\pi} + \frac{\pi}{4 R_a} (U_{ak} - U_{zk}) \sqrt{U_{ak} U_{zk}}\,. \tag{55}$$

We summarize the results obtained for the Z-reactances in Table III.

Table III

		
U_{ak}	$\frac{2}{\pi}(U-\omega\psi_t)$	$\frac{2}{\pi}R_a\omega Q_t$
U_{zk}	$\frac{2}{\pi}\omega\psi_t$	$\frac{2}{\pi}(U-R_a\omega Q_t)$
N_a	$\frac{U^2}{2R_a}\frac{\beta_g}{\pi}-\frac{\pi}{4R_a}(U_{ak}-U_{zk})\sqrt{U_{ak}U_{zk}}$	$\frac{U^2}{2R_a}\frac{\alpha_g}{\pi}+\frac{\pi}{4R_a}(U_{ak}-U_{zk})\sqrt{U_{ak}U_{zk}}$
N_z	0	0
N_s	N_a	N_a

The same values if

$$2\psi_t > \int_0^{\pi/\omega} u\,dt \quad \text{and} \quad 2Q_t > \int_0^{\pi/\omega} i\,dt \quad \text{and} \quad \beta_g = 0\,.$$

In that case the choke is open during the whole period, the condenser comports itself during the whole period as a closed switch.

On Fig. 16 are plotted the values of

$$U_a, U_z, \alpha_g, \beta_g \quad \text{and} \quad N_a$$

in the case of the choke as well as in that of the condenser. The two Z-reactances are each other's duals. The choke interrupts the circuit until it does not reach saturation, then it closes, with the condenser it is the contrary: until reaching saturation it conducts the current, after that it behaves like an open switch, it interrupts the circuit.

Table IV

U_{ak}	0	$\frac{2}{\pi}U$
U_{zk}	$\frac{2}{\pi}U$	0
N_a	0	$\frac{U^2}{2R_a}$
N_z	0	0
N_s	0	$\frac{U^2}{2R_a}$

VII. Resonance

Though by the aforesaid we became rather well acquainted with the irregular behaviour of the *Z*-elements, we will also try to start on the investigation of the resonance from the traditional definitions and to link the new phenomena with the known ones.

The resonant state of the circuit built up with linear reactive elements is characterized by the extreme value of the current or of the voltage — maximum

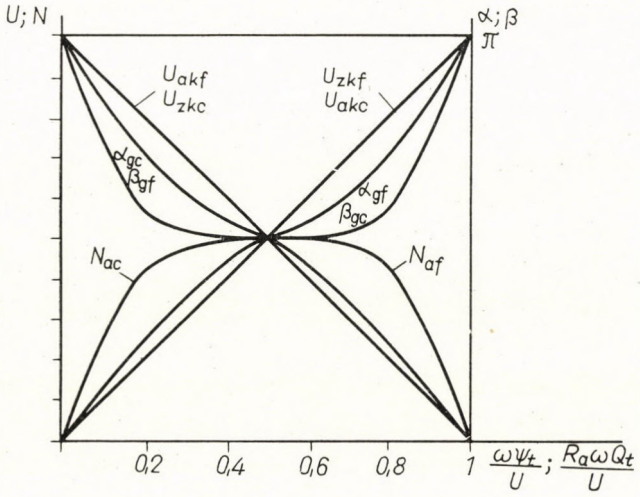


Fig. 16

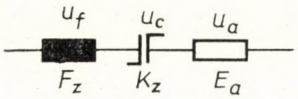


Fig. 17

or minimum — and by the cessation of the phase difference. The two characteristic cases of resonance are the series and the parallel resonance, the first is the opening, the latter the closing resonance circuit.

The closing circuit may be realized with a single *Z*-choke in the frequency range

$$\omega \geqq \frac{U}{\psi_t} \tag{56}$$

where the choke takes up the total voltage area, being $U_{ak} = 0$. So this does not happen at a determined frequency, but in a frequency range.

The opening circuit may be realized in the same way with a single *Z*-condenser in the frequency range

$$\omega \leqq \frac{U}{R_a Q_t} . \tag{57}$$

Connecting the Z-reactances in series according to Fig. 17, L_z takes up the voltage in the first part of the cycle. Following this the current flows freely as long as the condenser is not fully charged, from then on the reactance closes again (Fig. 18).

By this circuit the phase difference of the fundamental harmonic can be suppressed by placing the conductive section symmetrically in the centre. Then

$$\frac{2}{\pi}\,\omega\psi_t = \frac{2}{\pi}\,(U - R_a\,\omega Q_t - \omega\psi_t), \tag{58}$$

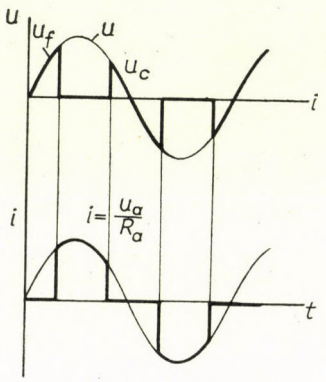


Fig. 18

from which

$$\omega = \frac{U}{2\,\psi_t + R_a\,Q_t}\,. \tag{59}$$

This state generally does not go with an extreme. We arrive at extreme values by diminishing the width of the current conducting section to zero. That occurs when $Q_t = 0$, i.e. the condenser is no longer a condenser, but simply the interruption of the circuit. In other words, by series Z-reactances the fundamental harmonics of current and voltage can be brought in phase when Equ. (59) is valid — that may be called resonance — but the extreme value representing closing cannot be realized.

Connecting the Z-reactances in parallel, at the beginning of the half cycle current is flowing through the condenser, at the end it flows through the choke (Figs 19 and 20). The symmetry condition is the equality

$$\frac{2}{\pi}\,\omega\psi_t = \frac{2}{\pi}\,(U - R_a\,\omega\,Q_t) \tag{60}$$

from which

$$\omega = \frac{U}{\psi_t + R_a Q_t} . \tag{61}$$

Here we obtain an extreme value, if the current conduction is continuous: when the condenser is charged, the choke just begins to conduct. The condition hereof is that

$$\frac{2}{\pi}(U - \omega\psi_t) + \frac{2}{\pi} R_a \omega Q_t = \frac{2}{\pi} U , \tag{62}$$

from which

$$R_a Q_t = \psi_t . \tag{63}$$

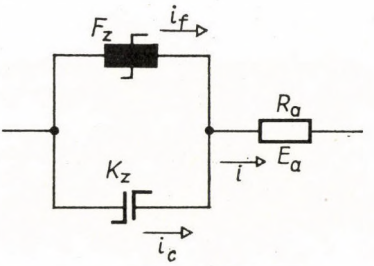


Fig. 19

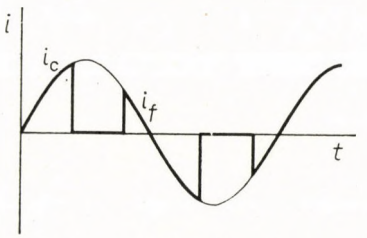


Fig. 20

In this case the circuit behaves as if it would consist of R_a alone, we obtain a completely conductive circuit. The condition (63) is, contrary to the usual resonance conditions, a stipulation independent of ω. The formal reason is that, in contrast to the traditional parameters C and L, we have worked with the values Q_t and ψ_t better fitting to the physical view. While C and Q_t are connected proportionally, L and ψ_t depend on each other reciprocally and thus ω appears with both ψ_t and Q_t as a multiplier.

REFERENCES

1. PHILIPPOW, E.: Nichtlineare Elektrotechnik. Akademische Verlagsgesellschaft (Geest & Portig K. G.), Leipzig 1963.
2. GÁDOR, L.: The Loading Capacity of Non-Linear Bridges. *Acta Techn. Hung.* **28** (1960), 281—308.

Aus Impedanzen mit Z-Kennlinien aufgebaute Stromkreise. Eine Gruppe der nicht linearen Glieder mit extremen Eigenschaften sind jene passiven Elemente der Stromkreise, zwischen deren charakteristischen Daten der Zusammenhang mittels *Z*-Kennlinien angegeben werden kann. Ihre in der Arbeit enthaltene Untersuchung erläutert qualitative Zusammenhänge, welche die Behandlung der nicht linearen Stromkreise erleichtern.

Circuits avec impédances à caractéristiques Z. Un groupe des éléments non-linéaires aux qualités extrêmes est constitué par les éléments des circuits passifs dont les données peuvent être reliées par des caractéristiques *Z*. Leur examen contenu dans l'étude met en vue des relations qualitatives facilitant le maniement des circuits non-linéaires.

Цепи с импедансами, имеющими Z-характеристику *(Л. Гадор)*. Нелинейными звеньями с крайними свойствами являются те пассивные элементы цепей, для которых среди характерных данных зависимость может быть дана при помощи Z-характеристик. Их изучение, описанное в статье, освещает такие зависимости, которые облегчают обращение с нелинейными цепями.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 177—188 (1968)

QUELQUES POINTS DE VUE POUR LE RÉGLAGE DE LA TENSION DU RÉSEAU

L. GÁDOR

CANDIDAT DES SC. TECHN.
INSTITUT DE RECHERCHES DE L'INDUSTRIE ÉLECTRIQUE, BUDAPEST

[Manuscrit présenté le 3 juillet, 1967]

La réduction par réglage des variations de tension se présentant aux différents points du réseau de distribution électrique est une exigence qualitative justifiée aussi du point de vue économique. L'énergie électrique utilisée à une tension différente de la nominale nuit à la bonne économie de l'utilisation; le dommage causé dépend de la quantité d'énergie consommée à une tension différente, en même temps que de l'écart de tension. Le critère du réglage se détermine en confrontant le dommage causé par l'écart de tension et les frais du réglage, sur la base de la statistique des variations de tension. La caractéristique de fonctionnement du régulateur type ATSZ développé par l'Institut de Recherches de l'Industrie Électrique suit la loi ainsi déterminée.

I. Introduction

Le système des centrales électriques met l'énergie électrique à la disposition des consommateurs à travers un réseau de distribution maillé, en transportant l'énergie «à domicile». L'énergie électrique doit être mise à disposition sur le lieu de consommation dans la quantité et la qualité voulues, pour satisfaire au mieux les exigences des consommateurs. Quant à l'exigence qualitative, celle-ci est satisfaite par le maintien à la valeur nominale des caractéristiques prescrites de l'énergie électrique, la fréquence et la tension.

En ce qui concerne la fréquence, le système complexe de la distribution électrique forme une unité à connexions rigides, ce qui permet de résoudre le maintien de la fréquence dans les centrales, lieux de production de l'énergie électrique, d'une façon appropriée et valable pour l'ensemble du réseau.

Du point de vue de la tension, le réseau chargé montre une image très inégale à cause des variations de tension le long des lignes; aux lieux de consommation les tensions ne sont, en effet, ni conformes à la valeur nominale ni uniformes, et cela même si ces valeurs nominales et cette uniformité à les tensions ont été préalablement assurées dans les centrales. Le but proposé n'est donc pas atteint en assurant par réglage, dans les centrales, la tension nominale ou quelque tension uniforme supérieure à la nominale, qui tiendrait compte de la chute de tension moyenne dans les lignes.

Le réglage à un niveau uniforme de la tension fournie par les centrales ne servirait pas non plus les intérêts de la répartition correcte des charges.

Les points d'alimentation, les points de consommation, le réseau qui les relie, les variations relativement irrégulières des charges aux différents points,

la variation de l'alimentation qui s'y ajuste de quelque façon forment ainsi un système complexe, où les courbes de niveau de la carte de distribution de la tension ne se situent que par endroits dans le plan de la tension nominale et passent ailleurs tantôt au-dessus, tantôt au-dessous de ce plan.

L'approvisionnement idéal en énergie de tension nominale consisterait à placer, à côté de chaque consommateur, un régulateur de tension. Il est évident qu'une telle décentralisation du réglage n'est pas économique et augmenterait outre mesure les frais de l'énergie électrique.

Dans la pratique, on approche le but désiré d'une façon plus économique en prenant les ondulations de la carte des tensions dans quelques points bien choisis et en les réglant pour une tension constante, ou mieux encore en portant les points fixés à des niveaux adaptés aux conditions de charge existant dans leur voisinage: à une tension plus élevée si la charge est plus forte et la chute de tension est plus grande sur les lignes conduisant aux consommateurs, et à une tension plus petite dans les périodes à charge moins élevée.

Nous venons de rejeter le réglage idéal complètement décentralisé de la tension, comme n'étant évidemment pas économique. Mais dans les considérations techniques et économiques, l'apparente évidence d'une constatation dictée par le simple bon sens ne saurait, bien entendu, remplacer le calcul. Peut-être serait-il tout de même rationnel de construire un régulateur auprès de chaque consommateur, ou le réglage serait-il superflu même dans les points choisis comme caractéristiques en raison de son manque d'économie. Il est en effet possible que la variation de la tension cause des dommages inférieurs au coût du réglage.

II. Effets économiques du réglage de la tension

Les conditions économiques des installations travaillant à une tension différente de la nominale ont été étudiées par M. T. Ailleret [1], dont nous résumons ici les idées en vue d'utiliser ses résultats comme point de départ. La qualité économique G d'une installation peut être considérée comme une différence, où la minuende est le service fourni par l'installation, tandis que le terme soustractif correspond d'une part à l'amortissement tenant compte des frais d'investissement, et d'autre part aux frais d'exploitation.

Au cas d'une installation électrique

$$G = G(s_1, s_2, s_3, \ldots, U), \tag{1}$$

où s_1, s_2, s_3 sont les caractéristiques techniques de l'installation, U étant la tension de service. L'installation est dimensionné pour un but et une tension donnés et en cas de solution satisfaisante, c'est avec ces données s_1, s_2, s_3, ..., U que l'installation fonctionnera le plus économiquement.

Il s'ensuit que

$$\begin{aligned} &\frac{\delta}{\delta s_1} G(s_1, s_2, s_3, \ldots, U) = 0\,, \\ &\frac{\delta}{\delta s_2} G(s_1, s_2, s_3, \ldots, U) = 0\,, \\ &\qquad \ldots \\ &\frac{\delta}{\delta U} G(s_1, s_2, s_3, \ldots, U) = 0\,. \end{aligned} \tag{2}$$

Supposons maintenant que l'installation fonctionne à la tension $U + \Delta U$ différente de la tension nominale U. A cette tension différente, l'économie de l'installation diminue et cette diminution peut s'écrire sous la forme

$$\Delta G = G(s_1, s_2, s_3, \ldots, U + \Delta U) - G(s_1, s_2, s_3, \ldots, U)\,. \tag{3}$$

En développant en série jusqu'au deuxième terme compris:

$$\begin{aligned} \Delta G = \Delta U \frac{\delta}{\delta U} G(s_1, s_2, s_3, \ldots, U) + \\ + \frac{1}{2} (\Delta U)^2 \frac{\delta^2}{\delta U^2} G(s_1, s_2, s_3, \ldots, U)\,. \end{aligned} \tag{4}$$

En vertu du système d'équations (2), le premier terme de la série est zéro, donc

$$\Delta G = \frac{1}{2} (\Delta U)^2 \frac{\delta^2}{\delta U^2} G(s_1, s_2, s_3, \ldots, U)\,, \tag{5}$$

ce qui signifie qu'en négligeant les termes d'ordre supérieur on aura à compter, en cas d'emploi d'une tension s'écartant de la nominale par ΔU, avec une diminution de l'économie proportionnelle à $(\Delta U)^2$.

A charges égales, un dommage à peu près égal est causé par l'élévation de 9 à 10% de l'écart de tension et l'exploitation à une tension différant de 4% de la nominale.

Le dommage causé par l'écart de tension dépend naturellement de la quantité des installations en service. Bien que le dommage apparaissant chez le consommateur dépende aussi du caractère de l'installation, son importance différant, par exemple, en cas d'une lampe à incandescence et d'un moteur de ventilateur, du point de vue statistique on ne commet pas d'erreur en le considérant simplement comme étant proportionnel à l'énergie consommée à la tension différente de la nominale.

III. Critère du réglage de la tension

Ce sont donc la tension et la charge variant stochastiquement qui déterminent ensemble les conditions de l'exploitation économique. En traitant de la détermination du critère de fonctionnement du régulateur, supposons pour le moment que la charge est constante, par exemple constamment nominale. Dans ce cas, le dommage causé par l'écart de tension est simplement proportionnel au carré de l'écart de tension et au temps; sa valeur pendant le temps T est

$$\Delta G = a \int_0^T (\Delta U)^2 \, dt, \tag{6}$$

où a est le facteur de proportionnalité.

Passons maintenant à des valeurs relatives

$$u = \frac{\Delta U}{U_n}, \tag{7}$$

où U_n est la tension nominale. Pour plus de simplicité, écrivons g au lieu de ΔG. Ainsi

$$g = a \int_0^T u^2 \, dt. \tag{8}$$

Pour le réglage, un régulateur à gradins est utilisé à peu près sans exception dans les réseaux. Les frais de réglage se composent de l'amortissement et des frais d'exploitation, resp. des pertes du régulateur; ils contiennent les frais d'investissement et d'entretien et les frais de fonctionnement directs.

L'amortissement et les frais d'exploitation sont reliés au fonctionnement. En effet, le régulateur n'effectuant pas de commutations est d'une durée pratiquement illimitée. Aussi les frais de réglage apparaissent-ils dans les frais des commutations.

Désignons par b les frais d'une commutation. L'établissement et l'exploitation du régulateur sont économiques si le dommage dû aux écarts de tension est réduit, grâce au réglage, d'une somme supérieure aux frais de réglage.

Soient u_e l'écart de tension relatif à l'entrée du régulateur, c'est-à-dire au côté non réglé et u_s l'écart relatif à la sortie, au côté réglé (fig. 1).

Le réglage diminuant l'écart de tension, la réduction du dommage causé par cet écart sera, pendant le temps T:

$$a \int_0^T u_e^2 \, dt - a \int_0^T u_s^2 \, dt = a \int_0^T (u_e^2 - u_s^2) \, dt. \tag{9}$$

Supposons que pendant ce temps, le régulateur devait commuter k_T fois pour réaliser u_s caractérisé par un écart plus faible. Le réglage est économique si

$$k_T\, b < a \int_0^T (u_e^2 - u_s^2)\, dt. \tag{10}$$

La question est de savoir comment il faut ajuster le régulateur pour que le nombre des commutations ne soit pas plus élevé que celui défini par l'inégalité (10).

Considérons un écart de tension ne durant qu'un temps relativement court $(t_2 - t_1)$, suivant la figure 2. Si pour cette impulsion d'écart de tension

$$a \int_{t_1}^{t_2} u^2\, dt < b\,, \tag{11}$$

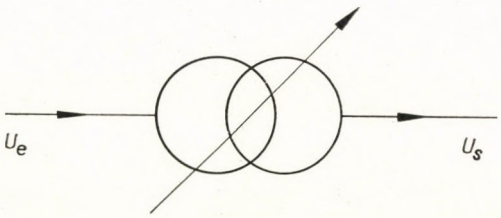


Fig. 1

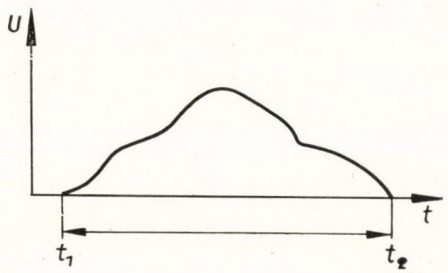


Fig. 2

alors une seule commutation coûte déjà plus que le dommage dû à l'écart de tension. Ceci nécessite évidemment l'utilisation d'une retardation, le régulateur ne devant pas fonctionner pour l'écart de courte durée de la figure 2. Ajoutons que la retardation n'est pas simplement un intervalle défini et variable, c'est le temps pendant lequel l'intégrale du côté droit de l'équation (8) atteint, à compter du commencement de l'écart, une valeur constante définie, par exemple la valeur b.

S'il n'en était pas ainsi, des deux impulsions de tension de la figure 3, de valeur égale

$$a \int u^2\, dt < b\,,$$

la première ne déclencherait pas, mais la deuxième déclencherait le fonctionnement du régulateur retardé de t_0, qui n'est pourtant pas justifié du point de vue de la rentabilité, ni dans l'un ni dans l'autre cas.

Est-il permis de commuter en chaque cas, si

$$a \int_{t_1}^{t_2} u^2 \, dt < b.$$

Soit l'écart de tension suivant la figure 4 tel que

$$a \int_{t_1}^{t_2} u^2 \, dt = b \tag{12}$$

et supposons que par la commutation, le régulateur supprime complètement l'écart. S'il n'y avait pas de réglage, le dommage serait

$$a \int_{t_2}^{t_3} u^2 \, dt .$$

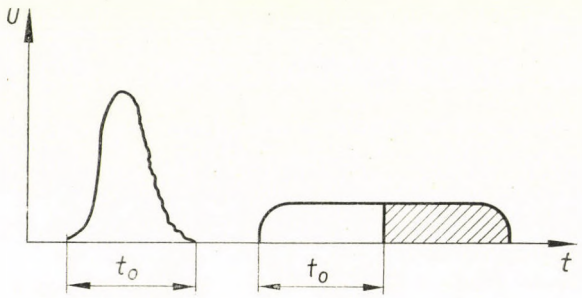


Fig. 3

Par suite du réglage fait au temps t_2, la surface de dommage hachurée est éliminée, mais les frais de la commutation apparaissent. La commutation est donc rentable si

$$a \int_{t_1}^{t_2} u^2 \, dt + b < a \int_{t_1}^{t_3} u^2 \, dt ,$$

ou

$$b < a \int_{t_2}^{t_3} u^2 \, dt . \tag{13}$$

Ce qui importe donc, ce n'est pas de connaître l'écart de tension préalable et le dommage qui en résulte, mais de déterminer le dommage que l'on subirait après le moment prévu pour la commutation, si celle-ci n'avait pas lieu.

Notre élément détecteur doit donc prévoir l'avenir, du moment qu'on prescrit comme condition première l'économie univoque du fonctionnement.

Nos installations techniques connaissent deux formes de prévision, à savoir l'extrapolation et la statistique. L'extrapolation elle-même est, en principe, une décision statistique, car c'est à partir de la tendance du processus,

du quotient différentiel qu'elle établit la probabilité d'une grandeur se précisant dans l'avenir. Les variations de tension des réseaux de distribution sont en général des processus lents, la différentielle est en général trop petite pour pouvoir être facilement utilisée pour le réglage. Néanmoins, un critère de réglage tenant compte de la différentielle est également imaginable, mais la prévision par la statistique promet de meilleurs résultats.

Supposons que les dommages causés par les écarts de tension forment un ensemble de répartition normale. Si en général

$$g_i = a \int_{t_i}^{t'_i} u^2 \, dt \tag{14}$$

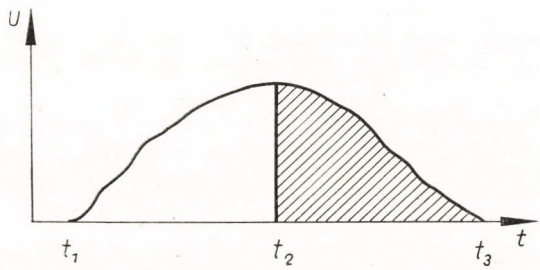


Fig. 4

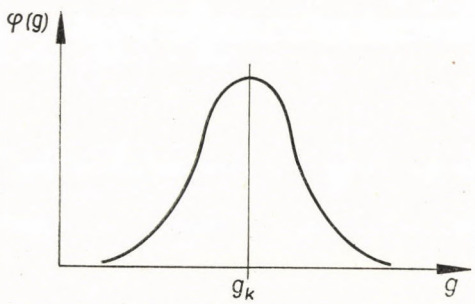


Fig. 5

est l'écart se présentant au $i^{\text{ième}}$ cas durant un temps $(t'_i - t_i)$ et le nombre n de tous les écarts examinés est assez grand, la probabilité de l'occurrence d'un g donné est

$$\varphi(g) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(g-g_k)^2}{2\sigma^2}}, \tag{15}$$

où g_k est la valeur moyenne et en même temps la plus fréquente de g (fig. 5) et σ l'écart quadratique moyen.

Retenons toujours l'hypothèse simplificatrice que le régulateur supprime l'écart de tension en un seul pas. Si nous choisissons les conditions du réglage

d'une façon telle que celui-ci n'intervienne pas même à une valeur élevée de g, alors le réseau fonctionne sans réglage, la totalité du dommage dû à l'écart se manifeste avec la même valeur que si le dommage moyen apparaissait à chaque occasion:

$$G' = ng_k. \tag{16}$$

En réduisant progressivement g, fixé comme condition du fonctionnement, le dommage G causé par l'écart de tension diminue et le nombre des commutations augmente (fig. 6). Dans le cas pris comme exemple, $g_k = 40$, $n = 1000$. Sur la figure, nous voyons la valeur de l'économie $G_{max} - G$ rapportée à une commutation, valeur qui croît avec le nombre des commutations.

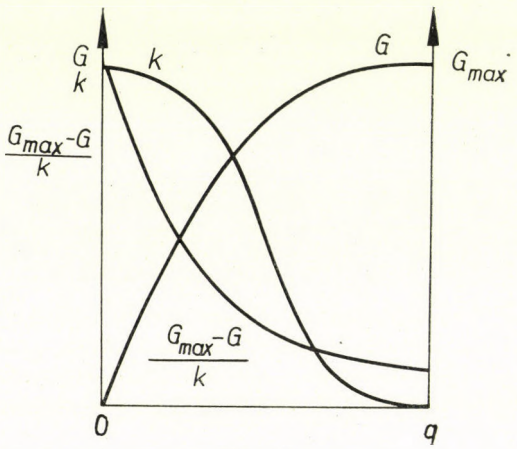


Fig. 6

Donc, s'il est économique de régler un ensemble d'écarts de tension donné, le mieux sera de régler chaque écart individuel avec un retard minimum. En ce cas, l'économie réalisable sera justement g_k par commutation. Rappelons que tout ceci n'est valable qu'à condition que le régulateur supprime complètement l'écart par une commutation unique et que la retardation ne soit jamais comprise simplement comme un temps, mais comme une durée dépendant de la tension et calculée à partir de $\int u^2 dt$.

La condition de la suppression par réglage de l'ensemble des écarts peut être formulée très simplement, en disant que le réglage est économique si

$$b < g_k. \tag{17}$$

Dans le réseau, nous pouvons compter en général avec deux ensembles indépendants, de répartition à peu près normale: l'ensemble des écarts transitoires de courte durée et la variation journalière régulière de la tension. Pour les

écarts courts et transitoires, la condition (17) n'est probablement pas satisfaite, c'est pourquoi il faut déterminer la condition du réglage d'une façon telle que la retardation choisie laisse ces écarts sans réglage et élimine seulement les variations lentes de la tension. Mais pour résoudre le réglage d'un réseau donné réel, il faut faire l'analyse statistique des variations de tension du réseau en question. Après des changements importants opérés dans le réseau, il faut répéter cette analyse.

IV. Réglage par gradins

En effectuant nos examens, nous avons jusqu'à présent admis la condition simplificatrice que le fonctionnement du régulateur supprime d'un seul coup l'écart de tension. Examinons maintenant le rôle de la gradation.

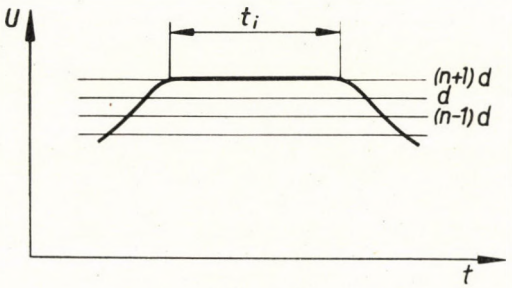


Fig. 7

Soit d la valeur relative des gradins de réglage. S'il existe un écart de tension, la tension s'arrête de temps à autre au niveau $(n + 1)\, d$ (fig. 7). Les durées t_i des écarts de grandeur $(n + 1)\, d$ peuvent, elles aussi, être considérées avec une bonne approximation comme ayant une répartition gaussienne.

Si la tension dévie de la nominale pendant un temps t_i d'une valeur correspondant au $(n + 1)^{\text{ième}}$ gradin, le dommage causé par l'écart de tension est

$$g_{i(n+1)} = a\,(n + 1)^2\, d^2\, t_i\,. \tag{18}$$

La valeur moyenne et en même temps la plus probable de $g_{i(n+1)}$ est

$$g_{k(n+1)} = a\,(n + 1)^2\, d^2\, t_k\,, \tag{19}$$

où t_k est la valeur moyenne et en même temps la plus probable de la durée de l'écart $(n + 1)\, d$.

Quand le réglage fonctionne, il rapproche d'un degré la tension de sa valeur nominale; à ce niveau le dommage causé par l'écart de tension est

$$g_{in} = an^2 d^2 t_i . \tag{20}$$

L'économie réalisée par le réglage effectué au commencement du temps t_i est

$$\begin{aligned} g_{i(n+1)} - g_{in} &= a(n+1)^2 d^2 t_i - a n^2 d^2 t_i \\ &= a(2n+1) d^2 t_i . \end{aligned} \tag{21}$$

Le réglage est rentable si les frais de réglage sont plus petits que l'économie

$$b < a(2n+1) d^2 t_i . \tag{22}$$

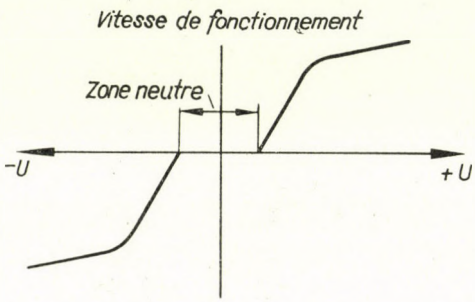


Fig. 8

En appliquant les idées énoncées au paragraphe III en rapport avec l'équation (17), nous pouvons dire que le réglage — et celui de chaque écart de tension — est rentable si

$$b < a(2n+1) d^2 t_k . \tag{23}$$

Dans cette relation t_k lui-même est soumis aussi aux lois de la répartition normale, les faibles écarts de tension subsistent pendant un temps plus long que les grands écarts. Dans le domaine des faibles écarts, on se trouve près du sommet de la courbe de répartition de Gauss et à quelques pourcents d'écart près, les t_k peuvent être considérés comme pratiquement égaux. Dans ce domaine — d aussi étant constant — le côté droit de l'inégalité croîtrait en proportion avec n. L'économie exige qu'on ne s'éloigne pas beaucoup de l'égalité, aussi faut-il diminuer le facteur de proportionnalité a dans la même mesure. Cela signifie l'augmentation de la vitesse de fonctionnement du régulateur et à conditions d'ailleurs inchangées, la diminution des intervalles entre commutations successives.

En cas d'écarts de tension plus considérables, la diminution de t_k compense l'augmentation de n, il n'y a donc aucune utilité à augmenter davantage la vitesse du fonctionnement. Il est donc recommandé de donner la forme de la figure 8 à la caractéristique de fonctionnement du régulateur. La zone neutre

assure l'absence des oscillations dans le fonctionnement du régulateur à gradins et la section à montée uniforme caractérise le domaine des faibles écarts, tandis qu'aux grands écarts, elle passe au fonctionnement à vitesse constante.

La figure 9 montre la caractéristique effectivement relevée du régulateur type ATSZ—3 développé par l'*Institut de Recherches de l'Industrie Électrique*, qui s'accorde bien avec l'allure de la courbe de la figure 8. En ajustant la vitesse de fonctionnement, il est possible de rechercher le mode de fonctionnement le plus favorable, par l'analyse statistique des conditions de tension dans le réseau à régler.

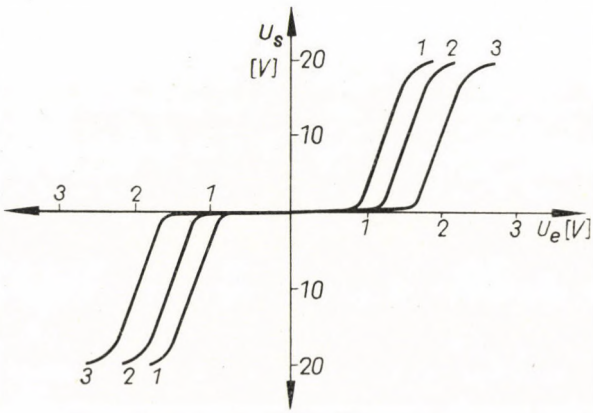


Fig. 9

Au cours des développements qui précèdent, nous n'avons pas tenu compte du fait qu'en dehors des écarts de tension, l'économie dépend aussi des charges. Cela fait entrer simplement un nouveau facteur variable dans la relation, facteur dont la prise en compte ne pose d'ailleurs aucune difficulté lors de la réalisation pratique. Cependant, sa considération totale ne serait pas utile, car elle signifierait qu'on abandonne complètement le réglage du réseau fonctionnant à vide. Si ce n'est pas ce qu'on veut, la charge ne doit être appliquée que sous la forme d'une correction modérée.

BIBLIOGRAPHIE

1. AILLERET, P.: L'introduction de la mesure dans la notion de la qualité du service d'une distribution d'électricité et son intérêt du point de vue de la conception des réseaux. *Bull. SFE* (1956).
2. LANGER, H.: Regelungstechnischer Einsatz des Transformators. *ETZ-A* **81** (1960).
3. GAUSSENS, P.—GAZALET, P.: Le réglage de la tension dans les réseaux de distribution. *Bull. SFE* (1960).
4. *Баркан, Я. Д.:* Некоторые вопросы автоматизации регулирования напряжения. *Электричество* (1963).
5. GÁDOR, L.: Elosztó hálózati feszültségszabályozás (Réglage des écarts de tension des réseaux de distribution). *Villamosság* (1964).

Some Points of View for the Regulation of the Mains Voltage. In electrical distribution networks it is a condition of quality and is also economically justified to reduce by regulation the voltage fluctuations in the different points of the network. The electrical energy consumed at voltages differing from the normal one is detrimental to economic use; the damage depends on the quantity of energy consumed at out-of-normal voltage and on the voltage deviation. The criterion for the voltage regulation can be determined by opposing, on the basis of the statistics of voltage variations, the damage caused by the voltage deviation and the costs of voltage regulation. The characteristic of the type ATSZ regulator developed by the *Research Institute of Electrical Industry* is conform to the laws determined in this way.

Einige Gesichtspunkte für die Regelung der Netzspannung. In den elektrischen Energieverteilungsnetzen ist es eine Qualitätsbedingung und auch ökonomisch begründet, das Ausmaß der in den verschiedenen Netzpunkten auftretenden Spannungsschwankungen durch Regelung zu verringern. Die mit einer von der Netzspannung abweichenden Spannung verbrauchte elektrische Energie ruiniert die Wirtschaftlichkeit des Verbrauchs; der verursachte Schaden hängt von der Menge der mit abweichender Spannung verbrauchten Energie und der Spannungsabweichung ab. Das Kriterium für die Regelung kann durch Gegenüberstellung des von der Spannungsabweichung verursachten Schadens und der Kosten für die Regelung bestimmt werden, auf Grund der Statistik der Spannungsschwankungen. Die Kennlinie des vom Forschungsinstitut der Elektroindustrie entwickelten Reglers Typ ATSZ folgt der so bestimmten Gesetzmäßigkeit.

Некоторые принципы регулирования сетевого напряжения *(Л. Гадор)*. Для электрораспределительных сетей в качестве качественного требования ставится задача уменьшения масштаба колебания напряжения, наблюдающегося в различных точках сети, что обоснованно также с экономической точки зрения. Электроэнергия, расходованная при напряжении, отклоняющемся от номинала, ухудшает экономичность потребления; размер ущерба зависит от количества энергии, расходованного при напряжении, отклоняющемся от номинала, и масштаба отклонения напряжения от номинала. Критерий регулирования можно определить путем сопоставления ущерба от отклонения напряжения от номинала и расходов регулирования на основе статистики колебания напряжения. Кривая характеристики работы регулятора, разработанного в Исследовательском институте электротехнической промышленности, типа *ATC* следует определенной таким образом закономерности.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 189—200 (1968)

THE EFFECT OF QUINONE ON THE SURFACE PROPERTIES OF GERMANIUM

A. LŐRINCZY – M. NÉMETH-SALLAY

RESEARCH INSTITUTE FOR TECHNICAL PHYSICS
OF THE HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST

and

I. C. SZÉP — J. TIHANYI

INDUSTRIAL RESEARCH INSTITUTE FOR ELECTRONICS "HIKI", BUDAPEST

Experiments were made on $p—n—p$ germanium transistors and examples to study the effect of quinone. The dependence of reverse-current and current-gain on the surface potential of transistors was investigated by means of the emitter and collector field effect. After quinone treatment the reverse-current characteristics improved, while the current-leakage, noise-factor and instability were reduced. The surface potential was shifted towards the positive. For quantitative checking of the physical properties of the surface, germanium samples were examined by field-effect measurements. Based on results of two different measurements suggestions are given relating to the processes taking place on the surface.

I. Introduction

It is well known that the presence of water on the surface of a semiconductor influences the electrical properties of the $p—n$ junctions [1, 2]. According to experience in the case of germanium semiconductor devices this influence can be favourable within a certain range of pressure (e.g. it increases the current-gain of $p—n—p$ transistors). Beyond a certain water concentration, however, its effect is harmful. The reverse current of the devices increases and becomes unstable, the break-down voltage decreases, when the partial pressure of water-vapour exceeds a few 10 Torr. At low reverse voltage, where mainly the generation of reverse current plays an important role, the water influences the reverse current through the surface recombination velocity (at the same time influencing the current-gain of the transistors, too) [3]. At higher reverse voltages a considerable increase of reverse current as well as instability and creep of the characteristics can be observed [4]. This so-called "leakage" current is the result of a very complicated mechanism. Although the physical origin of this phenomenon is not well understood so far, it is, however, obvious that its occurrence is mainly caused by the presence of water and by ion migration on the surface [5].

From the point of view of the reliability of the devices, for reasons mentioned above, it would seem to be of great importance to remove the bound water from the surface and to prevent the adsorption of water. The latter can be achieved by perfect hermetic sealing, but the removal of the adsorbed water raises more complicated problems. The ideal solution of the problem would

be the elimination of water to such an extent that the instability and its harmful effect would cease without causing substantial changes in the parameters of the semiconductor device.

The problem cannot be solved by simple drying. First, even the highest admissible temperature for Ge-devices is not sufficient to remove the water [6], second, on drying other parameters become impaired e.g. the current-gain of the transistors. In [7] quinone has been suggested for the adjustment of the favourable amount of water, assuming that the potential of the quinone-hydroquinone redox-system is established on the surface by using up water and producing oxygen according to the following reaction:

$$C_6H_4O_2 + H_2O = C_6H_4(OH)_2 + O.$$

According to this suggestion the favourable effect is caused by the joint action of residual water and oxygen, and the redoxpotential.

On the basis of above suggestion and of the data available [8] it was decided to make a detailed study of this process in order to clarify its mechanism. The influence of quinone was recently investigated by LYUZE and others [9] who applied an alcoholic solution of quinone to the germanium surface. According to their results, on quinone treatment on *n*-type samples *p*-type surface layers (inversion layer) developed, while the recombination velocity was found to be very low (5—50 cm/s), leading to an increase of the current-gain of transistors. Using fresh solution a change of 15—20% was obtained while with thermally aged solutions a change of 100% was observed.

The action of the quinone-hydroquinone system on *p*-type silicon samples was investigated by A. R. MOORE and H. NELSON, using its aqueous solution [8]. They found an increase of the effective life-time and a decrease of the recombination velocity. According to their interpretation the accumulation of holes in the surface layer prevents the minority charge carriers from reaching the surface, thereby changing the surface recombination velocity.

Several experiments were performed to adjust the favourable water content artificially, e.g. using materials, which establish a constant water vapour pressure (BaO, $CaSO_4$ and B_2O_3) [10]. The data available were found to yield insufficient information and the experimental methods used on transistors were too indirect for the study of the surface phenomena. Therefore, investigations of more exact nature were set as an aim.

II. Experimental Procedure

1. Preparation

It was considered necessary to determine the surface characteristics of germanium and germanium transistors, and to clarify the relationship between the electrical data of the transistors and the surface parameters, respectively.

The semiconductor devices used for the experiments were uncapsulated alloyed audiofrequency transistors, the size of n-type germanium samples was $1{,}5 \times 0{,}5 \times 0{,}05$ cm, with 20 ohmcm specific resistivity and [111] orientation. After polishing, the accuracy of the parallelism of the plates was within 1 μ, the end faces were provided with alloyed tin contacts, doped with antimony. (According to our measurements these contacts were ohmic even at the temperature of liquid nitrogen.) The transistors were chemically treated by alkaline anodic etching and subsequent washing and drying.

The mechanically damaged surface layer of the germanium samples was removed by etching in alkaline 30% H_2O_2 (100 cm^3 30% H_2O_2 + 1 cm^3 10% NaOH sol.) then rinsed in bidistilled water, subsequently in alcohol and afterwards dried.

The samples to be measured were mounted in a vacuum system under a pressure of 10^{-4} Hgmm, where they were dryed for abt. 1 hour at 150 °C. After the heat-treatment the characteristics of the initial state were measured at room temperature, later on — to eliminate any interference of a solvent — quinone in vapour form was introduced into the closed container, which was previously disconnected from the pump. The equilibrium state on the surface was obtained within abt. 1 minute. The quinone used was crystalline, of reagent grade.

2. *Methods of measurements*

The surface processes taking place on transistors can be examined by measuring two parameters: a, the reverse current; b, the current-gain.

The reverse current has a surface and a bulk component. For germanium $p-n$ junctions the surface component is generally dominating and the contribution of the bulk can be neglected. One part of the surface reverse current is caused by surface generation. The latter is determined by physical parameters of generation-recombination centers and by the surface potential [11]. The other part, developing only at a suitable value S of surface potential is the so called "channel" current associated with the surface inversion layer. Thus the surface reverse current depends on the surface potential, therefore, the reverse current can be changed by varying the surface potential. On the other hand, the surface potential can also be influenced by changing the voltage of an insulated field electrode placed onto the surface [12]. In our measurements the field electrode was placed in the vicinity of the reverse biased $p-n$ junction and the change of reverse current vs. surface field-electrode voltage was shown on the screen of an oscilloscope. From the oscilloscope curves we were able to evaluate the relation between the surface potential and the surface recombination velocity [13]. The current-gain depends on the surface recombination, thus, indirectly on the surface potential as well. Changing the voltage of the field-electrode placed in the vicinity of the emitter

$p-n$ junction, the collector current changes corresponding to the change of h_{fe}. Interpreting the oscilloscope pattern obtained by plotting the collector current vs. the field electrode voltage, qualitative information on the recombination velocity was obtained.

For quantitative examinations a large signal sinusoidal field-effect with photoconductivity modulation was used. Fig. 1 shows the design of the sample-holder. A thin mica plate was placed onto the polished brass plate, then the germanium sample, then again a mica plate, on top of which a glass plate was placed coated by a transparent gold layer. The whole system was clamped together by a spring, serving at the same time as electrical contact, connecting

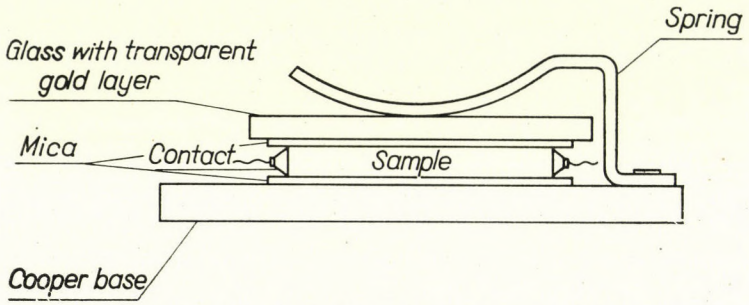


Fig. 1. Sample-holder with germanium slab

in parallel the plate capacitors. Their capacity was up to 200 pF, and a field voltage of 150 V maximal AC amplitude, 70 Hz frequency was applied to the sample.

3. Experimental results

The samples were examined by the reverse current and h_{fe} field effect methods as described above, before and after quinone treatment.

Fig. 2 shows a typical reverse current vs. field effect curve, whereas on Fig. 3 the typical h_{fe} vs. field voltage curve is shown. Fig. 3a refers to the state before the quinone treatment, Fig. 3b to that after the treatment.

The change of patterns, obtained by plotting the reverse current vs. field-effect may be interpreted according to Fig. 4. The curve, showing the resulting reverse current vs. surface potential consists of the generation and inversion reverse current. On Fig. 4 the ideal curve $I_{rev}(\Phi_s) = I_{rev}(V_{field})$ composed of the assumed I_s generation and I_{ch} inversion channel current is shown. The resultant is the solid line. In practice we were not able to cover such a broad Φ_s region, therefore, only a part of the curve is visible on the screen (curves A and B, which agree with the curves of Fig. 2a and Fig. 2b).

Comparing the reverse current-field effect curves, obtained both theoretically and experimentally the following conclusion could be drawn:

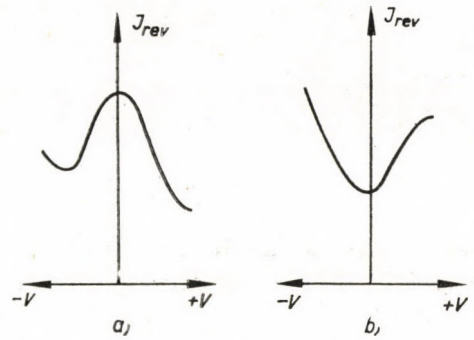


Fig. 2. Reverse current field effect curves: collector reverse current vs. field electrode voltage (on the collector): original surface (*a*), after quinone treatment (*b*)

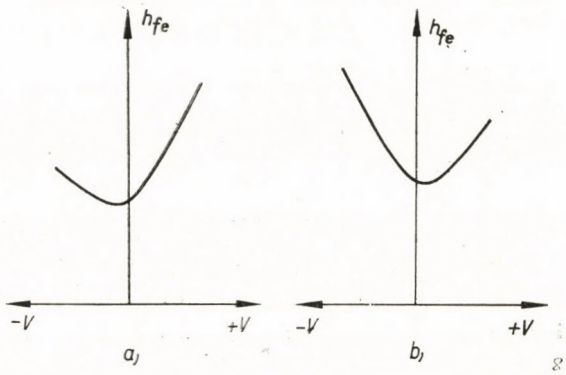


Fig. 3. h_{fe} field effect; collector current with grounded emitter at constant base current vs. field electrode (on the emitter) voltage: original surface (*a*), after quinone treatment (*b*)

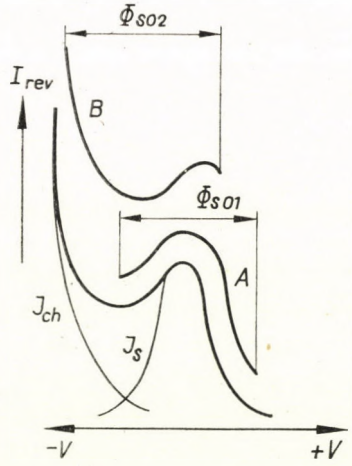


Fig. 4. The resulting reverse current is composed of the I_{ch} channel current and the I_s generation current; the experimental curves *A* and *B* might be fitted on the theoretical curve

a) The value of Φ_{s01} (surface potential in absence of field) before the quinone treatment has almost the same value as the maximum Φ_s of the generation current-surface potential curve;

b) as a result of the quinone treatment Φ_{s02} is shifted towards positive values.

Consequently, in the presence of quinone a negative charge appears on the surface, equivalent to the effect of a negative field electrode voltage. This negative charge causes an increase in the concentration of positive carriers in the vicinity of the surface of the semiconductor. The change, observed on

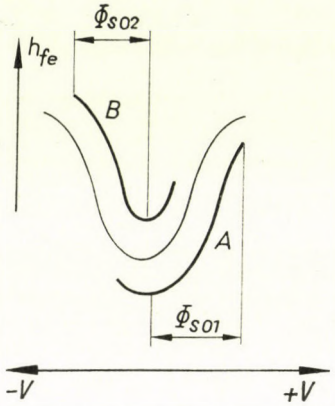


Fig. 5. Comparison between the ideal curve h_{fe} and the experimental curves (AB)

the h_{fe} field effect curves is in complete agreement with the conclusion drawn from the reverse current-field effect curves. Namely, the curve $h_{fe}(\Phi_s) \sim$ $\sim h_{fe}(V_{field})$ is appr. inversely proportional to the surface recombination velocity vs. surface potential function, i.e. the minimum current gain appears at the maximum of the recombination velocity (Fig. 5). On Fig. 5 the ideal curve $h_{fe}(\Phi_s) \sim h_{fe}(V_{field})$ has been plotted in arbitrary units. The place where the experimental field effect curves A and B should be fitted are indicated. Similar to Fig. 4 the displacement of Φ_{s0}, i.e. the presence of negative surface charge can be observed as well.

After the quinone treatment, the reverse current characteristics improved; the leakage current, the low frequency noise and instability diminished. Fig. 6 shows a typical change of reverse characteristics. Although the channel current appearing after the quinone treatment was higher than the low voltage current present before the quinone treatment, the stability considerably increased. Though both measurements confirm, that negative charge builds up on the surface, the changes due to quinone treatment could not be determined quantitatively. For determining the surface characteristics, measurements were made on germanium samples: the method used is described in [14].

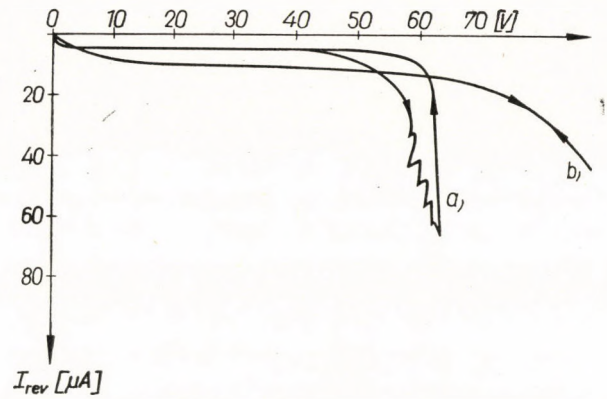


Fig. 6. Typical reverse characteristic change on quinone treatment at the germanium $p—n—p$ transistor collector-basis junction

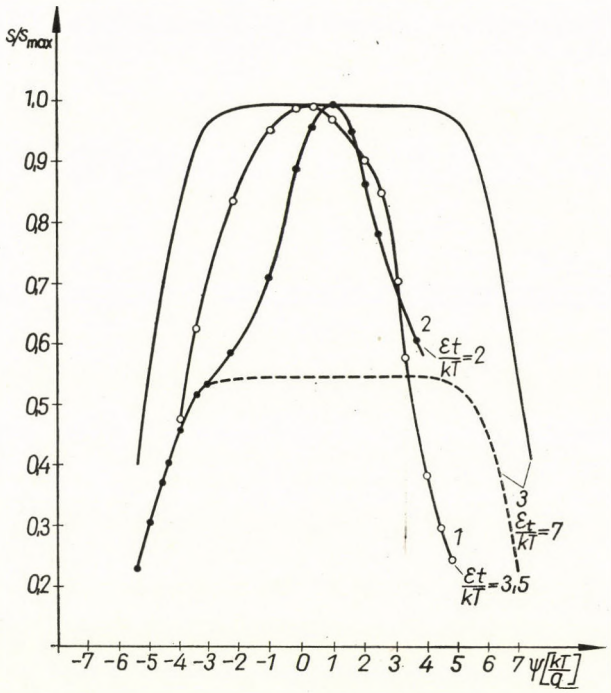


Fig. 7. Normalized surface recombination velocity (S/S_{max}) vs. surface potential: original surface (1), after quinone treatment (2), the renormalized curve of the new level (3)

The results of these measurements are shown in Fig. 7 and Fig. 8. Fig. 7 shows the curve obtained by plotting the normalized surface recombination velocity against the surface potential, while Fig. 8 shows the dependence of the concentration of charges, trapped by fast surface states, on the surface potential. Both figures show the original curve and that obtained after the quinone treatment.

On Fig. 7 curve *1* is that before the quinone treatment, curve *2* that after the treatment, while *3* (solid line) is the renormalized surface recombination curve of newly formed levels. The measurements prove that the condition of a single recombination level, resp. that of well-separated multiple-levels, assumed in [14] can be applied.

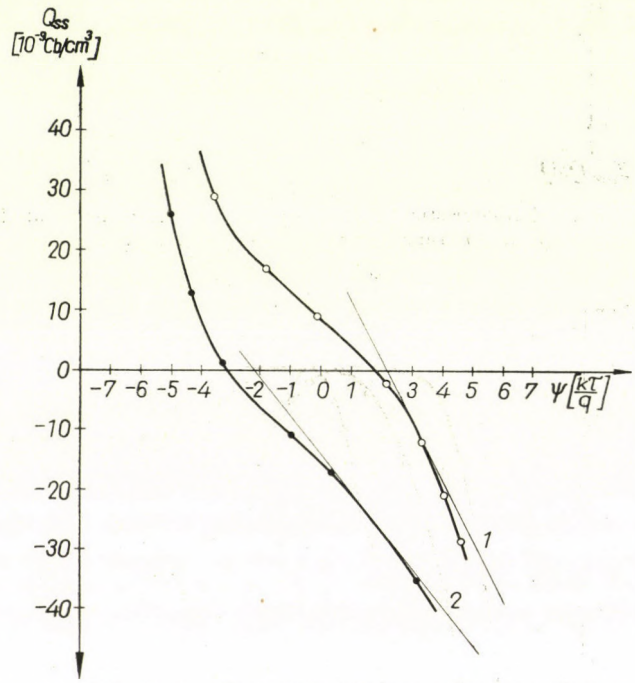


Fig. 8. Charges captured by fast surface states (Q_{ss}) as a function of the surface potential: original surface (1), after quinone treatment (2)

The surface parameters and their changes, determined from our measurements, are as follows:

a) due to the quinone treatment the surface recombination velocity generally diminished by abt. 50% (in case of the samples shown on the figures a decrease from 1000 cm/s to 500 cm/s is observed);

b) in the presence of quinone the maximum of the recombination curves shift towards more positive values. On the basis of these displacements the ratio of capture probabilities decreased from $\alpha_p/\alpha_n = 7$ to 1,5;

c) instead of the original surface level located at 3,5 kT, having a concentration of $2{,}4 \cdot 10^{11}$ cm^{-2}, two new levels appeared: one located at 2 kT, with concentration of $1{,}5 \cdot 10^{11}$ cm^{-2}, the other located at 7 kT with concentration of $2{,}5 \cdot 10^{11}$ cm^{-2} (the energy of the level is calculated from the center of the forbidden gap);

d) the value of the surface potential in absence of field after the quinone treatment shifts from $\Phi_{s01} = 1$ kT to $\Phi_{s02} = -3$ kT.

Based on these results on Fig. 9 we have tried to illustrate the conditions which are developed on the germanium surface. According to the experimental evidence before the quinone treatment, the surface layer of the germanium is slightly enriched in negative carriers, which indicates the presence of positive charges in the oxide layer covering the surface. After quinone treatment the positive carriers prevail in the surface layer of germanium (inversion occurs), indicating the presence of a negative charge in the oxide. Besides the change of the surface potential, the character of the recombination centers changes as well. The possible physical and chemical causes of these changes are discussed hereafter.

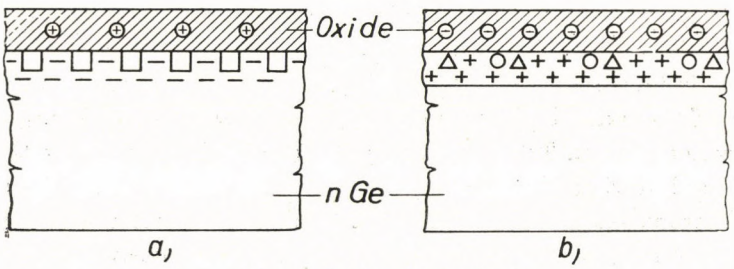


Fig. 9. Surface conditions:

a) ⊕ positive charge on the surface; — negative space charge; □ recombination center; $\alpha_p/\alpha_n = 7$, $E_t = 3{,}5$ kT, $N_t = 2{,}4 \cdot 10^{11}$ cm^{-2}

b) ⊖ negative surface charge; + holes in the inversion layer; △ recombination center; $\alpha_p/\alpha_n = 1{,}5$, $E_t = 2$ kT, $N_t = 1{,}5 \cdot 10^{11}$ cm^{-2}; ○ recombination center; $\alpha_p/\alpha_n = 1{,}5$, $E_t = 7$ kT, $N_t = 2{,}5 \cdot 10^{11}$ cm^{-2}

III. Discussion

In aqueous solutions used for surface cleaning of germanium a hydrated oxide layer developes, the structure of which is extremely disordered: it contains bound water molecules, hydroxilgroups, oxigen vacancies. The protons of the water molecules create a positive charge on the layer, effecting in the vicinity of the germanium surface a region enriched by electrons. Fig. 9a indicates this situation. The surface potential is established in accordance with the charge states. The role of the water molecules is not restricted to developing the surface potential only, but according to the investigations of A. V. RZHANOV [14], the effectiveness of the electron-hole recombination centers is diminished as well.

As a result of the applied heat-treatment in vacuum (150 °C, for 1 hour at 10^{-4} Torr), a considerable part of the water, present in the oxide layer was removed, the proof of which is the fact that the recombination velocity was

increased while the surface potential shifted towards the positive. All these phenomena may be explained by the decrease of the positive oxide charge caused by the water molecules and by the disappearance of the impeding effect on the activity of the recombination centers. By more an effective drying process, the complete inversion of the surface can be obtained: i.e. the oxide becomes negatively charged. This charge may result from electrons originating from the germanium, which fill the oxygen vacancies. If the sample contains water, the vacancies are occupied by the oxygen of the water molecules, whereas the positive charge is due to protons. As mentioned, the effect of the quinone treatment was examined on surfaces which had previously been heat-treated as described above. On these surfaces it was found that the surface potential was somewhat shifted towards the positive and the recombination velocity was increased. On quinone treatment the surface potential became more positive, whereas the recombination velocity decreased and new centers developed while the previous one disappeared.

These findings might lead to the thought that in the presence of quinone the charge of the oxide layer became more negative: corresponding to a state with the water absent. This effect might possibly be caused by the oxygen atoms released during the reaction of quinone and water, obtaining their missing electrons from the germanium. The quinone molecule itself is also capable of capturing electrons, thus causing the development of a layer on the oxide surface with a negative charge. The previous process consumes water molecules, therefore one part of the oxygen vacancies is able to capture electrons again, thus shifting the surface potential of germanium towards the positive. It seems likely that each of the three possibilities contributes to the development of the final state.

Theoretically the decrease of the recombination velocity could be explained by the behaviour of a complex, consisting of water molecule and active center, as suggested by RZHANOV, which on removing the water molecule, loses its recombination character. This would make the dehydrating effect of the quinone probable. It seems to disagree with this explanation that as the result of the quinone treatment, the recombination center observable at normal condition disappears and two new centers of different energy and capture cross section appear instead. In the new centers the capture probability of the two different charge carriers is nearby equal, the surface recombination velocity inevitably decreases, therefore, in the surface inversion layer, developed by the negatively charged oxide. Experiments performed so far do not give sufficient explanation in detail, e.g. how the quinone treatment influences the development of new centers and how it impedes the activity of the original center.

Our experiments seem to substantiate that the stabilization of the reverse current of transistors by quinone is due to the elimination of the action

of water, the decrease in recombination velocity being favourable as well. Therefore, the use of quinone for stabilizing semiconductor devices may be considered reasonable and well-founded.

LITERATURE

1. Lőrinczy, A.—Pataki, G.: *Acta Physica Hung.* **13** (1961), 67.
2. Lőrinczy, A.—Pataki, G.: *Acta Physica Hung.* **15** (1963), 57.
3. Wallmark, J. T.—Johnson, R. R.: *RCA Review* **18** (1957), 512.
4. Makoto Kikuchi: *Researches of the Electrotechnical Laboratory* (1960); No. 591, 1.
5. Atalla, M. M.—Bray, A. R.—Linder, R.: *Proc. IRE* **106** (1959); Part B (Int. Conv. on Transistor and Associated Semiconductor Devices) Supplement No 17, 1130.
6. Rzhanov, A. V.: *Soviet Solid State Physics* **3** (1961), 822.
7. Deutsches Patentamt 1054178 (7. August 1957).
8. Moore, A. R.—Nelson, R.: *RCA Review* **17** (1956), 5.
9. Lyuze et al.: *Proc. of the Conf. on the Semiconductor Surfaces*, Tomsk (1964), 79.
10. Gnaedinger, R. J.—Flashen, S. S.: *J. Phys. Chem.* **59** (1955), No. 67.
11. Grove, A. S.—Fitzgerald, D. J.: *Solid State Electronics* **5** (1966), 894.
12. Szép, I. C.—Tihanyi, J.: Symp. on Reliability on Electronics, Budapest (1964).
13. Beale, J. R. A.—Thomas, D. E.—Watkins, T. B.: see ref. 5, 1004.
14. Rzhanov, A. V.: *Trudi Fizicheskogo Inst. Lebedeva* (1964).

Einfluß von Chinon auf die Oberflächeneigenschaften von Germanium. Die Verfasser haben Versuche mit *p—n—p* Germanium-Transistoren durchgeführt, um den Einfluß des Chinons zu klären. An Transistoren wurden mit Hilfe des Emitter- und Kollektor-Raumeffekts die Abhängigkeit des Rückstroms und des Stromverstärkungsfaktors vom Oberflächenpotential untersucht. Unter der Wirkung von Chinon wurde die Rückstromcharakteristik verbessert, der Sickerverlust und das damit verbundene Geräusch und Instabilität gingen zurück. Das Oberflächenpotential verschub sich in Richtung nach den positiven Werten. Die Verfasser maßen zwecks quantitativer Bestimmung der physikalischen Kenngrößen der Oberfläche Germaniumproben mittels Raumeffektmessung. Auf Grund der Ergebnisse von zweierlei Messungen bringen die Verfasser Vorstellungen über die Natur der Vorgänge, die sich auf der Oberfläche abspielen.

L'influence de la quinone sur les qualités de surface du germanium. Les auteurs ont fait des essais sur transistors *p—n—p* au germanium et sur échantillons de germanium, en vue d'élucider l'influence de la quinone. A l'aide de l'effet spatial de l'émetteur et du collecteur, ils ont examiné sur des transistors, comment le courant inverse et le facteur d'amplification du courant dépendent du potentiel de surface. La caractéristique du courant inverse a été améliorée sous l'influence de la quinone, tandis que le courant de fuite et le bruit et l'instabilité qui y sont associés ont diminué. Le potentiel de surface s'est déplacé en direction des valeurs positives. Les auteurs ont mesuré l'effet spatial sur des échantillons de germanium, pour déterminer quantitativement les caractéristiques physiques de la surface. Des considérations relatives à la nature des phénomènes se déroulant à la surface sont présentées sur base des résultats des deux sortes de mesures effectuées.

Воздействие хинона на поверхностные свойства германия *(А. Лёринци, М. Немет-Шаллаи, И. Ц. Сеп и Й. Тихани).* Авторы провели опыты на *p—n—p* германиевых транзисторах и образцах германия с целью выяснения воздействия хинона. На транзисторах с помощью эмиттерного и коллекторного полевого эффекта исследована зависимость от поверхностного потенциала обратного тока и коэффициента усиления по току. Под воздействием хинона характеристика обратного тока улучшается, а ток утечки и связанный с последним шум и нестабильность падают. Поверхностный потенциал сдвинулся в сторону положительных значений. Авторы для количественного определения физических показателей поверхности исследовали образцы германия измерением полевого эффекта. Работа на основе результатов двух различных видов измерений дает соответствующие представления относительно природы процессов, протекающих на поверхности.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 201—218 (1968)

CALCULATION OF THE COEFFICIENT OF SELF-INDUCTION AND OF THE IMPEDANCE OF CYLINDRICAL COILS WITH OPEN IRON CORE

GY. TEVAN

CAND. OF TECHN. SC.
TECHNICAL UNIVERSITY OF HEAVY INDUSTRIES, MISKOLC (HUNGARY)

[Manuscript received July 31, 1967]

The paper gives an approximation formula for the calculation of the coefficient of self-induction and the impedance of symmetrically arranged cylindrical coils with open iron cores. The exterior magnetic field of the coil is replaced by the exterior field of an ellipsoid of rotation with convenient excitation and the excitation consumed by the air is determined on this basis.

I. Introduction

As is known, the coefficient of self-induction of cylindrical air core coils of finite length is

$$L = k\,\mu_0 \frac{A}{l} N^2$$

where A is the cross section of the coil, l its length, N the number of turns, $\mu_0 = 4\,\pi \cdot 10^{-7}$ Vs/Am and k the Nagaoka-coefficient (resulting of an elliptic integral) depending on the ratio l/d of the coil with diameter d. This formula is not valid for coils with an iron core. In the case of very slender iron core coils we may calculate with the formula

$$L = \frac{BA}{Hl} N^2$$

where B and H are the corresponding values of the magnetization curve. If the l/d ratio of the iron core coil is not "great" the latter relation cannot be applied, as part of the excitation corresponding to the air is not negligible.

II. Calculation of the coefficient of self-induction

If the part of the excitation corresponding to the iron core is marked Θ_v, the excitation corresponding to the air Θ_l and the mean flux related to one turn Φ, the coefficient of self-induction is:

$$L = \frac{\Phi N}{\frac{1}{N}(\Theta_v + \Theta_l)} = \frac{\Phi N^2}{\Theta_v + \Theta_l} .$$

Let us introduce the symbols

$$L_v = \frac{\Phi N^2}{\Theta_v} \tag{1}$$

and

$$L_l = \frac{\Phi N^2}{\Theta_l} , \tag{2}$$

then

$$\frac{1}{L} = \frac{1}{L_v} + \frac{1}{L_l} . \tag{3}$$

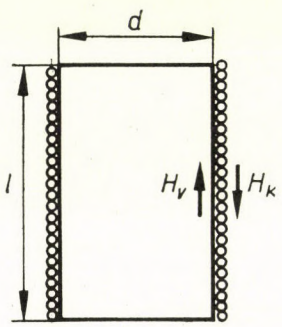


Fig. 1

Hence the coefficient of self-induction of the coil equals the resulting coefficient of self-induction of parallel coils with a coefficient of self-induction corresponding to the excitation of air and to the excitation of iron.

Let us content ourselves temporarily with discussing iron core coils where the lengths of the coil and the iron core are equal (Fig. 1). In view of the fact that the flux linked with the central turns is not much greater than that linked with the extreme turns, we do not commit a great error by taking the (linear) mean value of the longitudinal component of the magnetic field strength of the iron core (H_v), as being equal with the value H belonging to the point of the magnetization curve with the induction B_v, which can be calculated from the mean flux. Thus on the basis of relation (1)

$$L_v = \frac{\Phi N^2}{H_v l} = \frac{B_v A N^2}{H_v l} \tag{4}$$

where B_v and H_v are the coherent values of the magnetization curve.

For the excitation Θ_l consumed by the air we use the approximation that the longitudinal component of the magnetic field strength in the iron core is constant. From the law of excitation applied to the loop directly surrounding whichever part of the uniformly wound coil it follows then that the longitudinal component of the magnetic field strength on the exterior cylinder surface of the coil is also constant. If this is denoted H_k it is evident from the law of excitation written for the whole coil that the excitation of the air is $\Theta_l = H_k l$, further it is obvious that on the exterior surface of the cylindrical

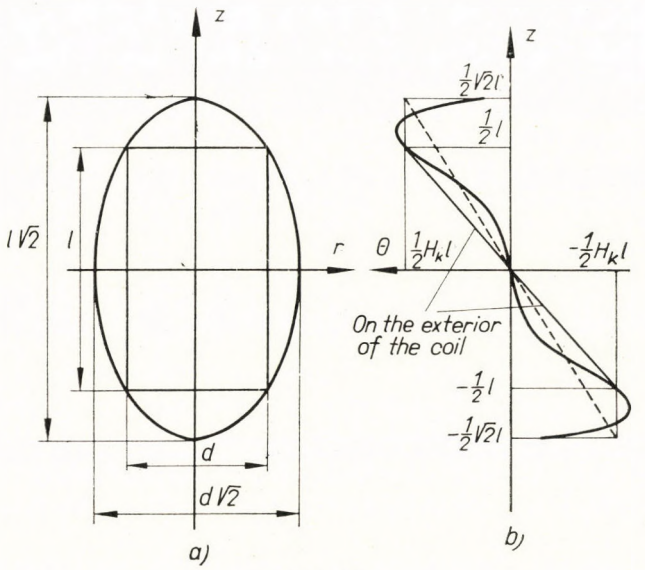


Fig. 2

coil the magnetic (scalar) potential varies longitudinally from $-1/2\, H_k l$ to $+1/2\, H_k l$ according to a linear law — if we assume the zero of the potential to be in the centre. Let us surround the cylindrical coil with such an ellipsoid of rotation where the axes of the basic ellipse are $l\sqrt{2}$ and $D\sqrt{2}$ (Fig. 2a).

On the partial ellipsoid surface $-l/2 \leq z \leq l/2$ the magnetic potential is somewhat smaller than the potential of the point with the same co-ordinate z on the exterior surface of the cylindrical coil; at the points $z = 0$, $z = \pm l/2$ there is no difference (Fig. 2b). On the 'ellipsoid caps' $z > l/2$ and $z < -l/2$ the magnetic potential increases with increasing z, then it decreases. The cause of the increase is that the derivative of the potential cannot jump at the values $z = l/2$ (or $z = -l/2$), that of the decrease is that the lines of forces straying from the end of the iron core to the upper surface of the 'cap' bend towards the axis and therefore the surface component of the field of forces is of opposite direction at the top of the 'cap' than at the bottom. From the

point of view of the determination of the connection between the flux and the excitation we do not commit a grave error by taking the potential on the ellipsoid for linearily distributed as represented on Fig. 2b by a dotted line. Thus, the magnetic potential on the surface of the ellipsoid is

$$\Theta = \frac{H_k l}{2} \cdot \frac{z}{l/\sqrt{2}} = \frac{1}{\sqrt{2}} H_k z \,.$$

As a flux cannot move off to the infinite the normal component of the magnetic field of forces on the 'infinite distant' spherical surface, i.e. the derivative of the potential in normal direction is zero. Thus, the potential of whichever straight is traversing the centre of the ellipsoid approaches a finite value as it moves off the centre.

Hence a limited scalar function is to be looked for with a prescribed value on the ellipsoid of rotation, defined outside the ellipsoid and satisfying the Laplace-relation. It is adviseable to solve this problem in a system of rotational elliptic coordinates. If $l > d$ is stretched, if $l < d$ is a flat, a system of rotational elliptic co-ordinates should be employed:

$$r = \frac{1}{\sqrt{2}} \sqrt{l^2 - d^2} \cos u \sinh v \,,$$

$$z = \frac{1}{\sqrt{2}} \sqrt{l^2 - d^2} \sin u \cosh v$$

and (5)

$$r = \frac{1}{\sqrt{2}} \sqrt{d^2 - l^2} \sin u \cosh v \,,$$

$$z = \frac{1}{\sqrt{2}} \sqrt{d^2 - l^2} \cos u \sinh v,$$

respectively.

The given ellipsoid is characterized by the co-ordinate $v = v_0$, for which

$$\sinh v_0 = \frac{d}{\sqrt{l^2 - d^2}}, \quad \cosh v_0 = \frac{d}{\sqrt{d^2 - l^2}}$$

and (6)

$$\cosh v_0 = \frac{l}{\sqrt{l^2 - d^2}}, \quad \sinh v_0 = \frac{l}{\sqrt{l^2 - d^2}},$$

respectively.

As the potential on this ellipsoid has the value $(1/\sqrt{2})\, H_k z$, the boundary condition can be written in the form of

$$(\Theta)_{v=v_0} = \frac{lH_k}{2} \sin u$$

and

$$(\Theta)_{v=v_0} = \frac{lH_k}{2} \cos u,$$

resp. or even shorter

$$(\Theta)_{v=v_0} = \frac{\Theta_l}{2} \sin u$$

and

$$(\Theta)_{v=v_0} = \frac{\Theta_l}{2} \cos u,$$

resp. where $\Theta_l = H_k l$ is the excitation falling to the share of the air. As the boundary condition and also the boundary surface are axially symmetrical, the solution is such as well. From the form of the Laplace equation expressed in rotationally symmetrical ellipsoid co-ordinates we get the partial differential equation

$$\frac{1}{\cos u} \frac{\partial}{\partial u} \left(\cos u \frac{\partial \Theta}{\partial u} \right) + \frac{1}{\sinh v} \frac{\partial}{\partial v} \left(\sinh v \frac{\partial \Theta}{\partial v} \right) = 0$$

and

$$\frac{1}{\sin u} \frac{\partial}{\partial u} \left(\sin u \frac{\partial \Theta}{\partial u} \right) + \frac{1}{\cosh v} \frac{\partial}{\partial v} \left(\cosh v \frac{\partial \Theta}{\partial v} \right) = 0,$$

resp.

Its general solution is:

$$\Theta = \sum_{n=0}^{\infty} P_n (\sin u) \left[A_n Q_n (\cosh v) + B_n P_n (\cosh v) \right]$$

and

$$\Theta = \sum_{n=0}^{\infty} P_n (\cos u) \left[A_n Q_n (j \sinh v) + B_n P_n (j \sinh v) \right],$$

respectively, where $P_n(x)$ is the Legendre function of the first kind and the n-th order, $Q_n(x)$ that of the second kind and the n-th order and $j = \sqrt{-1}$. As Θ must be limited in the exterior space of the ellipsoid and $P_n(x)$, except $P_0(x) = 1$, and x approach infinity together, $B_1 = B_2 = B_3 = \ldots = 0$. Further on the basis of the Equ. $P_1(x) = x$, the boundary condition may be written in the form of

$$(\Theta)_{v=v_0} = \frac{\Theta_l}{2} P_1 (\sin u),$$

or

$$(\Theta)_{v=v_0} = \frac{\Theta_l}{2} P_1 (\cos u),$$

respectively, and hence $A_0 = A_2 = A_3 = A_4 = \ldots = 0$, $B_0 = 0$, and

$$A_1 = \frac{\Theta_l}{2} \frac{1}{Q_1(\cosh v_0)}$$

and

$$A_1 = \frac{\Theta_l}{2} \frac{1}{Q_1(j \sinh v_0)},$$

respectively.

So the solution of our problem consists of one member only:

$$\Theta = \frac{\Theta_l}{2} P_1(\sin u) \frac{Q_1(\cosh v)}{Q_1(\cosh v_0)}$$

or

$$\Theta = \frac{\Theta_l}{2} P_1(\cos u) \frac{Q_1(j \sinh v)}{Q_1(j \sinh v_0)}.$$

Taking into account that $P_1(x) = x$ and $Q_1(x) = x \operatorname{Ar\,coth} x - 1$, $Q_1(jx) = jx \operatorname{Ar\,coth}(jx) - 1 = x \operatorname{arc\,cot} x - 1$, the potential is

$$\Theta = \frac{\Theta_l}{2} \frac{\cosh v . \operatorname{Ar\,coth}(\cosh v) - 1}{\cosh v_0 . \operatorname{Ar\,coth}(\cosh v_0) - 1} \sin u,$$

or

$$\Theta = \frac{\Theta_l}{2} \frac{\sinh v . \operatorname{arc\,cot}(\sinh v) - 1}{\sinh v_0 . \operatorname{arc\,cot}(\sinh v_0) - 1} \cos u, \tag{7}$$

resp.

The total flux issuing from the upper half of the ellipsoid (and flowing into the lower half) is evidently greater than the flux linked with the exterior turns of the cylindrical coil, but smaller than the flux linked with the central turns; a part of this latter flux flowing between the coil and the ellipsoid. The average flux related to one turn, Φ may thus be taken as approximately equal to the flux issuing from the upper half of the ellipsoid. In order to determine this we need the measure of the elements of arc along the co-ordinates u and v

$$\sqrt{g_{uu}} = \sqrt{\left(\frac{\partial r}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\right)^2}$$

and

$$\sqrt{g_{vv}} = \sqrt{\left(\frac{\partial r}{\partial v}\right)^2 + \left(\frac{\partial z}{\partial v}\right)^2}.$$

The normal component of the magnetic field strength is on the ellipsoid $v = v_0$

$$(H_v)_{v=v_0} = -\frac{1}{(\sqrt{g_{vv}})_{v=v_0}} \left(\frac{\partial \Theta}{\partial v}\right)_{v=v_0}.$$

The flux issuing from the upper part of the ellipsoid is

$$\Phi = \int_0^{\pi/2} \mu_0 (H_v)_{v=v_0} 2 (r)_{v=v_0} \pi (\sqrt{g_{uu}})_{v=v_0} du =$$

$$= -2\pi\mu_0 \int_0^{\pi/2} (r)_{v=v_0} \left(\frac{\partial \Theta}{\partial v}\right)_{v=v_0} \left(\sqrt{\frac{g_{uu}}{g_{vv}}}\right)_{v=v_0} du.$$

On the basis of (5) $g_{uu} = g_{vv}$, further on the basis of (5) and (6)

$$(r)_{v=v_0} = \frac{d}{\sqrt{2}} \cos u$$

and

$$(r)_{v=v_0} = \frac{d}{\sqrt{2}} \sin u,$$

respectively, hence the sought for flux is, applying the relation (7):

$$\Phi = \frac{d\pi\mu_0}{2\sqrt{2}} \Theta_l \frac{\coth v_0 - \sinh v_0 . \operatorname{Ar}\coth(\cosh v_0)}{\sinh v_0 . \operatorname{Ar}\coth(\cosh v_0) - 1}$$

or

$$\Phi = \frac{d\pi\mu_0}{2\sqrt{2}} \Theta_l \frac{\tanh v_0 - \cosh v_0 . \operatorname{arc}\cot(\sinh v_0)}{\sinh v_0 . \operatorname{arc}\cot(\sinh v_0) - 1},$$

respectively.

As

$$\operatorname{Ar}\coth(\cosh v_0) = \operatorname{Ar}\cosh\left(\frac{\cosh v_0}{\sqrt{\cosh^2 v_0 - 1}}\right) = \operatorname{Ar}\cosh(\coth v_0)$$

and

$$\operatorname{arc}\cot(\sinh v_0) = \operatorname{arc}\cos\left(\frac{\sinh v_0}{\sqrt{1 + \sinh^2 v_0}}\right) = \operatorname{arc}\cos(\tanh v_0),$$

further applying the equality (6) and (2)

$$L_l = N^2 \frac{d\pi\mu_0}{2\sqrt{2}} F\left(\frac{l}{d}\right), \tag{8}$$

where

$$F\left(\frac{l}{d}\right) = \frac{\frac{l}{d}\sqrt{\frac{l^2}{d^2} - 1} - \operatorname{Ar}\cosh\frac{l}{d}}{\frac{l}{d}\operatorname{Ar}\cosh\frac{l}{d} - \sqrt{\frac{l^2}{d^2} - 1}} \left(\text{if } \frac{l}{d} \geqq 1\right) \tag{8a}$$

and

$$F\left(\frac{l}{d}\right) = \frac{\arccos\frac{l}{d} - \frac{l}{d}\sqrt{1-\frac{l^2}{d^2}}}{\sqrt{1-\frac{l^2}{d^2}} - \frac{l}{d}\arccos\frac{l}{d}} \quad \left(\text{if}\,\frac{l}{d} \leq 1\right), \tag{8b}$$

respectively.

The diagram of the function $F(l/d)$ is shown in Fig. 3. When $l/d < 0{,}25$ we may calculate with the formula

$$F\left(\frac{l}{d}\right) = \frac{\pi}{2} + \left(\frac{\pi^2}{4} - 2\right)\frac{l}{d} \tag{8c}$$

and when $l/d > 10$ with the formula

$$F\left(\frac{l}{d}\right) = \frac{\dfrac{l}{d}}{\left(\ln\dfrac{2\,l}{d} - 1\right)\left(1 + \dfrac{d^2}{l^2}\ln\dfrac{2\,l}{d}\right)}\,. \tag{8d}$$

Hence the coefficient of self-induction of the coils with open iron core can be calculated with the relations (4), (8) and (3).

The obtained coefficient of self-induction belongs to the flux $B_v A$ and to the current $B_v AN/L$.

The a.c. impedance of the coil can be calculated — disregarding the active resistance of the conductor — from the formula

$$\frac{1}{\mathbf{Z}} = \frac{1}{jL_l\,\omega} + \frac{1}{\mathbf{Z}_v} = \frac{1}{jX_l} + \frac{1}{\mathbf{Z}_v}\,. \tag{9}$$

The impedance of the coil would be $\mathbf{Z}_v$, if there were a medium of infinite permeability in place of the air. In the case of a laminated iron core, of a sintered powder iron core or of a ferrit core, at sine voltage

$$\frac{1}{\mathbf{Z}_v} = \frac{pG}{(4{,}44\,B_m\,AfN)^2} + \frac{qG}{j\,(4{,}44\,B_m\,AfN)^2} \tag{9a}$$

where p and q are the active and reactive powers, resp. of the sheet iron (the sintered powder iron core, the ferrit core) per unit weight at frequency f and alternating stress with maximum induction B_m, G the weight of the core, A its cross section. If the iron core is compact, the absolute value of the impedance $\mathbf{Z}_v$ and the tangent of its phase angle for *strong magnetic fields* can be

calculated on the basis of L. R. NEJMAN's book: "Poverkhnostnyu effekt v ferromagnitnykh telakh" (1949) by the following formulae

$$Z_v = \frac{d\pi}{l}\left[1{,}63\,\frac{\varrho}{\delta_1}\left(1-\frac{\delta_1}{d}\right)\right]; \quad \tan\varphi = 0{,}76\left(1-0{,}45\,\eta\mu_1+0{,}8\,\frac{\delta_1}{d}\right) \quad (9b)$$

where ϱ is the specific resistance, μ_1 the permeability calculated from the effective value H_v of the magnetic field strength assumed to be on the surface of the iron core and from the induction belonging to it,

$$\delta_1 = \sqrt{\frac{\varrho}{\pi f \mu_1}}$$

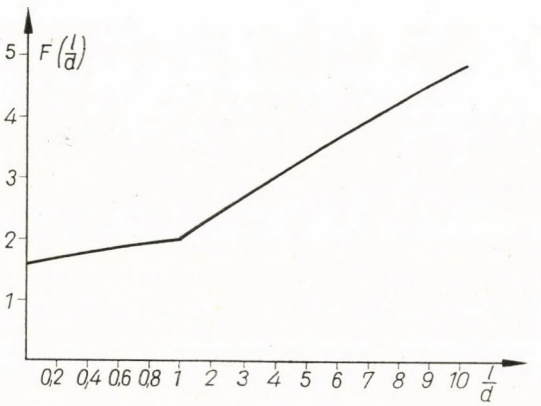


Fig. 3

and η is the hysteresis loss energy per unit volume, related to unit maximum induction. The unit of this latter is

$$\frac{VAs}{m^3}\cdot\frac{m^4}{V^2 s^2} = \frac{Am}{Vs}.$$

The formula (9b) can be employed for $\delta_1 < d/5$. To the impedance Z calculated from the formula (9) the active resistance of the conductor is to be added. Not only the impedance from Equ. (9b), but also that from Equ. (9a) depends on the magnetic field strength H_v, the first member of (9a) only to a small degree, but its second member essentially. The resulting impedance **Z** evidently belongs to the voltage

$$U = \frac{H_v l}{N} Z_v$$

and to the current

$$\frac{H_v l}{N}\cdot\frac{Z_v}{Z}.$$

III. Comparison of measurement and calculation

1. Measurement with d.c.

The measurements were carried out with two-coil iron cores with a diameter $d = 26$ mm and lengths corresponding to the ratios $l/d = 0{,}5$; 1; 2; 4; 10; 32. At the commutation of the d.c. conducted through the coil made of thicker wire the change of flux was measured by a flux meter connected to the multi-turn coil. (Both coils cover the iron core completely.) Adjusting the change of flux measured on the third measuring coil placed only on the central

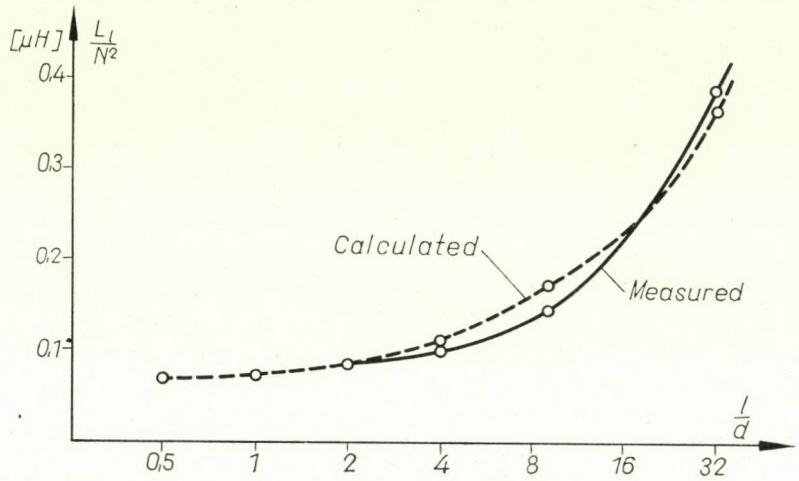


Fig. 4

part of the longest iron core to being equal to the changes of flux of the principal measurement, the excitation per unit length of the iron core was calculated from the current of the current coil (in the centre of the long iron core the magnetic field strength on the exterior surface of the current coil is negligible). The difference between the excitation determined by the principal measurement and the excitation absorbed by the iron core gives the excitation consumed by the air Θ_l, and from this the coefficient of self-induction L_l. Table I shows the measured values of L_l/N^2 and those calculated on the basis of formula (8), and Fig. 4 the corresponding curves.

Table I

l/d	0,5	1	2	4	10	32
L_l/N^2 [μH] (measured)	0,0663	0,0738	0,0872	0,103	0,146	0,39
L_l/N^2 [μH] (calculated)	0,0653	0,0726	0,0865	0,1114	0,1755	0,367

2. *Measurement with mains-frequency a.c.*

The impedance of the iron core coils was measured. The circuit diagram of the measurement is shown on Fig. 5.

The voltages were measured with a static voltmeter and with the iron core coils of ratios $l/d = 0{,}5;\ 1;\ 2$ an amplifier had to be used, too. By this measurement the impedance Z of the coil without active resistance and its phase angle φ can be determined at the adjusted current I. On the basis of Equ. (9):

$$\frac{1}{Z(\cos\varphi + j\sin\varphi)} = \frac{1}{jX_l} + \frac{1}{Z_v(\cos\varphi_v + j\sin\varphi_v)},$$

i.e.

$$\frac{1}{Z}\cos\varphi + j\left(\frac{1}{X_l} - \frac{1}{Z}\sin\varphi\right) = \frac{1}{Z_v}\cos\varphi_v - j\,\frac{1}{Z_v}\sin\varphi_v.$$

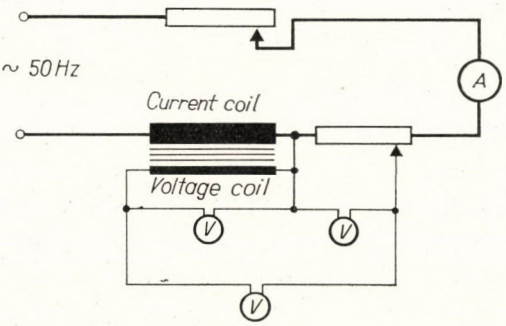


Fig. 5

If the imaginary part of the latter equation is divided by its real part and from the relation thus obtained X_l is expressed

$$X_l = \frac{Z}{\sin\varphi - \cos\varphi \cdot \tan\varphi_v}. \tag{10a}$$

Hence for the calculation of X_l the value of $\operatorname{tg}\varphi_v$ is necessary. This quantity depending on H_v (the surface value of the magnetic field strength) was measured with the circuit according to Fig. 5, with the current coil of the longest iron core and with the third coil placed only in the central part as a voltage coil, at various currents and has been represented in function of $H_v = IN/l$. The connection between I and $\mathbf{H}_v$ of the principal measurement is

$$I = \frac{IZ(\cos\varphi + j\sin\varphi)}{jX_l} + \frac{\mathbf{H}_v l}{N},$$

from which

$$H_v = |\mathbf{H}_v| = \frac{IN}{l}\left|1 - \frac{Z\sin\varphi}{X_l} - j\frac{Z\cos\varphi}{X_l}\right| = \\ = \frac{IN}{l}\sqrt{1 + \frac{Z^2}{X_l^2} - 2\frac{Z\sin\varphi}{X_l}} \tag{10b}$$

X_l was determined finally by the following iteration method: a $\tan\varphi_v$ value was assumed, from this and from Equ. (10a) a X_l value was calculated, and then from Equ. (10b) a H_v value. From the diagram of $\tan\varphi_v(H_v)$ the $\tan\varphi_v$ value belonging to it was choosen and this again was substituted into Equ. (10a), etc.

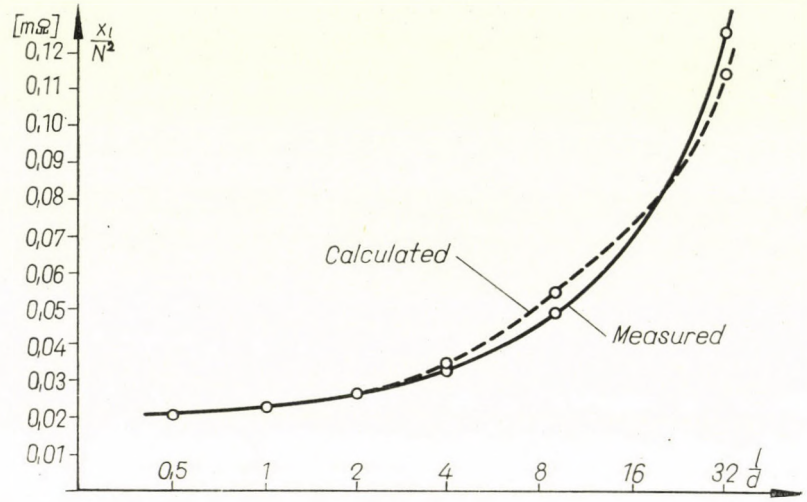


Fig. 6

The measured and calculated values X_l/N^2 are contained in Table II and the corresponding curves in Fig. 6.

Table II

l/d	0,5	1	2	4	10	32
X_l/N^2 [mΩ] (measured)	0,0216	0,0235	0,0273	0,0334	0,0497	0,127
X_l/N^2 [mΩ] (calculated)	0,0205	0,0228	0,0272	0,035	0,0551	0,1152

IV. Calculation of the coefficient of self-induction if the coil and iron core are of different dimensions

Only those cases are dealt with, where the planes of symmetry of the coil and that of the iron core coincide. If the iron core projects over the coil (Fig. 7a) the exterior magnetic field may be considered as the superposition of the exterior magnetic field of the two arrangements according to Fig. 7b. The resulting flux is the sum of the fluxes corresponding to the two arrangements:

$$\Phi = \Phi_l + \Phi_v = L_l \frac{\Theta}{N^2} + L_{lv} \frac{\Theta_v}{N^2} .$$

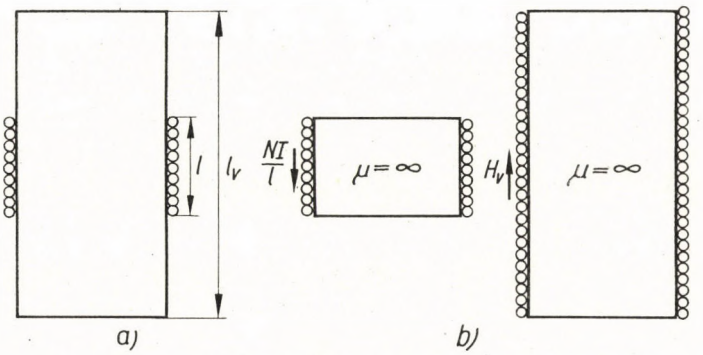


Fig. 7

But $\Theta = NI$, $\Theta_v = -H_v\, l_v$, further by replacing the relation (8)

$$\Phi = \frac{d\pi\mu_0}{2\sqrt{2}} F\left(\frac{l}{d}\right) NI - \frac{d\pi\mu_0}{2\sqrt{2}} F\left(\frac{l_v}{d}\right) H_v\, l_v .$$

On the other hand, on the basis of (4), writing l_v instead of l:

$$H_v\, l_v = \frac{\Phi N^2}{L_v}$$

thus

$$\Phi\left[1 + \frac{d\pi\mu_0}{2\sqrt{2}} F\left(\frac{l_v}{d}\right) \frac{N^2}{L_v}\right] = \frac{d\pi\mu_0}{2\sqrt{2}} F\left(\frac{l}{d}\right) NI$$

and thus for the coefficient of self-induction $L = N\Phi/I$ the following relation is obtained:

$$\frac{1}{L} = \frac{2\sqrt{2}}{d\pi\mu_0 F\,(l/d)\, N^2} + \frac{F(l_v/d)}{F(l/d)} \frac{1}{L_v} = \frac{1}{L_l} + \frac{1}{\dfrac{F(l/d)}{F(l_v/d)} L_v} . \qquad (11)$$

Likewise for the a.c. impedance **Z** not containing the active resistance of the coil

$$\frac{1}{\mathbf{Z}} = \frac{1}{jX_l} + \frac{1}{\frac{F(l/d)}{F(l_v/d)}\mathbf{Z}_v} . \tag{12}$$

The resulting coefficient of self-induction belongs to the flux

$$L_v \frac{H_v l_v}{N^2}$$

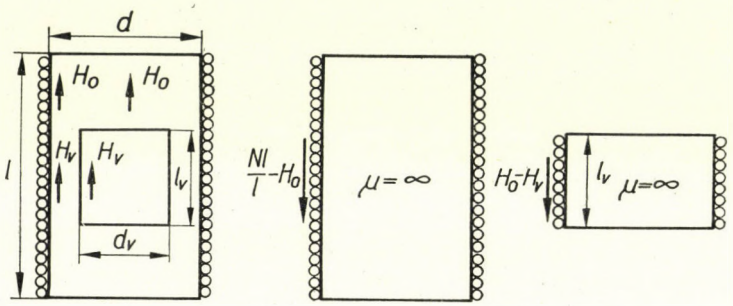


Fig. 8

and to the intensity

$$\frac{L_v H_v l_v}{NL} ,$$

and the resulting impedance belongs to the voltage

$$\frac{H_v l_v}{N} Z_v$$

and to the intensity

$$\frac{H_v l_v}{N} \cdot \frac{Z_v}{Z} .$$

It must be noted that in the deduction the fact that the coil and the iron core have the same diameter plays an essential part. Otherwise the summation of the fluxes Φ_t and Φ_v would not be justified.

If the coil and the iron core are arranged according to Fig. 8a, then applying the approximation that the axial component of the magnetic field strength inside the coil but outside the iron core is constant by sections, the exterior magnetic field may be assumed to be the superposition of the exterior

magnetic fields of the two arrangements according to Fig. 8b. The resulting flux is

$$\Phi = \frac{d\pi\mu_0}{2\sqrt{2}} F\left(\frac{l}{d}\right)(NI - H_0 l) + \frac{d\pi\mu_0}{2\sqrt{2}} F\left(\frac{l_v}{d}\right)(H_0 - H_i) l_i .$$

Taking into account that

$$\mu_0 \frac{d^2\pi}{4} H_0 = \Phi$$

and

$$\pi \frac{d^2 - d_v^2}{4} \mu_0 H_v + \frac{L_v H_v l_v}{N^2} = \Phi,$$

$$\Phi \left\{ 1 + \frac{4}{\mu_0 d^2 \pi} \frac{d\pi\mu_0}{2\sqrt{2}} \left[lF\left(\frac{l}{d}\right) - l_v F\left(\frac{l_v}{d}\right) \right] + \right.$$

$$\left. + \frac{\frac{d\pi\mu_0}{2\sqrt{2}} F\left(\frac{l_v}{d}\right) l_v}{\pi \frac{d^2 - d_v^2}{4} \mu_0 + \frac{L_v l_v}{N^2}} \right\} = \frac{d\pi\mu_0}{2\sqrt{2}} F\left(\frac{l}{d}\right) NI .$$

Introducing the symbols

$$L_0 = \mu_0 \frac{d^2\pi}{4l} N^2 \tag{13a}$$

and

$$L_{v0} = \mu_0 \frac{d^2 - d_v^2}{4 l_v} \pi N^2 . \tag{13b}$$

we obtain for the coefficient of self-induction $L = N\Phi/I$ the following relation:

$$\frac{1}{L} = \frac{1}{L_l} + \frac{1 - \frac{l_v}{l} \frac{F(l_v/d)}{F(l/d)}}{L_0} + \frac{\frac{F(l_v/d)}{F(l/d)}}{L_{v0} + L_v} . \tag{14}$$

This coefficient of self-induction belongs to the flux

$$(L_{v0} + L_v) \frac{H_v l_v}{N^2}$$

and to the intensity

$$\frac{(L_{v0} + L_v) H_v l_v}{NL} .$$

Likewise we obtain for the impedance $\mathbf{Z}$ not containing the active resistance of the coil the formula

$$\frac{1}{\mathbf{Z}} = \frac{1}{jX_l} + \frac{1 - \frac{l_v}{l}\,\frac{F(l_v/d)}{F(l/d)}}{jX_0} + \frac{\frac{F(l_v/d)}{F(l/d)}}{jX_{v0} + \mathbf{Z}_v}\,. \tag{15}$$

Here $X_0 = L_0\,\omega$ and $X_{v0} = L_{v0}\,\omega$. This $\mathbf{Z}$ value belongs to voltage

$$\frac{H_v l_v}{N}\,|j\,X_{v0} + \mathbf{Z}_v|$$

and to intensity

$$\frac{H_v l_v}{N}\,\frac{|j\,X_{v0} + \mathbf{Z}_v|}{Z}\,.$$

If $d_v = 0$, $L_v = 0$ and on the basis of (13a) and (13b) $L_{v0} = L_0\,l/l_v$, and so we get from (14) for the coefficient of self-induction L_1 of the empty coil the relation

$$\frac{1}{L_1} = \frac{1}{L_l} + \frac{1}{L_0}$$

From this

$$\frac{L_1}{L_0} = \frac{1}{1 + \frac{L_0}{L_l}}$$

that is, on the basis of (8) and (13a)

$$\frac{L_1}{L_0} = \frac{1}{1 + \frac{1}{\sqrt{2}\,\frac{l}{d}\,F\left(\frac{l}{d}\right)}} = k_1\left(\frac{l}{d}\right).$$

This coefficient k_1 should be equal to the Nagaoka-coefficient k. Table III contains the calculated value of k_1 and k for some l/d values. The principal reason of the divergency is that the approximation according to which the axial component of the magnetic field inside the coil, but outside the iron core is constant by sections gives a small error only, when the straightening effect exerted by the iron core on the magnetic field comes into existence.

Table III

l/d	0	0,01	0,1	0,2	0,25	0,5	1	4	10	∞
k_1	0	0,0218	0,1863	0,318	0,374	0,56	0,737	0,945	0,978	1
k	0	0,035	0,2033	0,3198	0,3654	0,5255	0,6884	0,9016	0,9588	1

From the last relation

$$F\left(\frac{l}{d}\right) = \frac{1}{\sqrt{2}\,\frac{l}{d}\left(\frac{1}{k_1} - 1\right)}\,.$$

With arrangements containing little iron compared to the inner volume of the coil it is, therefore, advisable to calculate in the formulae (14) and (15) with the value

$$G\left(\frac{l}{d}\right) = \frac{1}{\sqrt{2}\,\frac{l}{d}\left(\frac{1}{k} - 1\right)} \tag{16}$$

calculated from the Nagaoka-coefficient k instead of with the quantity $F(l/d)$.

REFERENCES

1. Simonyi, K.: Theory of Electricity. Tankönyvkiadó, Budapest 1958.
2. Neumann, L. R.—Kalantarov, P. L.: Theoretische Grundlagen der Elektrotechnik, Bd. I. Verlag Technik, Berlin 1952.
3. Ollendorff, F.: Potentialfelder der Elektrotechnik. Verlag von Julius Springer, Berlin 1932.

Berechnung des Selbstinduktionskoeffizienten bzw. der Impedanz von zylindrischen Spulen mit offenem Eisenkern. Der Artikel gibt eine Näherungsformel zur Berechnung des Selbstinduktionskoeffizienten bzw. der Impedanz von symmetrisch angeordneten zylyndrischen Spulen mit offenem Eisenkern. Für das äußere magnetische Feld der Spule wird das äußere Feld eines entsprechend erregten Drehellipsoids eingesetzt, und auf dieser Basis wird die auf die Luft entfallende Erregung bestimmt.

Calcul du coefficient de self-induction et de l'impédance des bobines cylindriques à noyau de fer ouvert. L'article donne une formule approchée pour le calcul du coefficient de self-induction, respectivement de l'impédance des bobines cylindriques symétriques à noyau de fer ouvert. Le champ magnétique extérieur est remplacé par le champ extérieur d'un ellipsoïde de révolution à excitation convenable et c'est sur cette base que l'excitation revenant à l'air est définie.

Расчет коэффициента самоиндукции или же импеданса цилиндрических обмоток с незамкнутым сердечником *(Д.Теван)*. В статье приводится приближенная формула, пригодная для расчета коэффициента самоиндукции или же импеданса симметричных цилиндрических катушек с незамкнутым сердечником. Внешнее магнитное поле катушки замещается внешним полем эллипсоида вращения с соответствующим возбуждением и на основе чего определяется возбуждение, падающее на воздух.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 219—226 (1968)

ДОЛГОВЕЧНОСТЬ И ТЕОРИЯ НАДЕЖНОСТИ

Ж. Б. ЛИНКОВСКИЙ

Подробно анализируются вопросы долговечности ремонтируемых машин и приборов в свете теории надежности.

Рассматривается связь долговечности ремонтируемых машин с основными понятиями теории надежности.

В теории надежности [1] одной из основных функций является надежность, под которой понимается вероятность безотказной (или нормальной) работы машины (или прибора), включенной при $t = 0$, на отрезке времени $[0; (t)]$. Обозначается она через $P(t)$ и $0 \leq P(t) \leq 1$. Для неремонтируемой машины, непрерывно работающей, надежность $P(t)$ является монотонно убывающей функцией времени t (рис. 1). Обозначим через $a = [P(t)]_{\text{доп}}$ — допустимое значение надежности $P(t)$, т. е. наименьшее значение надежности, которое можно допустить при использовании машин и приборов. Если $P(t) = F(t)$ и $t \geq 0$, $t = g\,[P(t)]$ — обратная функция, то долговечность численно можно характеризовать временем $T_{долгов}$, при котором достигается значение a у неремонтируемой машины:

$$T_{долгов} = g(a). \tag{1}$$

Часто выбирают $a = 0{,}5$. Например, в случае экспоненциального закона надежности $P(t) = e^{-\lambda t}$, где $\lambda > 0$ — некоторая постоянная, разная для каждого вида машины или прибора (т. н. интенсивность или «опасность» отказов), то имеем:

$$T_{долгов} = \frac{|\ln a|}{\lambda}. \tag{2}$$

Однако, у машины возможны ремонты или разладки. Поэтому в общем случае мерой долговечности машин служит время, в течение которого машина может безотказно работать (или работать в заданных условиях, например, точности, если говорить о приборе) с уровнем надежности не ниже допустимого a:

$$P(t) \geqslant a. \tag{3}$$

При экспоненциальном законе надежности среднее время $T_{ср}$ безотказной работы равно: $T_{ср} = \lambda^{-1}$, т. е. при $a \gtrsim 0{,}5$ долговечность меньше $T_{ср}$.

Из вышеизложенного следует, что для определения долговечности $T_{долгов}$ в условиях ремонта машины необходимо выяснить новый закон изменения надежности $P(t)$, который в общем случае может быть и прерывным. Таким образом, функция $P(t)$, определенная на конечном (или полубесконечном интервале $[0;\infty]$ — лишь для удобства математического описания) интервале $[0; T_{ф}]$, будучи однозначной функцией своего аргумента t, в некоторых точках области своего определения может претерпевать скачок, а на участках ремонта нельзя даже говорить о *существовании* этой функции. В самом деле, что такое ремонт? Ремонт производится либо в случае отказа или поломки машины, либо путем профилактического останова мащины

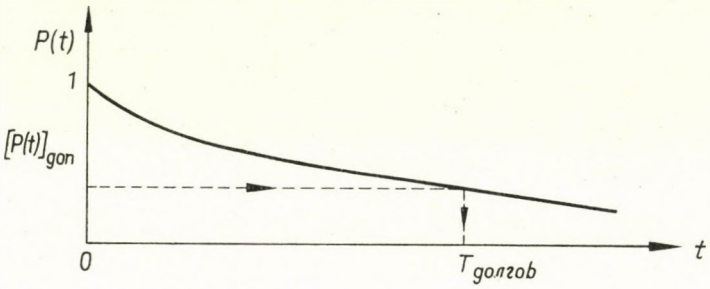


Рис. 1

для осмотра ее узлов и деталей, выявления тех или иных дефектов, или для замены наиболее сильно изнашивающихся деталей (блоков) новыми. Так, например, известно, что при длительных рейсовых пробегах автобусов на расстояние до 1500 *км* и больше, перед каждым рейсом практикуют замену основных подшипников двигателя, чтобы максимально повысить надежность (и безопасность) пробега с большой скоростью движения (до 120 *км/час*). В самом деле, никак нельзя говорить о надежности работы машины в процессе ремонта, так как тогда она разобрана, т. е. находится полностью в нерабочем состоянии, и значит, абсолютно неработоспособна в процессе ремонта продолжительностью $t_{рем}$. Итак, в течение ремонта нельзя говорить о существовании машины, способной выполнять свою функцию, а стало быть, и о надежности машины $P(t)$ в этот период ее существования. Это обстоятельство должно быть учтено при применении тех или иных математических методов для анализа надежности ремонтируемых машин, а следовательно, долговечности. Обозначим через $t^{(к)}_{1,рем}; t^{(к)}_{2,рем}$ — начало и конец K-ного ремонта, $t^{(к)}_{рем} = [t^{(к)}_{2,рем}; t^{(к)}_{1,рем}]$ — продолжительность отрезка времени K-ного ремонта, $K = 1, 2, 3, \ldots, N$. При этом, конечно, предполагается, что $t^{(к)}_{рем}$ не является случайной величиной. Практически $t^{(к)}_{рем}$ все же является некоторой случайной величиной $\eta_к$, зависящей от характера поломки или износа, от технических средств ремонта и квалификации об-

служивающего персонала и т. д. В дальнейшем мы будем оперировать со средним $M\eta_K$ (математическим ожиданием) этой случайной величины, которую коротко обозначим через $t_{рем}^{(к)}$, $K = 1, 2, \ldots, N$.

Пусть машина между ремонтами работает (или частично хранится, т. е. выключена), тогда график изменения надежности при количестве N ремонтов машины имеет примерный вид, изображенный на рис. 2. Следовательно, функция надежности $P(t)$ определена для $t \in [0; t_{1,рем}^{(1)}]$, и далее, для $t \in [t_{2,рем}^{(к)}; t_{1,рем}^{(к+1)}]$, $k = 1, 2, \ldots, N-1$, и для $t \geqslant t_{2,рем}$.

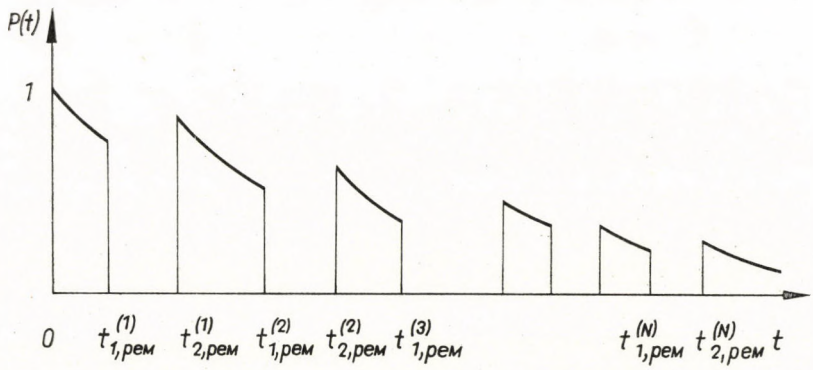


Рис. 2

Практически эти зависимости для каждого вида машины или прибора должны выясняться на основании многочисленных наблюдений и опыта эксплуатации большого парка машин. Анализ ремонтной практики показывает, что для $1 \leq k \leq n < N$ имеет место:

$$P[t_{2,рем}^{(к)}] > P[t_{1,рем}^{(к)}]\,;\ 1 \leq k \leq n < N, \tag{4}$$

для последующих ремонтов с номером ремонта $K > n$ наоборот:

$$P[t_{2,рем}^{(к)}] < P[t_{1,рет}^{(к)}]\,;\ k > n\,, \tag{5}$$

т. е. последняя группа ремонтов даже ведет к уменьшению послеремонтной надежности.

Вообще, надежность $P(t)$ для значений времени t, указанных выше, есть вероятность безотказной (или нормальной работы в заданных условиях) в течение следующих значений времени, принадлежащих следующим отрезкам времени t:

$$\Delta_1 = [0, t] \equiv t, \text{ для } 0 \leq t \leq t_{1,рем}^{(1)};$$

$$\Delta_2\, t_{1,рем}^{(1)} + [t - t_{2,рем}^{(1)}], \text{ для } t_{2,рем}^{(1)} \leq t \leq t_{2,рем}^{(2)};$$

$$\Delta_3 = t^{(1)}_{1,рем} + [t^{(2)}_{1,рем} - t^{(1)}_{2,рем}] + [t - t^{(2)}_{2,рем}];$$

$$\text{для } t^{(2)}_{2,рем} \leq t \leq t^{(3)}_{1,рем};$$

$$\Delta_k \equiv t^{(1)}_{1,рем} + [t^{(2)}_{1,рем} - t^{(1)}_{2,рем}] + [t^{(3)}_{1,рем} - t^{(2)}_{2,рем}] + \ldots +$$

$$+ [t^{(\kappa-1)}_{1,рем} - t^{(\kappa-2)}_{2,рем}] + [t - t^{(\kappa-1)}_{2,рем}], \text{ для } t^{(\kappa-1)}_{2,рем} \leq t \leq t^{(\kappa)}_{1,рем},$$

и, наконец,

$$\Delta_N = t^{(1)}_{1,рем} + \sum_{j=1}^{N-1} [t^{(j+1)}_{1,рем} - t^{(j)}_{2,рем}] + [t - t^{(N)}_{2,рем}],$$

$$\text{для } t \geq t^{(N)}_{2,рем}.$$

Возникает вопрос, можно ли теоретическим или полуэмпирическим путем оценить сверху величину послеремонтной надежности $P[t^{(j)}_{2,рем}]$, $j = 1, 2, \ldots, N$. Оказывается, в отдельных случаях это вполне возможно. Для этого необходимо вначале выяснить схему основного (последовательного) соединения отдельных блоков машины *в смысле надежности*, т. е. такого соединения, при котором выход из строя одного блока (элемента) приводит к выходу из строя всей машины в целом. Пусть в анализе структурной схемы машины было выявлено, что она состоит из m последовательно соединенных блоков (в смысле надежности) с разными надежностями $P_j(t)$, $j = 1, 2, \ldots, m$. Тогда, как известно, надежность $P(t)$ машины будет равна произведению надежностей ее блоков:

$$P(t) = \prod_{j=1}^{m} P_j(t). \tag{6}$$

Рассмотрим для простоты изложения первый профилактический ремонт, который был начат в момент времени $t = t^{(1)}_{1,рем}$ и закончен при $t = t^{(1)}_{2,рем}$. Машина была разобрана, и пусть был сменен только один блок с номером $j = l$, имеющий наименьшую надежность $\min_j P_j(t)$ в момент начала ремонта, т. е. произведено частичное «обновление» машины. Предположим, что качество ремонта в таком простейшем случае было отличным, т. е. соблюдены все требования технологии сборки и т. д. Тогда послеремонтная надежность $P[t^{(1)}_{2,рем}]$ машины равна:

$$P[t^{(1)}_{2,рем}] = P_l[t^{(1)}_{2,рем}] \cdot \prod_{\substack{j=1 \\ j \neq 1}}^{m} P_j[t^{(1)}_{2,рем}]. \tag{7}$$

Если новый l-ный блок был недавно выпущен заводом-изготовителем, то наилучший случай имеет место при принятии $P_l[t^{(1)}_{2,рем}] \approx 1$. В этом случае послеремонтная надежность выше доремонтной надежности:

$$P[t^{(1)}_{1,рем}] > P[t^{(1)}_{2,рем}]. \tag{8}$$

Пусть, для примера для каждого блока справедлив экспоненциальный закон надежности:

$$P_j(t) = e^{-\lambda_j t}, t \geq 0 ; \tag{9}$$

и новый блок был изготовлен еще до начала работы машины, с разницей во времени, равной t_l (t_l невелико), и находился на складе. Тогда в момент окончания ремонта он имеет надежность:

$$P_l[t^{(1)}_{2,рем}] = \exp\left\{-\lambda_l [t_l + t^{(1)}_{2,рем}]\right\}. \tag{10}$$

Тогда имеем:

$$P[t^{(1)}_{2,рем}] = e^{-\lambda_l [t_l + t^{(1)}_{2,рем}] - \sum\limits_{\substack{j=1 \\ j\neq 1}}^{m} \lambda_j t^{(1)}_{2,рем}}. \tag{11}$$

Возьмем несколько иной пример, пусть при (9) поставленный новый блок был взят по-прежнему со склада, и его надежность изменяется по усеченному нормальному закону (для больших t_l), т. е. определенному для $t \geqslant 0$:

$$P_l(t) = \frac{1 - \Phi\left(\dfrac{t-m}{\sigma}\right)}{1 - \Phi\left(-\dfrac{m}{\sigma}\right)}, \tag{12}$$

где $m > 0$, $\sigma > 0$ — некоторые параметры, $\Phi(z)$ — табулированная функция Лапласа [2]:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-v^2/2}\, dv. \tag{13}$$

Тогда

$$P[t^{(1)}_{2,рем}] = e^{-\sum\limits_{j+1, j\neq l}^{m} \lambda_l t^{(1)}_{2,рем}} \cdot \frac{1 - \Phi\left[\dfrac{t_l + t^{(1)}_{2,рем} - m]}{\sigma}\right]}{1 - \Phi\left(-\dfrac{m}{\sigma}\right)}. \tag{14}$$

Аналогично анализируются и другие случаи смены нескольких блоков при первом или последующих ремонтах. Но уже из этих формул видно, что время ремонта $t^{(1)}_{1,рем}$ также играет роль, так как в процессе ремонта у других блоков ($j \neq l$) снижается надежность $P_j(t)$, хотя они находятся в режиме хранения.

Можно привести более сложные формулы, учитывающие, что при хранении на участке $[t^{(1)}_{1,рем}; t^{(1)}_{2,рем}]$ надежность неремонтируемых блоков ($j \neq l$) изменяется не по экспоненциальному закону, а по нормальному закону.

Итак, если на отрезке времени $[0; t^{(1)}_{1,рем}]$ надежность отдельных блоков ($j \neq l$ и $j = l$) изменяется по экспоненциальному закону (9), а на участке $[t^{(1)}_{1,рем}; t^{(1)}_{2,рем}]$ надежность блоков с $j \neq l$ изменяется по усеченному нормальному закону ($t \geq 0$), то нетрудно получить следующее выражение:
для

$$P_{\substack{j \\ j \neq l}}(t) = P(t^{(1)}_{1,рем}) \cdot \frac{1 - \Phi\left(\dfrac{t - m_j}{\sigma_j}\right)}{1 - \Phi\left(\dfrac{t^{(1)}_{2,рем} - m_j}{\sigma_j}\right)}, \tag{15}$$

для $t \leq t^{(1)}_{2,рем}$,
и т. д.

В случае аварийной поломки производится смена нескольких блоков, а также различный ремонт оставшихся блоков. В этом случае ремонтные блоки числом $(m - s)$ надо характеризовать каждый своей послеремонтной надежностью $P_j[t^{(1)}_{2,рем}]$, и тогда для всей машины получаем следующую послеремонтную надежность:

$$P[t^{(1)}_{2,рем}] = \prod_{j=1}^{s} P_j[t^{(\alpha)}_{2\,рем}] \cdot \prod_{j=s+1}^{m} P_j[{}^{(1)}_{2,рем}]. \tag{16}$$

Заметим, что при профилактическом ремонте может также реализоваться случай отсутствия смен блоков, т. е. $S = 0$, а все (или часть блоков) ремонтируются и т. д.

Функциональный вид зависимости надежности после первого ремонта может быть различным, что зависит от многих обстоятельств. Например, можно хорошо аппроксимировать закон надежности в случае небольшого времени $t^{(1)}_{рем}$ и небольшого $t^{(2)}_{1,рем}$ (т. е. в случае начального периода эксплуатации) экспоненциальным законом в виде:

$$P(t) = P[t^{(1)}_{2,рем}] \cdot \mathrm{e}^{-\lambda_{1,рем}[t - t^{(1)}_{2,рем}]};\ t \geq t^{(1)}_{2,рем}. \tag{17}$$

Пусть в этом примере производится всего один ремонт. Тогда долговечность $T_{долгов}$ после одного ремонта определится из (17) (отсчет времени производится, начиная с момента $t = 0$ эксплуатации машины):

$$T_{долгов} = t^{(1)}_{2,рем} + \frac{\left|\ln\left\{a/P[t^{(1)}_{2,рем}]\right\}\right|}{\lambda_{1,рем}}. \tag{18}$$

Если $t^{(1)}_{1,рем}$ и время ремонта $t^{(1)}_{рем}$ большие, то в ряде случаев закон надежности после первого ремонта вполне приемлимо описывается усеченным нормальным законом:

$$P(t) = P[t^{(1)}_{2,рем}] \cdot \frac{\left[1 - \Phi\left(\frac{t-m}{\sigma}\right)\right]}{1 - \Phi\left[\frac{t^{(1)}_{2,рем} - m}{\sigma}\right]}\ ;\ t \geq t^{(1)}_{2,рем}, \tag{19}$$

тде $m > 0$, $\sigma > 0$ — соответствующие параметры. В этом случае долговечность $T_{долгов}$ равна:

$$T_{долгов} = m + \sigma x_p, \tag{20}$$

где xp — квантиль функции распределения (13), отвечающий заданному уровню вероятности p:

$$p = 1 - \frac{a}{P(t^{(1)}_{2\,рем})}\left\{1 - \Phi\left[\frac{t^{(1)}_{2,рем} - m}{\sigma}\right]\right\}. \tag{21}$$

Аналогично можно рассмотреть бóльшее число ремонтов и другие законы изменения надежности. Конечно, величина долговечности $T_{долгов}$ зависит от законов изменения надежности, т. е. в общем случае нет устойчивости количественной характеристики долговечности по отношению к различным законам надежности. Таким путем трактуется надежность ремонтируемых машин и приборов, и в частности, долговечность.

ЛИТЕРАТУРА

1. *Маликов, И. М.—Половко, А. М.—Романов, Н. А.—Чукреев, П. А.:* Основы теории и расчета надежности; изд. 2-е. Судпромгиз, Ленинград 1960.
2. *Крамер, Г.:* Математические методы статистики. ГИИЛ, Москва 1948.

Life and Theory of Reliability *(I. Linkovski-Conde).* The paper deals in detail, from the point of view of the theory of reliability, with the questions related to the life of the machines and the instruments to be repaired.

Lebensdauer und Theorie der Verläßlichkeit *(I. Linkowski-Conde).* Der Aufsatz behandelt im Lichte der Verläßlichkeitstheorie ausführlich die mit der Lebensdauer der auszubessernden Maschinen und Geräte (Instrumente) zusammenhängenden Fragen.

Durée de service et théorie de la sécurité de fonctionnement *(I. Linkovski-Conde).* A la lumière de la théorie de la sécurité de fonctionnement, l'étude traite en détail des questions concernant la durée de service des machines et instruments à réparer.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 227—238 (1968)

SOME QUESTIONS OF THE SYSTEM IDENTIFICATION BY STOCHASTIC SIGNALS*

F. CSÁKI
CORRESPONDENT MEMBER OF THE HUNGARIAN ACADEMY OF SCIENCES
and
J. GYÜRKI
AUTOMATION RESEARCH INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST

The paper shows the difficulties in connection with the practical application of theoretically simple identification methods based on correlation techniques of stochastic signals. The investigations are based on analogue computer and analogue correlator studies.

I. Introduction

The determination of control system parameters, as it is well known from the literature, raises many problems of instrumentation- and system-engineering character [1]. In case of the input of deterministic test signals the magnitude, form and duration of the signals may be chosen as a compromise between the accuracy, the rate of permissible disturbation of the system, the intensity of the additive noises and the allotted evaluation time. In many cases the application of step functions is not permitted. In such cases only the identification by stochastic signals or by statistical calculation can lead to results. The advantages of these methods are, that the informations concerning the static and dynamic characteristics can be obtained without the application of an artificial test signal, just by the observation of the system variables and the evaluation of the observed variations.

The identification of linear systems, or of those, which can be regarded in the region of the operating point as linear, may be performed on the basis of the stochastic processes' time- and frequency-domain characteristics. The auto- and cross-correlation functions contain informations mainly for the characteristics of the time domain (for the weighting function and its parameters), while the power density spectra for those of the frequency domain (for the transfer function and its parameters). It is known, that between these stochastic characteristics mutual transformation relations exist in the linear systems. The measurement, methods and equipment for the determination of the characteristics are also known [2, 3].

In practical investigations the accuracy and other measurement-technical limitations of the now available computing devices (correlators and spectrum analyzers), further the accuracy of calculations performed by the universal analogue and digital computers and other restrictions of the measuring technics

* Original has been included into the program of the *IFAC Symposium* on "Identification in Automatic Control Systems", Prague 1967.

must be taken into consideration. These factors make the doubtlessly existing advantages of the statistical identification processes in many cases practically unrealizable and permit the utilization of the results, the drawing of conclusions only with certain reservations.

This paper wishes mainly to summarize the practically arising problems, the experiences of our measurements, carried out by an actual equipment.

II. The identification processes

In the following $W(p)$ is the symbol of the transfer function of the control system under investigation, subscript i refers to the input signal, subscript o to the output signal, resp. the quantities in connection with them.

The fundamental relationships of the identification based on the characteristics of the frequency domain are [e.g. 2, 3, 9]:

$$\Phi_{co}(j\omega) = |W(j\omega)|^2 \Phi_{ii}(j\omega)\,, \tag{1}$$

$$\Phi_{io}(j\omega) = W(j\omega)\,\Phi_{ii}(j\omega)\,. \tag{2}$$

Because of the measurements of the cross-power-density spectra being problematical, only relationship (1) can be used directly for the practical investigations. The actual spectrum analyzers, by reason of their working principle, allow only the recording of approximate spectra. In devices of continuous operation the accuracy and usefulness of the power density spectra are being made very uncertain by the values of the frequency sweep, the proprieties — diverging from the ideal ones — of the characteristics of the applied band filter and of the averaging low-pass filter, the evaluation, based on stochastic samples with a finite time length. The point-by-point recording of power density spectra gives more accurate results, but then in the case of the evaluation of samples with a satisfactory length the counting time will be considerable.

From the power density spectra obtained by measurement the frequency characteristic $W(j\omega)$ can be determined by the use of (1). From this equation the coefficients of the transfer function can be obtained by iteration methods [3, 4]. The accuracy of these methods is rather bad, in spite of the extensive calculations to be carried out. Also the circumstance, that $\Phi_{oo}(j\,\omega)$ contains the noise spectrum, which arose within the system and which does not originate from the power density spectrum $\Phi_{ii}(j\,\omega)$ is strongly detrimental for the evaluation and the accuracy.

In the practice of control engineering only the identification processes based on the correlation functions are used almost exclusively, due to the fact, that on the one hand data of the time domain characteristics are directly obtained from them, on the other hand they contain numerous other quality characteristics (mean square value, average value) as well. From the correlation func-

tions also the power density spectra can be determined by the help of the Wiener-Khintshin theorem.

From the viewpoint of measurement technics the determination of the correlation functions does not mean any unsolved measuring task. The necessary mathematical operations can be carried out on analogue or digital computers, resp. on special correlators more or less precisely. The analogue or digital principle correlators differ — concerning their basic principle and construction — mainly in the method of realization of the time lag (artificial delaying lines, magnetic band delaying, Stieltjes-correlator, relay-correlator, etc.). From these principles of measurement limitations of the obtainable accuracy arise (e.g. quantization errors) which are being farther increased by the inaccuracies of the operations performed in general by electronic counting elements (multiplier devices diverging from the ideal, finite amplitude and frequency ranges). These errors can be specially evaluated with reference to some actual equipment exactly and their effect may be decreased or kept under some given value by means of more precise counting elements, which is a circuit problem. In many cases also the proper choice of the measurement conditions may help in eliminating them.

The relationships, existing between the correlation functions and the weighting function of the system, which are the basic relationships of the identification, are [e.g. 2, 3, 9]:

$$\varphi_{io}(\tau) = \int_{\infty}^{\infty} w(t)\, \varphi_{ii}(\tau - t)\, dt\,, \tag{3}$$

$$\varphi_{oo}(\tau) = \int_{-\infty}^{\infty} w(t)\, \varphi_{oi}(\tau - t)\, dt\,, \tag{4}$$

$$\varphi_{oi}(\tau) = \int_{-\infty}^{\infty} w(t)\, \varphi_{ii}(\tau + t)\, dt\,, \tag{5}$$

$$\varphi_{oo}(\tau) = \int_{-\infty}^{\infty} w(t)\, \varphi_{io}(\tau + t)\, dt\,, \tag{6}$$

$$\varphi_{oo}(\tau) = \int_{-\infty}^{\infty} w(t_1) \int_{-\infty}^{\infty} w(t_2)\, \varphi_{ii}(\tau - t_2 + t_1)\, dt_2\, dt_1. \tag{7}$$

For identification purposes relationships (3) and (4) as well as (7) seem to be expedient with respect to their simple interpretability.

It is known that the application of the white noise ($\varphi_{ii}(\tau) = \delta(\tau)$) for an input signal allows the direct determination of the weighting function. In this case:

$$\varphi_{io}(\tau) = w(\tau)\,, \tag{3*}$$

$$\varphi_{oo}(\tau) = \int_{-\infty}^{\infty} w(t_1)\, w(\tau + t_1)\, dt_1 = \int_{-\infty}^{\infty} w(t_2 - \tau)\, w(t_2)\, dt_2\,. \tag{7*}$$

But generally this case cannot be brought about, because if the $\varphi_{ii}(\tau)$ correlation functions originate from the observation of the system variables, they aren't of a white noise nature in general. If there is a possibility for the input of an outer disturbation (which decreases the measuring time and improves the reliability of the results) from a noise generator, even this signal can be regarded as "white" only within a finite frequency range, what presents itself as a farther source of error, if we use relationship (3*) and (7*) for the calculation. Therefore it is to the purpose to determine nearly in every case the weighting function $w(t)$ in the knowledge of

$$\varphi_{io}(\tau), \varphi_{ii}(\tau), \varphi_{oo}(\tau)$$

by the deconvolution method based on relationships (3)—(7) [5, 6].

III. Measurement results

The purpose of our investigations is to point out the limitations of the application of relationships (3)—(7). We shall examine mainly the questions of the resolving capability of the statistical methods, the accuracy of the correlation functions obtained by the evaluation of a signal sample with finite length and the relation of both of these. We have built the model of the control elements under investigation on an analogue computer. Similarly instead of by calculations, we determined the principled auto- and cross-correlation functions, referring to the ideal case also with the help of the analogue computer, by model building. These are output auto-, resp. cross-correlation functions belonging to an ideal white noise, or to an input noise of known autocorrelation which would be obtained by averaging of an infinite length. In the case of known system structure these functions may be computed analytically with the help of relationships (3)—(7). We avoided this analytical computation and the graphical representation on the basis of the obtained relationships by having modelled on the analogue computer a derived system, on whose input the application of the auto-correlation function of the input-noise resulted in obtaining for an output signal the auto-correlation function of the output signal, resp. the cross-correlation function of the input and output signals in a directly recordable form. The method of modelling, the interrelation of the derived system with the original one and other measurement-technical questions of the process are given in reference [7].

The correlation functions determined in the above manner (drawn in continuous lines in the enclosed figures) formed the basis of the resolving power investigations and were figuring as control curves to the correlation functions determined by the NORATOM AS ISAC analogue correlator (drawn in dotted lines in the figures).

The question of the resolving capability of the correlation functions is in close relation to the problem of the resolving capability of the weighting function. Under resolving capability the ability of distinguishing or identifying of systems is meant on the basis of the form of auto- and cross-

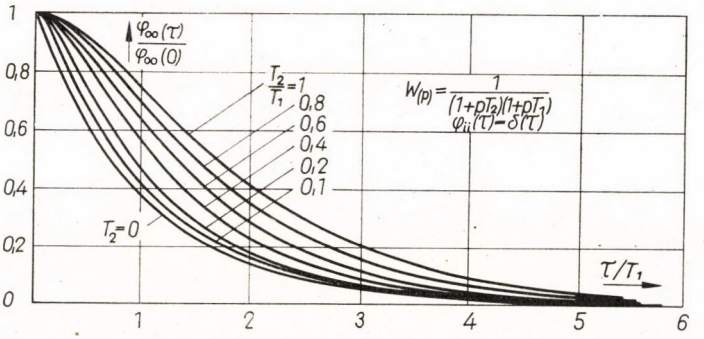


Fig. 1

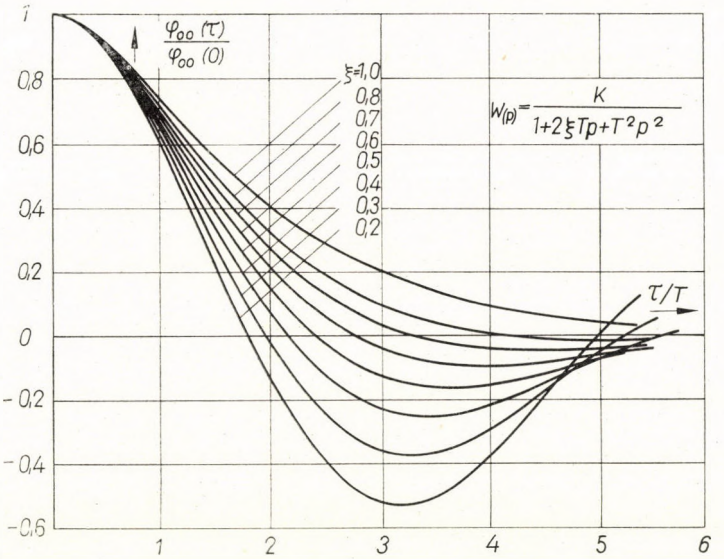


Fig. 2

correlation functions belonging to stochastic input and output signals of systems with parameters of different character and different numerical values and of certain easily determinable macroscopic characteristics (inflection point, points of zero-transition, points and magnitudes of extreme values) [6].

Fig. 1 and 2 show the auto-correlation functions of the output signals with ideal white noise input in the case of proportional two-capacity and oscillating two-capacity elements. The correlation functions are shown in normal-

ized form as parameters of the time constant ratio, resp. of the damping factor ξ

Fig. 3 and 4 show input and output cross-correlation functions belonging to known input noise auto-correlation function in the case of proportional elements with single and two time constants. All four figures contain principled correlation functions, (taken up on an analogue computer) aimed to illustrate the applicability, the resolving capability of relationships (7), resp. (3).

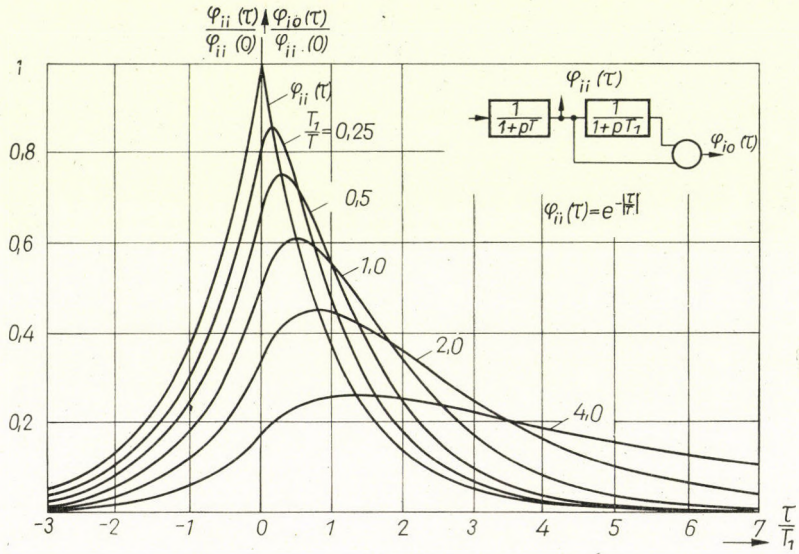


Fig. 3

The figures show, that the method renders a satisfactory resolving capability. In the case of a low τ shift the correlation functions differ strongly, what is advantageous from the measurement-technical viewpoint. In the case of a given system structure, because of the manifold possibilities of the input signal auto-correlation function it would be difficult to construct numerical relationships, practical rules between the system parameters and the macroscopic characteristics (inflection point, magnitude and place of the maximum) of the correlation functions. This was not our aim and is not even necessary, because on the basis of relationships (3)—(7) a digital computer program of arbitrary accuracy can be prepared for producing $w(t)$ by deconvolution from the correlation functions determined by measurements [6].

Another purpose of our investigations was to compare the correlation functions measured on the analogue correlator with the ideal correlation functions and to decide, in what measure the inaccurate measurement deteriorate the resolving capability The accuracy of the correlation functions determined by measurements is influenced by three factors:

a) The errors of the measuring device;

b) the signal sample of finite duration ($T \neq \infty$);

c) the variations of the input noise characteristics.

The errors of the measuring device can be for the most part measured and so their effect can be allowed for. The errors caused by the finite time observation and by the variations of the input noise statistical properties (non ideal and finite white noise) can be reduced by increasing the time of observation.

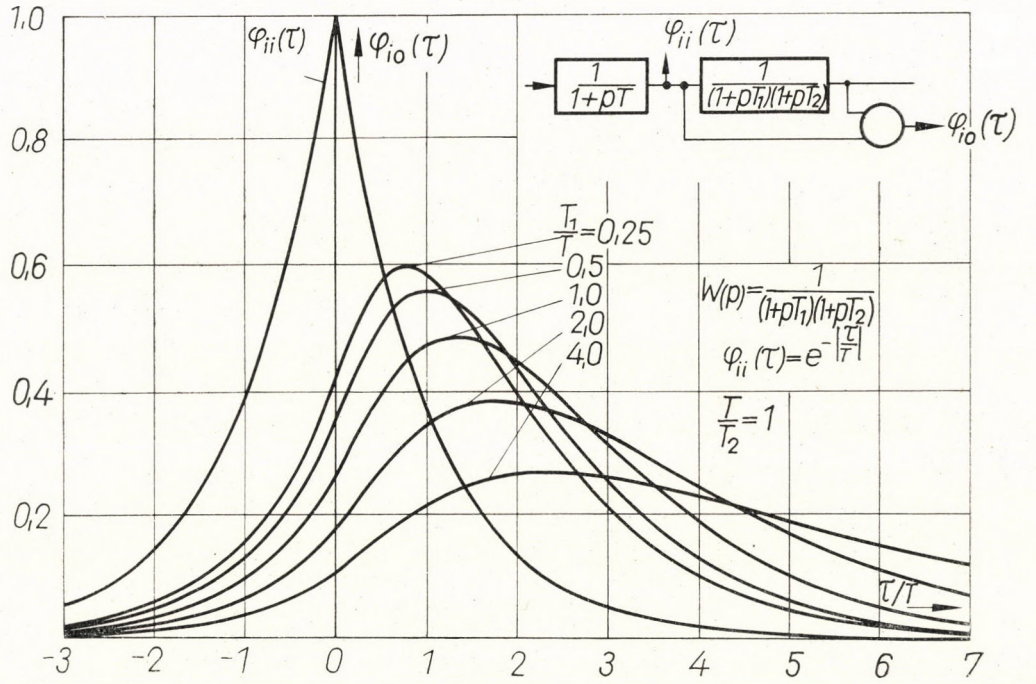


Fig. 4

The method of calculating the errors originating from the observation time being finite, in the case of correlation functions has been known for a relatively long time [3, 5, 8]. The application of these relationships is difficult, especially if we have no "a priori" knowledge whatever about the correlation functions determined by measurements (and about the system).

The scatter of the auto-correlation function value of process $x(t)$ in the case of observation time T is:

$$\sigma^2(\tau) = M\left[\varphi_T^2(\tau) - \varphi^2(\tau)\right] = \frac{1}{T^2} M\left[\left(\int_0^T y(t)\,dt\right)^2 - \varphi^2(\tau)\right] \tag{8}$$

where $y(t) = x(t)\,x(t+\tau)$, and M denotes the averaging.

With processes showing normal distribution the auto-correlation function of $y(t)$ can be computed from the correlation function of $x(t)$. So relationship (8) can be brought to the form more suitable for calculation:

$$\sigma^2(\tau) = \frac{2}{T^2} \int_0^T (T - \Theta) \left[\varphi^2(\Theta) + \varphi(\Theta + \tau)\, \varphi(\Theta - \tau)\right] d\Theta . \tag{9}$$

The application of relationship (9) claims a considerable work of calculation, especially if we assume, that only the points, determined by measurement of $\varphi(\Theta)$ are available. The calculation of error is also disturbed in some

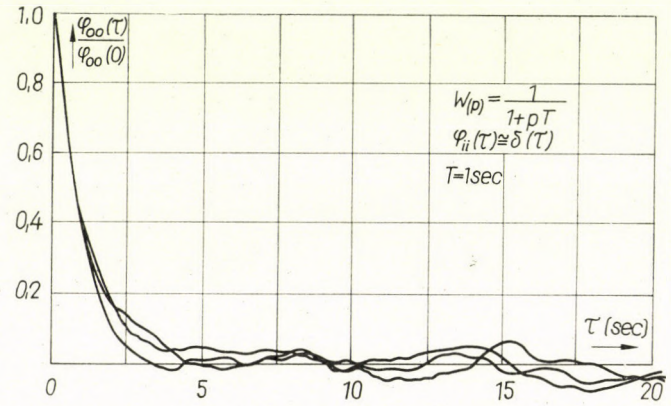


Fig. 5

measure by the fact, that $\varphi(\Theta)$ is inaccurate as well; it is just the scatter of this we wish to determine.

Fig. 5 shows auto-correlation functions determined on the basis of three different samples. In the case of a low value of τ the curves are of identical character and show a good correspondence with the ideal auto-correlation function. In the case of a higher value of τ the functions show a totally random course, what can be related to the finite length sample ($T = 800$ sec) and the properties of the input white noise being divergent from the ideal ones.

Fig. 6 shows output signal auto-correlation functions of systems with different frequency ranges at identical observation times ($T = 800$ sec). According to the figure the variations originating from the finite signal sample and from the changing of the input characteristics influence the resolving capability of the correlation functions little, if the information, which is useful from the viewpoint of the identification can be obtained from the range

$$|\varphi(\tau)| \geq 0{,}1\, \varphi(0) .$$

Fig. 7 shows the input auto-correlation function determined by measurement and the input and output noise cross-correlation function for the case of a proportional element with single time constant. This figure also shows, that

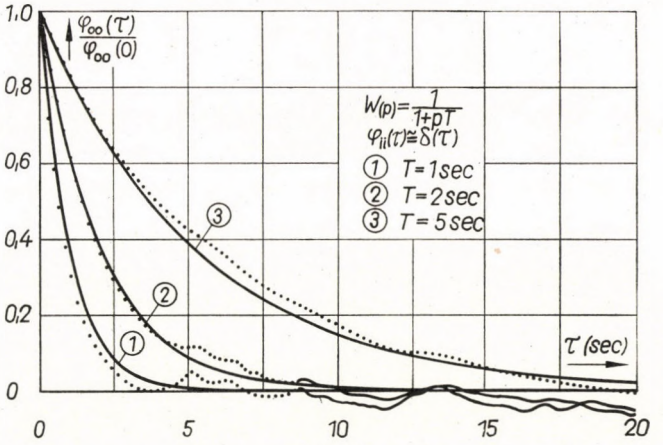


Fig. 6

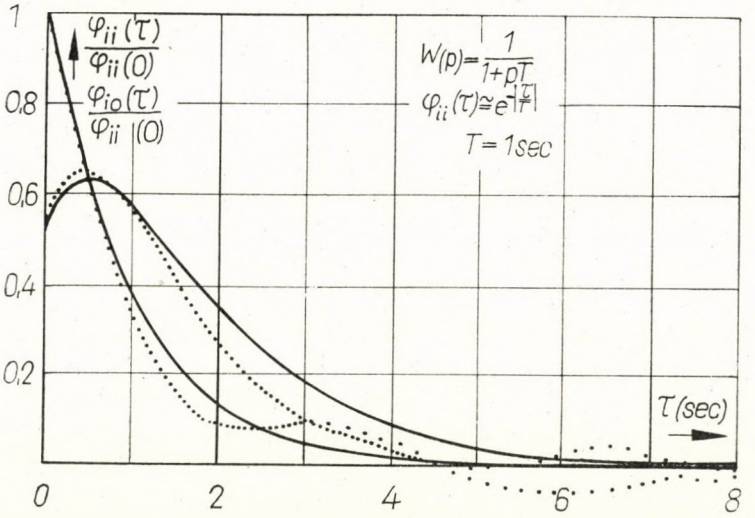


Fig. 7

only the function values, belonging to a low τ shift show such a regular course, which may form the basis of the identification.

Fig. 8 and 9 show the output signal auto-correlation functions of a single time constant proportional and of a double time constant oscillating element in the case of an input signal, which may be regarded as approximately a white noise. Comparison with the ideal curves show, that the variations arising in

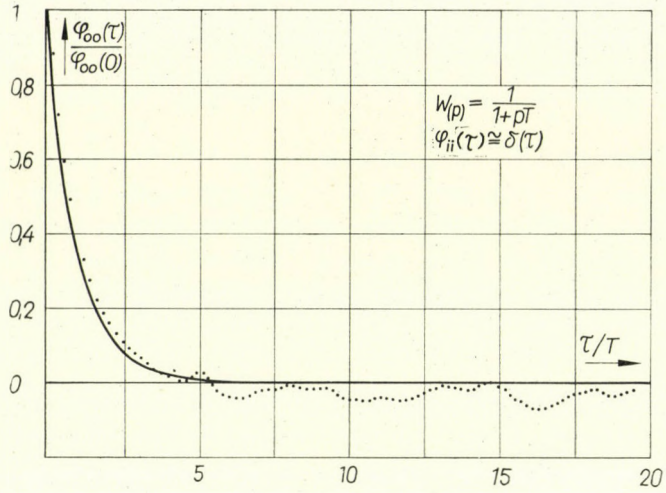


Fig. 8

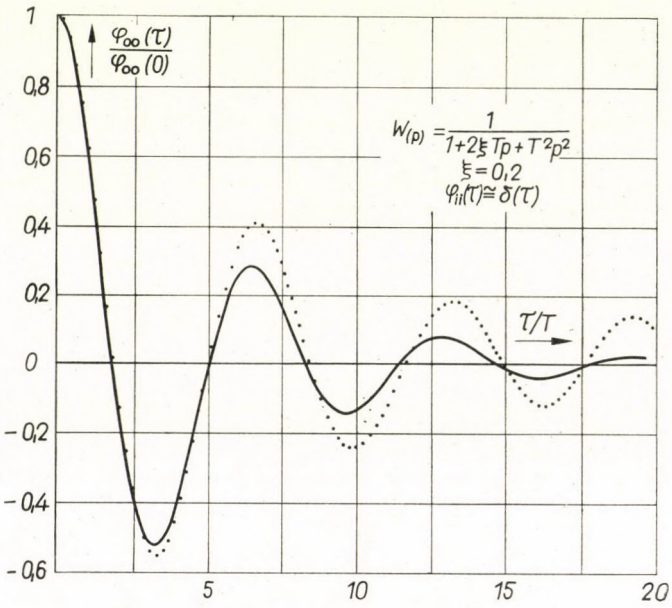


Fig. 9

consequence of the finite signal sample cause a much more important effect and it is mainly this circumstance, which renders the correlation function values belonging to a high τ value unusable.

IV. Conclusions

Relationships (1)—(7) and the literature of the system identification by stochastic signals prove unequivocally the principled possibilities of this method. Theoretically only the mathematical solution of the applying integral equations sets a problem. In practice these possibilities cannot be perfectly utilized by reason of the inaccuracies of the measuring devices, the non-stationary properties of the signals and the calculations based on the evaluation of the finite signal sample. It is these limitations we wished to point out in summarizing the experiences of our measurements, performed on an analogue correlator.

Our measurements proved, that the limitations of measuring-technical character, which lead in the first place to scatters arising in consequence of the evaluation of the finite signal sample, render the identification difficult and deteriorate its accuracy. Insofar as the identification can be performed by the help of the correlation function values belonging to low τ values or if there is a possibility of substituting the part of the function belonging to high τ values by an approximating function, the principled possibility of the statistical methods can be realized practically as well.

REFERENCES

1. Eykhoff, P.—van der Grinteen, P. M. E. M.—Kwakernaak, H.—Veltman, B. P. Th.: Systems modelling and identification, IFAC 1966.
2. Laning, J. H.—Battin, R. H.: Random processes in automatic control. McGraw-Hill 1956.
3. Middleton, D.: An Introduction to Statistical Communication Theory. McGraw-Hill 1960.
4. Senf, B.—Strobel, H.: Verfahren zur Bestimmung von Übertragungsfunktionen linearer Systeme aus gemessenen Werten des Frequenzgangs. Messen, steuern, regeln. 1961.
5. *Солодовников, В. В.:* Статистическая динамика линейных систем автоматического управления. Физмашгиз, Москва 1960.
6. Cuenod, M.—Sage, A.: Comparison of some methods used for process identification. IFAC Symposium, Prague 1967.
7. Gyürki, J.: Univerzális analóg számológép alkalmazása a korreláció analízisben. III. Automatizálási Kollokvium 1966.
8. *Кутин, Б. Н.:* О вычислении корреляционной функции стационарного случайного процесса по экспериментальным данным. *Авт. и тел.* **1** (1957), № 3.
9. Csáki, F.: Szabályozások dinamikája. Akadémiai Kiadó, Budapest 1966.

Einige Probleme der Systemidentifikation mit stochastischen Signalen. Die Arbeit zeigt, welche Schwierigkeiten bei der praktischen Durchführung einer theoretisch einfachen — auf der mit stochastischen Signalen durchgeführten Korrelationstechnik beruhenden — Systemidentifizierung entstehen. Die Untersuchungen beruhen auf der Modellierung mit Analogrechnern von einfachen Übertragungsgliedern und der Verwendung eines Analogkorrelators.

Quelques problèmes de l'identification d'un système par signaux stochastiques. L'étude montre les difficultés que soulève l'exécution pratique d'une identification de système théoriquement simple, basée sur la technique des corrélations par signaux stochastiques. Les examens reposent sur la représentation sur modèle — par calculatrice analogue — de simples membres de transfert et l'utilisation d'un corrélateur analogue.

Некоторые проблемы идентификации системы, выполненной при помощи стохастических сигналов *(Ф. Чаки* и *Й. Дьёрки)*. В статье показано, какие затруднения возникают при практическом осуществлении принципиально простой идентификации системы, основывающейся на корреляционной технике, выполняемой при помощи стохастических сигналов. Исследования основываются на аналоговом ЭВМ моделировании простых переносных членов и использовании коррелятора аналоговой системы.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (1—2), pp. 239—242 (1968)

THE EMISSION DEPENDENCE OF THE LENGTH OF THE CATHODE SPACE AND OF THE POSITIVE COLUMN

E. WINTER

MEMBE R OF THE HUNGARIAN ACADEMY OF SCIENCES
RESEARCH INSTITUTE FOR TECHNICAL PHYSICS OF THE HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST

and

J. BITÓ

INDUSTRIAL RESEARCH INSTITUTE OF ELECTRONICS, "HIKI", BUDAPEST

[Manuscript received December 14, 1967]

With the application of cathodes having different work functions a relation was determined in the case of low-pressure d.c. and a.c. discharges between the termionic emission current of the cathode and the length of the cathodic spaces. It was established that the change of the emission current influencing the volumetric phenomena in the cathode glow space has an effect on the whole cathodic space. The axial length of the Faraday dark space in which the energetic homogenization and maxwellization of electrons take place does not depend on the emission current.

I. Introduction

The task of the cathode-side of the discharges is to generate together with the cathode the necessary electron current for the discharge. This can be considered as a double task [1]: on the one hand, the necessary number of electrons must be generated, on the other hand, the produced electrons must be accelerated to the velocity necessary for the maintenance of the discharge with unchanged parameters. This double task is carried out in coordination with the cathode and cathodic spaces in many cases, however, at the expense of considerable loss of voltage and space. In the following we intend to outline one result of our experiments carried out for the explanation of the connection between the cathodic qualities and cathode space parameters. We sought for a connection between the thermionic emission of the cathode and the axial length of the cathodic spaces.

II. Experimental conditions, measuring methods

Our experiments were carried out in the case of oxide-cathode arc discharges, by d.c. and a.c. feeding. As filling gas argon of 3 torr pressure and beside it mercury of $6 \cdot 10^{-3}$ torr vapour pressure were applied and both the argon and the mercury arc pressure were held at a constant value during the

experiments. The experiments were carried out in the current range between 200—600 mA: but in the following only the results obtained at the constant 400 mA value of d.c. and a.c. discharge are given, namely, in the case of other examined currents similar connections were obtained.

In order to make our experiments unambiguous in common gas field 4—4 oxide cathodes of different emissivity were applied, and the oxide coating was carried on to tungsten double-spirals. Each of these tubes consisted of 3 straight sections converging into one point, and the cathodes of different emissivity were placed at the end of these sections. The diameter of the tubes was 40 mm, and the so formed discharge length — from one electrode to the other — 200 mm. In each discharge system there was one reference electrode at the wedge-shaped peak of the discharge, where the three discharge tubes converged.

The measurements were carried out at d.c. and 50 cps frequency a.c. discharges using numerous discharge systems.

The emission of the different cathodes was measured among conditions corresponding to the discharge parameters. The boundaries of the cathode-spaces were examined in the case of direct current with the application of a suitable background with a cathetometer, in the case of a.c with a stroboscopic-cathetometric system.

III. Results

In the case examined by us, the cathodic side is composed of two, easily separable parts: of the cathode light extending from the surface of the oxide cathode to the Faraday dark space and of the Faraday dark space extending to the cathode-side boundaries of the positive column. Corresponding to this we examined the axial shift of the cathode light, emerging before the cathodes of different emission current, but otherwise of the same structure, and working under the same conditions. The obtained results are shown on Fig. 1, where we gave in the function of i_e emission current the distance of the above-mentioned boundaries from the cathode surface. With the aid of cathodes of different emission coating and activated in a different way we succeeded in finding a connection in a relatively broad emission interval between the length of the cathode spaces and the thermionic current emitted by the cathode. On the basis of the curves on Fig. 1, two essential facts could be established: first it seems that the emission dependence of the cathode-spaces may be divided into two parts having a different character, on the other hand, it could be observed that in the first part — in the emission current range up to 360 mA — the total length of the cathode-spaces arising by the direct current discharge exceeds that of the cathode-spaces of alternating current.

In the lower emission current range — up to 360 mA — the total length of the fields diminishes in both cases (by direct and alternating current discharges) with the growing emission current: the detailed analysis of the curves, however, shows, that the diminution takes place almost exclusively because of the diminution of the axial length of the cathode light and the boundary of the Faraday dark space and together with this the boundary of the positive column is proportionally displaced without the change of the axial length of the Faraday dark space. This is valid in the case of direct and alternating current as well.

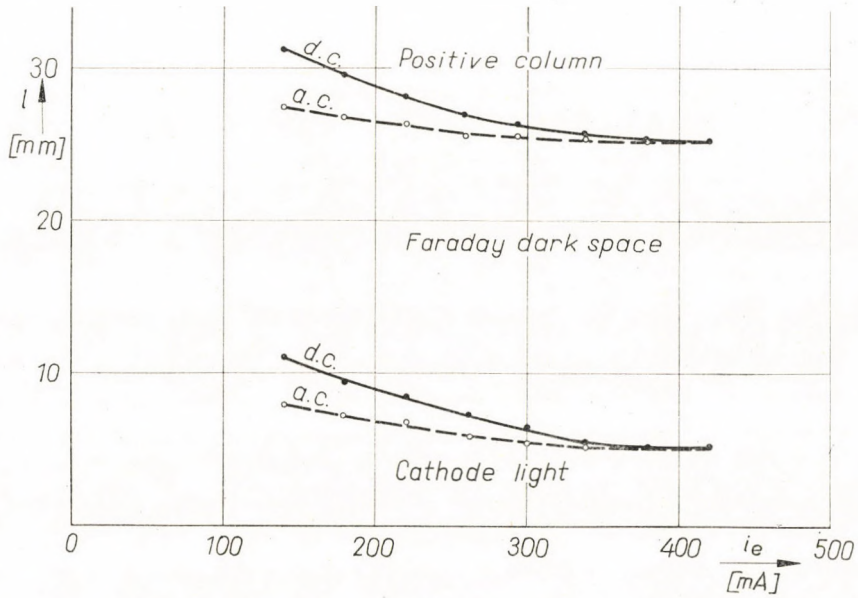


Fig. 1

Taking into account the properties of the cathode-spaces we can conclude that the greater the electron current, emitted by the cathode, at the examined conditions primarily influenced by the axial diminution of the cathode dark space the length of the cathode light range. When from the cathode a greater thermionic electron current is emitted, the length of the space of the cathode fall may diminish, as a smaller volumetric ionization is needed for the generation of the electron current necessary for the discharge. As can be seen this connection is valid only up to a certain emission current: the higher emission currents do not influence the length of the cathode light and the length of the cathode space in it. This minimal axial length is needed for the necessary acceleration of the generated electrons.

The energetic homogenization of the electrons and the formation of their Maxwell—Boltzmann distribution is performed by the Faraday dark

space, and taking this into account, it is clear that the alteration of the cathode emission does not essentially influence the length of this part of the field.

So the total length of the cathode space can be diminished in a certain emission current range by the increasing of the cathode emission through the diminution of the axial length of the cathode light also containing the negative glow space. Corresponding to this the positive column of the discharge may expand towards the direction of the cathode.

REFERENCES

1. Bitó, J.: *HIKI Közlemények* **3** (1966), 19.
2. Cayless, M. A.: *Brit. J. Appl. Phys.* **8** (1957), 331.

Die Emissionsabhängigkeit der Länge der kathodischen Räume und der positiven Säule. Die Verfasser finden einen Zusammenhang zwischen dem Emissionsstrom der Kathode und der Länge der kathodischen Gebiete in dem Falle einer Gleich- und Wechselstrom-Niederdruckentladung, bei Verwendung von Kathoden verschiedener Austrittsarbeit. Sie beweisen, daß die Änderung des Emissionsstromes durch die Beeinflussung der Volumserscheinungen in dem Kathodenlicht-Gebiet eine Wirkung auf den gesamten Kathodenraum ausübt. Die Axiallänge des Faraday-Dunkelraumes, der die energetische Homogenisation und Maxwellisation der Elektronen hervorruft, hängt nicht von der Größe des Emissionsstromes ab.

La dépendance de l'émission des espaces cathodiques et de la longueur de la colonne positive. Les auteurs utilisant des cathodes à différentes énergies de sortie déterminent une relation entre le courant d'émission cathodique et les dimensions des espaces cathodiques, en cas de décharges à courants continu et alternatif sous basse pression. Ils montrent que la variation du courant d'émission agit sur tout l'espace cathodique en influençant les phénomènes spatiaux dans le domaine lumineux cathodique. La dimension axiale de l'espace sombre de Faraday, effectuant l'homogénéisation énergétique et la maxwellisation des électrons ne dépend pas de la valeur du courant d'émission.

Эмиссионная зависимость длин катодных пространств и положительных столбцов *(Э. Винтер* и *Я. Бито)*. Применением катодов разной работы выхода, в случае разряда низкого давления постоянного и переменного тока, устанавливается соотношение между током эмиссии и размером катодных пространств. Показано, что изменение тока эмиссии имеет влияние на объемные процессы в пространстве катодного света, и поэтому действует во всем катодном пространстве. Размер Фарадеевого темного пространства, вызывающего энергетическую гомогенизацию и максвеллизацию электронов, не зависит от величины тока эмиссии.

INDEX

Huszár, I.: Berechnung des Kräftespiels von Schwalbenschwanzkommutatoren — Calculation of the System of Forces of Arch-Bound Commutators — Calcul du jeu de forces des collecteurs à queue d'aronde — *Хусар, И.:* Расчет работы коллекторов в виде ласточкина хвоста 1

Хатвани, Й.—Кардош, П.—Барат, Я.: Применение цифровых аналогов в устройствах промышленной автоматики — *Hatvany, J. — Kardos, P. — Barát, J.:* The Use of Counting Technique Instruments in Industrial Automation Equipments — Anwendung zähltechnischer Mittel bei industriellen Automatik-Anlagen — L'emploi des techniques opérationnelles dans les systèmes d'automatisation industrielle ... 31

Gertler, J.: Some Concepts on the Programming of Direct Digital Process Control Systems — Einige Probleme der Programmierung unmittelbarer Digital-Prozeß-Steuersysteme — Quelques problèmes de la programmation des systèmes de commande digitale directe des processus — *Гертлер, Я.:* Некоторые вопросы программирования систем непосредственного цифрового управления процессами... 55

Линковский, Ж. Б.: Средняя интенсивность отказов на заданном интервале времени работы аппаратуры — *Linkovski-Conde, J.:* The Mean Intensity of Instrument Breakdown in a Given Time Interval — Durchschnittliche Intensität des Schadhaftwerdens von Instrumenten in einem gegebenen Zeitintervall — L'intensité moyenne des pannes des instruments dans un intervalle de temps donné 67

Шомло, Я.: Исследование устойчивости цифрового регулятора — *Somló, J.:* Investigation of the Stability of a Digital Controller — Untersuchung der Stabilität eines digitalen Reglers — Examen de la stabilité d'un régulateur digital ... 75

Kemény, A. P.: Effect of Structural Irregularities on Avalanche Breakdown in p — n Step Junctions — Der Einfluß der strukturellen Irregularitäten von p — n Übergängen auf die volumetrische Lawinendurchbruchspannung — L'influence des irrégularités structurales des jonctions p — n abruptes sur la tension de rupture en avalanche spatiale — *Кемень, А. П.:* Воздействие нерегулярности структуры абруптных p — n переходов на объемное напряжение лавинного пробоя 101

Korach, M.—Fülöp, J.: Un effet thermique des fours-tunnel — A Heat Effect in Continuous Furnaces — Ein wärmetechnischer Effekt der Tunnelöfen — *Корах, М. — Фюлёп, Я.:* Один теплотехнический эффект туннельных печей 137

Gádor, L.: Circuits with Impedances Having Z-Characteristics — Aus Impedanzen mit Z-Kennlinien aufgebaute Stromkreise — Circuits avec impédances à caractéristiques Z — *Гадор, Л.:* Цепи с импедансами, имеющими Z-характеристику... 155

Gádor, L.: Quelques points de vue pour le réglage de la tension du réseau — Some Points of View for the Regulation of the Mains Voltage — Einige Gesichtspunkte für die Regelung der Netzspannung — *Гадор, Л.:* Некоторые принципы регулирования сетевого напряжения 177

Lőrinczy, A.—Németh-Sallay, M.—Szép, I. C.—Tihanyi, J.: The Effect of Quinone on the Surface Properties of Germanium — Einfluß von Chinon auf die Oberflächeneigenschaften von Germanium — L'influence de la quinone sur les qualités de surface du germanium — *Лёринци, А.—Немет-Шаллаи, М.—Сеп, И. Ц.—Тихани, Й.:* Воздействие хинона на поверхностные свойства германия 189

Tevan, Gy.: Calculation of the Coefficient of Self-Induction and of the Impedance of Cylindrical Coils with Open Iron Core — Berechnung des Selbstinduktionskoeffizienten bzw. der Impedanz von zylindrischen Spulen mit offenem Eisenkern — Calcul du coefficient de self-induction et de l'impédance des bobines cylindriques à noyau de fer ouvert — *Теван, Дь.:* Расчет коэффициента самоиндукции или же импеданса цилиндрических обмоток с незамкнутым сердечником 201

Линковский, Ж. Б.: Долговечность и теория надежности — *Linkovski-Conde, J.:* Life and Theory of Reliability — Lebensdauer und Theorie der Verlässlichkeit — Durée de service et théorie de la sécurité de fonctionnement 219

Csáki, F.—Gyürki, J.: Some Questions of the System Identification by Stochastic Signals — Einige Probleme der Systemidentifikation mit stochastischen Signalen — Quelques problèmes de l'identification d'un système par signaux stochastiques — *Чаки, Ф.—Дьёрки, Й.:* Некоторые проблемы идентификации системы, выполненной при помощи стохастических сигналов 227

Winter, E.—Bitó, J.: The Emission Dependence of the Length of the Cathode Space and of the Positive Column — Die Emissionsabhängigkeit der Länge der kathodischen Räume und der positiven Säule — La dépendance de l'émission des espaces cathodiques et de la longueur de la colonne positive — *Винтер, Э.—Бито, Я.:* Эмиссионная зависимость длин катодных пространств и положительных столбцов 239

Anhaltszahlen für die Wärmewirtschaft in Eisenhüttenwerken

Herausgegeben von der Energie- und Betriebswirtschaftsstelle des Vereins Deutscher Eisenhüttenleute (Wärmestelle Düsseldorf)

6., völlig neu bearbeitete Auflage 1968

Format DIN C 5. XIII/601 Seiten mit 344 Bildern und 392 Zahlentafeln

In Ganzleinen DM 132,—

Seit 45 Jahren das Standardwerk für Ingenieure aller Fachrichtungen, die mit den vielfältigen Fragen hüttenmännischer Energiewirtschaft zu tun haben. Die Weiterentwicklung der hüttenmännischen Wärmewirtschaft in den letzten 10 Jahren ist gekennzeichnet durch neue Stahlherstellungsverfahren, verbesserte Wärm- und Wärmebehandlungsöfen und den verstärkten Einsatz heizwertreicher Brennstoffe. Eine besondere energiewirtschaftliche Rolle spielt der Sauerstoff in den Sauerstoffaufblas-Konvertern bei der Stahlherstellung. Die Bedeutung der Wärmebehandlungsverfahren wuchs mit den steigenden Qualitätsansprüchen. Diesen Veränderungen in der Eisenhüttentechnik trägt die Neuauflage der »Anhaltszahlen« durch entsprechende Schwerpunkte in der Stoffauswahl Rechnung.

Stahl-Eisen-Liste

Liste der in der Bundesrepublik Deutschland hergestellten Eisenwerkstoffe

Herausgegeben vom Verein Deutscher Eisenhüttenleute

2. Auflage 1967. Format DIN B 5. VIII/436 Seiten. In Ganzleinen DM 68,—

Die STAHL-EISEN-LISTE hat zur Aufgabe

a) in die Vielfalt der Eisenwerkstoffe, die unter den verschiedensten Namen laufen, Ordnung zu bringen,

b) den nach Gebrauchseigenschaft, Erschmelzung und Verarbeitung günstigsten Werkstoff für die Verwendung festzulegen,

c) zur Verringerung der Werkstoffarten beizutragen,

d) Grundlagen für einheitliche Lieferbedingungen zu schaffen,

e) eine einheitliche Bezeichnung herbeizuführen.

Erzeugern und Verbrauchern, Ingenieuren und Kaufleuten, überhaupt allen, die mit Eisen und Stahl zu tun haben, wird auch die neue Auflage der STAHL-EISEN-LISTE ein unentbehrliches Nachschlagewerk sein.

VERLAG STAHLEISEN M. B. H., 4 DÜSSELDORF
POSTFACH 8229

verfahrenstechnik

Ein für den Fachmann klarer, eindeutiger Titel. Ein Programm — so umfassend, wie die technisch-wissenschaftliche Disziplin Verfahrenstechnik bedeutend ist, da sie die vielfältigen physikalischen und chemischen Prozesse der Stoffbehandlung und Stoffumwandlung in der gesamten Grundstoffindustrie sowie in zahlreichen Zweigen der verarbeitenden Industrie umfaßt.

Diese neue — und vornehmlich anwendungstechnische — Zeitschrift fördert

die grösstmögliche und ökonomisch beste verfahrenstechnische Nutzanwendung in der gesamten Industrie.

Eine solche Zeitschrift will genutzt sein! Wir empfehlen daher: Sichern Sie sich im Abonnement die regelmäßige Belieferung sofort nach Erscheinen. Sie entbehren dann keiner und vor allem nicht aktueller Information. Der

Abonnementspreis

ist, gemessen an der Qualität des Gebotenen, außerordentlich günstig. 12 Hefte jährlich kosten DM 54,—, zuzüglich Porto.

KRAUSSKOPF-VERLAG · D6500 MAINZ · LESSINGSTRASSE 12—14

Printed in Hungary

A kiadásért felel az Akadémiai Kiadó igazgatója — Műszaki szerkesztő: Farkas Sándor

A kézirat nyomdába érkezett: 1968. I. 9. — Terjedelem: 21,50 (A/5) ív, 120 ábra, 5 melléklet

68.64899 Akadémiai Nyomda, Budapest — Felelős vezető: Bernát György

Acta Techn. Hung. **61** (1968), 1—30

HUSZÁR, I.: *Calculation of the System of Forces of Arch-Bound Commutators*

The mechanical model of the arch-bound commutator can be represented by two one-dimensional spring systems. With the one the radial forces and deformations are considered, with the other the axial ones. The paper gives the relations by which the spring rates of the individual spring elements are determined. By clarifying the calculating method of the spring system the forces corresponding to the various states occurring in commutator service (assembly, rotation, warming, rotating warm state) may be determined. The stresses stated thus supply the data needed for dimensioning the commutator.

Acta Techn. Hung. **61** (1968), 31—54

HATVANY, J.—KARDOS, P.—BARÁT J.: *Technique Instruments in Industrial Automation Equipments*

After clarifying the concepts the paper deals with the dimensioning problems of the combined system. A counting procedure checked by reproducible measurements on an experimental installation is proposed instead of the completely empirical methods which were proposed earlier. For dimensioning and checking more complicated systems the only practicable way is simulation and evaluation carried out on a digital computer. The paper deals with the problem of a generally valid simulation program based on the employment of the theory of finite automates. It discusses the theoretic problems of the application of the chosen principle and covers its proof for the case of systems of the given type. The experimental checking of the computer simulation procedure was carried out through the simulation of the measuring arrangement outlined earlier. The results delivered by the simulation program give the same values as the results gained by direct messurement.

Acta Techn. Hung. **61** (1968), 55—66

GERTLER, J.: *Some Concepts on the Programming of Direct Digital Process Control Systems*

The paper deals with some programming problems of the centralized direct digital control of continuous industrial processes. It shows that the program language ALGOL 60 can be extended for process control purposes and introduces the concept of changing classes and changing forms. As standard procedure of the extended program language it summarizes the fundamental operations of direct digital process control and it contains considerations regarding the optimum form of the standard procedures. In order to simplify programming it introduces as a further extension of ALGOL 60 the concept of latent and substitution parameters and demonstrates the use of the extended language on a sample program.

Acta Techn. Hung. **61** (1968), 67—74

LINKOVSKI-CONDE, J.: *The Mean Intensity of Instruments Breakdown in a Given Time Interval*

The author introduces a new function for the reliability of an instrument which gives the mean intensity of breakdown for a given time interval. He deduces its relation with the other reliability functions and determines the confidence interval of the real values of the function. He examines the mean reliability for a given time interval and deals with its evaluation on the base of sampling.

Acta Techn. Hung. **61** (1968), 75—100

SOMLÓ, J.: *Investigation of the Stability of a Digital Controller*

The paper deals with the investigation of the stability of a non-linear digital controller by the describing function method. By simplifications it is possible to determine the analytical relations which serve to select the parameters guaranteeing stable working. Whether the selected parameters are suitable or not can be checked with the aid of the securate describing functions. In the paper the earlier results of the author are used. The investigation is connected with a definite problem of application.

Acta Techn. Hung. **61** (1968), 101—136

KEMÉNY, A. P.: *Effect of Structural Irregularities on Avalanche Breakdown in p—n Step Junctions*

Besides surface phenomena the avalanche breakdown voltage of transistors and diodes is reduced by geometric or resistivity irregularities along the junction. Recent theories based on the one-dimensional solution of the Poisson-equation do not hold for the practical case when a hemispherical or hemicylindrical curvature is "embedded" in the planar section of the junction. Though the problem is two-dimensional and seems not accessible to analytical solution, an approximation based on the known interrelation between the volumetric breakdown voltages BV_{CVO} and BV_{CEO} of transistors has been stated. High-current pulsed measurements eliminate disturbances due to surface avalanche. It was found that the higher the resistivity of the light doped side of the junction is (i. e. by high voltage devices), the larger is the breakdown-voltage reducing effect of volumetric irregularities; secondly, the breakdown-voltage descreases most if the radius of the embedded curvature (or irregularity) amounts to appr. one fifth of the depletion layer width.

Acta Techn. Hung **61** (1968), 137—154

KORACH, M.—FÜLÖP, J.: *A Heat Effect in Continuous Furnaces*

The authors establish the differential equations for the temperature distribution in continuous furnaces. The furnace is considered as a series of three units and thus a system of three times two differential equations is obtained, which are connected by suitable conditions. The differential equations are approximately solved by the relaxation method and the results are presented in the form of diagrams. The diagrams agree well with the results of experiences carried out in the laboratory and in operation and with increasing pushing velocity the core temperature increasingly lags behind the surface temperature. The method permits the approximate calculation of the temperature curves in continuous furnaces.

Acta Techn. Hung. **61** (1968), 155—176

GÁDOR, L.: *Circuits with Impedances Having Z-Characteristics*

A group of non-linear elements with extreme properties are those passive circuit elements between whose characteristic data the relation may be expressed by Z-type diagrams. Their examination shown in the paper brings into evidence qualitative relations which facilitate the handling of non-linear circuits.

Acta Techn. Hung. **61** (1968), 177—188

GÁDOR, L.: *Some Points of View for the Regulation of the Mains Voltage*

In electrical distribution networks it is a condition of quality and is also economically justified to reduce by regulation the voltage fluctuation in the different points of the network. The electrical energy consumed at voltages differing from the normal one is detrimental to economic use; the damage depends on the quantity of energy consumed at out-of-normal voltage and on the voltage deviation. The criterion for the voltage regulation can be determined by opposing, on the basis of the statistics of voltage variations, the damage caused by the voltage deviation and the costs of voltage regulation. The characteristic of the type ATSZ regulator developed by the Research Institute of Electrical Industry is conform to the laws determined in this way.

Acta Techn. Hung. **61** (1968), 189—200

LŐRINCZY, A.—NÉMETH-SALLAY M.—SZÉP, I. C.—TIHANYI, J.: *The Effect of Quinone on the Surface Properties of Germanium*

Experiments were made on *p-n-p* germanium transistors and examples to study the effect of quinone. The dependence of reverse-current and current-gain on the surface potential of transistors was investigated by means of the emitter and collector field effect. After quinone treatment the reverse-current characteristics improved, while the current-leakage, noise-factor and instability were reduced. The surface potential was shifted towards the positive. For quantitative checking of the physical properties of the surface, germanium samples were examined by field-effect measurements. Based on results of two different measurements suggestions are given relating to the processes taking place on the surface.

Acta Techn. Hung. **61** (1968), 201—218

TEVAN, GY.: *Calculation of the Coefficient of Self-Induction and of the Impedance of Cylindrical Coils with Open Iron Core*

The paper gives an approximation formula for the calculation of the coefficient of self-induction and the impedance of symmetrically arranged cylindrical coils with open iron cores. The exterior magnetic field of the coil is replaced by the exterior field of an ellipsoid of rotation whith convenient excitation and the excitation consumed by the air is determined on this basis.

Acta Techn. Hung. **61** (1968), 219—226

LINKOVSKI-CONDE, J.: *Life and Theory of Reliability*

The paper deals in detail, from the point of view of the theory of reliability, with the questions related to the life of the machines and the instruments to be repaired.

Acta Techn. Hung. **61** (1968), 227—238

CSÁKI, F.—GYÜRKI, J.: *Some Questions of the System Identification of Stochastic Signals*

The paper shows the difficulties in connection with the practical application of theoretically simple indenfication methods based on correlation techniques of stochastic signals. The investigations are based on analogne computer and analogue correlator studies.

Acta Techn. Hung. **61** (1968), 239—242

WINTER, E—BITÓ, J.: *The Emission Dependence of the Length of the Cathode Space and of the Positive Column*

With the application of cathodes having different work function a relation was determined in the case of low-pressure dc. and ac. discharges between the termionic emission current of the cathode and the length of the cathodic spaces. It was established that the change of the emission current influencing the volumetric phenomena in the cathode glow space has an effect on the whole cathodic space. The axial length of the Faraday dark space in which the energetic homogenization and maxwellization of electrons take place does **not** depend on the emission current.

The Acta Technica publish papers on subjects of technical sciences in English, German, French and Russian.

The Acta Technica appear in parts of varying size, making up volumes.

Manuscripts should be addressed to

Acta Technica
Nádor u. 7.
Budapest V.
Hungary

Correspondence with the editors and publishers should be sent to the same address.

The rate of subscription to the Acta Technica is 165 forints a volume. Orders may be placed with "Kultúra" Foreign Trade Company for Books and Newspapers (Budapest I., Fő utca 32. Account No. 43-790-057-181) or with 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
Nádor u. 7.
Budapest V.
Hongrie

Toute correspondance doit être envoyée à cette même adresse.

Le prix de l'abonnement est de 165 forints par volume.

On peut s'abonner à l'Entreprise pour le Commerce Extérieur de Livres et Journaux «Kultúra» (Budapest I., Fő utca 32. Compte courant No. 43-790-057-181) ou à l'étranger chez tous les représentants ou dépositaires.

«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 рукописи следует направлять по адресу:

Acta Technica
Nádor u. 7.
Budapest V.
Венгрия

По этому же адресу направлять всякую корреспонденцию для редакции и администрации.

Подписная цена «Acta Technica» — 165 форинтов за том. Заказы принимает Предприятие по внешней торговле книг и газет «Kultúra» (Budapest I., Fő utca 32. Текущий счет № 43-790-057-181) или его заграничные представительства и уполномоченные.

94,— Ft

Reviews of the Hungarian Academy of Sciences are obtainable at the following addresses:

ALBANIA
Ndermarja Shtetnore e Botimeve
Tirana

AUSTRALIA
A. Keesing
Box 4886, GPO
Sydney

AUSTRIA
Globus Buchvertrieb
Salzgries 16
Wien I

BELGIUM
Office International de Librairie
30, Avenue Marnix
Bruxelles 5
Du Monde Entier
5, Place St. Jean
Bruxelles

BULGARIA
Raznoiznos
1, Tzar Assen
Sofia

CANADA
Pannonia Books
2, Spadina Road
Toronto 4, Ont.

CHINA
Waiwen Shudian
Peking
P. O. B. 88

CZECHOSLOVAKIA
Artia
Ve Směčkách 30
Praha 2
Poštovná novinová služba
Dovoz tisku
Vinohradská 46
Praha 2
Mad'arská Kultura
Václavské nám. 2
Praha I
Poštová novinová služba
Dovoz tlače
Leningradská 14
Bratislava

DENMARK
Ejnar Munksgaard
Nörregade 6
Copenhagen

FINLAND
Akateeminen Kirjakauppa
Keskuskatu 2
Helsinki

FRANCE
Office International de Documentation et Librairie
48, rue Gay Lussac
Paris 5

GERMAN DEMOCRATIC REPUBLIC
Deutscher Buch-Export und Import
Leninstraße 16
Leipzig 701
Zeitungsvertriebsamt
Clara Zetkin Straße 62
Berlin N. W.

GERMAN FEDERAL REPUBLIC
Kunst und Wissen
Erich Bieber
Postfach 46
7 Stuttgart S.

GREAT BRITAIN
Collet's Holdings Ltd.
Dennington Estate
London Rd.
Wellingborough. Northants.
Robert Maxwell and Co. Ltd.
Waynflete Bldg. The Plain
Oxford

HOLLAND
Swetz and Zeitlinger
Keizersgracht 471—487
Amsterdam C
Martinus Nijhof
Lange Voorhout 9
The Hague

INDIA
Current Technical Literature
Co. Private Ltd.
India House OPP
GPO Post Box 1374
Bombay I

ITALY
Santo Vanasia
Via M. Macchi 71
Milano
Libreria Commissionaria Sansoni
Via La Marmora 45
Firenze

JAPAN
Nauka Ltd.
92, Ikebukur O-Higashi 1-chome
Toshima-ku
Tokyo
Maruzen and Co. Ltd.
P. O. Box 605
Tokyo-Central
Far Eastern Booksellers
Kanda P. O. Box 72
Tokyo

KOREA
Chulpanmul
Phenjan

NORWAY
Johan Grundt Tanum
Karl Johansgatan 43
Oslo

POLAND
RUCH
ul. Wronia 23
Warszawa

ROUMANIA
Cartimex
Str. Aristide Briand 14—18
București

SOVIET UNION
Mezhdunarodnaya Kniga
Moscow G—200

SWEDEN
Almquist and Wiksell
Gamla Brogatan 26
Stockholm

USA
Stechert Hafner Inc.
31. East 10th Street
New York, N. Y. 10003
Walter J. Johnson
111, Fifth Avenue
New York, N. Y. 10003

VIETNAM
Xunhasaba
19, Tran Quoc Toan
Hanoi

YUGOSLAVIA
Forum
Vojvode Mišića broj 1
Novi Sad
Jugoslovenska Knjiga
Terazije 27
Beograd

10. V. 1968. Index: 26.025

ACTA TECHNICA

ACADEMIAE SCIENTIARUM HUNGARICAE

ADIUVANTIBUS
I. BARTA, E. BÖLCSKEI, P. O. GESZTI, A. LÉVAI

REDIGIT
M. MAJOR

TOMUS 61 FASCICULI 3—4

AKADÉMIAI KIADÓ, BUDAPEST
1968

ACTA TECHNICA

A MAGYAR TUDOMÁNYOS AKADÉMIA MŰSZAKI TUDOMÁNYOK OSZTÁLYÁNAK KÖZLEMÉNYEI

SZERKESZTŐSÉG ÉS KIADÓHIVATAL: BUDAPEST V., ALKOTMÁNY U. 21.

Az Acta Technica német, angol, francia 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
Budapest V., Nádor u. 7.

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Az Acta Technica előfizetési ára kötetenként belföldre 120 forint, külföldre 165 forint. Megrendelendő a belföld számára a „Akadémiai Kiadó"-nál (Budapest V., Alkotmány utca 21. Bankszámla 05-915-111-46), a külföld számára pedig a „Kultúra" Könyv- és Hírlap Külkereskedelmi Vállalatnál (Budapest I., Fő utca 32. Bankszámla: 43-790-057-181) vagy annak külföldi képviseleteinél és bizományosainál.

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

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

Die zur Veröffentlichung bestimmten Manuskripte sind an folgende Adresse zu senden

Acta Technica
Nádor u. 7.
Budapest V.
Ungarn

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

Abonnementspreis pro Band: 165 Forint, Bestellbar bei dem Buch- und Zeitungs-Außenhandels-Unternehmen »Kultúra« (Budapest I., Fő utca 32. Bankkonto Nr. 43-790-057-181) oder bei seinen Auslandsvertretungen und Kommissionären.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (3–4), pp. 243—260 (1968)

THE YIELD STRENGTH IN HOMOGENEOUS ALLOYS OF α-IRON

J. A. VERŐ

MEMBER OF THE HUNGARIAN ACADEMY OF SCIENCES

[Manuscript received September 11, 1967]

Making use of existing concepts to explain the strengthening effect of solute atoms and of experimental results of previous authors, an attempt was made to establish a systematic connexion between atomic dimensions and the strengthening effect of the alloying elements of α-iron. In substitutional solid solutions of α-iron the specific strengthening effect, expressed as the increase of the yield strength in kp/mm²-s produced by one atomic % alloying element was found to amount to $130 \cdot (\Delta D)^2$ or eventually $100(\Delta D)^2$. The experimental data, however, show a considerable scatter, probably as a consequence of secondary effects which were not taken into account, especially in early experiments. Interstitial solute atoms, chiefly carbon in ferrite and martensite displace the nearest iron atoms and distort the slip planes in a similar manner, as substitutional solutes do. Substituting the distance of displacement into the formula $130(\Delta D)^2$ it is found that one atomic % dissolved carbon raises the yield strength of α-iron by 87 kp/mm² which is in excellent agreement with the existing experimental results. The very same strengthening effect is obtained, when the carbon atoms are assumed to occupy octahedral or tetrahedral sites in the lattice of α-iron; these are the possible sites of small interstitial atoms in the generally accepted lattice model of martensite. The experimental data seem to support Fleischer's concept on the influence of carbon on the flow stress and yield strength, predicting a proportionality of the flow stress to the square root of the dissolved carbon content.

I. Introduction

Solute atoms inserted into the lattice of a metal occupy a greater or smaller space, than those of the base metal and therefore both the lattice and all the planes of the solvent metal become distorted around the site of the foreign atoms. Slip on distorted planes, however, requires a higher stress, as if some additional frictional resistance were to be surmounted. In other words, solute atoms increase the critical shear stress of the base metal and all other mechanical properties connected with this stress, the limit of elasticity, the yield strength, but the tensile strength and the hardness also. It is easily understandable that the change of all these properties must be connected with the size and number of local lattice distortions, i.e. with the dimensions and the quantity of solute atoms.

A relatively simple relation could be established between the dimensions and quantity of solute atoms and the increase of the critical shear stress for homogeneous alloys of some metals, chiefly for those of copper [1]. The effect of solute atoms on mechanical properties of α-iron was experimentally examined,

several times over, but no attempt was made to systematize the measured values. On the contrary, a few years ago it was concluded [2] that "*There appears to be no systematic influence of atomic number or of the change of lattice parameter produced by the solute on its strengthening effect and the cause of the differences must be sought in the local reactions between solvent and solute.*" This conclusion clearly indicates that the interpretation of the existing experimental results obtained with ferrous alloys must offer some difficulties.

A special difficulty arises from the circumstance that small atoms such as carbon and nitrogen dissolve interstitially in α-iron. Though these atoms also produce a distortion of the lattice, in the homogeneous phases of α-iron containining these elements, in ferrite and martensite, further strengthening mechanisms can also be active in raising the critical shear stress and the yield strength. By most authors, however, it is accepted that the major effect on the properties of these phases is exerted by the lattice distortion caused by the solute atoms.

In the following chapters it is attempted to find a systematic connexion of the dimensions of the solute atoms and the change of the yield strength in homogeneous alloys of α-iron. The substitutional and interstitional solid solutions will be treated from a common point of view, on the basis of the space requirement of the solute atoms. It should, however, be pointed out that the relations will be dealt with merely phenomenologically, omitting any discussion or criticism of theoretical concepts of strengthening mechanisms.

II. The yield strength of substitutional solid solutions of α-iron

Making reference to early investigations of v. GÖLER and G. SACHS [3] it is accepted that in a continuous series of solid solutions of two given metals the critical shear stress goes through a maximum at about 50 ÷ 50 atomic per cent (Fig. 1). For dilute alloys of such a system, i.e. for small solute contents the curve is considered as a straight line the slope of which is given according to [1] by

$$\frac{d\tau}{dc} = k_1 \left(\frac{1}{a} \cdot \frac{da}{dc} \right)^n \tag{1}$$

where the left side is the increase of the critical shear stress produced by one atomic % alloying element, whereas on the right side the term in brackets is the relative change of the lattice parameter a pertaining likewise to one atomic % of solute; k_1 and the exponent n are constants which can be determined from experimental results.

In a number of homogeneous alloy systems the gradient $d\tau/dc$ was found experimentally also sufficiently constant up to about 5 atomic % of solute element, in most cases, however, it is not known how far the curve in

the Fig. 1. is really a straight line. For alloys of α-iron the connexion of the resistance to flow with an amount of solute atoms is less certainly established. In most cases only few values were actually measured and these show some spread also. Thus, the experimental results do not allow to conclude with certainty whether there is a perfectly straight part of the shear stress — concentration curve at all and how long it is. In order to bypass this uncertainty only alloys containing one atomic % solute elements will be taken into account

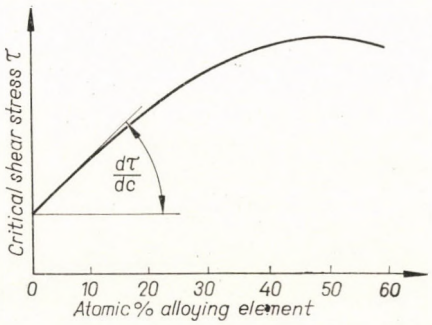


Fig. 1. The critical shear stress as a function of concentration in a homogeneous alloy system

disregarding, whether the line connecting the points presenting the flow stress of pure iron and that of the alloy containing one atomic % solute is the tangent of the flow stress-concentration curve or not.

There are hardly any data to be found about the critical shear stress of monocrystalline α-iron containing solute atoms, too. Most authors determine the yield strength of polycrystalline samples by tensile testing and derive herefrom the "flow stress in shear" by halving it according to the relation

$$\sigma_y = \frac{\tau_f}{\sin 45° \cdot \cos 45°} = 2\,\tau_f \tag{2}$$

where σ_y is the yield strength, τ_f the resolved flow stress in shear and 45° the angle included by the maximum shear stress with the axial tensile stress.

The yield strength, however, is proportional both to the critical shear stress and to the flow stress in shear; the same is true also for their change:

$$\frac{d\sigma_y}{dc} = k_2 \frac{d\tau}{dc} \tag{3}$$

and the relative change of the lattice parameter is proportional to the difference ΔD of atomic diameters of solvent and solute, respectively:

$$\frac{1}{a} \cdot \frac{da}{dc} = k_3 \cdot \Delta D\,. \tag{4}$$

By combining the equations (3) and (4) with (1) we obtain

$$\frac{d\sigma_y}{dc} = K \cdot (\Delta D)^n \tag{5}$$

where K is a constant incorporating the constants k_1, k_2 and k_3.

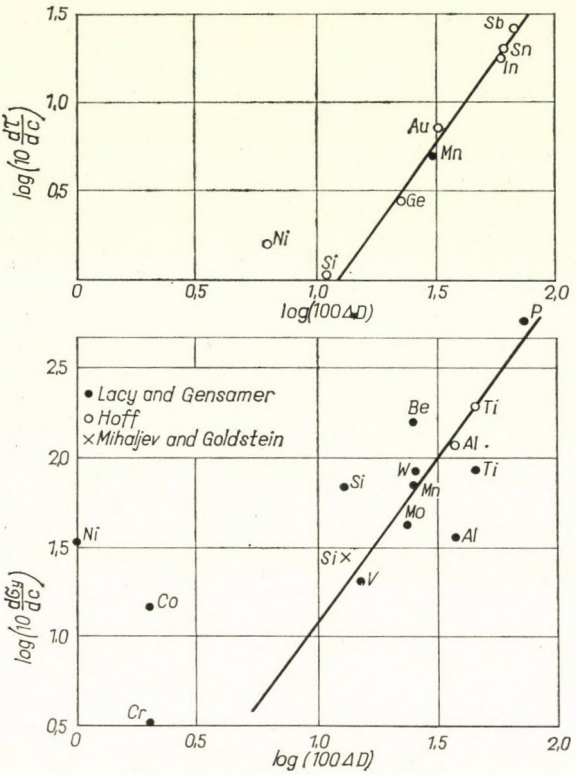


Fig. 2. The critical shear stress of homogeneous copper alloys *(above)*; the yield strength of ferrous alloys versus the difference of atomic diameters of solvent and solute *(below)*

In the first step the exponent n must be determined for the substitutional solid solutions of α-iron. For this purpose and also for comparison's sake in the upper part of Fig. 2 the values of $d\tau/dc$ found by [4] in copper alloy systems are shown as a function of $\log(\Delta D)$. Herein the points presenting the specific effect of dissolved elements on the critical shear stress lie sufficiently near to a straight line, but except the point of nickel. The slope of the straight line gives for the value of the exponent $n = 2$ in the equation (1).

It should be pointed out that in this diagram the alloying element nickel shows the greatest deviation from the hardening effect corresponding

to its atomic diameter. The diameters of the other alloying elements which appear in this diagram greatly differ from that of copper.

In the lower diagram of Fig. 2 the specific values $d\sigma_y/dc$ were plotted versus log (ΔD) according to equation (5). The values ΔD were calculated from the nearest approach of atoms in the lattice [5] and compiled in Table I.

Table I

Atomic diameter of iron and its alloying elements and derived values

Element	Atomic diameter D [Å]	Difference of diameter, ΔD [Å]	(ΔD^2) [10^{-4} Å^2]	log (100 · ΔD)
Fe	2,48	—	—	—
Al	2,86	0,38	1440	1,58
Be	2,22	0,26	670	1,415
Co	2,50	0,02	4	0,30
Cr	2,50	0,02	4	0,30
Cu	2,56	0,08	64	0,90
Mn	2,24	0,24	575	1,38
Mo	2,72	0,24	575	1,38
Ni	2,49	0,01	1	0,00
P	1,74	0,76	5500	1,87
Si	2,35	0,13	170	1,115
Ti	2,92	0,44	1940	1,645
V	2,63	0,15	225	1,175
W	2,74	0,26	676	1,415
C	1,54	—	—	—
N	1,52	—	—	—

The data given in [5] in [kX] units were not converted into [Å] units, because the difference is not significant in the present work.

The values of $d\sigma_y/dc$ used to construct the lower diagram of Fig. 2 were selected from those collected in Table II. Chiefly the experimental results of [6] were adopted and only a few additional points representing the effect of aluminium, silicon and titanium were plotted according to the data of P. HOFF [8].

When compared with the upper diagram relating to copper alloys, the scatter of points in the lower diagram is seen at once to be more pronounced. The farthest points from the bulk array of them relate to the elements cobalt, chromium and nickel, the atomic diameter of which hardly differs from that of iron. The strengthening effect of these elements was found by [6] to be

Table II

The specific strengthening effect of substitutional solutes

Element	Increase of the yield strength according to the source [kp/mm²-s per atomic per cent]				
	[6]	[7]	[8]	[9]	[10]
Al	2,8	3,5	12	—	—
Be	11,2	16	—	—	—
Co	2	1,5	—	—	—
Cr	0,5	1,5	6,5	—	5,5
Cu	—	—	20	—	9
Mn	7,4	5,2	4,5	9,8	8,5
Mo	4,2	7	35	38	—
Ni	5	5	4	—	4,5
P	56	28	12	45	30
Si	7	8,4	8	1,8	2,8
Ti	8,5	11,2	20	90	66
V	2	3,5	15	—	35
W	—	8,4	—	—	—
Column	I	II	III	IV	V

relatively strong. A high value has been given for silicon too, while two other reactive elements, aluminium and titanium proved considerably less effective than would be expected from their atomic dimensions. If, however, the additional points showing results of P. HOFF are accepted as eventually correct ones, then a straight line can be drawn through the bulk of points relative to the elements vanadium, silicon, molybdenum, manganese, tungsten, aluminium, titanium and phosphorus; the point for chromium lies perhaps on the same line; this element raises the yield strength of iron by less than 0,3 kp/mm², but the exact value is not known. Recently it was found that small additions of chromium tend to lower the yield strength of iron [11]; this negative hardening effect of chromium may be due to scavenging other trace elements, primarily some strongly strengthening, interstitially dissolved carbon and nitrogen.

The straight line best fitting between the probably most reliable points runs parallel to that drawn in the upper diagram; since the coordinates of both diagrams are the same, the exponent n in equation (5) has probably also the value of 2.

Some additional remarks must be made to enhance understanding the scatter of the experimental data collected in Table II and of the points plotted both in Figs 2 and 3.

First it must be borne in mind that in 1944, when the experiments of LACY and GENSAMER were carried out, it was not yet possible to use materials of high purity or of precisely known composition and the experimental alloys could not be melted under strongly controlled conditions, as is possible nowadays. Thus, impurities present in iron could readily lead to reactions with reactive alloying elements. Oxydes of silicon or aluminium formed as coarse inclusions certainly have a different hardening effect from that of the com-

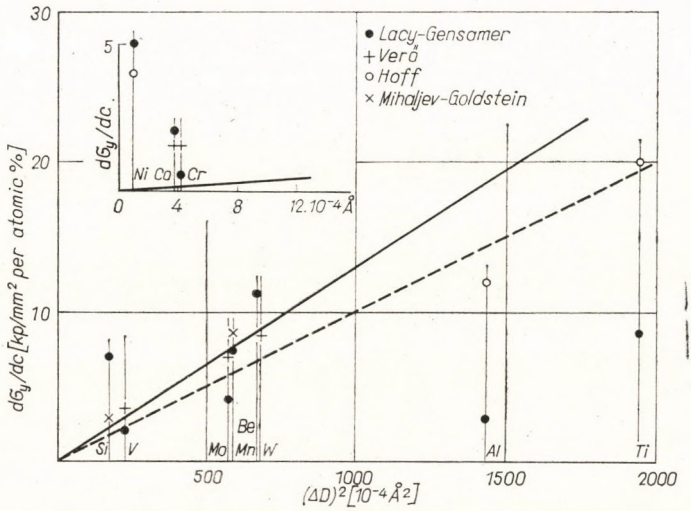


Fig. 3. The specific increase of yield strength versus the square of the difference of atomic diameters

ponents in the dissolved state. Aluminium, titanium and some other elements form nitrides which precipitate in fine dispersion under certain circumstances. Very little was known about these processes in 1944 and therefore they could not be taken into account.

Another additional effect of alloying elements may be due to variations in the grain size of ferrite as a consequence of the change of transformation temperature. In alloys containing, for example, nickel ferrite forms from austenite at a considerably lower temperature, than in pure iron; the grain size of ferrite is then finer, its yield strength, however, higher, than corresponding to the lattice distortion. The exceedingly strong effect found for cobalt, chromium and nickel is probably due to secondary effects of this type.

Since it was assumed that some results of [6] can bear an unavoidable error, it was thought advisable to draw attention to further experimental data also. Unfortunately, however, no further comprehensive examination of ferritic solid solutions was carried out in the last 20 years from which the effect of solute metals on the strength in pure alloys could be learnt and so it was

necessary to make use of data relating to less pure alloys and commercial steels, too.

In column II of Table II there are data published by the author in Hungarian only [7]; these resulted from tests carried out on alloys of commercially pure metals. In column III we can find data originating from experiments of HOFF [8] carried out with steels containing about 0,15% C.

In the last two columns of Table II there are values listed which originate from formulae proposed to calculate the yield strength of steel from its composition. Though in steels most alloying elements also exert other effects than hardening by solution in ferrite, these formulae or more precisely the factors expressing the specific effect of the alloying elements certainly help to estimate, whether a value obtained with purer binary alloys should be lowered or increased in order to obtain a probably more correct value.

ROGERS [9] processed the data characterizing the composition and mechanical properties of about 8400 heats of mild steel by statistical methods. It was found that the yield strength is increased by the following amounts, if one weight % of them is added: manganese by 14 028 psi, silicon by 5160 psi, phosphorus by 118 400 psi, molybdenum by 31 700 psi and titanium by 150 000 psi. These data, converted into kp/mm²-s and related to one atomic % are listed in column IV.

MIKHALYEV and GOLDSTEIN [10] published the following formula

$$\sigma_y\,[\mathrm{kp/mm^2}] = 12{,}4 + 28 \cdot \mathrm{C} + 8{,}4 \cdot \mathrm{Mn} + 5{,}6 \cdot \mathrm{Si} + 5{,}5 \cdot \mathrm{Cr} + \\ + 4{,}5 \cdot \mathrm{Ni} + 36 \cdot \mathrm{V} + 77 \cdot \mathrm{Ti} + 55 \cdot \mathrm{P} \qquad (6)$$

to calculate the yield strength of steel plates; the symbols of the alloying elements mean their amount in weight per cents. It was claimed that 50% of the results obtained by means of this formula are correct within ± 1 kp/mm² and 75% of them within ± 2 kp/mm². The factors of the formula, related again to one atomic %, are listed in column V.

The strengthening effect of most elements appears to be very different according to the five authors. In principle the data of LACY and GENSAMER in column I could be regarded as acceptable, because these were the results of controlled laboratory experiments which were, however, carried out many years ago. It seems justified to doubt, whether the effect of titanium given as 8,5 or 11,5 kp/mm² per atomic % may be correct, if the same effect in many heats of steel was found to be six to ten times greater. As another example silicon must be mentioned; the strengthening effect of this element was found with binary alloys to be 7 to 8,4 kp/mm², while in thousands of heats of steel it proved to raise the yield strength only by 1,8 to 2,8 kp/mm², though this element exerts its hardening effect both in binary Fe-Si-alloys and in steels

practically in the same way, by going into solution in ferrite. Similar thoughts arise also in relation to other alloying elements.

In Fig. 3 the specific increase of yield strength is shown depending from $(\Delta D)^2$ according to equation (6). Though not all values listed in Table II were plotted in this diagram, there is yet a considerable scatter of the points. The straight line drawn between the points has been located according to the following considerations.

The line must run through the origin of the coordinate system, though the points lying nearby tend to require here another, higher location of it. It is, however, perfectly clear that the specific increase of the yield strength pertaining to the difference $(\Delta D)^2 = 0$ must likewise be zero, because otherwise alloying iron with itself would also exert a hardening effect. A further probably correct and reliable part of the line runs through the points presenting the effect of the elements vanadium, manganese, molybdenum and tungsten. The equation of this line can be written as

$$\frac{d\sigma_y}{dc} = 130 \cdot (\Delta D)^2. \tag{7}$$

It is claimed that this equation gives the probable value of the specific effect of the alloying elements on the yield strength of iron, in spite of the fact that the experimental values relating to elements with $(\Delta D)^2 < 0{,}02$ seem to be greater and those relating to elements with $(\Delta D)^2 > 0{,}10$ are in most cases smaller, than indicated by the line.

As an alternative another line can be drawn, if the data relating to the elements titanium and phosphorus are considered to be the most reliable; then we have the broken straight line corresponding to the equation

$$\frac{d\sigma_y}{dc} = 100 \cdot (\Delta D)^2. \tag{8}$$

At present it is hardly possible to discern which of the equations (7) or (8) is the more accurate. For this purpose it appears advisable to repeat the experiments of the previous authors, making use of the better conditions existing today.

Alloying elements with an atomic diameter which greatly differs from that of the iron atom dissolve in α-iron only in small amounts and, on the other hand, to a great extent soluble elements affect the yield strength only fairly. Calculations carried out with the data of alloy systems of α-iron showed that the yield strength of substitutional solid solutions may not be greater than 40 to 45 kp/mm², if that of pure iron is accepted to be about 3 kp/mm².

III. The yield strength of interstitial solid solutions of α-iron

If there are uncertainties and controversial points when dealing with the substitutional solid solutions, the same is exceedingly true for the other kind of solid solutions, for the interstitial phases of α-iron containing carbon and nitrogen, i.e. for ferrite and martensite. As was recently emphasized by BAIRD [11] the experimentally observed effect of interstitials on the yield stress of iron at near the ambient temperature is linear with respect to solute

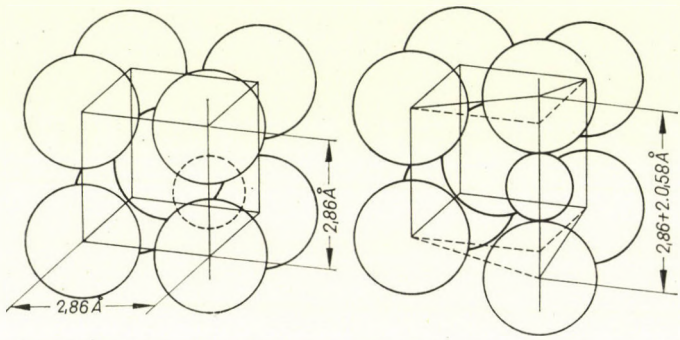


Fig. 4. Insertion of a carbon atom into an octahedral site of the iron lattice and the distortion of the latter produced thereby

content, whereas *Fleischer*'s treatment predicts a square root dependence of the yield stress on solute content [12]. In addition there are further theoretical concepts, too.

Based on the first suggestions made by KURDJUMOV the lattice model of interstitial solid solutions of α-iron was constructed and described by PETCH [13] and LIPSON and PARKER [14]. The model is a result of spatial considerations taking into account the probable locations of the atoms in the lattice and, of course, the known dimensions of the lattice of such phases. Though the correctness of the model has not yet been experimentally proven, it is generally accepted and quoted by every author in the most recent literature also [15].

According to this model the interstitial atoms occupy the octahedral sites having coordinates 00 1/2 (see Fig. 4). In the middle of the edge the free space between two neighbouring iron atoms has the diameter 2,86 — 2,48 = = 0,38 Å, where 2,86 Å is the lattice parameter and 2,48 Å the atomic diameter of iron. If a carbon atom having a diameter of 1,54 Å enters into a hollow space measuring 0,38 Å only, the neighbouring iron atoms must be dislocated in the direction of the edge by a distance of (1,54 — 0,38) : 2 = 0,58 Å. Thus the insertion of a single carbon atom produces a distortion of the lattice

and of the slip plane also, as if two iron atoms had been substituted by solute atoms with a diameter by 0,58 Å greater than that of iron atoms. Making use of equations (7) and (8), respectively, we obtain that one atomic % of interstitially dissolved carbon raises the yield strength of iron by

$$\Delta\sigma_y = 130 \cdot 2 \cdot 0{,}58^2 = 87 \text{ kp/mm}^2, \tag{9}$$

or by

$$\Delta\sigma_y = 100 \cdot 2 \cdot 0{,}58^2 = 67 \text{ kp/mm}^2. \tag{10}$$

In martensites containing smaller quantities of carbon the interstitial atoms may occupy tetrahedral sites also where there is a greater hollow space with a diameter of $0{,}252a = 0{,}72$ Å. Each inserted carbon atom displaces the four neighbouring iron atoms by the half amount of the difference 1,54 — — 0,72 = 0,82 Å and so we have as the increase of the yield strength produced by one atomic % carbon

$$\Delta\sigma_y = 130 \cdot 4 \cdot \left(\frac{0{,}82}{2}\right)^2 = 87{,}3 \text{ kp/mm}^2 \tag{11}$$

and with the factor 100 standing in equation (10) 67,1 kp/mm². Thus, carbon atoms situated in octahedral and tetrahedral positions apparently have the same strengthening effect.

In order to compare the calculated results with the experimentally measured ones, first some theoretical concepts to predict the dependence of the yield strength on the concentration must be briefly discussed because the experimental results are always evaluated according to one or another concept.

The early concept of MOTT and NABARRO [16] postulating a linear increase of the yield strength with the carbon content in solution seems to be discarded in recent publications. The newer theories consider the distorted portion of the iron lattice surrounding each inserted carbon atom as "precipitates". It is assumed that the stress required to force a dislocation to pass through such precipitate particles as pinning points in a slip plane is greater than the stress which produces slip along an undistorted plane. The additional stress increase the yield strength of iron according to OROWAN [17] by the amount

$$\Delta\sigma_y = cG\frac{b}{t} \tag{12}$$

where c is a constant value of about 0,25, G the shear modulus, b the Burgers-vector of the dislocation and t the distance of neighbouring pinning points i.e. of the precipitate particles.

In the case of interstitial solid solutions this distance means the distance of solute atoms and is consequently inversely proportional to the cube root of the number of interstitials present in a unit volume of the alloy or to the square root of the number of foreign atoms occurring in a unit area of a slip plane. The number of atoms in both considerations, of course, can be expressed by atomic fractions or by atomic per cents.

According to one concept the increase of yield strength is assumed to be proportional to the cube root of a parameter expressing the quantity of the precipitate. A formula of this kind was developed by ANSELL [18] according to which the yield strength of interstitial solid solutions is estimated to have the value

$$\sigma_y = \frac{G}{K} f^{1/3} \tag{13}$$

where G is the shear modulus and f the volume fraction of the precipitate and K a constant value of about 100 for martensites. The fraction f of the precipitate can be considered as being proportional to the carbon content in atomic %, if the distorted portions of the lattice have the same dimensions around each inserted carbon atom. This formula, however, has hardly ever been used by other investigators.

The distance t of obstacles interfering in a slip plane with the movement of dislocations was taken into account as being inversely proportional to the square root of atomic concentration by FLEISCHER and HIBBARD [12]. Accordingly the yield strength of a homogeneous alloy of α-iron containing C atomic % carbon is given by

$$\sigma_y = \sigma_0 + mC^{1/2} \tag{14}$$

where σ_0 is the yield strength of pure iron and m a constant factor often expressed as a fraction of the shear modulus G. In a similar equation the flow stress in shear is defined as

$$\tau_f = \tau_0 + m_1 C^{1/2}. \tag{15}$$

The latter formulae (14) and (15) were adopted by ROBERTS and OWEN [19] when they evaluated their own experimental results relative to the yield strength of martensitic ferrous alloys and in addition the results of other investigators, too. The data were summarized in a table enumerating the factors m_1 as the gradient of the flow stress in shear versus the square root of the atomic fraction of interstitial solute. The results obtained by WERT [20] were esteemed to be the most reliable ones, because WERT determined the quantity of carbon being really in solution by measurement of internal friction, whereas

the other investigators simply assumed the amount of dissolved carbon was the same as the total carbon content of the alloy. This assumption, however, is not justifiable, because an amount of carbide may remain undissolved and dissolved carbon can precipitate both during quenching and also afterwards. Undissolved or dissolved and then precipitated carbide has, of course, another strengthening effect than carbon in solution.

The gradient of the line presenting the flow stress depending on the square root of the atomic fraction of dissolved carbon is given according to WERT as $0{,}05 \cdot G$ at 293 °K. The shear modulus G is a function of the absolute temperature:

$$G\ [\mathrm{lb/in^2}] = (13{,}30 - 0{,}0030\ T) \cdot 10^6$$

and thus at $T = 293$ °K

$$G\ [\mathrm{kp/mm^2}] = (13{,}30 - 0{,}003 \cdot 293) \cdot 10^3 \cdot 0{,}7 = 8700\ \mathrm{kp/mm^2}.$$

One atomic % means as atomic fraction 0,01 and so the increase of the resolved flow stress in shear produced by this quantity of carbon amounts to

$$\Delta\tau_f = m_1 C^{1/2} = 0{,}05 \cdot 8700 \cdot \sqrt{0{,}01} = 43{,}5\ \mathrm{kp/mm^2} \qquad (16)$$

according to equation (15), while the increase of the yield strength is twice as great [see equation (2)]:

$$\Delta\sigma_y = 2\,\Delta\tau_f = 87\ \mathrm{kp/mm^2} \qquad (17)$$

in excellent accordance with the result obtained in equations (9) and (11).

If in a repeated investigation of the strengthening effect of solute atoms would prove equation (7) to be accurate, then the lattice distortion caused by carbon atoms might be made responsible for the total hardening effect of this interstitially dissolving element. If, however, equation (8) would be found to be correct, then a part of the total hardening effect of carbon must be ascribed to other effects, to that of a fine grain size, of twinning and of precipitation from the dissolved state. At present, however, it is hardly possible to distinguish, which of the calculated values is the correct one.

WERT found that the increase of the flow stress in shear is about the same in ferrite and in martensite, if considered as a function of the carbon content. Therefore, experimental data relating to martensitic test pieces can be used in addition to the very scarce values determined with pure binary alloys. It was hoped that the character of the connexion of the yield strength with the solute concentration would be better understandable, if further data

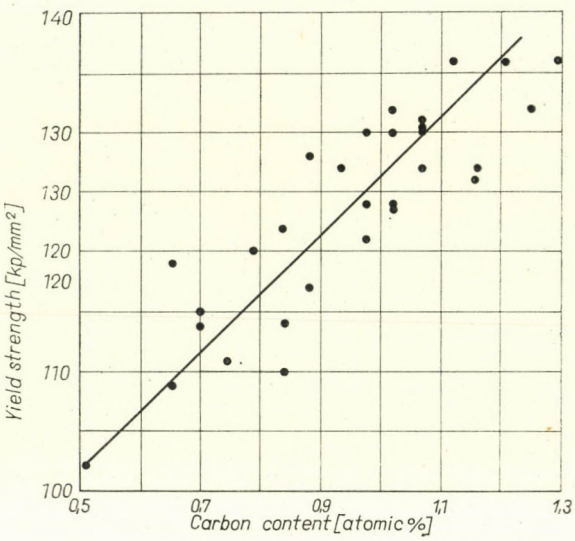


Fig. 5. The yield strength of slightly alloyed martensitic steels as a function of the carbon content according to [21]

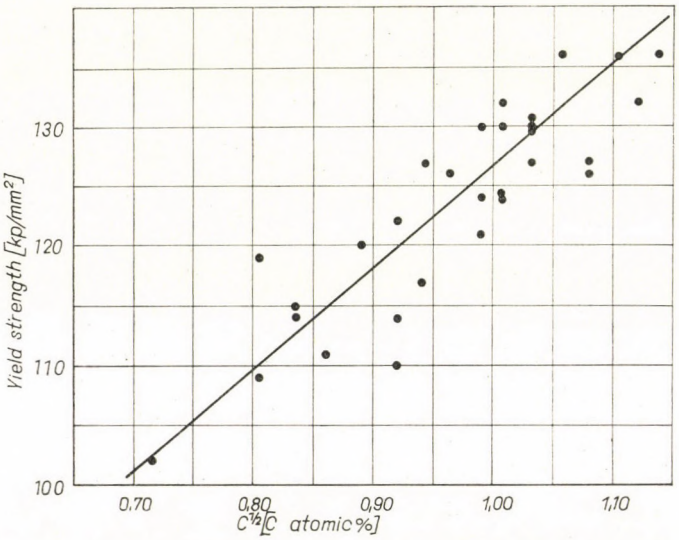


Fig. 6. The same data, as in Fig. 5, but as a function of the square root of the carbon content

could be taken into account, even if these were rather of technical, than of laboratory quality.

BUSBY et al [21] examined quenched steel samples which contained 0,10 to 0,30% C and in addition 3 to 5% manganese, chromium, nickel and molybdenum to secure a sufficient hardenability. The purpose of these experiments was to find the limit up to which steels can be used as constructional materials in the martensitic condition. Their results relating to the yield strength are plotted as a function of carbon content in Fig. 5 and as a function of the square root of carbon content in Fig. 6; the carbon content is always meant in atomic per cents.

In both diagrams the average trend in the increase of the yield strength with the carbon content is presented by straight lines drawn through the scattered points. The equation of the line shown in Fig. 5 is written as

$$\sigma_y = 78 + 48{,}5 \cdot C_{\text{at}\cdot\%} \tag{18}$$

while that of the line in Fig. 6 is

$$\sigma_y = 40 + 87{,}5 \cdot C_{\text{at}\cdot\%}^{1/2}. \tag{19}$$

The latter equation (19) is surely more correct, not because the factor 87,5 herein agrees very closely with the corresponding value 87 derived from data of lattice distortion, but rather with regard to the size of the first term in the two equations. This term means the yield strength of the carbon-free material containing a few per cents of metallic alloying elements. The yield strength of such an alloy is by no means as high as 78 kp/mm² in equation (18), but 40 kp/mm² standing in equation (19) is a much more acceptable value.

The hardness too is related in some way to the flow stress and in addition the hardness of plain carbon and alloy steels in the martensitic condition is very precisely known. If the Rockwell C hardness numerals determined by HODGE and OREHOSKY [22] are converted into Vickers hardness values, the diagram shown in Fig. 7 results. The increase of hardness in relation to the carbon content is obviously not linear. One atomic % = 0,215 weight % dissolved carbon increases the hardness of α-iron from 45 to 550 kp/mm², i.e. by about 500 kp/mm². The yield strength of plain Fe-C-alloys amounts from 16 to 18 per cents of the hardness value. The increase of the hardness by 500 kp/mm² accordingly indicates an increase of the yield strength by 80 to 90 kp/mm², likewise in good agreement with the result obtained in other ways.

The atomic diameter of nitrogen is a little smaller than that of iron, 1.52 Å; its strengthening effect therefore is expected to be somewhat less strong than that of carbon. One atomic % of this element increases the yield strength

of iron by 84,5 kp/mm², if equation (7) with the factor 130 is used in calculation. In general, the known values of the effect of nitrogen on the yield strength appear to support this conclusion.

Finally it must be mentioned that the data presented in Figs 5, 6 and 7 seem to support the correctness of Fleischer's concept predicting a linear connexion of the yield strength with the square root of the concentration of interstitially dissolved atoms, otherwise it was, however, beyond the scope of the

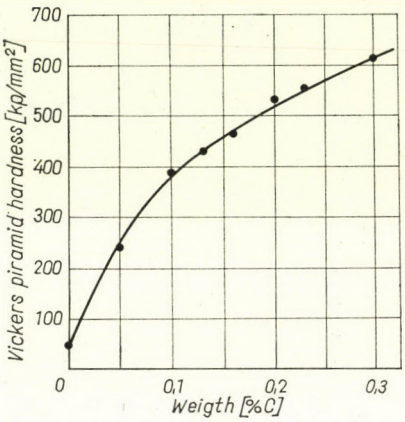


Fig. 7. The hardness of martensite as a function of the carbon content, converted values of [22]

present investigation to prove or to disprove one or another concept relating to the mechanism of the hardening effect of interstitial atoms.

As an outstanding result of the calculations based on atomic dimensions it should be mentioned that the strengthening effect of carbon obtained by means of formulae (7), (9) and (11) with the constant factor 130 agrees very closely with experimental results determined under various conditions. This finding seems to support the correctness of the factor 130 in contrast to the factor 100 in formula (8). If so, however, the whole strengthening effect of carbon is due to lattice distortion produced by interstitial atoms and nothing of it remains to be ascribed to other mechanisms assumedly active in martensite, such as fine grain size and twinning.

REFERENCES

1. HONEYCOMBE, R. W. K.: *Progr. in Metals Science* **9** (1961), 95.
2. ALLEN, N. P. in: "Iron and its dilute alloys" (edited by SPENCER, C. W. and WERNER, F. R.); 1963, 292.
3. v. GÖLER, F.—SACHS, G.: *Z. Physik* **55** (1929), 581.
4. LINDE, J. O.—EDWARDSON, ST.: *Ark. Fys.* **8** (1954), 511.

5. ASM Metals Handbook; 1948, 21.
6. LACY, E. C.—GENSAMER, M.: *Trans. ASM* **32** (1944), 88.
7. VERŐ, J.: Az ipari vasötvözetek metallografiája (Metallography of technical ferrous alloys), Vol. I; 1960, 459.
8. HOFF, P.: Mitt. Kohle- u. Eisenforschungs G. m. b. H. Dortmund, 1938; see also FOLKHARD, E.: *Schweizer Archiv* **21** (1955), 183.
9. ROGERS, W. T.: *Trans. ASM* **43** (1951), 1126.
10. Михалев, М. Ш.—Голдштейн, М. Й.: Стал **18** (1958), 942.
11. BAIRD, J. D.: *JISI* **205** (1967), 718.
12. FLEISCHER, R. L.—HIBBARD, W. R.: *Nat. Phys. Lab. Symposium* No. 15; 1963, 262.
13. PETCH, N. J.: *JISI* **147** (1963), 221.
14. LIPSON, H.—PARKER, A. M. B.: *JISI* **149** (1944), 123.
15. HABRAKEN, L.—DE BROUWER, J. L.: De ferri metallographia, Vol. I; 1966, 219.
16. MOTT, N. F.—NABARRO, F. R. N.: Rep. on Conf. on "Strength of Solids". Phys. Soc. London; 1948, 1.
17. OROWAN, E.: in "Dislocations in Metals". *AIME* (1954), 69.
18. ANSELL, G. J.: in "Precipitation from iron-base alloys" (edited by SPEICH, G. R. and CLARK, J. B.) 1965, 258.
19. ROBERTS, M. J.—OWEN, W. S.: *ISI Spec. Rep.* (1965); No. 93, 171.
20. WERT, C.: *Trans. AIME* **188** (1950), 1242.
21. BUSBY, C. C.—HAWKES, M. F.—PAXTON, H. H.: *Trans. ASM* **47** (1955), 135.
22. HODGE, J. M.—OREHOSKY, M. A.: *Trans. AIME* **167** (1946), 627.

Einfluß der gelösten Legierungselemente auf die Fließgrenze des α-Eisens. Die Fließgrenze des α-Eisens wird von den darin substitutionell gelösten Legierungsmetallen um $d\sigma_f/dc = 130(\Delta D)^2$ kp/mm² erhöht, wo für ΔD der Unterschied der Atomdurchmesser in Å einzusetzen ist. Derselbe Ausdruck gibt auch, auf Grund des allgemein akzeptierten Gittermodells, die Wirkung des interstitionell gelösten Kohlenstoffs richtig, in guter Übereinstimmung mit den Versuchswerten, an. Ein Atomprozent gelösten Kohlenstoffs erhöht die Fließgrenze des Eisens um $\Delta\sigma_F = 130 \cdot 2 \cdot 0{,}58^2 = 87$ kp/mm². Mehrere Anzeichen sprechen dafür, daß dieser Wert nur auf Legierungen mit 1 Atomprozent Kohlenstoffgehalt bezogen werden kann, im allgemeinen jedoch scheint die Beziehung $\sigma_F = 40 + 87{,}5 \cdot \sqrt{C_{\text{At}\%}}$ zu gelten.

L'influence des éléments d'alliage dissous sur la limite d'écoulement du fer α. La limite d'écoulement du fer α est augmentée par ses éléments d'alliage solubles par substitution, de $d\sigma_F/dc = 130(\Delta D)^2$ kp/mm² où ΔD doit être remplacé par la différence des diamètres atomiques exprimée en Å. Sur la base du modèle de réseau généralement accepté, la même expression donne l'influence du carbone dissous interstitiellement, en bon accord avec les données expérimentales. Un pour cent atomique de carbone dissous augmente la limite d'écoulement du carbone de $\Delta\sigma_F = 130 \cdot 2 \cdot 0{,}58^2 = 87$ kp/mm². Il semble que cette valeur n'est valable que pour des alliages contenant un pour cent atomique de carbone, et c'est la relation $\sigma_F = 40 + 87{,}5 \cdot \sqrt{C_{\text{at}\%}}$ qui est valable en général.

Воздействие растворенных легирующих элементов на условный предел текучести α-железа *(Й. Верэ)*. Условный предел текучести α-железа увеличивается растворяющимися в нем субституционно легирующими металлами на $d\sigma_F/dc = 130\ (\Delta D)^2$ *кг/мм²*, где вместо ΔD необходимо подставить разность атомных диаметров в Å. Это же выражение дает также (на основе общепринятой модели кристаллической решетки) действие углерода, растворяющегося интерстиционно, в удовлетворительном совпадении с экспериментальными данными. Один атомный процент растворенного углерода увеличивает условный предел текучести железа на $\Delta\sigma_F = 130 \cdot 2 \cdot 0{,}58^2 = 87$ *кг/мм²*. Ряд симптомов указывает на то, что эта величина может быть отнесена только к сплаву с содержанием одного атомного процента углерода, а вообще же действительной кажется зависимость $\sigma_F = 40 + 87{,}5 \cdot \sqrt{C_{\text{атомн.}\%}}$.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (3—4), pp. 261—276 (1968)

GEOMETRIE DER ABWÄLZFRÄSER MIT GRUNDSCHNECKEN EVOLVENTENFÖRMIGER AUSFÜHRUNG*

I. LIPKA

ENTWICKLUNGSINSTITUT DER WERKZEUGMASCHINENWERKE, HALÁSZTELEK (UNGARN)

[Eingegangen am 5. Oktober, 1965]

Vom Verfasser wurde ein Kriterium abgeleitet, auf dessen Grund entschieden werden kann, ob ein mit irgendeinem Abwälzwerkzeug hergestelltes Zahnradprofil eine gewöhnliche Kreisevolvente darstellt, oder eine hiervon abweichende Kurve. Dieses Evolventen-Kriterium entscheidet über den Evolventencharakter des abgewälzten Zahnprofils anhand der Erfüllbarkeit einer partiellen Differentialgleichung. In der vorliegenden Arbeit wendet der Verfasser das Evolventen-Kriterium bei dem Zahnprofil an, das im Wälzverfahren mit dem Evolventen-Schneckenfräser hergestellt worden ist. Es wird nachgewiesen, daß ein Zahnprofil, das mit einem Schneckenfräser hergestellt wird, dessen Grundschnecke eine Evolventenschnecke ist, keine gewöhnliche (spitze) Kreisevolvente darstellt.

I. Einleitung

In der vorausgegangenen Arbeit, mit dem Titel »Wälztheorie von Zahnrädern mit einem Evolventenprofil« [1] wurde ein analytisches Kriterium abgeleitet, das zur Entscheidung darüber dienen soll, ob das Zahnprofil eines Zahnrades, welches mit irgendeinem Wälzwerkzeug hergestellt ist, eine gewöhnliche Kreisevolvente oder aber eine von der Kreisevolvente abweichende Kurve darstellt. In der vorliegenden Arbeit wird die für die Untersuchung des Charakters der Zahnprofilevolvente ausgearbeitete Methode bei Zahnprofilen angewandt, welche mittels Wälzfräser in Evolventen-Grundschneckenausführung hergestellt wurden. Bei der Untersuchung des Zahnprofils, das mittels Wälzfräser hergestellt wurde, ist die Anwendbarkeit des genannten Evolventenkriteriums von entscheidender Bedeutung; dieses Kriterium bildet eine notwendige Bedingung dafür, daß die durch Abwälzung gewonnene Profilkurve eine gewöhnliche Kreisevolvente sei. Dieses Kriterium wird jetzt, um es bei einem mit Wälzfräser hergestellten Zahnprofil verwenden zu können, auf folgende Weise formuliert:

Die $(x; y)$ Koordinatenebene soll mit der Stirnebene (mit der Ebene senkrecht zur Radachse) des durch Abwälzung herzustellenden Zahnrades zusammenfallen. Die Durchdringungskurvenschar der Fräserfläche und Stirn-

* In den Fragen, die in der vorliegenden Arbeit behandelt werden und noch den Gegenstand einer Diskussion bilden, erwartet die Redaktion eine Stellungnahme der Fachkreise.

ebene — welche infolge der Drehbewegung des Fräsers entsteht — sei in der mit der Stirnebene zusammenfallenden $(x; y)$ Koordinatenebene durch das folgende zweiparametrige Gleichungssystem dargestellt:

$$x = X(t; \Phi) + a;$$

$$y = Y(t; \Phi),$$

wobei a die Abszisse des Fräserzentrums in einem $(x; y)$ Koordinatensystem bedeutet, dessen Anfangspunkt O mit dem Mittelpunkt des zu erzeugenden Zahnrades zusammenfällt. Wenn nun das erzeugte Zahnprofil eine gewöhnliche Kreisevolvente ist, die von einem Kreis mit einem Radius r_a abgewickelt wurde, dann erfüllen die Funktionen $X(t; \Phi)$, $Y(t; \Phi)$ mit zwei Veränderlichen folgende partielle Differentialgleichung:

$$\begin{vmatrix} \frac{\partial X(t;\Phi)}{\partial \Phi} & \frac{\partial Y(t;\Phi)}{\partial \Phi} \\ \frac{\partial X(t;\Phi)}{\partial t} & \frac{\partial Y(t;\Phi)}{\partial t} \end{vmatrix}^2 - r_a^2 \varkappa^2 \left[\left(\frac{\partial X(t;\Phi)}{\partial t} \right)^2 + \left(\frac{\partial Y(t;\Phi)}{\partial t} \right)^2 \right] = 0,$$

wobei $\varkappa$ das Verhältnis der simultanen Drehwinkel vom Fräser und dem zu verzahnenden Rad ist, und wobei man ferner voraussetzt, daß die Funktionaldeterminante nicht identisch verschwindet.

Für die Anwendung des Kriteriums muß man die Durchdringungskurven der Fräserfläche und der Stirnebene kennen, d. h. die Parameterdarstellung des ebenen Schnittes der evolventenförmigen Schraubenfläche, wofür man das Gleichungssystem der Schraubenfläche ableiten soll.

II. Ableitung des Gleichungssystems der evolventenförmigen Schraubenfläche

Wenn die gerade Erzeugende A, die sich von der z-Achse in der Entfernung k befindet und mit dieser Achse nicht parallel ist, sich um die Achse z mit einer konstanten Winkelgeschwindigkeit dreht (Bild 1a) und sich dabei gleichzeitig in der Richtung der z-Achse mit konstanter Geschwindigkeit bewegt, dann entsteht eine geradlinige Schraubenfläche.

Der Fußpunkt Q der Normal-Transversale der Erzeugenden A und der z-Achse ist ein Punkt der Schraubenlinie, die auf dem Zylinder mit dem Radius k beschrieben wird, und seine Koordinaten lauten:

$$x = k \cos \varphi;$$

$$y = k \sin \varphi;$$

$$z = r \varphi.$$

Der Faktor r hat hier die Bedeutung:

$$r = \frac{S}{2\pi},$$

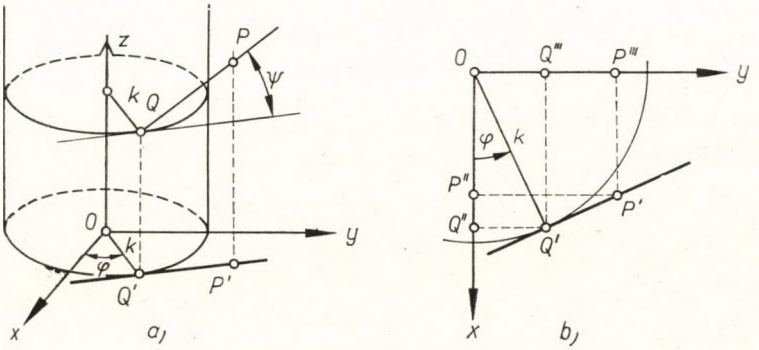


Bild 1

wobei S die Steigung der Schraubenfläche, d. h. den Weg bedeutet, den die Erzeugende in Richtung der z-Achse während einer vollen Umdrehung um die z-Achse von der Größe 2π zurücklegt. Die Erzeugende A soll mit der zur z-Achse senkrechten Ebene einen konstanten Winkel ψ bilden, und man betrachte den auf der Erzeugenden, von dem Fußpunkt Q in der Entfernung t liegenden Punkt P:

$$QP = t.$$

Die senkrechten Projektionen der Punkte P und Q auf die $(x; y)$ Ebene seien P' und Q'. Da $P'Q' = t \cos \psi$ ist, so ist dem Bild 1b unmittelbar zu entnehmen, daß der Punkt P die folgende x-Koordinate hat:

$$x = QQ'' - P''Q'' = k \cos \varphi - t \cos \psi \sin \varphi;$$

in ähnlicher Weise ist:

$$y = OQ''' + Q'''P''' = k \sin \varphi + t \cos \psi \cdot \cos \varphi;$$

und zuletzt die z-Koordinate des Punktes P:

$$z = Q'Q + QP \sin \psi = r \cdot \varphi + t \sin \psi.$$

Demnach, wenn φ und t veränderliche Parameter sind, so ist das Gleichungssystem in Parameterform von der Gesamtheit der Erzeugenden, d. h. der geradlinigen Schraubenfläche:

$$\begin{aligned} x &= k\cos\varphi - \cos\psi \cdot \sin\varphi \cdot t, \\ y &= k\sin\varphi + \cos\psi \cdot \cos\varphi \cdot t, \\ z &= r \cdot \varphi + \sin\psi \cdot t. \end{aligned} \tag{1}$$

Hierbei bedeutet k den Radius des Kehlzylinders der Schraubenfläche. Wenn $k = 0$ ist, dann bezeichnet man die Schraubenfläche als eine spirale Schraubenfläche. Sofern $k \neq 0$ und der Stirnschnitt der Schraubenfläche, d. h. der zur z-Achse senkrechte ebene Schnitt eine Schleifenevolvente darstellt, dann nennt man die Schraubenfläche eine Konvolute. Wenn dagegen der Stirnschnitt eine gewöhnliche (sog. spitze) Kreisevolvente ist, so ist die Schraubenfläche eine evolventenförmige Schraubenfläche (s. [2], S. 27). Nun schreibe man das Gleichungssystem in Parameterform der im Stirnschnitt der Schraubenfläche, in der $z = 0$ Ebene entstehenden ebenen Kurve an. Wenn $z = 0$ ist, so wird aus der dritten Gleichung von (1):

$$t = -\frac{r \cdot \varphi}{\sin\psi},$$

und schreibt man diesen Ausdruck in die Ausdrücke von x und y, so wird das Gleichungssystem in Parameterform der Stirnschnittkurve:

$$\begin{aligned} x &= k\cos\varphi + \varphi r \cot\psi \cdot \sin\varphi\,; \\ y &= k\sin\varphi - \varphi r \cot\psi \cdot \cos\varphi\,. \end{aligned} \tag{1'}$$

Dies ist tatsächlich das Gleichungssystem der Kreisevolvente, und zwar einer Kreisevolvente, die von einem Grundkreis mit dem Radius

$$r \cdot \cot\psi$$

abgewälzt wurde. In dem Fall, wenn

$$k < r\cot\psi$$

ist, so ist der unter (1′) angegebene Stirnschnitt eine Schleifenevolvente; ist dagegen

$$r\cot\psi < k,$$

dann erzeugt das Gleichungssystem unter (1′) eine gestreckte Evolvente. Wenn jedoch

$$r \cot \psi = k,$$

d. h., der Radius des Grundzylinders gleich dem Radius k des Kehlzylinders ist, dann bedeutet der Stirnschnitt eine gewöhnliche (spitze) Kreisevolvente und die Schraubenfläche wird als eine Evolventen-Schraubenfläche benannt. Sei mit g der Radius des Grundzylinders bezeichnet, also:

$$r \cot \psi = g.$$

Dann gilt für die Evolventen-Schraubenfläche: $k = g$ und nach dem Gleichungssystem (1′) der im Stirnschnitt entstehenden gewöhnlichen Kreisevolvente:

$$x = g \cos \varphi + g \varphi \sin \varphi;$$

$$y = g \sin \varphi - g \varphi \cos \varphi,$$

wobei für g die Beziehung

$$r = g \tan \psi \tag{2}$$

besteht. Nun schreibe man das Gleichungssystem in Parameterform der Evolventen-Schraubenfläche (Evolventen-Schnecke) an. Da für diese Schraubenfläche $k = g$, ferner aus (2)

$$\cos \psi = \frac{g}{\sqrt{r^2 + g^2}}; \quad \sin \psi = \frac{r}{\sqrt{r^2 + g^2}}$$

ist, so wird das Gleichungssystem der Evolventen-Schraubenfläche nach (1) folgende Form annehmen:

$$x = g \cos \varphi - \frac{g}{\sqrt{r^2 + g^2}} \sin \varphi \cdot t;$$

$$y = g \sin \varphi + \frac{g}{\sqrt{r^2 + g^2}} \cos \varphi \cdot t;$$

$$z = r \cdot \varphi + \frac{r}{\sqrt{r^2 + g^2}} \cdot t$$

oder, wenn statt des Entfernungs-Parameters t der Parameter

$$\frac{t}{\sqrt{r^2 + g^2}} = t \frac{\sin \psi}{r}$$

eingeführt wird, den wir erneut mit t bezeichnen, so kann man das Gleichungssystem in Parameterform der Evolventen-Schraubenfläche (Evolventenschnecke) in folgender Form anschreiben:

$$x = g \cos \varphi - g \sin \varphi \cdot t;$$

$$y = g \sin \varphi + g \cos \varphi \cdot t;$$

$$z = r \cdot \varphi + r \cdot t,$$

wobei φ und t die Parameter, g den Radius des Grundzylinders der Schnecke bedeuten; $r = S/2\,\pi$ (S ist die Steigung der Schraubenfläche).

III. Bestimmung des ebenen Schnittes der Evolventen-Schraubenfläche

Man drehe die Achse T (die mit der z-Achse zusammenfällt) der Evolventen-Schraubenfläche im negativen Sinn um den Winkel $(90° + \gamma)$ um die x-Achse, wo der spitze Winkel γ den auf dem Teilzylinder vorhandenen Steigungswinkel der Schraubenfläche bedeutet. In diesem Fall zeigt das Bild 2 die Lage der T-Achse.

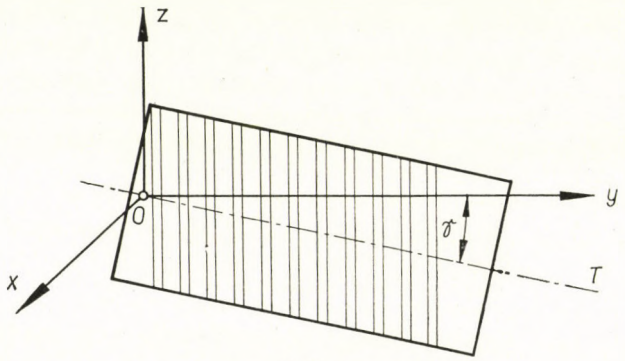


Bild 2

Wenn γ den Steigungswinkel bedeutet, der zur Schraubenfläche auf dem Teilzylinder mit dem Radius R vorhanden ist, d. h.

$$\tan \gamma = \frac{r}{R}$$

(wo $r = S/2\pi$), so ist in der geneigten Lage der Schraubenfläche die $(x; y)$-Ebene senkrecht zur Schraubenlinie mit dem Steigungswinkel γ, die auf dem Zylinder mit dem Radius R liegt. In diesem Fall liefert die Ebene $z = 0$ einen Normalschnitt der Schraubenfläche, und diese Ebene ist gleichzeitig die Stirnebene des abzuwälzenden Zahnrades, in der das abgewältze Zahnprofil untersucht werden soll. Das Gleichungssystem der Schraubenfläche, die sich in einer

geneigten, im Bild 2 dargestellten Lage befindet, erhält man aus dem Gleichungssystem, das im Abschnitt *II* abgeleitet wurde, wenn man in der Koordinaten-Ebene $(y; z)$ die folgende Rotation anwendet:

$$y' = y \cos \alpha - z \sin \alpha;$$

$$z' = y \sin \alpha + z \cos \alpha,$$

wobei jetzt

$$\alpha = -(90° + \gamma)$$

ist.

Mit dieser Rotation, wenn man die y' und z' Koordinaten erneut mit y und z bezeichnet, gewinnt man für die gedrehte Schraubenfläche mit der Achse T das folgende Gleichungssystem:

$$x = g \cos \varphi - g \sin \varphi \cdot t;$$

$$y = -(g \sin \varphi + g \cos \varphi \cdot t) \sin \gamma + (r\varphi + rt) \cos \gamma;$$

$$z = -(g \sin \varphi + g \cos \varphi \cdot t) \cos \gamma - (r\varphi + rt) \sin \gamma.$$

Man drehe die mit diesem Gleichungssystem erzeugte Schraubenfläche um die T-Achse um den Winkel Φ; dann erhält man folgendes Gleichungssystem:

$$x = g \cos (\varphi + \Phi) - g \sin (\varphi + \Phi) \cdot t;$$

$$y = -[g \sin (\varphi + \Phi) + g \cos (\varphi + \Phi) \cdot t] \sin \gamma + (r\varphi + rt) \cos \gamma;$$

$$z = -[g \sin (\varphi + \Phi) + g \cos (\varphi + \Phi) \cdot t] \cos \gamma - (r\varphi + rt) \sin \gamma.$$

Man bestimmt die Schnittkurve dieser Fläche mit der Ebene $z = 0$, und diese Kurve bedeutet nichts anderes als die Durchdringungskurve der Fräserfläche und der Stirnebene, wenn die Fräserfläche um ihre T-Achse um den Winkel Φ gedreht worden ist. Wenn im vorhergehenden Gleichungssystem $z = 0$ ist, so wird aus der dritten Gleichung

$$t = -\frac{g \sin (\varphi + \Phi) \cdot \cos \gamma + r\varphi \sin \gamma}{g \cos (\varphi + \Phi) \cdot \cos \gamma + r \sin \gamma}.$$

Setzt man nun diesen Wert in die Ausdrücke von x und y ein, und führt man die Reduktionen durch, so ergibt eine einfache Rechnung folgende Beziehungen:

$$x = g \frac{r \sin \gamma [\cos (\varphi + \Phi) + \varphi \sin (\varphi + \Phi)] + g \cos \gamma}{r \sin \gamma + g \cos \gamma \cdot \cos (\varphi + \Phi)};$$

$$y = gr \frac{\varphi \cdot \cos (\varphi + \Phi) - \sin (\varphi + \Phi)}{r \cdot \sin \gamma + g \cos \gamma \cdot \cos (\varphi + \Phi)}. \qquad (3)$$

Das letztere Gleichungssystem erzeugt in Parameterform die Durchdringungskurve der Evolventen-Schraubenfläche — welche um den Winkel Φ gedreht und mit der Achse geneigt ist — und der Stirnebene $z = 0$ des Zahnradkörpers (der Parameter ist φ). Wenn man auch den Wert von Φ als veränderlich betrachtet, so bedeuten die Beziehungen unter (3) das Gleichungssystem der Durchdringungskurvenschar, die infolge der Fräserdrehung entsteht. Die unter (3) angegebenen Funktionen X und Y mit zwei Veränderlichen entsprechen den im Punkt I angeführten, im Evolventen-Kriterium auftretenden Funktionen $X(t; \Phi)$ und $Y(t, \Phi)$, mit deren Hilfe auch dort die Durchdringungskurvenschar der Fräserfläche und Stirnebene mit dem Unterschied hergestellt wurde, daß dort der Anfangspunkt des Koordinatensystems $(x; y)$ mit dem Mittelpunkt des zu verzahnenden Rades, in (3) dagegen mit dem Fräserzentrum zusammenfällt.

Das Gleichungssystem der Durchdringungskurven wird noch etwas umgeformt, wonach der Wert $r = g \tan \psi$, entsprechend dem Ausdruck (2) in die Gleichungen eingesetzt und anstelle von $(\varphi + \Phi)$ die neue Veränderliche

$$t = \varphi + \Phi$$

eingeführt wird, wobei

$$\varphi = t - \Phi$$

ist.

Demnach nimmt das Gleichungssystem (3), wenn man anstelle von x und y, im Einklang mit den vorangehenden Bezeichnungen, X und Y schreibt, folgende Form an:

$$\begin{aligned} X &= g \frac{\tan \psi \cdot \sin \gamma \cdot \cos t + \tan \psi \cdot \sin \gamma (t - \Phi) \cdot \sin t + \cos \gamma}{\tan \psi \cdot \sin \gamma + \cos \gamma \cdot \cos t}; \\ Y &= g \cdot \tan \psi \cdot \frac{(t - \Phi) \cos t - \sin t}{\tan \psi \cdot \sin \gamma + \cos \gamma \cdot \cos t}, \end{aligned} \tag{4}$$

wobei von Φ und t, Φ den Scharparameter bedeutet.

IV. Untersuchung eines abgewälzten Zahnprofils, hergestellt durch einen Schneckenfräser mit Grundschnecken evolventenförmiger Ausführung

Im vorigen Abschnitt wurde die unter (4) angegebene Durchdringungskurvenschar gewonnen, bei der jetzt das abgeleitete und im Abschnitt I angeführte Evolventenkriterium angewendet wird. Hierfür werden die partiellen Ableitungen der Funktionen X und Y berechnet. Wenn der Kürze halber N den Nenner des unter (4) angegebenen Bruch-Ausdruckes bezeichnet,

$$N = \tan \psi \sin \gamma + \cos \gamma \cdot \cos t,$$

dann erhält man die partiellen Ableitungen:

$$X_t = \frac{g\,[\cos\gamma\cdot\sin t + (t-\Phi)\tan\psi\cdot\sin\gamma]\,(\cos\gamma+\tan\psi\cdot\sin\gamma\cdot\cos t)}{N^2};$$

$$Y_t = -\frac{g\tan\psi\cdot\sin t\,[\cos\gamma\cdot\sin t + (t-\Phi)\tan\psi\cdot\sin\gamma]}{N^2},$$

$$X_\Phi = -\frac{g\tan\psi\cdot\sin\gamma\cdot\sin t}{N}; \quad Y_\Phi = -\frac{g\tan\psi\cdot\cos t}{N};$$

wobei die Indizes t und Φ die partiellen Ableitungen nach den Veränderlichen t bzw. Φ bedeuten. Mit diesen Werten wird die in dem Kriterium enthaltene Funktionaldeterminante:

$$\begin{vmatrix} X_t & X_\Phi \\ Y_t & Y_\Phi \end{vmatrix} = -\frac{g^2\tan\psi\,[\cos\gamma\cdot\sin t + (t-\Phi)\tan\psi\cdot\sin\gamma]}{N^2}. \tag{5}$$

Ferner ist die quadratische Summe, welche ebenfalls im Evolventen-Kriterium bzw. in der partiellen Differentialgleichung vorkommt:

$$X_t^2 + Y_t^2 = \frac{1}{N^4}\cdot g^2\,[\cos\gamma\cdot\sin t + (t-\Phi)\tan\psi\cdot\sin\gamma]^2\,[(\cos\gamma + \\ + \tan\psi\sin\gamma\cos t)^2 + \tan^2\psi\sin^2 t].$$

Wenn man mit diesen Ausdrücken die partielle Differentialgleichung bildet, gewinnt man das Evolventen-Kriterium in folgender Form:

$$\begin{vmatrix} X_t & X_\Phi \\ Y_t & Y_\Phi \end{vmatrix}^2 - \varkappa^2 r_a^2 (X_t^2 + Y_t^2) = \\ = \frac{1}{N^4}\cdot g^2\,[\cos\gamma\cdot\sin t + (t-\Phi)\tan\psi\cdot\sin\gamma]^2\,\{g^2\tan^2\psi - \\ - \varkappa^2 r_a^2\,[(\cos\gamma+\tan\psi\sin\gamma\cos t)^2 + \tan^2\psi\sin^2 t]\} = 0.$$

Diese Gleichung wird offensichtlich erfüllt, wenn der Ausdruck

$$\cos\gamma\sin t + (t-\Phi)\tan\psi\sin\gamma$$

in der eckigen Klammer verschwindet. In diesem Fall verschwindet jedoch auch die Funktionaldeterminante unter (5), obzwar sie nach der Annahme, die in dem Evolventen-Kriterium enthalten ist, von dem Wert 0 abweichend sein muß.

Das Verschwinden der unter (5) angeführten Determinanten bedeutet gerade soviel, daß man jenen Zweig der Diskriminantenkurve von der Durchdringungskurvenschar betrachtet, der aus der Menge der singularen Punkte besteht und daher nicht zur Hüllkurve gehört. Demnach muß zur Erfüllung des Evolventen-Kriteriums gefordert werden, daß der zweite Faktor der obigen Beziehung verschwinde, d. h. die folgende Gleichung identisch bestehe:

$$g^2 \tan^2 \psi - \varkappa^2 r_a^2 [(\cos \gamma + \tan \psi \sin \gamma \cdot \cos t)^2 + \tan^2 \psi \cdot \sin^2 t] = 0. \quad (6)$$

Da in dieser Gleichung die Variable t, vorkommt, so wird die linke Seite der Gleichung lediglich bei bestimmten (diskreten) Werten von t verschwinden.

Man untersuche, unter welchen Bedingungen die linke Seite der Gleichung (6) für alle Werte von t verschwindet. Wenn in die Gleichung (6) der Wert

$$\sin^2 t = 1 - \cos^2 t$$

eingeführt wird und man die linke Seite der Gleichung nach Potenzen von $\cos t$ ordnet, so erhält man für $(\cos t)$ folgende Gleichung zweiten Grades:

$$\varkappa^2 r_a^2 \cos^2 \gamma \cdot \tan^2 \psi \cdot \cos^2 t - 2 \varkappa^2 r_a^2 \tan \psi \cdot \sin \gamma \cdot \cos \gamma \cdot \cos t +$$
$$+ g^2 \tan^2 \psi - \varkappa^2 r_a^2 \cos^2 \gamma - \varkappa^2 r_a^2 \tan^2 \psi = 0.$$

Diese Gleichung wird dann für alle Werte von $\cos t$ erfüllt, wenn seine Koeffizienten verschwinden. Die Koeffizienten von $\cos^2 t$ und $\cos t$ sind dann 0, wenn

$$\cos \gamma = 0$$

ist; das konstante Glied verschwindet im Fall von

$$\cos \gamma = 0$$

nur dann, wenn:

$$g^2 \tan^2 \psi - \varkappa^2 r_a^2 \tan^2 \psi = 0,$$

d. h., weil $\tan \psi \neq 0$ ist, wenn:

$$g^2 - \varkappa^2 r_a^2 = 0$$

ist.

$\cos \gamma = 0$ wird jedoch bei $\gamma = 90°$ erfüllt, und dies bedeutet nach Bild 2, daß die T-Achse der Schraubenfläche mit der z-Achse zusammenfällt. In diesem Fall steht die $z = 0$ Ebene senkrecht auf der Achse der Schraubenfläche und schneidet die Fläche in einer gewöhnlichen (spitzen) Kreisevolvente, da die Stirnschnittkurve der Evolventen-Schraubenfläche eine spitze Kreis-

evolvente ist. In diesem Fall sind also die im Stirnschnitt entstehenden Durchdringungskurven Kreisevolventen, und dies entspricht einer Erzeugung des Evolventen-Zahnprofils mittels Stoßrades, wobei die Hüllkurve der Kreisevolventenschar das abgewälzte Zahnprofil liefert. Da bei der Zahnprofilerzeugung mittels des Stoßrades hinsichtlich des Verhältnisses $\varkappa$ der simultanen Drehwinkel von Werkzeug und Werkstück folgende Beziehung besteht:

$$\varkappa = \frac{g}{r_a},$$

wobei g den Radius des Stoßrades und r_a den des Grundkreises am zu erzeugenden Zahnrad bedeutet, so ist:

$$g^2 - \varkappa^2 r_a^2 = 0,$$

demnach verschwindet auch das konstante Glied der vorher angeschriebenen Gleichung zweiten Grades, als Voraussetzung der identischen Erfüllung der Beziehung unter (6). Es ist bekannt, daß die Evolventenschnecke (Schraubenfläche) eigentlich ein schrägverzahntes Stirnrad ist, bei dem der Schrägungswinkel β gleich dem Ergänzungswinkel vom am Teilzylinder befindlichen Steigungswinkel γ ist, d. h.

$$\beta = 90° - \gamma.$$

Da der Schrägungswinkel bei Geradzähnen $\beta = 0$ und demnach, entsprechend der vorigen Formel $\gamma = 90°$ ist, entspricht dem geradverzahnten Stoßrad der Fall $\gamma = 90°$, wobei das unter (6) angeführte Kriterium erfüllt ist, da das abgewälzte Zahnprofil wirklich eine gewöhnliche Kreisevolvente darstellt.

Mit Rücksicht darauf, daß der Steigungswinkel bei Wälzfräsern $\gamma \neq 90°$ (bei eingängigen Fräsern $\gamma_{max} = 6°$, bei zweigängigen $\gamma_{max} = 10°$) und in dieser Weise $\cos \gamma \neq 0$, verschwindet die Gleichung unter (6) nicht identisch. Dieser Umstand bedeutet gerade soviel, daß das im Stirnschnitt erzeugte Zahnprofil *keine spitze Kreisevolvente* sein wird. Wenn man jedoch aus der Gleichung unter (6) — welche in ($\cos t$) zweiten Grades ist — einen Wert $t = t_0$ bestimmt, bei dem die Gleichung erfüllt ist, dann wird die linke Seite der Gleichung (6) in einer sehr kleinen Umgebung der Stelle, bei der $t = t_0$ ist, sehr kleine Werte annehmen. Daher wird die Bedingung (6) in einer sehr kleinen Umgebung der Stelle t_0 »annähernd« erfüllt sein, und so kann man annehmen, daß ein kleiner Abschnitt des erzeugten Zahnprofils, der gerade der Stelle t_0 und einer kleinen Umgebung von t_0 entspricht, in gewissen Fällen eine gewöhnliche Kreisevolvente annähern wird. Deswegen in »gewissen« Fällen, da das Kriterium (6) gerade die notwendige Bedingung dafür aus-

drückt, daß aus dem erzeugten Profil eine Kreisevolvente wird, und in dieser Weise ist die Erfüllung der Gleichung (6) allein noch kein Beweis dafür, daß das erzeugte Zahnprofil in der Tat eine Kreisevolvente ist.

V. Untersuchung eines durch evolventenförmigen Schneckenfräser erzeugten Zahnprofils unter Berücksichtigung des Fräservorschubs

Im Rahmen der im vorigen Punkt angestellten Überlegungen wurde die Vorschubbewegung des Fräsers parallel mit der Werkstückachse nicht berücksichtigt. Die Überlegungen im Abschnitt IV entsprechen daher dem Fall, bei dem der Fräserzahn, infolge des Fräservorschubs, in voller Tiefe in das Material des Werkstückes eindringt, und es beginnt die Abwälzung des Zahnprofils in seiner endgültigen Form in einem Stirnschnitt des Werkstückes. Im weiteren Verlauf wird die durch den Fräser ausgeführte Wälzbewegung in ihrer Gesamtheit betrachtet, und es wird angenommen, daß der sich um seine T-Achse drehende Fräser (Bild 2) auch noch in Richtung der z-Achse eine geradlinige Vorschubbewegung ausführt. In diesem Fall, da der Drehwinkel des Fräsers um seine T-Achse mit Φ bezeichnet wurde, verändert sich das im Abschnitt III abgeleitete Gleichungssystem der Schraubenfläche in der Weise, daß man dem Ausdruck für die z-Koordinate eine lineare Funktion

$$c_0 + \Phi\, c_1$$

des Drehwinkels Φ hinzufügen muß (c_0, c_1 sind Koeffizienten). Daher wird das Parameter-Gleichungssystem der Schraubenfläche, welche um den Winkel Φ gedreht, und um das Stück ($c_1\, \Phi$) verschoben worden ist, folgende Form annehmen:

$$x = g \cos(\varphi + \Phi) - g \sin(\varphi + \Phi) \cdot t;$$

$$y = -[g \sin(\varphi + \Phi) + g \cos(\varphi + \Phi) \cdot t] \sin\gamma + (r\varphi + rt) \cos\gamma;$$

$$z = -[g \sin(\varphi + \Phi) + g \cos(\varphi + \Phi)t] \cos\gamma - (r\varphi + rt) \sin\gamma + c_0 + c_1 \Phi.$$

Genau so wie im Abschnitt III wird auch hier die Schnittkurve dieser Fläche mit der $z = 0$ Ebene bestimmt. Wenn im obigen Gleichungssystem $z = 0$ ist, so ergibt sich aus der dritten Gleichung:

$$t = \frac{c_0 + c_1 \Phi}{r \sin\gamma + g \cos\gamma \cos(\varphi + \Phi)} - \frac{g \sin(\varphi + \Phi) \cos\gamma + r\varphi \sin\gamma}{r \sin\gamma + g \cos\gamma \cdot \cos(\varphi + \Phi)}.$$

Wenn man den Wert t in die Ausdrücke für x und y einsetzt, und die Reduktionen durchführt, und ferner den Wert $r = g \tan\psi$ und die Veränderliche $t =$

$= \varphi + \Phi$ einführt ($\varphi = t - \Phi$), dann die für x und y in dieser Weise gewonnenen Ausdrücke mit ξ bzw. mit η bezeichnet, so erhält man folgende Beziehungen:

$$\xi = X - \frac{c_0 + c_1 \Phi}{N} \sin t;$$

$$\eta = Y - \frac{(c_0 + c_1 \Phi)(\sin \gamma \cdot \cos t - \cos \gamma \cdot \tan \psi)}{N}, \qquad (4^*)$$

wobei X und Y die unter (4) angegebenen Ausdrücke bezeichnen, und

$$N = \tan \psi \cdot \sin \gamma + \cos \gamma \cdot \cos t$$

ist.

Der Drehwinkel Φ des Fräsers in den Gleichungen (4*), welche die Durchdringungskurve der Stirnebene $z = 0$ und der Fräserfläche ergeben, nimmt die Form an:

$$\Phi = 2 \pi n + \Phi',$$

wobei n eine ganze Zahl ist, und für Φ' gilt:

$$0 \leqq \Phi' < 2 \pi.$$

Wegen der Periodizität der Funktionen $\sin t$ und $\cos t$ kann man für den Winkelparameter t folgende Beziehung anschreiben:

$$t = 2 \pi n + t',$$

$$0 \leqq t' < 2 \pi$$

und so gilt für die in (4) auftretenden Differenzen in den Ausdrücken von X und Y:

$$t - \Phi = t' - \Phi'; \quad (0 \leqq t' < 2 \pi; \quad 0 \leqq \Phi' < 2 \pi)$$

daß

$$0 \leqq t - \Phi < 2 \pi$$

ist.

Daraus folgt, daß die Ausdrücke X und Y in (4) lediglich Koordinaten von Punkten liefern, die auf einem endlich langen Stück der Schraubenfläche (am Fräser) liegen. Nun wird das Evolventen-Kriterium bei der Durchdringungsschar von (4*) angewandt. Die hierfür erforderlichen Werte der partiellen

Ableitungen nach t und Φ sind folgende:

$$\xi_t = X_t - \frac{\cos\gamma + \tan\psi \cdot \sin\gamma \cdot \cos t}{N^2}(c_0 + c_1\Phi);$$

$$\eta_t = Y_t + \frac{\tan\psi \sin t}{N^2}(c_0 + c_1\Phi),$$

$$\xi_\Phi = X_\Phi - \frac{c_1 \sin t}{N};\ \eta_\Phi = Y_\Phi - \frac{c_1(\cos t \cdot \sin\gamma - \tan\psi \cdot \cos\gamma)}{N},$$

wobei die Werte der partiellen Ableitungen X_t, Y_t, X_Φ, Y_Φ am Anfang des Abschnittes IV aufzufinden sind. Mit diesen partiellen Ableitungen wird das Evolventen-Kriterium durch folgende Identität geliefert:

$$\begin{vmatrix} \xi_t & \xi_\Phi \\ \eta_t & \eta_\Phi \end{vmatrix}^2 - \varkappa^2 r_a^2 (\xi_t^2 + \eta_t^2) = 0,$$

wobei

$$\varkappa = \frac{1}{z};\ r_a = \frac{z}{2}\cos\alpha;\ \varkappa r_a = \frac{\cos\alpha}{2}$$

sind. (Hierbei ist z die Zähnezahl des zu verzahnenden Rades.)

Die angeschriebene Gleichung muß für alle Werte von Φ und t erfüllt sein, bei denen

$$\begin{vmatrix} \xi_t & \xi_\Phi \\ \eta_t & \eta_\Phi \end{vmatrix} \neq 0$$

ist. Man wähle folgende Werte für Φ und t:

$$\Phi = -\frac{c_0}{c_1},\ t = \frac{\pi}{2},$$

dann sind die partiellen Ableitungen:

$$\xi_t = \frac{g\left[\cos\gamma + (t - \Phi)\tan\psi \cdot \sin\gamma\right]\cos\gamma}{\tan^2\psi \cdot \sin^2\gamma};$$

$$\eta_t = -\frac{g\tan\psi\left[\cos\gamma + (t - \Phi)\tan\psi \cdot \sin\gamma\right]}{\tan^2\psi \cdot \sin^2\gamma};$$

$$\xi_\Phi = -\frac{g\tan\psi \cdot \sin\gamma + c_1}{\tan\psi \cdot \sin\gamma};\quad \eta_\Phi = \frac{c_1\cos\gamma}{\sin\gamma}$$

und damit ergibt sich der Wert der Determinanten zweiter Ordnung:

$$\begin{vmatrix} \xi_t & \xi_\Phi \\ \eta_t & \eta_\Phi \end{vmatrix} = -g \frac{g \tan\psi + c_1 \sin\gamma}{\tan^2\psi \cdot \sin^2\gamma} [\cos\gamma + (t - \Phi) \tan\psi \cdot \sin\gamma] \neq 0,$$

wo jetzt:

$$t - \Phi = \frac{\pi}{2} + \frac{c_0}{c_1}$$

ist. Außerdem gilt noch:

$$\xi_t^2 + \eta_t^2 = g^2 \frac{\cos^2\gamma + \tan^2\psi}{\tan^4\psi \cdot \sin^4\gamma} [\cos\gamma + (t - \Phi) \tan\psi \cdot \sin\gamma]^2,$$

daher muß bei der Erfüllung des Evolventen-Kriteriums folgende Gleichung bestehen:

$$\begin{vmatrix} \xi_t & \xi_\Phi \\ \eta_t & \eta_\Phi \end{vmatrix}^2 - \varkappa^2 r_a^2 (\xi_t^2 + \eta_t^2) =$$

$$\left(t = \frac{\pi}{2}; \Phi = -\frac{c_0}{c_1}\right)$$

$$= \frac{1}{\tan^4\psi \cdot \sin^4\gamma} \cdot g^2 [\cos\gamma + (t - \Phi) \tan\psi \cdot \sin\gamma]^2$$

$$\left[(g \tan\psi + c_1 \sin\gamma)^2 - \frac{\cos^2\alpha}{4} (\cos^2\gamma + \tan^2\psi)\right] = 0.$$

Da der Ausdruck in der ersten eckigen Klammer nicht Null ist, so kann die Gleichheit nur dann erfüllt sein, wenn der andere Faktor verschwindet, d. h., wenn:

$$(g \tan\psi + c_1 \sin\gamma)^2 - \frac{\cos^2\alpha}{4} (\cos^2\gamma + \tan^2\psi) = 0$$

ist. In dieser Gleichung bedeutet c_1 den Vorschub des Fräsers, der einer Drehung um eine Winkeleinheit entspricht und dessen Wert verändert werden kann. Daher kann die obige Gleichung zweiten Grades (für c_1) auch für einen Vorschubwert von c_1 nicht bestehen, der von den beiden Wurzeln der Gleichung abweicht.

Damit wurde nachgewiesen, daß ein Zahnprofil, welches durch Abwälzung mit einem Evolventen-Schneckenfräser hergestellt wird, von der gewöhnlichen Kreisevolvente abweicht.

SCHRIFTTUM

1. Lipka, I.: Wältztheorie von Zahnrädern mit Evolventenprofil. *Acta Techn. Hung.* **60** (1968).
2. Szeniczei, L.: Csigahajtóművek (Schneckentriebwerke); Budapest 1957.

The Geometry of Hobs with Involute Basic Worm. The author deduced a criterion for deciding wether a gear profile produced by some generating tool is an ordinary involute of a circle or a curve differing from it. This involute criterion decides by the fulfilment or a partial differential equation the character of the generated tooth profile. In the present paper the author applies the involute criterion to a gear profile generated by a milling cutter with an involute basic worm. He proves that a tooth profile generated by such a cutter is not an ordinary (pointed) involute of a circle.

Géométrie des fraises vis-mère à vis de base à développante. L'auteur déduit un critère permettant de déterminer si le profil d'engrenage produit par un outil générateur est une développante circulaire ordinaire ou une courbe différente de celle-ci. Ce critère indique la caractère de développante du profil engendré d'après la satisfaction d'une équation différentielle partielle. L'auteur applique ce critère de développante au profil de denture engendré par une fraise à vis de base à développante. Il montre que le profil de dent engendré par une fraise ayant une telle vis de base ne sera pas une développante circulaire ordinaire (à denture pointue).

Геометрия обкаточных червячных фрез с эвольвентным червяком *(И. Липка).* Автором выведен критерий для выяснения вопроса, является ли профиль зубчатого колеса, изготовленного при помощи какого-либо обкаточного инструмента, простой круговой эвольвентой, или же отклоняющейся от нее кривой. Этот критерий эвольвентности на основе удовлетворения некоторого парциального дифференциального уравнения решает характер эвольвентности обкатанного профиля зубьев. Автор в этой работе применяет критерий эвольвентности для профиля зубчатых колес, обработанных обкаткой при помощи червячной фрезы с эволвентным червяком. В работе показано, что обкаткой с червячной фрезой с таким червяком профиль зуба не будет обычной круговой (острой) эвольвентой.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (3—4), pp. 277—350 (1968)

THE PROPER CONDUCTION OF THE FLUX OF MECHANICAL ENERGY WITH SPECIAL REGARD TO THE SELF-BALANCED MOVABLE CLUTCH

B. SZŐKE

[Manuscript received May 2, 1966]

Up to now, it was not possible to find, in the technical literature such a special formula which has now been deduced by the author on plain theoretical basis inasmuch as the proposed equation serves as a conclusive explanation of the strange occurrence experienced with a series of mill rolls, when namely, the feed-back output disappears at a gear ratio approaching 1 : 1 or ∞ : 1. The author demonstrates a special grouping of the coupling in series of mill frames having rolls with various diameters in which the feed-back of the energy-flux does never charge the driving gear but its circulation occurs within the roller mill system: thus the whole train of mechanism (gears, clutches) has to bear a less load. Further, the self-balancing movable clutch is considered. By theoretical analysis it is shown that this type of clutch represents a newly modelled machine part that serves as an insulator of oscillations and not as a damper.

List of symbols

PART *A)*

b	working gap between rolls;
c	spring rate [cm/kp];
d	maximum grain diameter to be drawn-in;
e	moment lever-arm on free wheeling roll for crushing;
h	free length of roll;
i	$v/v_1 > 1$, ratio of the surface speeds in one pair of rolls;
k	$= D_1/D = R_1/R < 1$, proportion of the diameter of the small roll to the diameter of the large roll (in case of equal diameters the subscript denotes the slow roll);
k_s	shear strength;
l_0	free length of the spring;
n	$r \cdot p \cdot m$ of the fast roll;
n_1	$= n/i$, $r \cdot p \cdot m$ of the slow roll;
r	pitch circle radius of the pinion (at the fast roll);
r_1	pitch circle radius of the large wheel (at the slow roll);
r_c	roll journal radius;
s	spring rigidity [kp/cm];
t	$= R + R_1 + b$, axial distance;
v	velocity, especially the surface velocity of the fast roll;
v_1	surface velocity of the slow roll;
w	relative velocity;
A	power input;
A_1	loss of power at the direct drive;
A_2	friction loss in the bearings (the second cipher in the subscript indicates the bearing in question);
A_3	friction loss at the driving gear;
A_4	friction loss at the scrapers;
A_5	power required by the feeding rolls;
A_a	power required for crushing;
A_z	power required for crumbling;

A_{δ} = $A_a + A_z$, power required for total grinding;
B_{max} = $d \cos \varrho$, max. width of the grinding tape;
B = $2l \cos \varphi$, width of the grinding tape fed in;
D roll diameter (fast roll);
D_1 roll diameter (slow roll);
F algebraic sum of friction forces during grinding;
L = $A + W$, piled-up power;
M torque of the fast roll;
M_a moment of the crushing force at the fast roll;
Ma_1 moment of the crushing force at the slow roll;
M_1 torque due to journal friction, reduced to radius R;
N radial force acting on the roll;
P power value [HP];
Q tooth pressure;
R radius of the fast roll;
R_1 radius of the slow roll;
S (crushing) shear force;
S_1 friction force;
S outer similarity point of the roll circles (for equal roll diameters S_∞);
T resultant of elementary forces traversing the similarity point;
W feed-back power, when friction loss is neglected;
W_s feed-back power, when friction loss is taken into account;
Z the outer resistance acting on the slow roll as a force reduced to the cylindric surface, when the teeth flanks become changed;
ε = λ/l_0, specific compressibility of the spring;
ε_r relative angular acceleration when teeth flanks become changed;
ϑ back lash angle at the large wheel;
λ compression of the elementary spring;
μ friction coefficient between rolls;
μ_c friction coefficient of the journal;
φ varying angle value, absolute;
ω angular velocity of the fast roll;
ω_1 angular velocity of the large gear (and of the slow roll);
ω_ϑ angular backlash overtake velocity of the large wheel during the change-over tooth of the flanks;
ω_r = $\omega_\vartheta - \omega$, relative angular velocity of the slow roll in relation to the fast roll during the change-over of the tooth-flanks;
Θ moment of inertia of the masses rotating together with the slow roll.

PART *B)*

a_1 relative acceleration of the masses free-sliding in guide-way I of the floating disk;
a_2 relative acceleration of the masses free-sliding in guide-way II of the floating disk;
a_{1c} Coriolis acceleration of the free-sliding masses in guide-way I;
a_{2c} Coriolis acceleration of the force-sliding masses in guide-way II;
b distance of the roller centre from the edge of the guide-way (Fig. 18a);
d bore in the roller;
d_1, d_1' pair of opposite pins in the coupling lever k_1;
d_2, d_2' pair of opposite pins in the coupling lever k_2;
e axial distance (eccentricity) of parallel shafts, or normal transversal distance between shafts in skew position;
h_1, h_2 width of the two guide-ways in the floating disk;
h distance of the intersection point of shafts from the floating disk;
h_{min} minimum allowable distance between roller centres to prevent interference;
k_1 coupling arm on shaft t_1;
k_2 coupling arm on shaft t_2;
l floating disk;
l width of the roller;
m mass of the floating disk (and its mass point);
m_1 free-sliding mass in guide way I;
m_2 free-sliding mass in guide way II;
m_c total mass of the rolling bearing and the pin-extension;
p Hertzian stress;

r moment arm (distance of pin centre d_1 from shaft centre t_1);
s minimum distance between the cylindrical surfaces of the rollers when no interference is to occur between the rollers of the two guide ways;
s_1 safety distance when the roller is not to slip into the other guide-way;
s_1 relation displacement of free-sliding mass m_1 in guide-way I;
s_2 safety distance when the roller is not to slip from its own guide way;
s_2 relation displacement of free-sliding mass m_2 in guide-way II;
s_3 safety distance's value made uniform;
t_1 driving shaft;
t_2 driven shaft;
v_1 relative velocity of the free-sliding mass m_1;
v_2 relative velocity of the free-sliding mass m_2;
z axial forced displacement of the floating disk;
C centrifugal force of the floating disk;
D_f shaftneck end-diameter;
D outer diameter of the roller bearing (actually with a reference subscript);
E resultant of the peripheral and the centrifugal forces;
E balance weight (equalling the floating disk's weight);
F peripheral force;
F_{1r} mass force as given by the acceleration of the force-sliding mass m_1;
F_{2r} mass force as given by the acceleration of the force-sliding mass m_2;
F_{1s} carrying force of the mass m_1;
F_{2s} carrying force of the mass m_2;
F_{1c} Coriolis force of mass m_1;
F_{2c} Coriolis force of mass m_2;
F_r resultant of forces F_{1r} and F_{2r};
F_s resultant of forces F_{1s} and F_{2s};
F_c resultant of forces F_{1c} and F_{2c};
G net weight of the roller and the pin extension;
M $= Fr$, moment transferred;
O_1 centre of shaft t_1;
O_2 centre of shaft t_2;
O geometric centre of the floating disk assumed as identic with its mass point m;
R roller radius (having the respective subscript);
α angle of axes in intersecting or skew position;
σ_t allowable tensile stress of the roller pin;
σ_t allowable tensile stress of the shaft neck;
φ rotation angle of shaft;
ω angular acceleration of shaft.

PART *A)*: MECHANICAL REGULARITIES THAT HAVE TO BE CONSIDERED FOR THE PURPOSE OF A COUPLING IN SERIES

I. Introduction

(Some practical experiences)

In the practice of mechanical engineering it is commonly known that newly developed details, seemingly of simple design, instead of proving satisfactory, bring about surprising failures during operation in spite of the most carefully elaborated construction and perfect workmanship. It happened during the years 1928—1932 that the director of the *EMAG* (First Hungarian Factory of Agricultural Machines), Mr. KALMÁR charged one of the most capable engineers, to study a new type of grain-washing machine having a hori-

zontal shaft. Although the mechanical calculation was carried out carefully, the materials well selected, the machine built with a high workmanship, the operators and engineers on the machine put into operation did not experience but fracture of the gear wheels; afterwards, when a chain drive was built in instead of the gear wheels, break down of the chain occurred systematically as well. The whole technical staff was confronted with incomprehensible phenomena. Later on, when in 1932, the year of economic depression accompanied by the dismissal of experts the present writer being the single designer in the factory, had been studying to construct a grain washer without a steam operated drying column attachment, only a drum-model with a vertical axis [1] could be found as a realizable first project.

In our times, the answer to this problem is already found. It is the reflux of energy that becomes fedback onto the driving mechanism and amounts to a *multiple of the motor output*, by which the inevitable fracture of the driving parts of the drum-type grain washer is induced.

For driving the pair of rolls of a mill-frame, the geartype positive coupling that ensured the differing surface speed of the rolls was in common use already in the 19th century. However, it happened only in the thirties of the 20-eth century that resarchers [2] took notion of the paradoxical fact that the wear of gears presented itself on the tooth flanks which would not act when the pinion of the motor imparted its energy to the larger toothed wheel of the roll, but if the flux of energy would be inverted: the large wheel operating as a driver, and the pinion as a driven gear; in other words as if the large wheel operating with a less number of rotations could feed back the energy to the pinion [3].

It was T. GOMBÁS, retired engineer (formerly at the *Ganz-MÁVAG Works*) who informed the present writer about the odd experience made at that time when for want of gear cutting machine tools gears were not machined but cast. This happened at the factory of the *Ganz-Danubius Works* where they used to lap in the chill-cast herringbone gears by special wearing-in machines before assembling; namely, after a certain service period the wear due to operation was to be found on the originally rough flanks that had not been lapped in, and not on the flanks already run in previously.

Experiments carried out during decades at the workshops of the American *Allis-Chalmers Factory* remained fruitless when they tried to solve the problem how to bring about the coupling in series of roll mills; obviously, the researchers did not perceive the special occurrence of a reverse reflux of energy.

II. The problem to be solved

When milling roll-pairs are coupled in series, one can observe as a supplementary result that the individual belt-drive that involves the risk of accidents ceases to be applied. In this connection the main requirement can be set up as

follows; all the disadvantageous features usually experienced in the function, and generally observed in the design of the original type roller mills should be eliminated.

Another important requirement comprises the possibility of *correctly adjusting the working gap* between the rolls [33]. When, due to some faulty design, the suspension of the spring loaded bearing boxes is inappropriate [1, 4, 5] and when on account of the variation of load the circumferential force of the gear or the chain drive varies too [3] a correct adjustment of the working gap for a longer period becomes impossible, the grading passages of the mill will be charged ever more by parts of the grist, and thus the efficiency of the whole mechanism is impaired. Consequently, the essential and crucial point of the coupling in series of the grinding rolls is ambodied by a *movable clutch* the function of which is not to transfer to the adjacent bearings some additional charge, if this originates from some elastic deformation or from the kinematic train of the clutch. The clutch should be designed as a device that transmits the torque in the absence of a tilting moment, and by means of a *plain couple of forces and not of some unilateral peripheral force* which might raise reaction forces in the bearings induced by power transfer.

Should the required type of clutch be developed in a form that can easily be assembled, an additional advantage should be obtained as the whole gear drive plus all the reserve gears could be dispensed with the storage of which would otherwise be necessary because of repeated maintenance (grinding, overhauling, fluting). On the other hand, according to up-to-date operational principles, the application of toothed wheels in a mechanism where neither the axial distance nor the parallelity of the axes can be kept constant, is not motivated at all.

Further, it is indispensable to get a true perception and a thorough comprehension of the entire *flux of mechanical energy conducted through the grinding rolls*. First this requires that the energy flux in a single pair of rolls should be analysed and the theoretical principles should be clarified.

Only if the above investigations are made, is it possible to find the answer to the question of how the mechanical parts of the whole machinery could be protected from overcharging, especially, by the correct selection of the roll diameters, by an adequate grouping of the roll-pairs, and by an appropriate arrangement of the driving gears. Grinding mills that comply with these specifications are cheaper, less voluminous, and work, in addition, by consuming less energy.

III. Recalling of some fundamental principles

Power-input-output measurements carried out according to up-to-date measuring methods show us a correct picture of the whole course of the flux of

energy [6—8]. As against these facts, theoretical examinations indicate contradictory results. Thus, the author feels it necessary to recall some fundamental principles relating to some general mechanical laws.

a) Any mechanical flux of energy can be considered as the product of two quantities, one of which can be characterized as an idle factor, and the other one as the effective factor. Speaking of thermal energy, the energy flux is directed always from the body having the higher *temperature*, towards the body having the lower temperature. In communicating vessels the liquid stream is always directed from the vessel having the *higher fluid-column* towards the vessel containing the fluid at a lower level. In the formula

$$L = F \cdot v$$

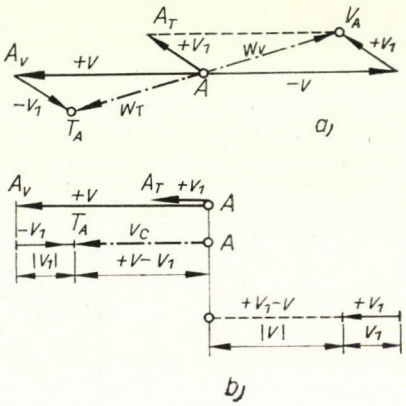


Fig. 1. Velocity in the effective factor of the energy flux: relative velocities when directions deviate (*a*); relative velocities when directions are identic (*b*)

(mechanical Effect is the product of Force and Velocity), it is the *Velocity* v that represents the effective factor.

Let us imagine two bodies mowing at different speeds close to each other, one rubbing against the other. It is always the body moving with a greater velocity that transfers energy to the other one having the lower velocity.*

For simplicity's sake, in order to demonstrate the general case, we will consider the example of a tramway and a cart moving at different slopes. The velocity of the tramway will be indicated as v and that of the cart as v_1 and $v_1 < v$.

With reference to Fig. 1a, some point A of the tram moves during the unit of time to point A_v namely in compliance with the tram-velocity $\overrightarrow{AA}_v = v$. Simultaneously, the point on the cart being in contact with the trampoint A moves to A_T, according to the cart-velocity $v_1 = \overrightarrow{AA}_T$. Thus the relative veloc-

*In this imaginary example, the velocity values are assumed as being kept constant.

ity of the tram-point A alongside the cart is

$$\vec{w}_T = +\vec{v} - \vec{v}_1 = \overrightarrow{AT}_A = w_T ,$$

in other words, a scratched line $\overline{AT}_A$ is delineated during the unit of time on the side-surface of the cart by the tram-point A. Consequently, the effect L imparted to the cart is

$$L = N\mu(\vec{v} - \vec{v}_1) = N\mu \cdot w_T$$

where N denotes the pressing force exerted on the contact point in a normal direction and μ is the friction-coefficient.

Again, the relative velocity of the cart-point A alongside the tram is

$$\vec{w}_v = +\vec{v} - \vec{v}_1 = \overrightarrow{AV}_A ,$$

i.e. the scratch $\overline{AV}_A$ is delineated on the side surface of the tram. Consequently, the energy delivered by the tram is

$$-L = N\mu(\vec{v}_1 - \vec{v}) = -N\mu\vec{w}_v .$$

Let us assume that the two vehicles move in a parallel direction. It follows that instead of vectorial addition, a linear (arithmetic) addition of velocities takes place (Fig. 1b). Denoting the ratio of velocities by

$$i = \frac{v}{v_1}$$

and the tangential force

$$F = N\mu ,$$

we obtain

$$L = F(v - v_1) = Fv\left(1 - \frac{v_1}{v}\right) = Fv\frac{i-1}{i} . \qquad (1)$$

b) Referring to other works of the author [9, 10], and with a view to Fig. 2, the thesis holds true according to which the force that acts upon a spherical grain being in a drawing-in position between two parallel cylindric rolls, coincides with a straight line that goes through the *outer similarity-point* of the cylinders. This force is the resultant of the normal force N which is perpendicular to the cylindric roll surface and the friction force $F = N \cdot \tan \varrho$, where ϱ denotes the friction angle between the material and the cylinder surface. The centres of the spherical grains belonging to various friction angles *are located on a conical section* [1, 11, 12].

c) At points A and B equal normal forces belong to the straight line that traverses the outer similarity point, and equal elementary forces act in this line on the cylindric roll surface; this corresponds to the case, when two men hold a compressed helical spring in a compressed state. The pressing force does not change when the centre part of the spring is propped by a wall. In general, when the spring force is normal to the cylindric surface, the geometric locus of

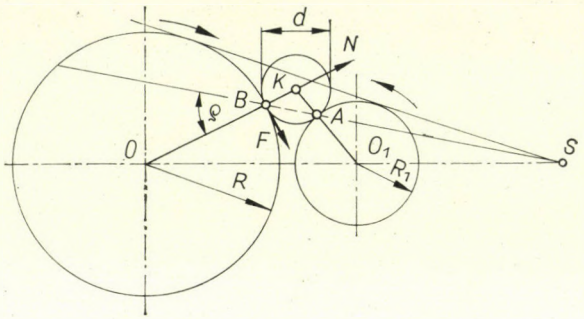


Fig. 2. Geometric relationship and rule of a spheric grain to be drawn-in

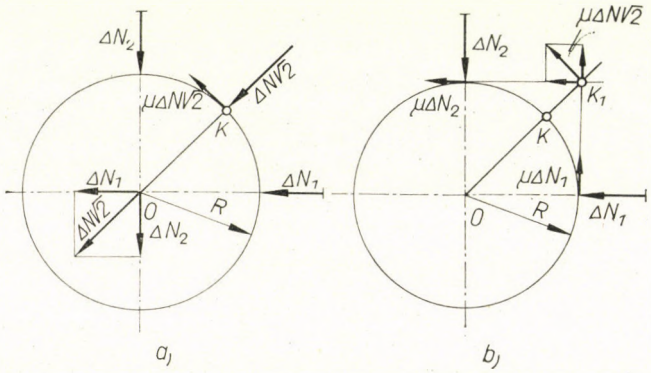


Fig. 3. Moments of friction forces: $\Delta M = R\,\mu\Delta N\,\sqrt{2}$ is the moment of the resultant of elementary friction forces (*a*); the elementary friction forces $\Delta M_1 = R\sqrt{2}\,\mu\Delta N\,\sqrt{2} = 2\,R\,\mu\Delta N$ (*b*)

the "spring centres" is a conical section, and when the rolls are parallel, this section is represented by a straight line. On the other hand, when the springs are compressed, the work exerted by the rolls has to be calculated separately for each roll.

d) According to a statement established by the author in other studies [13, 14], the moment of the friction force calculated from the vectorial resultant of the normal forces is always less than the algebraic sum of the moments of the elementary friction forces. Thus, we would follow a wrong aspect when we consider forces but acting on a single grain or masspoint, respectively.

In correspondence with load conditions as illustrated in Fig. 3, and assuming the equality of the forces

$$|\Delta N_1| = |\Delta N_2| = \Delta N$$

according to Fig. 3a, the elementary moment ΔM of the friction-force induced by the resultant $\overrightarrow{\Delta N \sqrt{2}}$ (as the vectorial sum $\overrightarrow{\Delta N_1} + \overrightarrow{\Delta N_2}$ of the elementary forces) is expressed by

$$\Delta M = \Delta N \sqrt{2}\, \mu\, \overline{OK} = \Delta N \sqrt{2}\, \mu\, R.$$

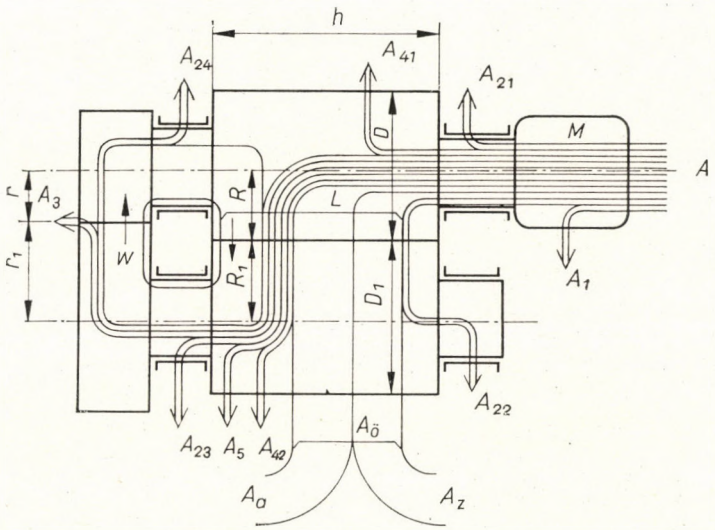


Fig. 4. Schematic stream of the energy-flux in a pair of grinding rolls

Denotations: M drive of the fast roll; W feed-back power by which the driving gear is loaded; A = input; $L = A + W$ piled-up power; A_0 = grinding power; A_1 = loss of drive; A_2 = friction loss in bearings; A_3 = friction loss of gears; A_4 = friction loss of scrapers; A_5 = power required for the drive of feeding rolls.

Again, the algebraic sum of the moments of the friction forces induced by the elementary forces is (according to Fig. 3b)

$$\Delta M_1 = 2\, \Delta N \mu\, R > \Delta M.$$

The validity of this inequality can easily be perceived when we clearly bear in mind that the resultant $\Delta N\ \mu \sqrt{2}$ of the two elementary forces acts on the arm $OK_1 > OK$, and OK_1 equals $R \sqrt{2}$; consequently:

$$\Delta N \mu \sqrt{2}\, OK_1 = \Delta N \mu \sqrt{2}\ R \sqrt{2} = 2\, \Delta N \mu\, R = \Delta M_1.$$

IV. Character of the flux of energy

As a result of several measurements carried out by some researchers [6—8] the energy flux passing through a pair of grinding rolls can be drawn according to the diagram shown in Fig. 4. Symbols in this Figure are:

A input, covering the folloving partial items;
A_1 loss in the motor, or the belt drive, respectively;
A_2 $A_{21} + A_{22} + A_{23} + A_{24}$ = friction loss in the bearings;
A_3 friction loss in the gear drive;
A_4 $A_{41} + A_{42}$ power required by the scrapers;
A_5 power required by the drive of the feeding roll;
A_z power required for crushing*;
A_a power required for crumbling.*

As a rule, when speaking of the milling trade, the notion of grinding includes the grading operation, too. Again, for our purposes, we will denote as grinding operation the simultaneous process of crushing and crumbling. In other words, for the present example, the sum of the items $A_z + A_a = A_{\ddot{o}}$ can be considered as the useful output of the whole grinding system. At the present state of technology, the friction loss A_2 (in bearings and gears) can be considered negligible as compared to $A_{\ddot{o}}$. Nevertheless, when the flux conditions as shown in the diagram (Fig. 4) are duly considered, even by neglecting A_2, the input A does not suffice in keeping the mill in operation. We have to take care of an additional input in order to cover the plus load W that represents some idle or dead power by which certain parts of the mechanism are charged, when the flowing direction of this flux points *from the effectively driving gear to the driven gear*. Thus, the transmission of the motor-output — according to point *a)* of Chapter III — is not performed by the pinion having radius r (being attached to the roll having radius R). Just because of $R > r$, the transmission is effected through the grist friction *on the cylindric roll surface having the greater radius* R, and rotating with the greater surface speed. Since the radius of the gear fixed onto the axis of the slow roll ($r_1 > R_1$) it is the larger wheel having the greater radius r_1 that assumes the function of a driving gear. Thus, the piled-up effect

$$L = A + W$$

will be the power that represents the total load of the fast-roll.

By overcoming the mechanical friction, the inner friction and the cohesion of the material, parts of the input power are transformed into heat and exert a heating effect on the machine, the grist and the environment.

Let us assume another arrangement of the driving mechanism according to the scheme in Fig. 5. Here the driving torque of the motor is transmitted,

* In this paper, we adhere to the following definitions: *Crushing*, disintegration by plain compressive forces; *crumbling*, disintegration by plain shearing forces; *grinding*, the simultaneous compound operation including both crushing and crumbling.

instead of the larger driving roll, to the larger toothed wheel. As is shown by this diagram, the flux of energy differs from what is delineated in Fig. 4. The total driving energy is transmitted to the fast-roll-axis not through the slow roll but by the larger tooth wheel. Consequently, the driving gear and the axis of the fast roll are charged by the additional feed-back power W. When compared with the former arrangement, it is obvious that the driving gear and the fast roll axis are subject to a much greater load; and this happens in spite of the fact that grinding conditions and also the grinding output $A_{\ddot{o}}$ are the same as before.

Thus, the safe operation of the whole mechanism is highly promoted when the nature of the energy flux is duly considered in its connection with the coupling in series of the pairs of rolls.

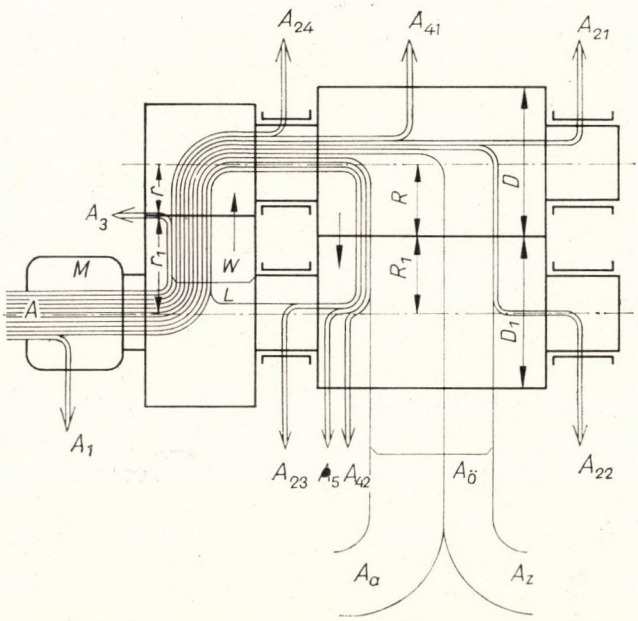


Fig. 5. Schematic stream of the energy-flux in a pair of grinding rolls

Denotations: M drive of the slow roll; $L = A + W$ piled up power by which the driving gear is loaded.

Up to now, an acceptable motivation of this special reflux of energy in grinding mechanisms could not be found in litarary sources. In the present state, the author tried to reveal the interrelation of these occurrences [3]. In the following chapters, the sequence of ideas as expounded in the above quoted source is more distinctly elaborated, and the corresponding calculation method is developed.

V. Crushing

1. Idealized conditions

The effect exerted by rolls running at the same surface speed is only the crushing of the grain drawn-in, but not crumbling. In literary sources, the case of drawing-in of grains only of the same diameter is being dealt with. In his former studies [3, 9, 10] author has already stated the formula that determines the maximum diameter of a spherical grain which can be drawn-in between rolls having different diameters. With the symbols as follows:

$2R$	diameter of the larger roll;
$2R_1$	$= 2k \cdot R$, diameter of the smaller roll;
$k < 1$	proportionality factor;
b	working gap;
ϱ	friction angle between the roll surface and the grain surface

we obtain

$$d = \frac{1}{\cos\varrho}\sqrt{(k+1)^2\cos^2\varrho\,R^2 + 4k\sin^2\varrho\,R^2 + 2(k+1)bR + b^2} - (k+1)R. \quad (2)$$

When the larger roll represents a concave surface, the term $(k + 1)$ has to be substituted by $(k - 1)$. For rolls with equal diameters (Fig. 6a) $k = 1$, and by using $2R = D$,

$$d = \frac{D(1 - \cos\varrho) + b}{\cos\varrho} \quad (3)$$

the knowledge of the maximum grain-diameter that can be drawn-in is indispensable because, at a given gap, the forces arising in the operating zone are the larger, the greater the drawing-in capacity which is characterized by the chord-value

$$\overline{AB} = d \cdot \cos\varrho = \boldsymbol{B}.$$

In what follows, we intend to give a theoretical survey of the play of forces in conditions which are created by the occurring phenomena. Thus, for simplicity's sake, by neglecting the journal friction, we whall follow the path of an elastic tape considered as a model running between the rolls (Fig. 6a).

The greatest width of this tape is the chord $\overline{AB} = \boldsymbol{B}$, belonging to the friction angle ϱ. With an arbitrary angle $\varphi < \varrho$ the elastic tape becomes thinner. To make the formula more simple, the spring rigidity s should be considered as a linear function of the length in the same way as by a cylindric helical spring.

Let us assume the same working conditions for each roll. Such a state of running can be realized, when the axis O_1 is driven by a special motor running

at the same speed as is the axis O. In the diagram of Fig. 6a the forces acting on the elementary surfaces of the rolls, and in Fig. 6b the forces acting on the elementary surfaces of the tape are shown.

As is seen in Fig. 6b, the elementary springs TJ and T_1J of the imaginary tape which are directed at the angle φ, are subjected to the pressing action of the elementary normal forces dN and dN_1; by these, at points T and T_1 of the

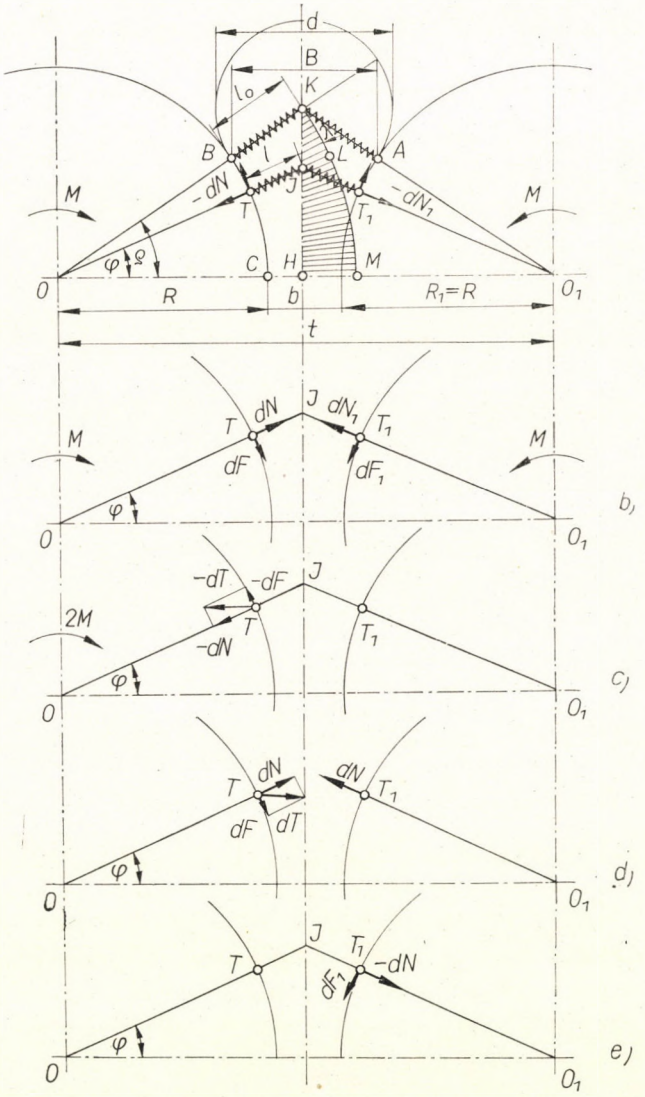


Fig. 6. Force effects on an elastic tape model between rolls rotating at equal speeds: Rolls driven individually — forces acting on the rolls (*a*) and forces acting on the tape (*b*); one single roll driven — forces acting on the directly driven roll (*c*), forces acting on the tape (*d*) and forces acting on the roll driven by the tape (*e*)

cylindrical roll surface, the elementary friction forces dF and dF_1 are induced, which again act as drawing in forces on the tape.

In compliance with the symmetric drive arrangement, the moment

$$dF \cdot R = dF_1 \cdot R_1$$

are separately exerted by each of the two motors.

Let us assume that only axis O is driven from without (Fig. 6c) and roll R_1 is driven only by friction forces by the roll R. Since the bearing friction is neglected, the axis O_1 takes the speed equalling the speed of the axis O, because both rolls, being in contact with the tape through friction, can be considered as having a toothing of infinitely fine teeth.

When the friction on the bearings and journals is active, this breaking effect causes but infinitely small difference of speed, inasmuch as the axis O_1 is retarded as compared to axis O. This condition has clearly been observed at those times when no positive driving mechanism was applied to the pairs of rolls. It has to be noted that in this case the speed ratio approaches the proportion 1 : 1, there is no feed-back action arising, and it was necessary to introduce the positive driving system just in order to secure a greater difference of speed between the rolls.

2. The torque induced on the arc for drawing-in

Let the "spring length" $l_0 = \overline{BK}$, as appears at the point of tape corresponding to the friction angle ϱ, be considered as a free length (Figs 6a and 6b) that only involves the action of the net weight which induces the initial friction force. Corresponding to the angle $\varphi < \varrho$, the elementary spring becomes shortened by $\lambda = JL$, (and $\lambda = l_0 - l$); now we calculate with the spring rigidity s [kp/cm] as the reciprocal value of the spring rate c [cm/kp] in order to obtain the elementary normal force that acts on the elementary surface in the environment of point T on the cylindrical surface:

$$dN = \lambda \cdot sRhd\varphi,$$

where h denotes the length of the cylinder. Using the new denotation

$$OH = R + \frac{b}{2} = \frac{1}{2}(D + b) = \frac{t}{2}$$

and with $\lambda = \overline{OL} - \overline{OJ}$, we obtain

$$\lambda = \frac{t}{2\cos\varrho} - \frac{t}{2\cos\varphi},$$

and by substituting the formula for λ into the equation of dN we obtain

$$dN = \frac{t}{2} Rhs \left(\frac{1}{\cos \varrho} - \frac{1}{\cos \varphi} \right) d\varphi \tag{4}$$

as the elementary pressing force of an elementary spring. From this the expression of the elementary friction force is:

$$dF = dN \tan \varrho = \frac{t}{2} Rhs \tan \varrho \left(\frac{1}{\cos \varrho} - \frac{1}{\cos \varphi} \right) d\varphi . \tag{5}$$

By taking into account the relationship

$$\frac{t}{2} = R + \frac{b}{2}$$

and by neglecting the net weight of the tape that is already quite unimportant here, we obtain the integral force:

$$F = R \left(R + \frac{b}{2} \right) h \cdot s \cdot \tan \varrho \left(\int_0^{\varrho} \frac{d\varphi}{\cos \varrho} - \int_0^{\varrho} \frac{d\varphi}{\cos \varphi} \right);$$

$$F = R \left(R + \frac{b}{2} \right) h \cdot s \cdot \tan \varrho \left\{ \left[\frac{\varphi}{\cos \varphi} \right]_{\varphi=0}^{\varphi=\varrho} - \left[\ln \tan \left(\frac{\pi}{4} + \frac{\varphi}{2} \right) \right]_{\varphi=0}^{\varphi=\varrho} \right\};$$

$$F = R \left(R + \frac{b}{2} \right) h \cdot s \cdot \tan \varrho \left[\frac{\varrho}{\cos \varrho} - \ln \tan \left(\frac{\pi}{4} + \frac{\varrho}{2} \right) \right]. \tag{6}$$

In Fig. 6c, the elementary forces acting on the driven roll, and in Fig. 6d, those acting on the taps, further in Fig. 6e those acting on the driving roll are shown.

Taking into account the explanation given in point IIIc the input from without, equals $A = 2\ FR\ \omega$, and expressed in horsepower: $P = 2\ FRn\ \pi/30 \cdot 75$ from which the torque induced along the draw-in arc of one cylinder is

$$F \cdot R = R^2 \left(R + \frac{b}{2} \right) \cdot h \cdot s \cdot \tan \varrho \left[\frac{\varrho}{\cos \varphi} - \ln \tan \left(\frac{\pi}{4} + \frac{\varrho}{2} \right) \right]. \tag{6'}$$

As an example, we will assume the friction coefficient for $\varrho = 17°$ as is usual for the case between cast iron and grain. We find the values

$$\text{arc } 17° = 0{,}29671,$$

$$\cos 17° = 0{,}95630,$$

$$\tan 17° = 0{,}30573;$$

consequently

$$\frac{\varrho}{\cos\varrho} = \frac{0{,}29671}{0{,}95630} = 0{,}3123\,,$$

$$\ln\tan(43^\circ + 8^\circ\,30') = \ln\tan 53^\circ\,30' =$$

$$= \ln\tan 1{,}35142 = 0{,}3084\,;$$

$$\left[\frac{\varrho}{\cos\varrho} - \ln\tan\left(\frac{\pi}{4} + \frac{\varrho}{2}\right)\right] = 0{,}3123 - 0{,}3008 = 0{,}0115\,,$$

$$0{,}0115 \cdot \tan 17^\circ = 0{,}0115 \times 0{,}30573 = 0{,}00352\,,$$

$$F = 0{,}00352\,R\left(R + \frac{b}{2}\right)h \cdot s\,. \tag{6a}$$

Thus the idealized loose roll operating without bearing friction, is rotated by the grist at the same speed as is the roll driven by the motor. Further, the moment arm of specific crushing work done by one single roll can be calculated.

3. Calculation of the moment arm

According to point III. 2 the elementary force $-dT$ resulting from the elementary normal force $-dN$ acting on the roll having a radius R, and from its induced friction force $-\mu dN$ traverses the outer similarity point of the profile circles having the radii R and R_1.

This resulting force is balanced by the force dT acting on the tape (Figs 6a and 6b). With rolls having equal radii, the similarity point is located in infinity, thus, the forces $-dT$ and $+dT$ are parallel to the plane of the axes.

Since the moment of the normal components $(-dN)$ traversing the centre O equals zero, the sum of the moments exerted by the elementary resultant forces $(-dT)$ will equal the sum of moments exerted by the elementary tangential forces (dF). Thus, when the resultant $-T$ of the elementary forces $-dT$ is calculated, and we divide the known rotating moment (of the force F) by this force-value $-T$, we obtain the arm value e belonging to the force T.

According to the diagram:

$$-dT = \frac{dN}{\sin\varphi}\,.$$

We substitute the value dN taken from equation (4):

$$-dT = \frac{t}{2}R \cdot h \cdot s\left(\frac{1}{\cos\varrho \cdot \sin\varphi} - \frac{1}{\cos\varphi \cdot \sin\varphi}\right)d\varphi,$$

and because of

$$\sin\varphi\cos\varphi = \frac{\sin 2\varphi}{2},$$

we obtain

$$-dT = \frac{t}{2} Rhs \left(\frac{1}{\cos\varrho \sin\varphi} - \frac{2}{\sin 2\varphi} \right) d\varphi;$$

$$-T = \frac{t}{2} Rhs \left(\frac{1}{\cos\varrho} \int_0^{\varrho} \frac{d\varphi}{\sin\varphi} - 2\int_0^{\varrho} \frac{d\varphi}{\sin 2\varphi} \right) =$$

$$= \frac{t}{2} Rhs \left\{ \frac{1}{\cos\varrho} \left[\ln \tan \frac{\varphi}{2} \right]_0^{\varrho} - 2 \left[\frac{1}{2} \ln \tan \frac{2\varphi}{2} \right]_0^{\varrho} \right\}.$$

By taking into account that

$$\frac{t}{2} = R + \frac{b}{2}$$

we obtain

$$-T = R\left(R + \frac{b}{2}\right) hs \left(\frac{1}{\cos\varrho} \ln \tan \frac{\varrho}{2} - \ln \tan \varrho \right). \tag{7}$$

Thus, by making use of (6′) and (7), the moment arm e equals

$$e = \frac{FR}{T} = \frac{R \tan\varrho \left[\dfrac{\varrho}{\cos\varrho} - \ln \tan \left(\dfrac{\pi}{4} + \dfrac{\varrho}{2} \right) \right]}{\dfrac{1}{\cos\varrho} \ln \tan \dfrac{\varrho}{2} - \ln \tan \varrho}.$$

Consequently, the moment arm e is a function of R and ϱ, and assumes a minute value, as is to be seen from the example herebelow.

Numerically, taking $\varrho = 17°$:

$$\frac{1}{\cos\varrho} = \frac{1}{0{,}95630};$$

$$\tan \frac{\varrho}{2} = \tan 8° 30' = 0{,}14945;$$

$$\ln 0{,}14945 = -1{,}905$$

$$\frac{1}{\cos\varrho} \ln \tan \frac{\varrho}{2} = \frac{-1{,}905}{0{,}95630} = -1{,}994;$$

$$\ln \tan \varrho = \ln 0{,}30573 = -1{,}185$$

$$\frac{1}{\cos\varrho} \ln \tan \frac{\varrho}{2} - \ln \tan \varrho = -1{,}994 + 1{,}185 = -0{,}809$$

and substituting

$$\frac{t}{2} = R + \frac{b}{2}$$

we have

$$T = 0{,}809\, R\left(R + \frac{b}{2}\right) hs. \tag{7a}$$

Making use of the equations (6′) and (6a), the moment of the tangential friction-forces is:

$$FR = 0{,}00352\, R^2 \left(R + \frac{b}{2}\right) h s \tag{6b}$$

and the moment arm e of the resulting force T is:

$$e = \frac{FR}{T} = \frac{0{,}00352\, R}{0{,}809} = 0{,}0049\, R.$$

E.g. for a roll having $R = 125$ mm, $e = 0{,}612$ mm.

VI. Crumbling

1. When one of the rolls does not rotate

Let us imagine that the roll having a radius R_1, which in the first example, was freely rotating (driven by the friction force of the grist) is now in a fixed position (Fig. 6a).

Of course, the resistance that has to be overcome by the roll driven from without now differs from what was determined in the first example. In this first case, the work to be done was represented by the task of pressing the material together. This crushing or breaking action was to be done by overcoming the inner friction of the grain, and/or causing some elastic deformation that could partly be restored later on. When doing the work of crumbling, some part of the moving "tape" is peeled off by the fixed roll, viz. the cohesive force has to be overcome, too. Thus, the fixed roll has to bear a load in which the shearing resistance is included, too.

Usually, shear-force is calculated as the product of the area to be sheared and the shear-resistance k_s calculated for the unit of area. The smaller the gap b, the more the probability can be assumed that the shearing operation takes place in the middle plane HK.

When the specific reduction of length reaches the value

$$\varepsilon = \frac{HM}{BK} = \frac{\lambda}{l_0} = 1,$$

then the value

$$b = 0.$$

When the value of b is finite, then $\varepsilon < 1$; and when $b = \boldsymbol{B}$, with $\varphi = 0$, and no shearing effect is present, then $\lambda = 0$ and $\varepsilon = 0$, too.

Consequently, for a correct calculation of the shearing force, one has to reckon with the correction factor:

$$\varepsilon = \frac{HM}{BK} = \frac{\dfrac{t}{2\cos\varrho} - \dfrac{t}{2}}{\dfrac{t}{2\cos\varrho} - R} = \frac{t(1-\cos\varrho)}{t - D\cos\varrho}.$$

Now, because of

$$t = D + b,$$

this factor is expressed by

$$\varepsilon = \frac{(D+b)(1-\cos\varrho)}{(D+b) - D\cos\varrho} = \frac{(D+b)(1-\cos\varrho)}{D(1-\cos\varrho) + b} =$$

$$= \frac{D+b}{D + \dfrac{b}{1-\cos\varrho}} = \frac{1 + (b/D)}{1 + \dfrac{b}{D(1-\cos\varrho)}}.$$

Further we have

$$D(1-\cos\varrho) = 2R(1-\cos\varrho) = \boldsymbol{B} - b$$

and thus we obtain

$$\varepsilon = \frac{1 + (b/D)}{1 + b/(\boldsymbol{B} - b)} = \frac{\boldsymbol{B} - b + (\boldsymbol{B} - b)(b/D)}{\boldsymbol{B} - b + b} =$$

$$= \frac{(\boldsymbol{B} - b)D + (\boldsymbol{B} - b)b}{\boldsymbol{B} \cdot D} = \frac{(\boldsymbol{B} - b)(D + b)}{\boldsymbol{B} \cdot D}.$$

Because of

$$\overline{HK} = \frac{t}{2}\tan\varrho = \frac{D+b}{2}\tan\varrho$$

the shearing force calculated for a cylinder having the length h:

$$S = \overline{HK}\,\varepsilon k_s h = \frac{hk_s \tan\varrho\,(D+b)(\boldsymbol{B} - b)(D+b)}{2\,\boldsymbol{B} \cdot D} =$$

$$= \frac{hk_s \tan\varrho\,(D+b)^2(\boldsymbol{B} - b)}{2\,\boldsymbol{B} \cdot D}. \tag{8}$$

At the value $b = 0$

$$S = \frac{hk_s \tan\varrho\, D^2 \cdot B}{2\,B \cdot D} = hk_s R \tan\varrho. \tag{8a}$$

2. *The rolls being positively engaged, rotate at different speeds*

Let us assume a plain crushing operation according to the points V. 1 and V. 2; then by neglecting the bearing friction, the required input equals:

$$A = 2\,R\,\omega\,F = R(\omega + \omega)F.$$

When due to a positive drive that serves to modify the speed, the roll having the radius R_1 rotates at a different angular speed $\omega_1 = \omega/i$, (i being the speed ratio), and the work of crushing is:

$$R = (\omega + \omega_1)\,F = R\left(\omega + \frac{\omega}{i}\right)F = R\omega\,\frac{i+1}{i}\,.$$

Again, the crumbling operation is proportional to the retardation (pro unit of time) of the slow roll, viz:

$$R = (\omega - \omega_1)\,S = R\omega S\,\frac{i-1}{i}\,.$$

Calculating the required input for a simultaneous crushing and crumbling

$$A = R\omega\,\frac{F(i+1) + S\,(i-1)}{i}\,. \tag{9}$$

The value in HP at the r.p.m. (n) equals:

$$P = R\,\frac{n\pi}{30 \cdot 75}\cdot\frac{F(i+1) + S(i-1)}{i}\,. \tag{9a}$$

As regards the peripheral force composed of two parts:

$$F\left(\frac{i+1}{i}\right) + S\left(\frac{i-1}{i}\right)$$

it must be borne in mind that only the *second member* depends on the retardation due to the braking effect as mentioned before and on the energy feed-back caused.

VII. Feed-back

1. *Determination of the feed-back power*

The crumbling force acting on the slow roll having a radius $R_1 = R$ (Fig. 4) is:

$$S\,\frac{i-1}{i}$$

the operating moment of which is

$$M_{a1} = RS \frac{i-1}{i} . \tag{10}$$

Corresponding to this moment, the peripheral force is, at the large toothed wheel having the radius r_1:

$$\frac{RS}{r_1} \cdot \frac{i-1}{i} .$$

By this peripheral force acting on the radius r of the pinion, the rotating moment that works on the axis O of the fast roll is

$$\frac{RS}{r_1} \cdot \frac{i-1}{i} \cdot r$$

and knowing that

$$\frac{r}{r_1} = \frac{1}{i}$$

we obtain the moment exerted on the fast roll by the crumbling force:

$$M_a = RS \frac{i-1}{i^2} , \tag{11}$$

from which the feed-back power is

$$W = \frac{RS(i-1)\,\omega}{i^2} = \frac{RS(i-1)}{i^2} \cdot \frac{n\pi}{30} , \tag{12}$$

when $i = 1$, $W = 0$.

Consequently, with a large ratio i, W becomes minute.

Let us find the extreme (maximum) value of W, and the corresponding value of i:

$$\frac{dW}{di} = \frac{d}{di}\left(RS\omega \frac{i-1}{i^2}\right) = RS\omega \frac{i^2 - (i-1)\,2\,i}{i^4} = 0$$

from which

$$i^2 - 2\,i^2 + 2\,i = 0,$$

viz.

$$i = 2 .$$

Using equation (10) we have:

$$W_{\max} = \frac{RS}{4}\omega = RS\,\frac{n\pi}{120}\,. \tag{12a}$$

As for the actual case with $b = 0$ we obtain: (see Equ. 8a)

$$W_{\max} = hk_s \tan\varrho\, R^2\,\frac{n\pi}{120}\,.$$

2. The piled-up power at the fast roll

According to Fig. 4, the power attacking the fast roll is larger than the input value A. The piled up power according to (9) and (12) is expressed as

$$L = A + W = R\omega\left[\frac{F(i+1)+S(i-1)}{i} + \frac{S(i-1)}{i^2}\right],$$

or

$$L = R\omega\left[\frac{F(i+1)\,i+S(i-1)\,i}{i^2} + \frac{S(i-1)}{i^2}\right] =$$

$$= R\omega\,\frac{F(i^2+i)+S(i^2+i)+S(i-1)}{i^2} =$$

$$= R\omega\,\frac{F(i^2+i)+S(i^2-1)}{i^2} = R\omega\left[\frac{i^2(F+S)}{i^2} + \frac{Fi-S}{i^2}\right],$$

viz.

$$L = R\omega\left(F+S+\frac{F}{i}-\frac{S}{i^2}\right). \tag{13}$$

Looking for the extreme value:

$$\frac{dL}{di} = R\omega\left(-\frac{F}{i^2}+\frac{2\,S}{i^3}\right) = 0\,,$$

viz. with

$$\frac{2\,S}{F} = i$$

we obtain:

$$L_{\max} = R\omega\left(F+S+\frac{F^2}{2\,S}-\frac{SF^2}{4\,S^2}\right) = R\omega\left(F+S+\frac{F^2}{4\,S}\right)$$

or

$$L_{\max} = R\omega S\left(\frac{F}{2\,S}+1\right)^2. \tag{13a}$$

3. *Effect of the bearing friction*

As has already been mentioned in point V. 1, according to experiences, the r.p.m. of the loose roll is somewhat smaller than that of the driven roll. Taking into account our afore-explained statements referring to the feed-back energy-flux, the retardation of the loose roll can also be determined.

Let us denote by N the force which in the axial plane presses the rolls together, by μ_e the friction coefficient and by r_c the journal radius, so we obtain as the braking moment of the bearing friction:

$$M_c = \mu_c N r_c .$$

In compliance with what has been stated before, the actual moment that is exerted by the crumbling force on the surface of the loose roll having a radius $R_1 = R$, is, according to Equ. (10):

$$S \frac{i-1}{i} R = \mu_c N r_c$$

from where we have

$$SRi - SR = iN\mu_c r_c ,$$

and

$$i(SR - N\mu_c r_c) = SR$$

or

$$i = \frac{SR}{SR - N\mu_c r_c}$$

and with $b \to 0$, according to (8a), and by μ denoting the friction coefficient between the rolls: we obtain

$$S = k_s R \cdot \mu \cdot h$$

and

$$\frac{n}{n_1} = i = \frac{k_s R^2 \mu h}{k_s R^2 \mu h - N\mu_c r_c} .$$

4. *The effect of the friction on a gear-driven pair of rolls*

The toothed wheels for mill rolls individually driven differ from the normal gears inasmuch as the total tooth heigh and the tooth space are much greater than the respective values of normal toothing. This special form is necessary because both the axial distance and the parallel position of the rolls might be changed during operation, and then of course, the gears have to be kept engaged without interruption.

When the serial coupling suggested by the author is applied for roll mills, the unfavourable running condition as described herebelow, cannot occur. Again, there exists a special method of serial coupling, viz. when only the shafts in the fixed bearing are coupled by means of clutches, whereas after newly adjusted rolls during operation i.e. in a position of widened axial distance or a biased position (when parallelity of the axes is done away with) the engaging of the rolls is carried out by means of the toothed gear, which method cannot be accepted as an up-to-date solution.

When the working flanks of the toothed wheels are changed, the over-dimensioned tooth space causes some undesired clacking. Namely, as long as the slow roll is exposed only to the friction effect, the pinion acts as the driving gear. Again, at the very moment when the gap is filled with grist, the large wheel, which is originally driven, begins to function as the driver. This change over can take place only if the large wheel performs an advance equalling the flank clearance angle ϑ viz. if its other tooth flank assumes the angular velocity $\omega_1 = \omega/i$ corresponding to the constant angular velocity ω of the pinion, after which the other pair of flanks comes into contact with each other.

We now apply the following denotations:

Z the sum of the friction resistances reduced to the radius R;
F the friction force induced by the normal pressure force N exerted by the fast roll;
Θ the moment of inertia of the masses rotating together with the slow roll;
ω_ϑ angular overtake velocity of the large wheel (and the slow roll) during the change-over period of tooth-flanks;
ω_r $= \omega_\vartheta - \omega$, the relative angular velocity of the slow roll as related to the fast roll, during the changeover period of tooth-flanks

so we obtain for the torque:

$$FR = N\mu R = ZR + \Theta \frac{d\omega_r}{dt},$$

and for the relative angular acceleration:

$$\frac{d\omega_r}{dt} = \varepsilon_r = \frac{R(F - Z)}{\Theta} = \frac{R(N \cdot \mu - Z)}{\Theta}.$$

Taking into account a short finite period of time, all the members on the right side of the equation can be considered as constant including the angular acceleration, wherefore the period of rotation needed to perform the advance-angle ϑ that corresponds to the tooth space, is

$$t = \sqrt{\frac{2\vartheta}{\varepsilon_r}} = \sqrt{\frac{2\vartheta \cdot \Theta}{R(N\mu - Z)}} = \sqrt{\frac{2\vartheta\Theta}{R(F - Z)}}.$$

This expression is instructive, first, because with a suddenly raised force F, obviously the time t is extremely short, and an impact is inevitable, second,

because, with a large inertia moment Θ and with a small difference $F-Z$, the time, when the other flanks are not contacting each other yet, might of course be rather long.

VIII. Remarks concerning some erroneous statements to be found in some literary sources

Methods followed by other researchers differ from that applied by the author here mainly inasmuch as they think it permissible to neglect some interrelations when they establish the calculation of the grinding input power. E.g. in a rectangular triangle, instead of the height belonging to the hypotenuse, the adjacent perpendicular side is taken, or else instead of some trigonometric function value, they reckon with the angular arc; when they deal with a progression, except for the first member, all others are neglected etc. The more important are deviations that have to be regarded as being mistaken or misunderstood principles.

a) The theoretic crushing formula established by AFANASJEV [2] includes two erroneous assumptions. First, when calculating the drawing-in force, the principle laid down here in point III. d) is left out of consideration: instead of the elementary friction forces, he refers to the imaginary friction force induced by the resultant; on the other hand, he deducts the force having a minus sign, leaving out of account that this negative force is acting not on the grist but on the roll.

b) There are textbooks, where the following denotations are used:

N driving power,
N_k feed-back power,
r pitch circle radius of the pinion,
r_1 pitch circle radius of the large toothed wheel;

and the formula is alleged to be

$$N_k = \frac{N \cdot r}{r_1 - r},$$

by trying to prove it by the motivation.
With the symbols

P force acting on the roll surface having the radius R;
Q force acting at the pitch circle of gears;
v surface velocity of the fast roll

the forces being in equilibrium:

$$Q = P \frac{R}{r_1}. \tag{I}$$

Balance of power

$$N = Pv - Qv\frac{r}{R}. \tag{II}$$

By substituting P from (I):

$$N = Q\frac{r_1}{R}v - Qv\frac{r}{R} = Qv\left(\frac{r_1}{R} - \frac{r}{R}\right) = Qv\frac{r_1 - r}{R} \tag{III}$$

or

$$Q = \frac{NR}{v(r_1 - r)}$$

and the feed-back power is

$$N_k = Qv\frac{r}{R}.$$

By substituting Q from (III):

$$N_k = \frac{NR}{v(r_1 - r)} \cdot v\frac{r}{R} = \frac{Nr}{r_1 - r}.$$

This formula is undoubtedly, without any real physical sense; namely, in the case of $r_1 = r$, the calculated value becomes ∞, whereas all the observations made during effective operation prove that with equal radii the feed-back disappears (it becomes zero).

Neither the modification suggested by Dr. Ing. O. Mooc is appropriate to make things clearer. According to his paper [6]

$$N_k = \frac{N \cdot r}{1{,}4\, r_1 - r}.$$

With this formula, when $r_1 = r$, the obtained value is

$$N_k = \frac{N}{1{,}4 - 1} = \frac{N}{0{,}4} = 2{,}5\ N$$

a value that again differs from zero.

Mooc himself affirms the observation, obtained by experimental measurements, that the feed-back power assumes the zero value when the ratio v/v_1 equals infinite or one.

When analysing the functions, and delineating the respective curves, Mooc found the locus of the maximum value as being at $i = 1{,}25$ (and not $i = 2$); now, when we take into account the friction effects acting on the roll surface, we can state that the critical value of i will be below 2. Reduced to the

roll radius R, the total friction force will be denoted by S_1, as resulting from both gear and a bearing friction. Consequently, the compound braking moment of both the crushing and the friction forces, acting on the slow roll (having the radius $R_1 = R$) is:

$$RS\frac{i-1}{i} + S_1 R,$$

(S being the shear force for crushing). From this, the peripheral force at the large toothed wheel having a radius r_1 is

$$\frac{R}{r_1}\left(S\frac{i-1}{i} + S_1\right)$$

and the moment induced at the axis O of the fast roll is:

$$\frac{r}{r_1} R\left(S\frac{i-1}{i} + S_1\right) = \frac{1}{i} R\left(S\frac{i-1}{i} + S_1\right) = R\left(S\frac{i-1}{i^2} + \frac{S_1}{i}\right).$$

This gives for the value of the feed-back power:

$$W_s = R\omega\left(S\frac{i-1}{i^2} + \frac{S_1}{i}\right) = R\omega S\left(\frac{i-1}{i^2} + \frac{S_1}{S\cdot i}\right) \qquad (12')$$

and

$$\frac{dW}{di} = R\omega S\left[\frac{i^2 - (i-1)\,2\,i}{i^4} - \frac{S_1}{S\cdot i^2}\right] = 0\,;$$

or

$$\frac{2i - i^2}{i^4} = \frac{S_1}{S\cdot i^2}, \quad \frac{2\,i - i^2}{i^2} = \frac{S_1}{S},$$

$$\frac{2}{i} - 1 = \frac{S_1}{S}, \quad \frac{2}{i} = \frac{S_1}{S} + 1\,;$$

$$i = \frac{2}{(S_1/S) + 1} < 2$$

the denominator being larger then 1. Thus, our theorem proves to be in full coincidence with the experimental findings of MOOG.

c) In his work [8] on the energy-flux of milling rolls, J. PRATIQUE tried to express the grinding input value, and his formula is more similar to our resulting equation. According to his statement, the feed-back power is:

$$W = (-TR + Ne)\,\omega_1$$

with the following interpretation:

R	diameter of the roll;
T	force being normal to the plane of the axes, and acting on the gravity centre of the grist;
N	force being parallel to the axes and acting on the gravity centre of the grist;
e	distance of the grist gravity centre above the axis-plane;
ω_1	the angular velocity of the slow roll.

This formula holds true when $\omega_1 = 0$, i.e. when $i = v/v_1 = \infty$; again, it is meaningless for $v/v_1 = 1$. As we have seen in point III. d, there is no use to deal with the problem when only one single mass-point is taken into consideration.

d) The importance of the feed-back power is obvious and decisive in fatigue-testing machines for gears. This problem is dealt with in a paper by D. MAROS [16]; not speaking about some clerical errors, he commits the mistake of neglecting the direction of the energy flux in question.

e) A regrettable error crept into the text of the author's paper quoted under [8]. On page 77, the value of the piled up power should read $A = A + W$ (instead of the erroneous formula $A - W$).

f) Proportion of the feed-back power and the input.

According to (12′):

$$W_s = R\omega \left(S \frac{i-1}{i^2} + \frac{S_1}{i} \right) = R\omega \frac{i(S+S_1) - S}{i^2} ;$$

and according to point *b)* of the present chapter, we can find, at the locus

$$i = \frac{2\,S}{S+S_1}$$

the extreme value

$$W_{s\max} = R\omega \frac{(S+S_1)^2}{4\,S} . \tag{12″}$$

According to (9), the input is

$$A = R\omega \frac{F(i+1) + S(i-1)}{i} = R\omega \left(\frac{F-S}{i} + F + S \right)$$

and its value at locus i that corresponds to the extreme value $W_{s\,\max}$

$$A = R\omega \frac{(F-S)(S+S_1) + 2\,S(F+S)}{2\,S} . \tag{9′}$$

We shall explain in what follows that the quotient p of the values (12″) and (9′) is a positive number that may be larger than one, when the value of

S_1 is large enough:

$$\frac{W_{s\,\max}}{A} = p = \frac{(S+S_1)^2}{2\,[(F-S)(S+S_1)+2\,S(F+S)]} > 1.$$

Taking

$$F = qS\,,$$

we obtain:

$$p = \frac{S^2 + 2\,SS_1 + S_1^2}{2\,S^2(1+3q) + 2\,SS_1(q-1)}\,.$$

We shall show that

$$S^2 + 2\,SS_1 + S_1^2 > 2\,S^2(1+3q) + 2\,SS_1(q-1)\,,$$

viz.

$$S_1^2 + 2\,SS_1(2-q) > 2\,S^2(1+3q) - S^2.$$

Dividing this expression by S^2, and by due reduction of the right side:

$$\left(\frac{S_1}{S}\right)^2 + 2\left(\frac{S_1}{S}\right)(2-q) > 1 + 6q.$$

Now the left side shall be completed to a full square, in other words the value:

$$(2-q)^2 = 4 - 4\,q + q^2$$

shall be added to both sides:

$$\left[\frac{S_1}{S} + (2-q)\right]^2 > 5 + 2q + q^2$$

or

$$\frac{S_1}{S} + 2 - q > \sqrt{5 + 2q + q^2}$$

i.e.

$$\frac{S_1}{S} > \sqrt{5 + 2q + q^2} + q - 2\,. \tag{12b}$$

The quotient $F/S = q$ can be considered as a parameter which depends first on the quality of the material. The value S_1 can be varied by changing the clamping force that is exerted by the bearings onto the journals; this corresponds to the definition according to which S_1 represents the friction force reduced to the diameter R of the slow roll as a "driving roll" resulting from the friction effects on both gear and bearing.

Numerically, we find with $q = 0$ for the plain break-operation:

$$W_{s\,\max} > A\,,$$

namely when

$$\frac{S_1}{S} > \sqrt{5} - 2 = 2{,}24 - 2 = 0{,}24\,;$$

with

$$q = 1/2, \quad W_{s\,\max} > A$$

namely when

$$\frac{S_1}{S} > \sqrt{5 + 1 + 1/4} + \frac{1}{2} - 2 = \frac{5}{2} - \frac{3}{2} = 1$$

with

$$q = 1, \; W_{s\max} > A,$$

namely when

$$\frac{S_1}{S} > \sqrt{8} - 1 = 2{,}83 - 1 = 1{,}83\,;$$

and with

$$q = 2, \; W_{s\max} > A,$$

namely when

$$\frac{S_1}{S_2} > \sqrt{13} = 3{,}6\,.$$

IX. Energy-flux variation depending on the system of coupling in series

1. Experiments and measuring results of the Corn Trust Research Institute

As has already been seen in the explanation given in point IV by comparing of Figs 4 and 5, the load exerted on the gear drive greatly increases by mounting the motor directly on the slow roll, even in the case of single pair of rolls. Some numerical data concerning the difference between the systems of Fig. 4 and Fig. 5 can be found when we take into account the measuring results obtained from experiments that were carried out by very careful and meticulous efforts of the *National Research Institute for Milling and Cornstoreing Trade* (to day: "Research Institute of the Corn-Trust"). According to a paper by J. KOLOSTORI: [7]

"*when speaking of smooth rolls, the value of the power ratio W/A is about 2 when the gear ratio $i = 1{,}5$ namely by neglecting the friction losses of bearings and gears; where W denotes the feedback power refed from the slow roll to the fast roll.*"

This input $[A - (A_1 + A_{21})]$ *was 6,6 kW; whereas there was a feed-back power of 12,5 kW from the slow roll reversed to the fast roll, that makes 1,9 times the original input to the fast roll. In this state, the rolls were overtightened. In a less tight position of the rolls, this ratio makes out to be* $8{,}7/4{,}5 = 2{,}1$.

For a pair of smooth rolls, two important observations were experienced:

a) *When a succesful grinding operation is effected then the feed-back power that accumulates backwards from the slow roll to the fast roll is considerably greater than with a pair of fluted break-rolls. Whereas the ratio W/A makes out, with fluted rolls, abt. 40%, the same ratio might amount, with smooth rolls, to more than 200%.*

b) *The flour produce is the more valuable, the greater the feed-back power to the fast roll is as compared to the input.*

With reference to the quoted passage, it is obvious that in a mill coupled according to the arrangement as shown in Fig. 4 when at a certain given tightening of the rolls the gear is subjected to a total load of 12,5 kW, the same gear, with a coupling system according to Fig. 5, should bear a total load of 12,5 + + 6,6 = 19 kW *with the same degree* of tighthening as before. In other words, when corresponding to the motor output the designers would calculate the gear and the attached mechanical parts for a load of 6,6 kW, very astonishing occurences might occur during the very first operation after starting.

3. Special features of grinding rolls having different diameters

Already in his former papers [9, 10] author has discussed the expectable advantages of pairs of rolls with different diameters. He has hinted at the possibility [17] of creating an arrangement in which the pairs of both the fluted (the breakers) and the smooth (the grinder) rolls are driven by the same gear i.e. with one and the same gear ration.

Author learned from an unpublished manuscript [18], that the advantages of a pair of rolls having different diameters had already been known, in the last century by A. MECHWART,* who in his engineering practice had built roller mills of this type. Again, in view of present tendencies, the problem of energetic properties of rolls having different diameters — which up-to-now was not taken into consideration yet by researchers — is far more important then the probable advantages of a single pair of mill rolls. To this end, let us compare the different two methods of couplings in series: on the one hand, the coupling system as is usually carried out in the milling industry (Fig. 7), and on the other hand, our new method [19] by which due attention is paid to an appropriate conduction of the energy-flux (Fig. 8).

* Lived 1834—1907, collaborator of *Ganz Works*, inventor.

In these figures, the full line denotes the flux of the input power delivered by the motor, and the rolled line denotes the flux of the feed-back power. In compliance with measuring results, our assumption comprises a grinding energy of 5 kW needed by each pair of rolls and a feed-back power of 9 kW.

As is shown by operating experiences, clutches and the attached parts are the most liable to be damaged in the case of being overloaded: in this

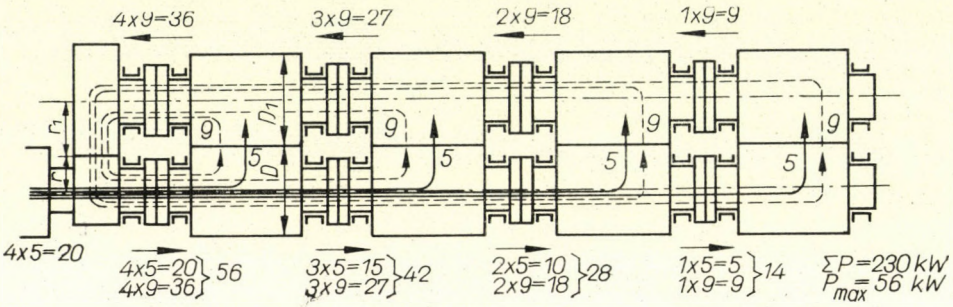


Fig. 7. Schematic stream of the energy-flux in pairs of mill rolls when the coupling in series is carried out according to a traditional system

Denotations:
Full line — grinding power; dotted line — feed-back power; arrows — flux direction; ciphers — partial power components required for the rolls, the clutches and the gears.

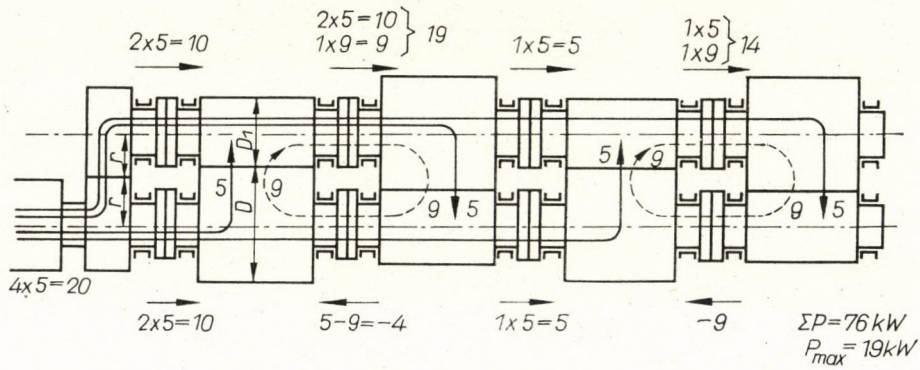


Fig. 8. Application of the system for coupling in series invented by the author and J. KOLOSTORI (schematic stream of energy flux in mill rolls having *unequal* diameters)

relation due attention is paid when delineating Figs 7 and 8 precisely to demonstrate the actual loads numerically in order to make their comparison possible.

When a series of 4 pairs of rolls is coupled together, the method illustrated in Fig. 7 involves the following load figures: the highest loaded clutch has to bear 56 kW, and the 8 clutches are subjected altogether to 230 kW, with the gear charged with 36 kW. The method shown in Fig. 8, when the constructional overall length and the number of the constituting component parts are the

same, the maximum individual load of a clutch amounts to 19 kW, and the total makes out 76 kW, which represents only abt. 1/3-rd of the former. Similarly, the driving gear is loaded only by 10 kW. The essential point of this invention is featured by the following claim as laid down in the patent specification:

"*Milling work comprising at least two adjacent pairs of rolls, with pairs working at the same axial distance, the adjacent rolls being coupled in series with each other, the gear ratio of the two serial axes being fixed to 1 : 1, and with different diameters at least on two coaxial rolls.*"

When the true trend of the energy flux is duly taken into account, then an appropriate solution of load distribution can be found even for an arrangement with rolls of equal diameter working as a pair in the same frame [20].

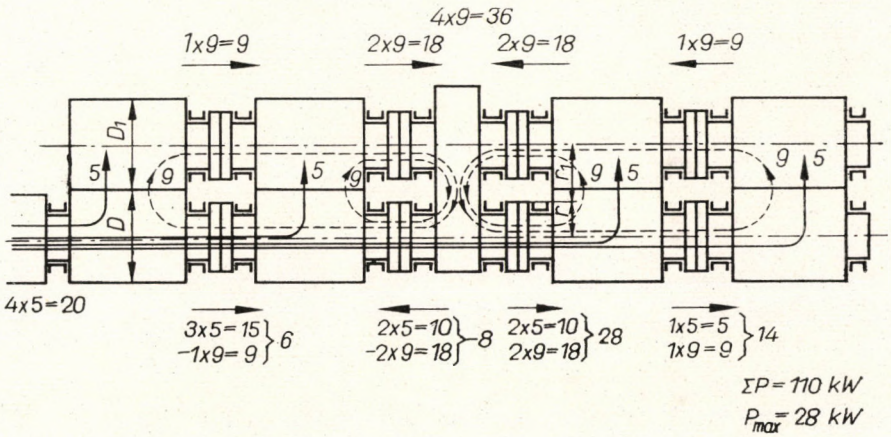


Fig. 9. Application of the system for coupling in series invented by J. KOLOSTORI and the author (schematic stream of the energy-flux in mill rolls having *equal* diameters)

This example is shown in Fig. 9, where the delineated milling work consists of the same component parts as in Fig. 7. And the actual loads amount to the following values: max. single clutch load 28 kW viz. the half value of the corresponding load according to Fig. 7. Total load 110 kW, viz. less than the half of the corresponding item according to Fig. 7.

For a mill frame in which the working rolls of the same pair have equal diameters, the essential feature of the invention is determined by the patent specification in the following claim clause:

"*Milling work, in which at least two pairs of rolls, each consisting of one fast roll and one slow roll, are mounted, by means of coupling in series of the fast rolls, and the slow rolls, respectively, the driving gear being coupled to the axis of the fast roll-axis, characterized by at least one pair of rolls located between the main drive and the gear work destined to create inequal speeds.*"

PART *B)* HOW TO REALIZE THE COUPLING IN SERIES?

I. Author's preliminary experiments carried out to improve the movable clutch

In Chapter II of Part I already some requirements which are indispensable for a well operating serial coupling of the mill rolls have been outlined. The present writer felt compelled, already in the early fourties, partly to make studies in order of how to make clutches that can be put into and out of gear is a faster and a more secure way [21, 22], partly to try to improve to movable type clutch; these endeavours were incited by emergency cases learned from the press e.g. forced standstill of the rigid type axial clutches caused by some sinking of the building; this happened in a rubberfactory, in the electric plant of *Kelenföld* (on the turbogenerator group), in the *Rózsa Mill Works Miskolc*, etc.

Author has already described in other papers [23, 24] why and at which points inadequate operating of presently known elastic and movable type of axial clutches might occur. In the present chapter some of the preliminary experiments should be referred to by means of which new constructional solutions have been found.

Author was much impressed, just at the beginning of his career, by a particular experience; in those times, it happened that oil dripped from the wormgear-housing of a new type suction-gas engine. The most intricate sealing of the bearing was not suitable to make this occurrence cease. Author found a new-fashioned solution; he had the dripping oil lead into a collector-chamber and had the oil conducted back (through a hole) into the housing. In former times, author succeeded in applying this principle to the solution of many practical problems. Just in the same way, this same principle was chosen by the author as the basic idea when he looked for a solution to the problem of how to prevent the transference of displacement and vibration phenomena of the driving shaft onto the driven shaft.

In our present case, the two steps according to the said principle, has been realized by the author as follows: First, a type of clutch was to be found which should be suitable during the transmission of the torque, not to transmit, at the same time, some tilting moment onto the adjacent bearings, and which should also be able to bear in itself all the vibrations that are induced by the relative displacements of the shafts. The realization invented by the author was named at this time as "vibration clutch". Second, there was to be found how the vibrating component part should create its own equilibrium (should be balanced automatically)?

The device that had been designed according to the aforesaid two principles, was denominated as the *self-balancing movable clutch*. The schematic view

had been already given in a former paper of the author [23]; a more detailed description will follow herein.

The first type of a vibrating clutch was designed by the author for the purpose of a lathe directly driven by an electric motor; this re-design was necessary because the well-known type of a rubber-pin coupling did not prove satisfactory. In Fig. 10, the scheme of the first type of vibrating clutch is shown. The photos of its component parts are to be seen in Fig. 11.

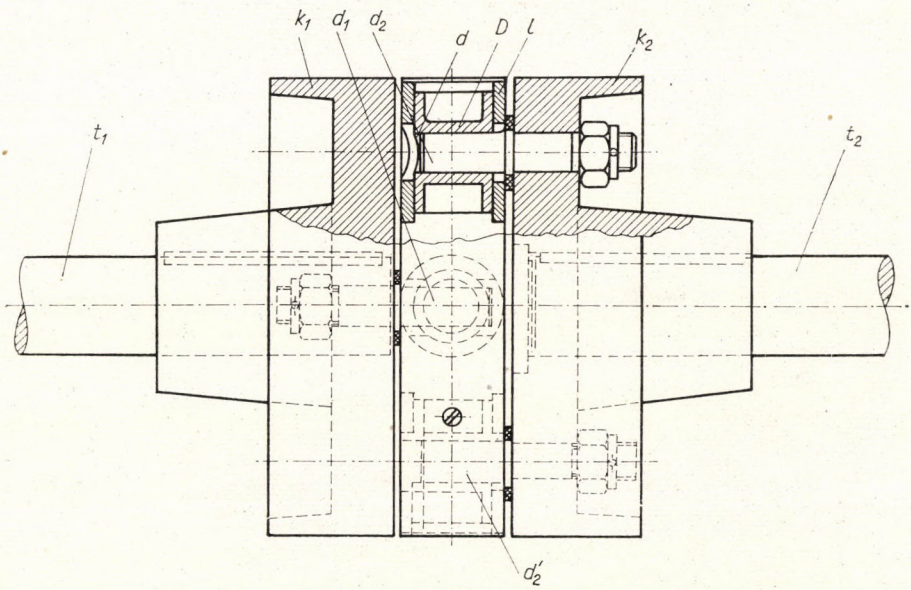


Fig. 10. Movable clutch, piston type schematic design

On the coupling disk k_1 keyed on the driving shaft t_1, there are two pins d_1 and d_1', symmetrically mounted in axial direction. Mounted on the driven shaft t_2, the coupling disk k_2 with two pins d_2 and d_2' is of the same design as disk k_1.

In the free floating disk type member l there are two pairs of holes perpendicular to each other with two pair of pistons D which can freely rotate in these holes. Again in each one of the pistons D there is a bore d, into which, according to the position of the pistons, one of the pins (d_1 and d_1', and d_2 and d_2') is located.

This type of coupling had been applied in a certain number of machines [25]. In Fig. 12, two vibration curves are shown. These diagrams were delineated by a vibration-testing instrument type "*Ascania*" in 1954; these tests were carried out by the *Hungarian Institute for Developing Machine Tools* (which bore in 1954 an other name: *Model-Machines Factory*).

The upper curve represents the vibrations of a CSEPEL Diesel engine (bore of 350 mm, power 51 HP, rpm 1500) by which a generator, mounted on

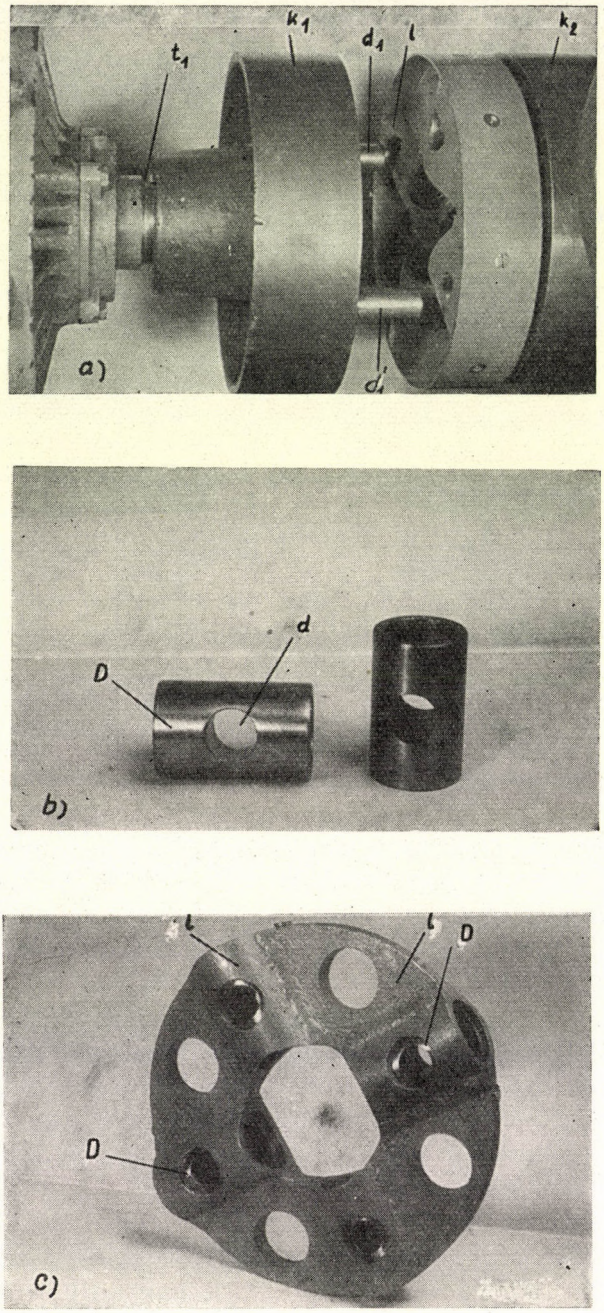


Fig. 11. Movable clutch, piston type: Mounted on the machine (*a*); *detail* the piston *D* with bore *d* (*b*); floating member *l* with four pistons (*c*)

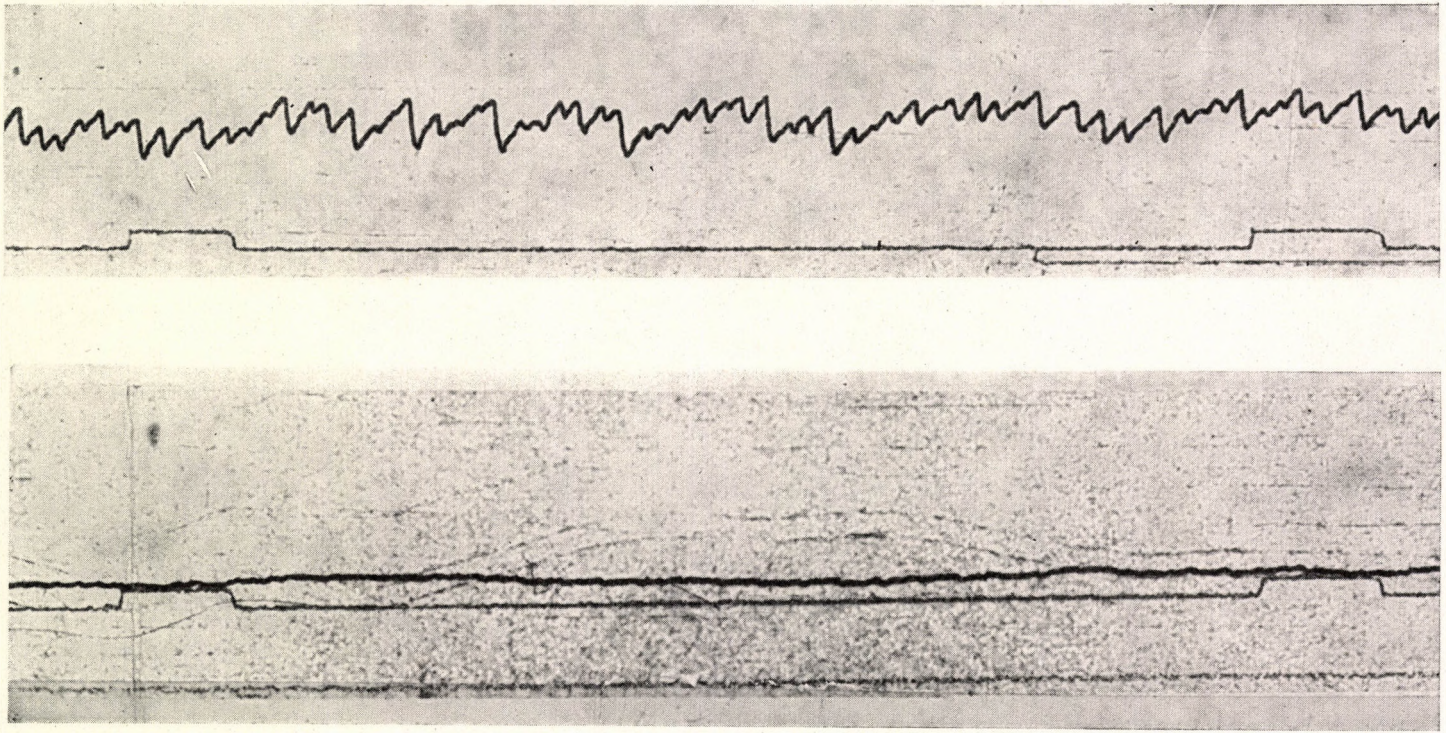


Fig. 12. Diagram of oscillations: the curves refer to the driving and the driven side of a piston type movable clutch. (The upper curve refers to the driver *Diesel Engine Csepel* 51 HP, 1500 r.p.m.; amplitude of abt. 100 μ in the horizontal plane, time sign. s; the lower curve refers to the dynamo driven amplitude abt. 10 μ)

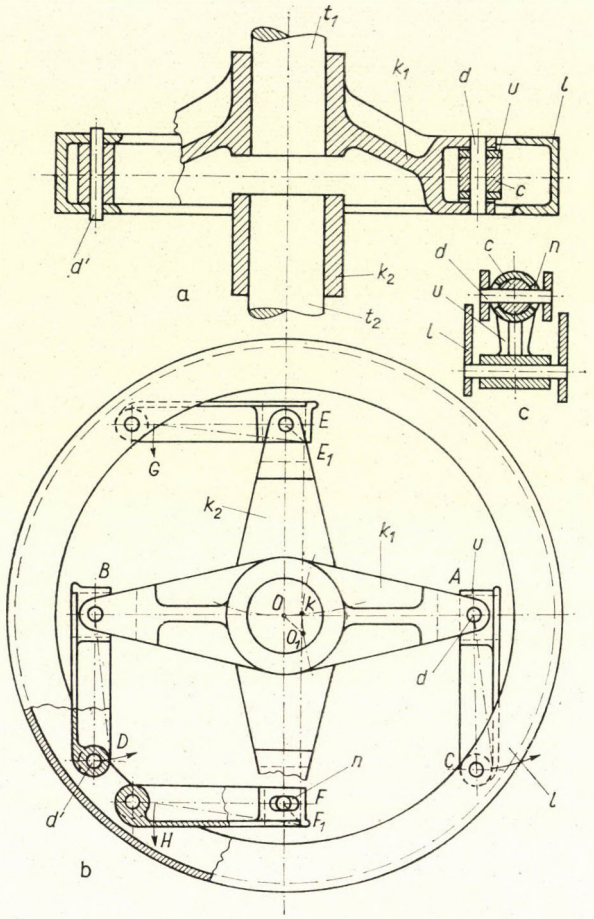


Fig. 13. Knuckled type self-aligning clutch

the same concrete fundament was driven; the lower curve represents the vibrations observed at the driven generator. Both instruments, one on the engine-body, the other on the shield of the generator, carried out measurements in the horizontal plane. On the figures the time intervals of 1 sec are indicated. The deviation between the two vibrations can easily be observed; the max. amplitude of the engine was about 100 μ, that of the generator only about 10 μ. This shows that the vibrations of the generator correspond to a free running.

This type of coupling has the great advantage of a fair self-adaptation to inaccuracies of mounting. A less favourable feature is represented by the fact that the sliding friction of pistons D required a special lubricating device.

The self-aligning clutch type shown in Fig. 13, representing a system of self-aligning articulated members was designed by the author to the end of eliminating sliding friction. The main distinct feature of this system is

that all the coupling elements are located in the same plane, wherefrom the overall axial length is short and no tilting moments can arise [26].

Each of the two 2-armed levers k_1 and k_2 is keyed on the driving and the driven shaft, respectively. (Fig. 13). The two arm-ends of each 2-armed lever are linked by means of a rod-member u to the intermediate ring-joint l. The linking articulation between the lever-arm k and the rod-member u is realized as a double wrist; the pin d located in the bearing holes of arm k, passes through the pivoting hole of the cylindric joint c with is located in the cylindric holes of the rod head of u, having longitudinal slots n in which joint c and pin d can perform a slight circular displacement around the axis of joint c. When the relative position of the two shafts t_1 and t_2 is not truly coaxial, but the axes form a slight angle, no tension or jamming can occur in the clutch mechanism: the joints c and pins d can be freely displaced in the slots n when the joint c is forced to pivot around its own axis. The connecting links n are connected to the ring by the pins d. The four pins $ABCD$ and also the other four $EFGH$ are located at the four corners of a parallelogram because the distances in the lever-arms (AB and EF respectively) are equal to the corresponding distances in the ring (CD and GH respectively).

In order to have an axial displacement performed by the two shafts t_1 and t_2, the articulated joints in the ring l at C and D, and G and H respectively, are of the same design as shown in Fig. 13 referring to the double-wrist type jointings at A and B, and E and F, respectively.

By a clutch of the described type, the transmission of a torque is carried out, not by a single peripheral force, but by a true couple. A model unit of this type of clutch was put into service working under the load of a torque of about 30 kpm; in an eccentric position arbitrarily chosen as 10 mm for the two shafts coupled together. The unit had to work in a day and night service in the *Gizella Mill Works* from the spring of 1956. The examination of the wear could not be realized because the whole model unit disappeared during the counter-revolutionary troubles in the autumn of 1956.

II. Roller type movable clutch

The best way to eliminate sliding friction is a design that incorporates some rolling elements like a railway wheel running on the rails.

According to Figs 14 and 15, the two coupling lever-members k_1 and k_2 are keyed to the shafts t_1 and t_2. On each two-arm lever there are pairs of pins $d_1 - d_1'$ and $d_2 - d_2'$, respectively, in symmetrical position and with their axes parallel to the shafts. On the extension of these pins, there are rollers $D_1 - D_1'$ and $D_2 - D_2'$ or roller bearings mounted. These rolling elements

are guided in two pairs of recesses which are grooved into the intermediate floating disk l, one pair of recesses having its centre line perpendicular to the centre line of the other pair of recesses: pair of recess h_1, with its centre line x; and pair of recess h_2, with its centre line y. Along the longitudinal direction

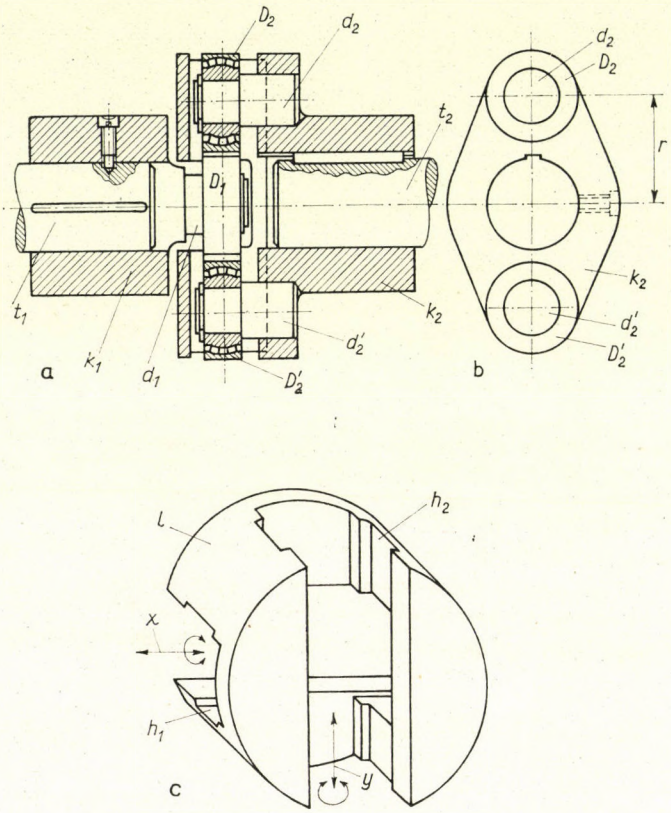
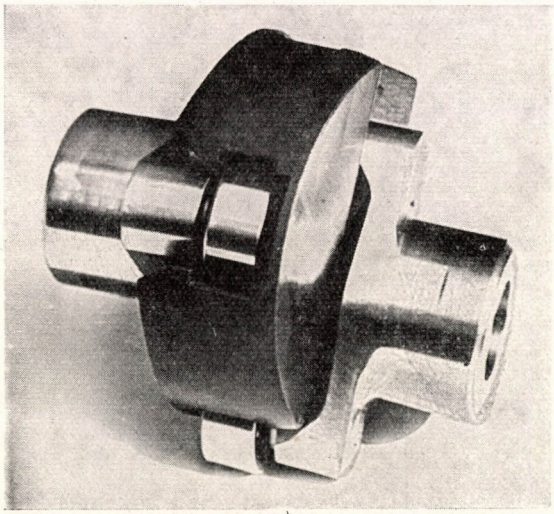


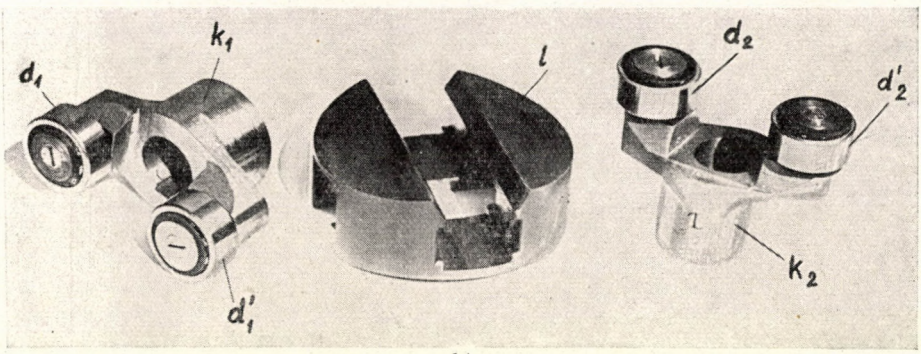
Fig. 14. Roller type movable clutch: longitudinal section (*a*); side-view (*b*); axonometric picture of the floating member (*c*)

of the recess (i.e. in radial direction of the clutch) the levers k and the shafts t_1 can be displaced, while the roller bearings are rolling on the guiding sidewalls of the recess. At the same time the self-aligning roller bearings can pivot slightly around the recess centre line x or y. E.g. the coupling lever-arm k_2 performs an axial shifting displacement along the shaft t_2. This clutch operates free from all kinds of stress, with skew shafts the axes of which are neither coaxial, nor parallel. The centre lines x and y of the pairs of bearings $D_1 - D_1'$ and $D_2 - D_2'$ are strictly parallel i.e. they determine a plane. Consequently, no tilting moment can arise, just contrary to a coupling that consists of two cardan joints and a telescopic shaft. In addition, our clutch requires far less space than a cardan joint.

Taking the shaft t_1 as fixed, the shaft t_2 can be shifted in a parallel position to itself, in every spatial direction; in other words, the shaft t_2 can assume ∞^2 positions. Since, the *ray-bundle* around a parallel position includes ∞^2 directions, the number of all the possible positions of the shaft makes ∞^4. Again, when the axial displacement of the shaft t_2 is controlled, by a sliding



a)



b)

Fig. 15. Roller type movable clutch: assembled (*a*); three main component parts (*b*)

wedge or by a dilatational fluted bush, the total number of possible displacements amounts to ∞^5, without exerting some force effect onto shaft t_1. Only a rotary movement can be transmitted from one shaft to the other: the torque becomes transmitted at the very starting of rotation.

As can be seen from diagrams in Fig. 16, a true couple (and not a single peripheral force) comes into play in transmitting the torque from the driving shaft to the driven one, when the recesses in the floating disk are, due to ma-

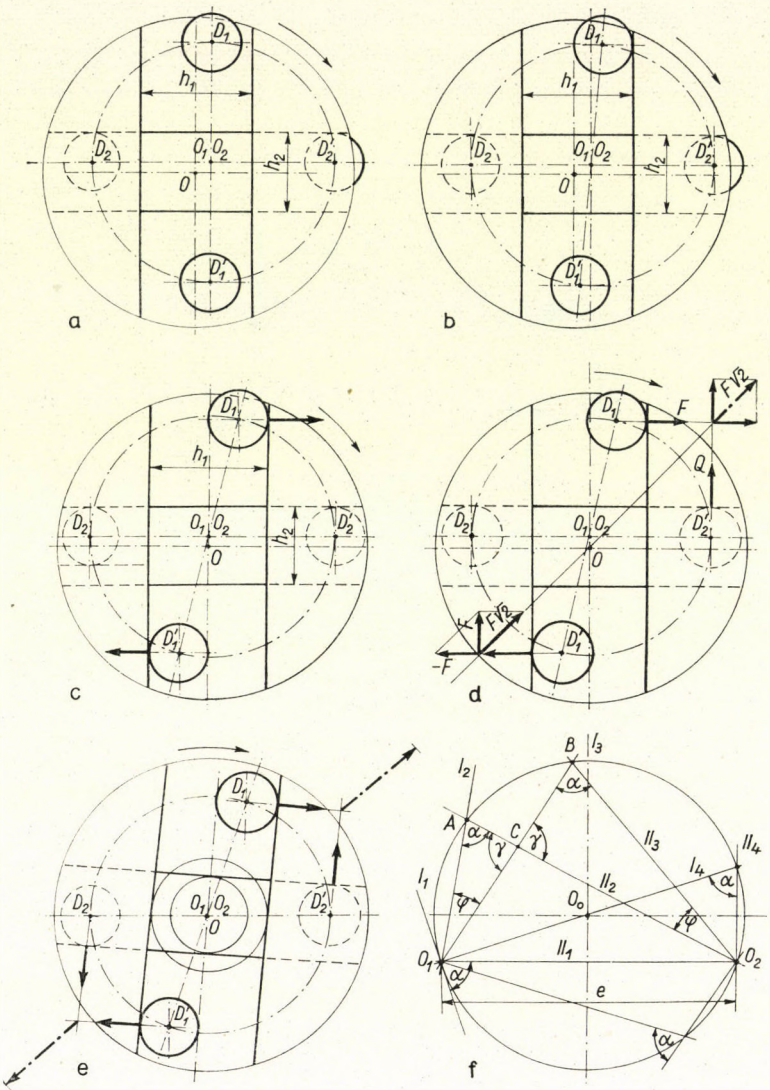


Fig. 16. The moment transfer is affected by a plain couple even when the width of the guide-ways is different and when the guide-ways' centre lines are not perpendicular to each other: The floating member (geometric and gravity centre in *O*), pulled down by its weight, rests on the pair of rolls D_2, D_2' *(a)*; start of rotation, D_1 begins to touch the floating member *(b)*; the side-wall of the recess is contacted by the roller D_1*(c)*; by the forces induced, the floating member is drawn-in into working position *(d)*; working position in which both the driving clutch-arm and the driven one are subjected to a plain couple (the balanced pulling forces acting on the floating member are denoted by a dot and dash line) *(e)*; the angular speed of the driving and the driven shaft is the same even when the recess is not in a rectangular position, or when shafts, are in a parallel position at an eccentricity *e(f)*

chining deviations, of unequal width or in a not quite true perpendicular position to each other. The particular phases are shown in Figs 16a, 16b, 16c, 16d, 16e and 16f.

Namely Fig. 16a: Starting position; unequal width $h_1 > h_2$; the floating disk, pulled down by its own weight, rests on the rollers D_2 and D_2'; the centre of the disk is O. Fig. 16b: The shaft starts rotating, and its roller D_1 comes into contact with the recess-wall in the disk, whereby this disk rotates by rolling on the roller D_2 and D_2' in the right hand direction as long as the roller D_1' in the lower position also comes into contact with the other wall of recess h_1 (see Fig. 16c). In Fig. 16d is shown that at this very moment a force F is exerted by the floating disk and pushes the roller D_2'.

Consequently, the resulting force $F\sqrt{2}$ that represents the vectorial addition of the original force F exerted by the roller D_1, and of the reaction-florce ($Q = |F|$) exerted by the roller D_2', pulls the floating disk. At the same time the force $(-F)$ exerted by roller D_1' and the aforementioned force $F\sqrt{2}$ result in a force $+F$ acting in a vertical direction upwards and brings the floating disk by lifting into contact with roller D_2.

In Fig. 16e the floating disk is lifted into its normal working position. One can state that both the driving and the driven shafts are each subjected to a true couple. The floating disk is subjected to a plain stress exerted by two pulling forces like a chain link. Concerning the workmanship of the floating disk the main requirement is the parallity of the recess-walls.

Provided the driving and the driven shafts are in parallel position, the centre O of the floating disk will rotate — at a double r.p.m. of the shaft's speed — on the periphery of a circle that represents a Thales-circle having a diameter equal to the eccentricity $e = \overline{O_1 O_2}$. Again, when the two recesses in the floating disk are not perpendicular to each other, the centre of the floating disk rotates on a circle for which the eccentricity e is not the diameter, but a chord (see Fig. 16f). Nevertheless, when the parallelity of the two shafts is maintained, the equality of the angular velocities of the two shafts is equally maintained. As can be seen from the diagram in Fig. 16f, a rotating displacement at an angle φ of recess I around the centre 0_1 involves a rotating displacement of recess II around the centre 0_2 at the same angle φ. This is proved by the similarity of the two triangles:

$$\triangle O_1AC \sim \triangle O_2BC.$$

Even when the arms r are not equal in either recess of the floating disk, the driven shaft is rotated by true couples. The difference consists in the mechanical stress of the floating disk; namely, the disk load is not a tensile stress caused by two equal and opposite forces; there are two equal couples which represent a torsional load.

1. *Effects of forces*

By the forces represented as a pair of forces of the torque, each of the driving pins is subjected to a bending stress. With a great r.p.m. the centrifugal forces of the pins and rollers are to be taken into account as producing some additional bending stress. In relation to its own axis, each of the rollers and the pins keeps a symmetrically fixed position; wherefore, independently of the eccentricity and speed the centrifugal forces are kept in balance. Conditions are quite different as regards the floating disk.

Again, when the disk centre point is — in relation to the shafts — in eccentric position, the single points of the disk describe quite different curves. Any single point of the recess centre-line moves along a Pascal-type conchoidal curve [1]. On the other hand, the centre O of the disks moves on a circle having a diameter e at an angular velocity of $2\,\omega$, when ω is the angular speed of the shafts. Accordingly, when the mass is imagined as concentrated in point O as the centre of gravity, the centrifugal force is:

$$C = m\frac{e}{2}4\,\omega^2 = 2\,me\omega^2. \tag{14}$$

The load of this force is always borne by the rollers. In the particular case when the whole force is borne by only two rollers, one single pin is loaded by the force

$$me\omega^2$$

which cannot be neglected when the values of e and ω are important.

In a general position, the load of the centrifugal force of the floating disk is distributed on all four pins.

At a rotation angle φ, this distribution between the two pairs of pins is expressed for the one pair of pins:

$$C_1 = C\sin\varphi = 2\,me\,\omega^2\sin\varphi$$

and for the other pair of pins (phase-difference $\pi/2$)

$$C_2 = C\cos\varphi = 2\,me\,\omega^2\cos\varphi.$$

In other words, the load pulsation of the centrifugal force of the disk mass reaches its maximum value $C_{\max}$ 4 times during one rotation, namely, twice on the driving pins, and twice on the driven pins.

2. *Principles of dimensioning*

When eccentricity is small, when the floating disk is light, and the angular speed is but moderate, then the centrifugal force of the floating disk can be disregarded for the purpose of dimensioning; its value might be taken into consideration when the calculated dimension values are checked. In the plane of coupling each pair of pins is subjected to two forces that act in the perpendicular direction. These are: the peripheral force F due to the torque M to be transmitted, and the centrifugal force which depends on the r.p.m., on the mass m_c of the rolling bearing and on the moment-arm r (Fig. 14b).

Now, calculating with a weight of G kp, and in view of

$$n = \frac{60}{2\pi}\omega = 9{,}55\,\omega$$

and the mass being

$$m_c = \frac{G}{9{,}81}\,\text{kg}$$

the centrifugal force is

$$C = m_c\, r\, \omega^2 = m_c\, r \frac{n^2}{91{,}2} = \frac{G n^2 r}{9{,}81 \cdot 91{,}2} = \frac{G n^2 r}{900}\,.$$

In other words, the value of C increases with increasing r, and on the other hand, the peripheral force

$$F = \frac{M}{2\,r}$$

decreases with increasing r; we can find a certain value of r for which the resultant force

$$E = \sqrt{F^2 + C^2}$$

assumes the minimum. When there are no contradictory constructional motives, the respective value of r should be maintained. Such a contrivance will represent a type of clutch working under minimum stress conditions:

$$y = E^2 = F^2 + C^2 = \frac{M^2}{4\,r^2} + m_c^2\,\omega^4 \cdot r^2\,,$$

$$\frac{dy}{dr} = -\frac{2\,M^2}{4\,r^3} + 2\,m_c^2\,\omega^4\,r = 0$$

and

$$r = \frac{1}{1{,}41\omega}\sqrt{\frac{M}{m_c}} = \frac{9{,}55}{1{,}41\,n}\sqrt{\frac{M}{m_c}}\,.$$

Denoting the weight of the roller bearing and the pin extension by $G = 9{,}81\ m_c$, we obtain

$$r = \frac{9{,}55 \cdot 3{,}13}{1{,}41\, n} \sqrt{\frac{M}{G}} = \frac{21{,}2}{n} \sqrt{\frac{M}{G}} . \tag{15}$$

Numerical example. Initial values: $M = 15$ mkp; $n = 2880$/min; $G = 0{,}4$ kp. It is

$$r = \frac{21{,}2}{2880} \sqrt{\frac{15}{0{,}4}} = \frac{21{,}2}{2880} \sqrt{37{,}5} = \frac{21{,}2 \cdot 6{,}12}{2880} = \frac{130}{2880} = 0{,}0452$$

viz. the optimum value of the moment-arm should be

$$r = 4{,}52 \text{ cm} \approx 5 \text{ cm};$$

consequently

$$F = \frac{M}{2\,r} = \frac{1500}{10} = 150 \text{ kp}.$$

The centrifugal force is

$$C = \frac{Gn^2 r}{900} = \frac{0{,}4 \cdot 2880^2 \cdot 0{,}05}{900} = \frac{0{,}4 \cdot 830 \cdot 5}{9} = 184 \text{ kp}.$$

and

$$E = \sqrt{150^2 + 184^2} = \sqrt{22\,500 + 33\,856} = \sqrt{56\,356} = 237 \text{ kp}.$$

Provided the mill has a slow running speed, the centrifugal force can be disregarded, wherefore *ONE* pin of a pair is loaded by the peripheral force

$$F = M/2r.$$

The shear stress of the pin:

$$\tau_c = 1{,}4 \frac{M}{2\,r \left(\frac{d^2 \pi}{4} \right)} .$$

The allowable tensile stress values σ_t and σ_c are determined:

For the pin (with diameter d):

$$\sigma_c = 2\,\tau_c ,$$

and

$$\frac{d^2 \pi}{4} \sigma_c = 1{,}4 \frac{M}{r} ; \tag{16}$$

for the shaft (with diameter D_t)

$$M = \frac{D_t^3 \pi}{32} \sigma_t .$$

By substitution into equation (16):

$$\frac{d^2\pi}{4}\sigma_c = 1{,}4\,\frac{D_t^3\pi}{32\,r}\sigma_t$$

and

$$d^2 = \frac{1{,}4\cdot 4}{32}\,\frac{D_t^3}{r}\,\frac{\sigma_t}{\sigma_c} = \frac{0{,}7}{4}\,\frac{D_t^3}{r}\,\frac{\sigma_t}{\sigma_c}\,. \tag{17}$$

With the assumed values we take for the shaft $\sigma_t = 600$ kp/cm²; and for the pin (with regard to a tensile stress in combination) $\sigma_c = 400$ kp/cm², so we obtain:

$$d^2 = \frac{0{,}7}{4}\cdot\frac{3}{2}\,\frac{D_t^3}{r} = \frac{2{,}1}{8}\,\frac{D_t^3}{r} \approx \frac{1}{4}\,\frac{D_t^3}{r}$$

and

$$d = \frac{1}{2}\sqrt{\frac{D_t^3}{r}}\,. \tag{17a}$$

When instead of r, the value of d is predetermined:

$$r = \frac{1}{4}\,\frac{D_t^3}{d^2}\,. \tag{17b}$$

For the calculation of the Hertzian stress occurring along a linear contact of the roller, the formula used in the construction of rolling bearings can be applied, which refers to the contact line between plane and cylinder with the following denotations:

F	peripheral force [kp] transmitted by the roller;
D	diameter [mm] of the roller;
l	width [mm] of the roller;
c	constant value equalling 89,1.

The Hertzian formula is:

$$p = c\sqrt{\frac{F}{Dl}}\ \text{kp/mm}^2$$

for which, using a steel roller with hardened surface, the allowable limit is of 20 000 kp/mm².

At a greater speed, the pin d can be calculated on the basis of bending stress under a bending load located at the middle plane of the roller. The radial load is given by the sum of the two centrifugal forces (due to the pin-extension and to the bearing), the tangential load is the sum of the peripheral force and of the centrifugal half force of the floating disk.

After the dimensions of the pin being calculated, the rolling bearing can be chosen on the basis of the static load. It is to be borne in mind that the rated static load specified for rolling bearings might be permitted at the three-fold — fourfold normal (limit) value when the bearing is subjected only to an oscillating movement [27]. In other words, the working conditions of the rolling bearing in the movable clutch are much more favourable when the shafts are not coaxial (only parallel, or in a skew position), as compared to truly coaxial shafts. Whereas in the latter case in the rolling bearing always subjected to load are the same balls or rollers, in the former case the load bearing element changes.

In this special case of load, the task of lubrication consists only of corrosion prevention. It is not necessary to build a housing for the bearing: a simple grease-bay is sufficient [28].

3. *Checking the form*

When the outer diameter (D) of the housing and the tolerated eccentricity (e) are known, it is indispensable to examine whether there exists or not any interference between the two bearings which carry out their rolling-oscillating movement in the two recesses perpendicular to each other.

Let the initial position be characterized by a vertical plane in which the axes of the shaft O_1 and that of the roller C_1 are located (Fig. 17a), and by a horizontal position of the straight line $\overline{O_2C_2}$. When the axis O_1 assumes a new position by rotating at an angle φ, the same angle of rotation is performed by the axis O_2, and the centre of the floating disk rotates from point O_1 onto point O. Let point O_1 be the centre of a system of rectangular coordinates. Let us consider path h equalling the distance $\overline{C_1'C_2'}$ when the roller centres C_1 and C_2 rotate as the end points of the radii $r = \overline{O_1C_1} = \overline{O_2C_2}$ around their respective centres O_1 and O_2, so we have (in compliance with the denotations in Fig. 17a)

$$x_1 = r \sin \varphi; \quad x_2 = -r \cos \varphi + e,$$

$$y_1 = r \cos \varphi; \quad y_2 = r \sin \varphi;$$

$$h^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2$$

viz.:

$$h^2 = (r \sin \varphi + r \cos \varphi - e)^2 + (r \cos \varphi - r \sin \varphi)^2,$$

$$y = h^2 = [r(\sin \varphi + \cos \varphi) - e]^2 + r^2 (\cos \varphi - \sin \varphi)^2.$$

Let us determine the value of the angle φ, at which the path h assumes

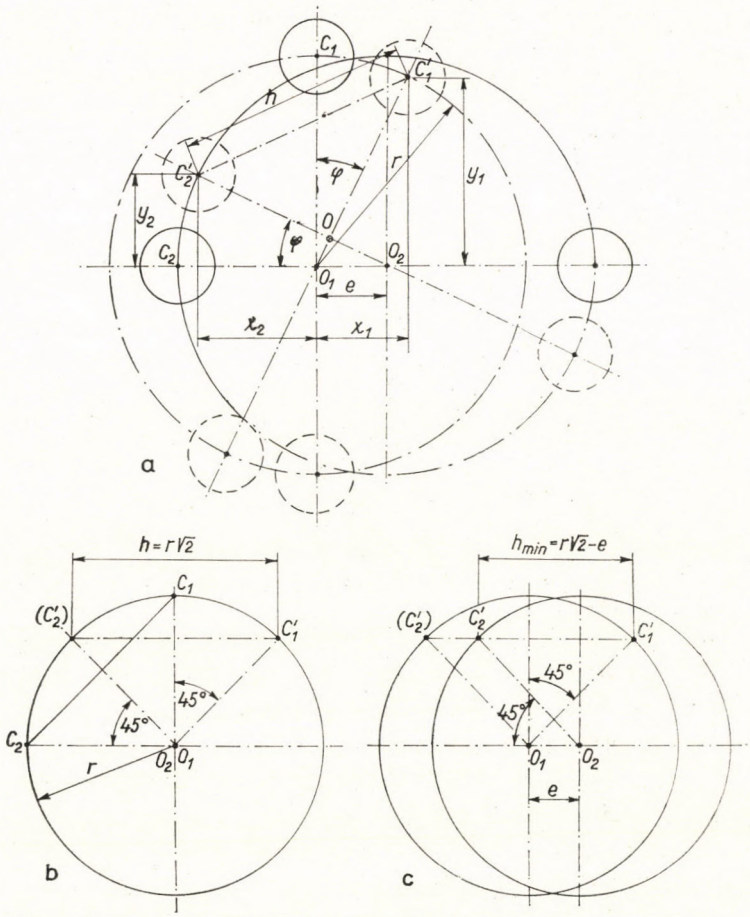


Fig. 17. Determination of the minimum distance between rollers: Rollers' distance, general position *(a)*; rollers' distance contact in coaxial position *(b)*; roller's distance assumes the minimum at $\varphi = 45°$, when shafts are stifted into an eccentric position *(c)*

a minimum.

$$\frac{dy}{d\varphi} = 2\,[r\,(\sin\varphi + \cos\varphi) - e]\,r\,(\cos\varphi - \sin\varphi) +$$

$$+ 2\,r^2\,(\cos\varphi - \sin\varphi)\,(-\sin\varphi - \cos\varphi) = 0\,,$$

$$2\,r^2\,(\sin\varphi + \cos\varphi)\,(\cos\varphi - \sin\varphi) - 2\,er\,(\cos\varphi - \sin\varphi) -$$

$$- 2\,r^2\,(\sin\varphi + \cos\varphi)\,(\cos\varphi - \sin\varphi) = 0\,,$$

viz.:

$$2\,er(\cos\varphi - \sin\varphi) = 0.$$

Obviously, the expression between the brackets is 0, only when

$$\sin \varphi = \cos \varphi,$$

i.e. $\varphi = 45°$ or $\varphi = 235°$.

At $\varphi = 45°$, we obtain

$$h_{\min} = \frac{2r}{\sqrt{2}} - e = r\sqrt{2} - e. \tag{18}$$

On the other hand, this statement can be justified by a simple visua observation. Namely, a coaxial position provided, the distance h between the roller axes is, (see Fig. 17b) independently of the value of the rotation angle, equal $h = r\sqrt{2}$, by including the case of $\varphi = 45°$. After shifting the shaft centre O_1 from its original position onto O_2 (according to Fig. 17c), the roller centre C_2 changed its position from (C_2') to C_2', and $\overline{O_1O_2}$ (Fig. 17c) equals $e = \overline{(C_2')C_2'}$; consequently, with

$$\varphi = 45°, h_{\min} = r\sqrt{2} - e.$$

When the centre lines of the recesses form an optional angle α different from 90°, the minimum value $h_{\min}$ will occur in a special position, in which the bisectrix of α is perpendicular to the eccentricity e.

In Fig. 18a, the rollers are illustrated in a position, in which an additional safety gap s is between each other. Using the same symbols as denoted in Fig. 18a, we can write:

$$h_{\min} = C_2C_1 = r\sqrt{2} - e = D_1 + s \tag{18a}$$

and the maximum available value of the eccentricity e

$$e_{\max} = r\sqrt{2} - (D_1 + s). \tag{18b}$$

E.g. with $r = 5$ cm, $D_1 = 5{,}4$ cm, $s = 0{,}2$ cm; we obtain $e_{\max} = 1{,}41 \cdot 5 - 5{,}6 = 7{,}05 - 5{,}6 = 1{,}45$ cm.

From the diagram according to Fig. 18a, the displacement b of the roller centre C_2 in the direction of its own centre line can be determined:

$$b = OC_2 - R_1 = \frac{h_{\min}}{\sqrt{2}} - R_1 = \frac{2R_1 + s}{\sqrt{2}} - R_1,$$

viz.:

$$b = R_1(\sqrt{2} - 1) + \frac{s}{\sqrt{2}}.$$

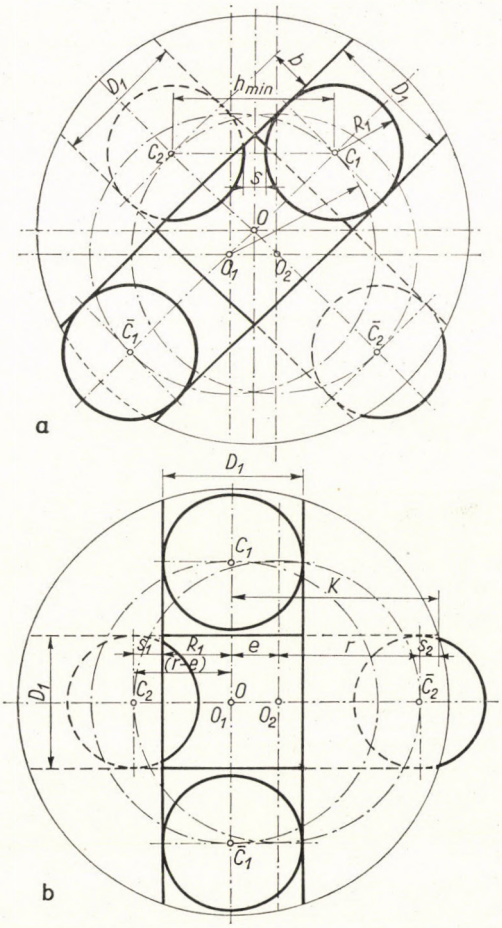


Fig. 18. Rollers' limit-positions: Ensures gap s between rollers *(a)*; safety gaps s_1 and s_2 at the guide-way *(b)*

With the numerical values of the above example:

$$b = 2{,}7 \cdot 0{,}41 + \frac{0{,}2}{1{,}41} = 1{,}107 + 0{,}141 = 1{,}248 \text{ cm}.$$

In Fig. 18b the initial position of the rollers is characterized by a safety gap s_1 in the extreme inner position, and s_2 in the extreme outer position.

As seen from the diagrams

$$s_1 = r - (e + R_1). \tag{19}$$

With the same numerical value as before:

$$s_1 = 5 - (1{,}45 + 2{,}7) = 5 - 4{,}15 = 0{,}85 \text{ cm}.$$

and

$$s_2 = K - (e + r). \tag{19a}$$

If we intend to make $s_1 = s_2$, then

$$r - e - R_1 = K - e - r,$$

viz.:

$$K = 2\,r - R_1. \tag{20}$$

With the numerical values as before

$$K = 2 \cdot 5 - 2{,}7 = 10 - 2{,}7 = 7{,}3 \text{ cm}.$$

Let us check the above calculation. Overall length of the recess is

$$2\,K = 2(e + s_1) + 2\,r = 14{,}6 \text{ cm};$$

and

$$2(e + s_1) = 14{,}6 - 10 = 4{,}6 \text{ cm}.$$

With $2\,e_{\max} = 2{,}9$ cm, we obtain

$$2\,s_1 = 4{,}6 - 2{,}9 = 1{,}7 \text{ cm}$$

and

$$s_1 = s_2 = 0{,}85 \text{ cm}.$$

In compliance with what has been expounded so far, assembling in *axial direction* can be made possible by means of a sliding key mounted on one shaft-end (or on its fixed sleeve) that allows for a necessary diletation. When we adhere to the possibility, of disconnecting the clutch together with the driving or the driven part in *lateral direction* without the necessity of loosening a single bolt or screw, then the following requiremment should be satisfied

$$(2\,r - D_1) > D_1$$

as seen from the diagram in Fig. 18b.

Again, when at the right hand side and the left hand side of the recess, a safety gap s_3 (on each side) should be maintained, the following equation has to be satisfied:

$$R_1 + s_3 + R_1 = s_3 + D_1 = r,$$

viz.:

$$s_3 = r - D_1$$

and

$$D_1 = r_{\min} - s_3\,. \tag{21}$$

In order to prevent any interference (see Fig. 18b)

$$D_1 = r\sqrt{2} - e_{\max} - s\,.$$

When the requirement is specified that no interference could occur between the rollers at the eccentricity e_{max}, and the dismounting in lateral direction should be made possible with a clutch having the minimum radial dimension, the following condition is determined by Equations (18b) and (21):

$$D_1 = r_{min}\sqrt{2} - e_{max} - s = r_{min} - s_3$$

viz.:

$$r_{min}\sqrt{2} - 1 = 0{,}41\, r_{min} = e_{max} + s - s_3 . \tag{22}$$

When the safety gaps are chosen as equal values:

$$s = s_3 ,$$

then the maximum allowable eccentricity is

$$e_{max} = 0{,}41\, r_{min} . \tag{22a}$$

With the numerical values of the above described clutch:

$$s_3 = r - D_1 = 5 - 5{,}4 = 0{,}4 \text{ cm};$$

with this, the lateral dismounting is impossible. In order to comply with such a condition, only the value $r = 5{,}6$ cm would correspond to the assumed $s_3 = 0{,}2$ cm.

When for the arm r the greater value of 5,6 cm is chosen, then according to Fig. 18b

$$e_{max} = r\sqrt{2} - (D_1 + s) = 1{,}41 \cdot 5{,}6 - (5{,}4 + 0{,}2) = 2{,}3 \text{ cm}$$

should be realized, instead of the value 1,45 cm as calculated before. Should we make full use of the greater eccentricity, then the dimension K should be determined, according to Equation (20), to 8,5 cm instead of 7,3 cm.

III. Shafts having skew axes or intersecting ones

As can be seen from Fig. 17a, in the case of parallel shaft-axes, the extreme position of the intersection point of the centre lines of the two recesses in the floating disk is located on the axis of that shaft, the rotating lever arm of which is perpendicular to the planes containing the shafts t_1 and t_2.

Let the one shaft be rotated around the distance between the axes as a normal transversal straight line, then the former statement should be modified as follows: With shafts having skew axes, the extreme position of the centre of the floating disk is located on the axis of that shaft the instantaneous lever arm of which is perpendicular to the normal transversal straight line.

Consequently, the floating disk performs a tilting movement at an angle α around the normal-transversal straight line when the shaft-axes are in a skew position (Fig. 19a), and the tilting movement is carried out around the straight

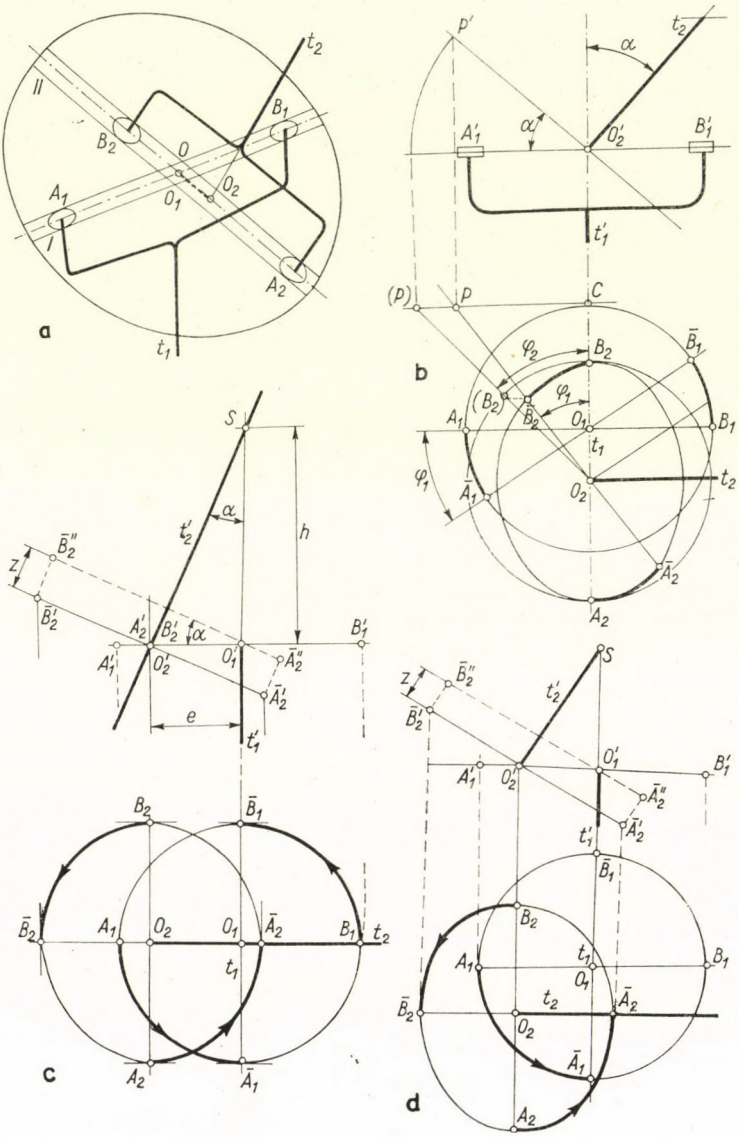


Fig. 19. Tilting and sliding movement of the floating member: With shafts in skew position, the centre *O* of the floating member has its locus on the axis of that shaft, the coupling arm of which assumes a momentaneous position of its CL perpendicular to the normal transversal line *(a)*; determination of the fluctuation of the angular velocity *(b)*; determination of the forced axial sliding of the floating member with inter secting shafts *(c)*; determination of the forced sliding with shafts in skew position *(d)*

line that passes through the intersection point, and is perpendicular to the plane formed by the axes of the shafts when they intersect each other, in which case the angle α is the angle formed by these axes.

1. Speed fluctuation

Let us consider the sketch in Fig. 19b, comprising a front-view and a view from above of a clutch with shafts having skew axes, in a position in which the normal transversal straight $\overline{O_1O_2}$ is represented at its true length.

As can be seen in the view from above the path of the rollers A_1 and B_1 of the shaft t_1 is the circle with the centre point O_1 and the diameter $\overline{A_1B_1}$. The path of the rollers A_2 and B_2 of the shaft t_2 is the circle with the centre point O_2, and the diameter $\overline{A_2B_2}$, represented here as an ellipse, with the centre point O_2 and the large diameter $\overline{A_2B_2}$.

Corresponding to what has been said before and to the initial extreme position in Fig. 19a, the centre point of the floating disk is located in O_1. When shaft t_1 is rotated at the angle φ_1, to be seen in its real magnitude in the view from above (Fig. 19b), the roller centres assume the new positions at points $\bar{A}_1\,\bar{B}_1$.

Since the rotating of the integer member by which the rollers $A_2\,B_2$ are supported, is only possible in a plane perpendicular to the shaft-axis t_2, this arm is to be seen, after having rotated, in its front-elevation, on the straight line $P'\,O_2'$ perpendicular to t_2'. In the view from above, the position after rotation $\bar{A}_2\,\bar{B}_2$ of the rollers is found as the endpoints of a diameter of the ellipse, on a straight line drawn from point O_2 as a perpendicular to the diameter $\bar{A}_1\,\bar{B}_1$. Namely, the centre line of the guide way $A_2\,B_2$ traverses point O_2, and can only be rotated in a plane perpendicular to the circle diameter $\bar{A}_1\,\bar{B}_1$. In this way, both projections of the centre line of the guide way belonging to shaft-axis t_2 are known. When rotating this centre line around the normal transversal into the plane perpendicular to the axis t_1, the rotating angle φ_2 of axis t_2 can be determined.

Let us consider the turned in projection (P) in the 1st drawing plane that belongs to an arbitrarily chosen point P, and

$$\sphericalangle\,\varphi_2 = \sphericalangle\,(P)\,O_2\,C\,.$$

According to Fig. 19b, we can write:

$$PC = CO_2 \tan\varphi_1\,,$$

$$(P)C = CO_2 \tan\varphi_2$$

viz.:

$$\frac{\tan\varphi_2}{\tan\varphi_1} = \frac{(P)C}{PC} = \frac{P'O_2'}{PC}\,.$$

Since

$$\frac{P'O_2'}{PC} = \frac{1}{\cos\alpha},$$

we obtain

$$\tan\varphi_2 = \frac{\tan\varphi_1}{\cos\alpha}. \tag{23}$$

With shafts intersecting each other, or with shafts in a skew position, the difference between the rotation arcs in our clutch is exactly the same, as that in a cardan-joint of shafts having the same angle of inclination α.

According to the calculating method for cardan-joints (see [21, 32]), we can write:

$$\omega_2 = \omega_1 \frac{\cos\alpha}{\sin^2\varphi + \cos^2\varphi \cdot \cos^2\alpha} \tag{24}$$

and at the position $\varphi = 0$

$$\omega_{2\max} = \frac{\omega_1}{\cos\alpha}. \tag{24a}$$

2. *Axial sliding*

The position of the meridional central plane of the floating disk in space is constant, whether the shafts are coaxial or parallel. Similarly, when the central meridional plane of the floating member *traverses the intersection point of two intersecting shafts, or when this plane traverses the normal transversal straight line* of two shafts in skew position, the floating member is compelled to perform a tilting movement, but *no axial gliding occurs.* Again, when the central meridional plane of the floating disk does not traverse the intersection point, neither does it transverse the normal transversal straight line (Fig. 19c), not only tilting, but an *axial gliding* is also inevitable.

Corresponding to the rotation angle $\varphi_1 = 0$, and $\varphi_1 = 90°$ of the shaft t_1, the position of the floating member is delineated, for intersecting axes, in Fig. 19c, and for axes in skew position, in Fig. 19d. From the diagram it can be clearly seen, that the central meridional plane of the floating disks is compelled to occupy the position $\bar{A}_2'' \bar{B}_2''$ instead of the position $\bar{A}_2' \bar{B}_2'$, equalling a sliding displacement z in the direction of axis t_2, provided that the rotating lever arm of the clutch is rigidly connected to shaft t_1, and that the axial sliding motion is made possible on the axis t_2.

According to the symbols denoted in Fig. 19c, the eccentricity is

$$e = h \cdot \tan\alpha$$

and the axial displacement is

$$z = e \sin \alpha = h \tan \alpha \cdot \sin \alpha = h \frac{\sin^2 \alpha}{\cos \alpha} . \tag{25}$$

When the mill frames in question are coupled in series, then the probable value of the distance h is the half length of a roll, because in general, both bearings can be adjusted. In a less favourable case, h equals the whole length of the roll measured from the clutch up to the outer bearing. The inclination angle of the axes of the adjacent rolls varies between $\alpha = 0° \div 2{,}5°$.

According to equations (24a) and (25) the probable values of the fluctuation of the angular velocity and the axial displacement can be calculated, when $\alpha = 2{,}5°$ and $h = 600$ mm (with $\cos 2°30' = 0{,}99937$); this is for the fluctuation of the angular speed:

$$\frac{1}{100} \frac{\omega_{2\max} - \omega_1}{\omega_1} = \frac{1 - \cos \alpha}{100 \cos \alpha} \approx 63 \cdot 10^{-7} \%,$$

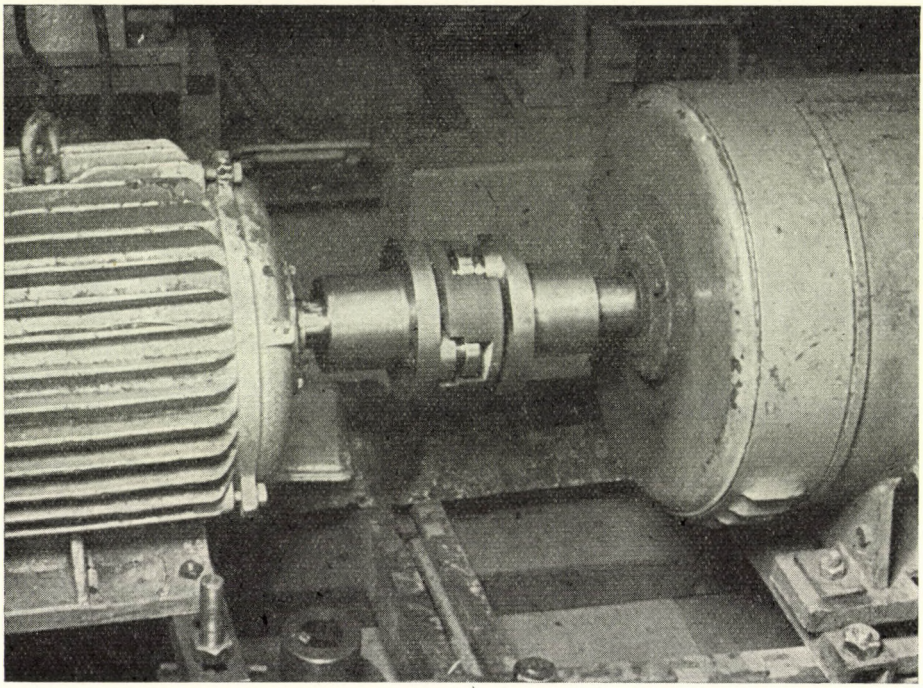
and for the axial displacement:

$$z = \frac{h \sin^2 \alpha}{\cos \alpha} = \frac{600 \cdot 0{,}03548^2}{0{,}99937} = \frac{600 \cdot 0{,}00126}{0{,}99937} \approx 0{,}75 \text{ mm}.$$

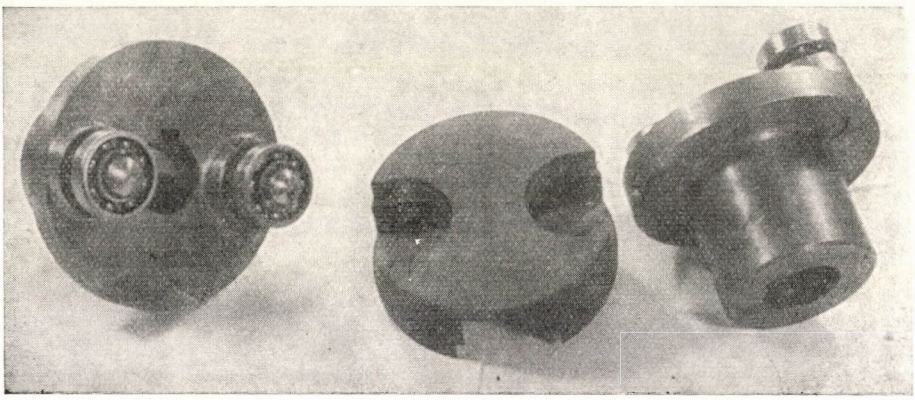
As can be seen from this calculation, it is worth-while to make the sliding track of the rollers in the floating disk some tenth of mm broader, because when sliding is carried out on the quide way of the roller, a *rolling along the plane that is perpendicular to the sliding direction, causes the friction to be greatly restricted.* In this connection, a clutch-arm built with a sliding key, mainly facilitates the assemblage.

IV. Recent experiences

As author had the opportunity to learn how leaders in the industry formed their opinion on the application of movable roller-type clutches, regrettably they are not able to recognize the importance of a better efficiency that can be obtained when no additional load occurs on the bearings, provided the clutch works by a plain couple. In this way, people in the *Sopiana Works*, where couplings and clutches are made, did not at all undertake to study this suggestion made by the author. When assembling the prototype of a newly constructed mill frame, they were rather prepared (in the *EMAG Factory of Agricultural Machines*) to insert an additional but superflous unit of trans-



a)



b)

Fig. 20. Roller type movable clutch: Mounted on a laboratory electric-motor *(a)*; three main component parts *(b)*

mission into the main drive, instead of a movable clutch, in order to economize reserve gear wheels. It was not difficult to predict that this type of clutch would fail and had to be discarded [29].

Although the application of movable clutches was suggested by the author for the purpose of coupling mill frames in series already in 1955, leaders of the interested plants preferred to insert a double cardan-joint, which as

known involves detrimental service conditions, that overcomes the apparent advantage of a short overall length [30].

The idea of applying the roller type movable clutch was the more welcomed by the research institutes.

The first model of roller type movable clutches (Fig. 20) were designed and built in 1961/62 by the *Institute for Development of Heavy Current Engineering*. This model served in the Institute's Laboratories for the transmission of torques of various magnitude. On the basis of these experiments, a new set of this type of clutch was built into medical appliances in order to save the adjacent bearings.

It was the department leader of the *National Research Institute for the Milling and Corn-Storing Trade* (new entitled: *Research Institute of the Corn-Trust*) I. KOLOSTORI, M. E. whose attention was attracted to the suggestions made by the present writer in 1955, and to his publications edited later, especially to that under III. Mr. KOLOSTORI asked for the cooperation of the author in the task of solving the problem of mill frames to be coupled in series. Mr. KOLOSTORI spent much time and labour as assistant collaborator in the designing and construction of the special type of clutch for the said purpose [31]. In Figs 21, 22 and 23, the experimental grinding appliance and the system of coupling the mill frames in the *Bicske Mill Works* are illustrated. The more rational conduction of the flux of energy proved effective mainly at the smooth rolls.

Four units of a movable clutch were built in as trial pieces; floating disks of unhardened steel were built into two of the clutches, and in the other two units floating disks of plastic "Danamid" were mounted. In a similar way, two clutches were fitted with ball bearings, and two others had self-aligning roller bearings. In Fig. 24, the detailed design of this type of clutch is shown that is built for the transfer of a torque 20 mkp, at a speed of $n = 550$ rpm.

There was no wear at all to be found on the ball and roller bearings of these clutches that were kept operating for a full year, from December 1964, in day and night service. On the floating disks of unhardened steel, trough type indentations were observed, as was expected. Again, on the floating disks of plastic, the mirror finish of the working surfaces remained unchanged. By the slight elastic deformation of the pastic material, *the Hertzian stresses were suddenly reduced.* After having the steel disks replaced by plastic ones, all the four clutches are still running.

V. New peculiar design of the roller type movable clutch

As can be seen from what has been explained above, the roller type movable clutch prevents oscillations from being transferred from one shaft to the other, because all kinds of vibrations are taken over by the floating disk. Again, the extra centrifugal force which comes from the own vibrations

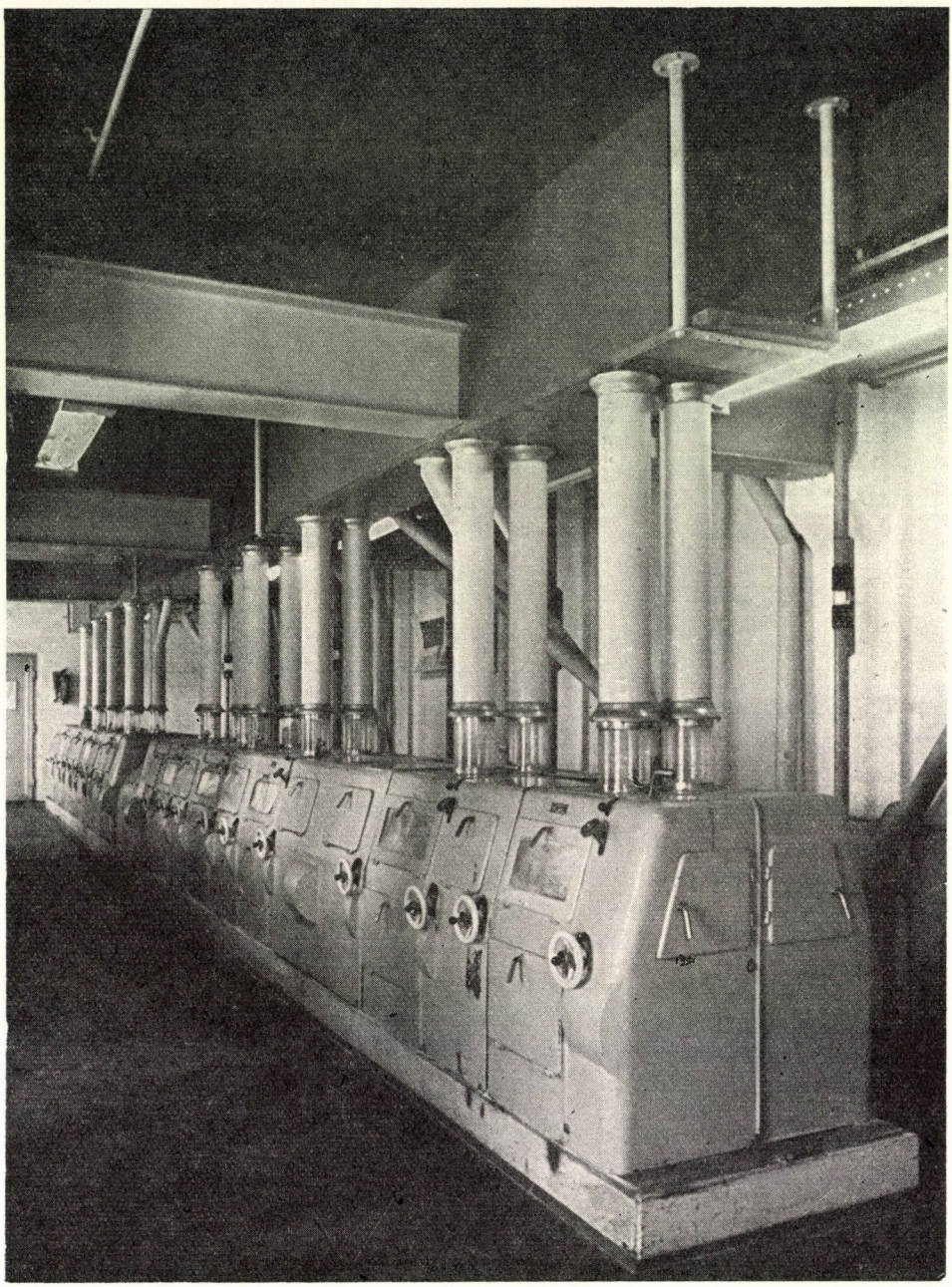


Fig. 21. Grinding equipment for testing purposes: Series of mill frames in the *Bicske-mill*

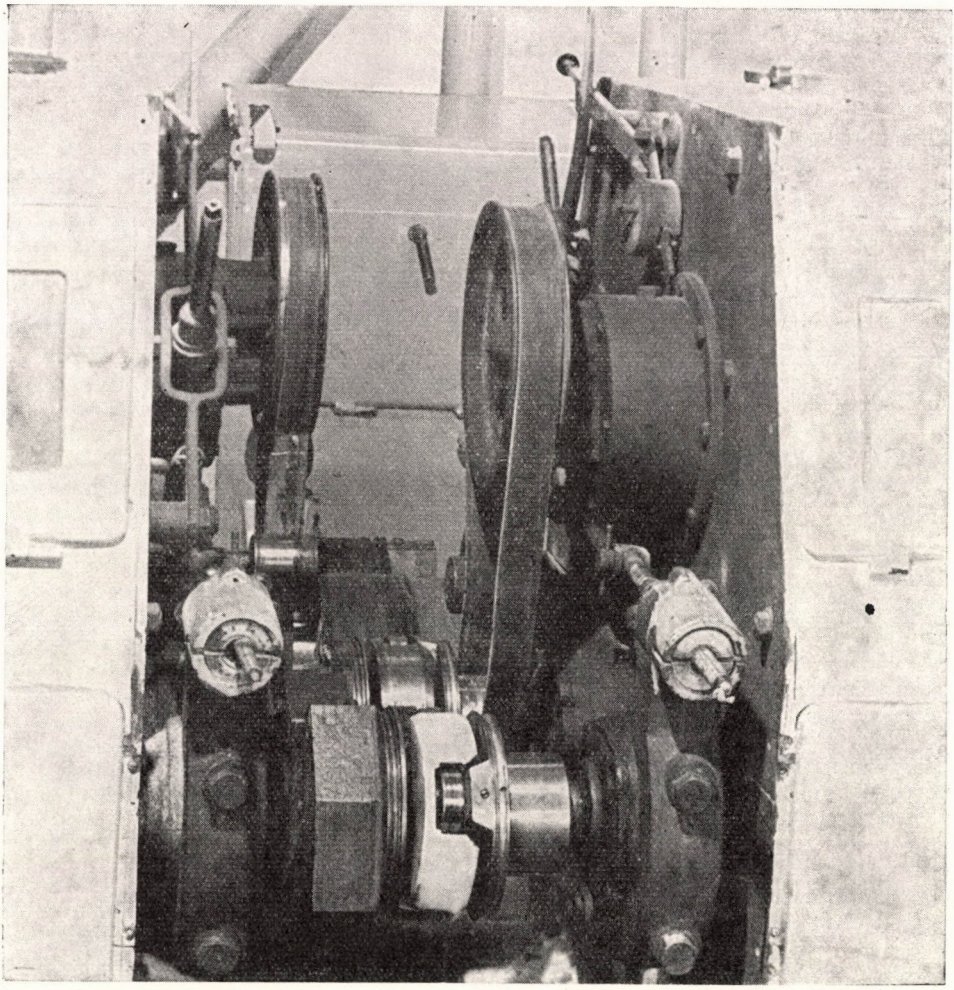


Fig. 22. Grinding equipment for testing purposes: Adjacent rolls coupled by a roller type movable clutch

of the floating disk can be automatically balanced in the case of shafts in parallel position, if only the bearings are protected from any additional load originating from the natural vibrations of the clutch, that can be induced by variation of speed or of eccentricity, during operation.

1. Balancing by means of a new member inserted into the kinematic chain

The basic idea of this new design is a second dummy floating disk of the same mass as the working disk, mounted in a special position in which

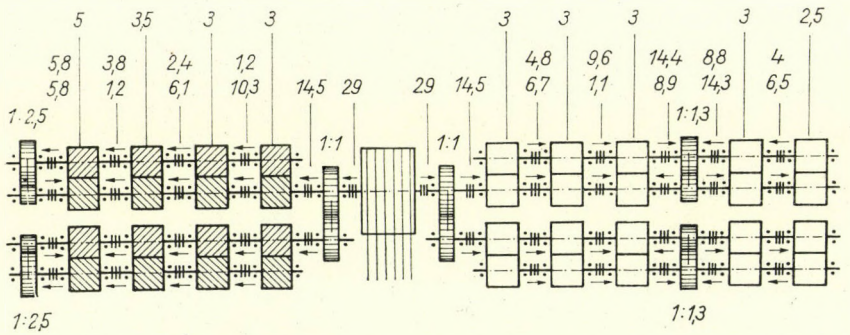


Fig. 23. Grinding equipment for testing purposes: Mill frames coupled in series. On the left side, the break rolls, on the right side the grinding rolls are positioned (direction of energy-flux indicated by arrows; ciphers refer to the power required for rolls, clutches, and gears)

the gravity centre of the dummy disk rotates on the same circle having a diameter e, being the same as that of the working disk, but at a phase displacement of 180°.

As is shown in Fig. 25a, we shall consider a rotation of the lever arm k_1 at an angle of φ, by assuming a position of the parallel shafts characterized by the intersection points of their respective axes with the plane of the drawing paper located at points O_1 and O_2. The centre lines of the guiding ways I and II of the working floating disk l are perpendicular to each other; consequently, when traversing the shaft-centres the mass-point of the working disk is located at that one of the shaft axes which is intersected by the guide way centre line being at this moment perpendicular to the axial distance $O_1 O_2 = e$. (NB. the symbol m denotes the mass of the floating disk *and* its gravity centre that coincides with the geometric centre.) In the diagram as delineated in Fig. 25a, the initial position of the mass point m is in the centre O_2.

From the initial position after an angular displacement at the angle φ, the position of the working floating disk m' is determined by the new position I' and II' of the guiding way centre lines. The angular displacement of the mass-point m around the midpoint O of the eccentricity e equals 2φ, wherefore its angular speed is 2ω and the centrifugal force generated is:

$$C = m \frac{e}{2} (2\omega)^2 = 2 me\omega^2 .$$

Now, in the balancing dummy disk there are two guiding ways having centre lines perpendicular to each other; their relative position with reference to the first turning lever-arm is characterized by a position in which the centre line a is perpendicular to I, and the centre line b is perpendicular to II. After the shaft made the angular displacement at angle φ, the mass-centre E of the balancing disk moves along the circle having a diameter e, and reaches the po-

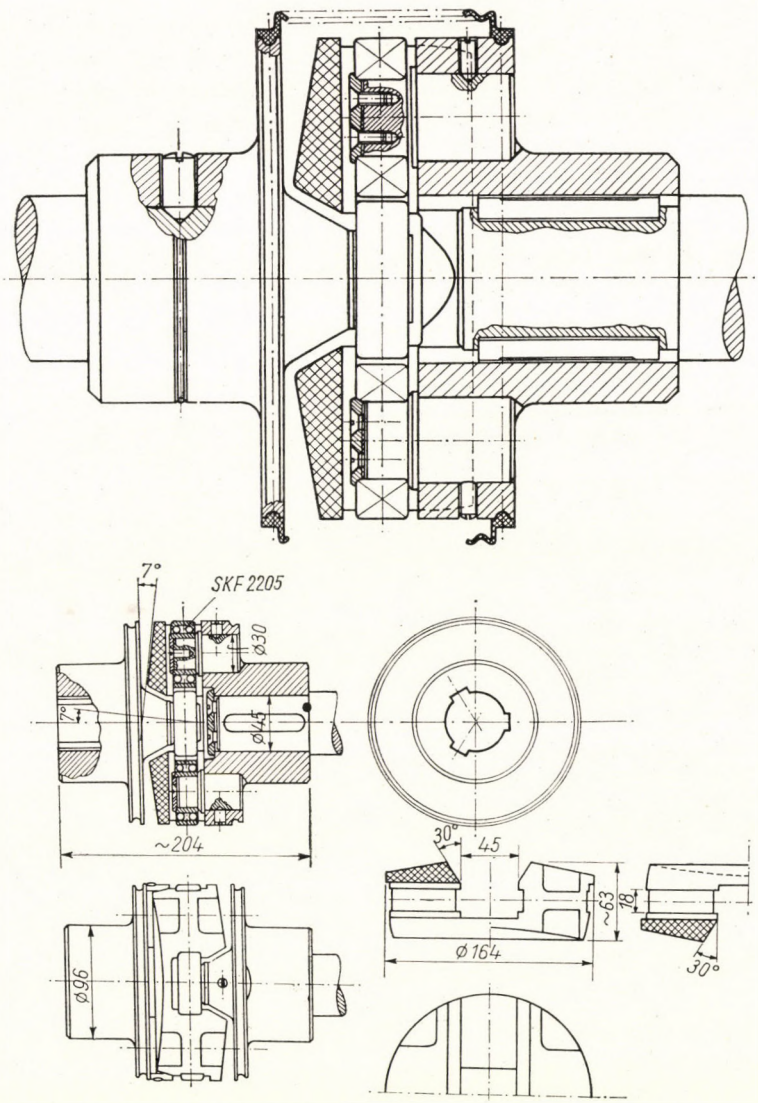


Fig. 24. Design sketch of the clutch developed for a coupling in series of the mill frames

sition E' opposite to point m'; thus, when $m = E$, the centrifugal force $(-C)$ acts at point E'.

Constructional details are shown in Fig. 23b, The by-lever arm b of one turning lever-arm that joints the balancing floating lever-arm is not fixed to the shaft, but it can rotate freely around the shaft-axis; by this contrivance it could be ensured that no constrained force could be generated even during assembling neither by inaccuracy of mounting, nor when the shaft axes are in a skew position. In this way, the phase difference of 180° is kept constant

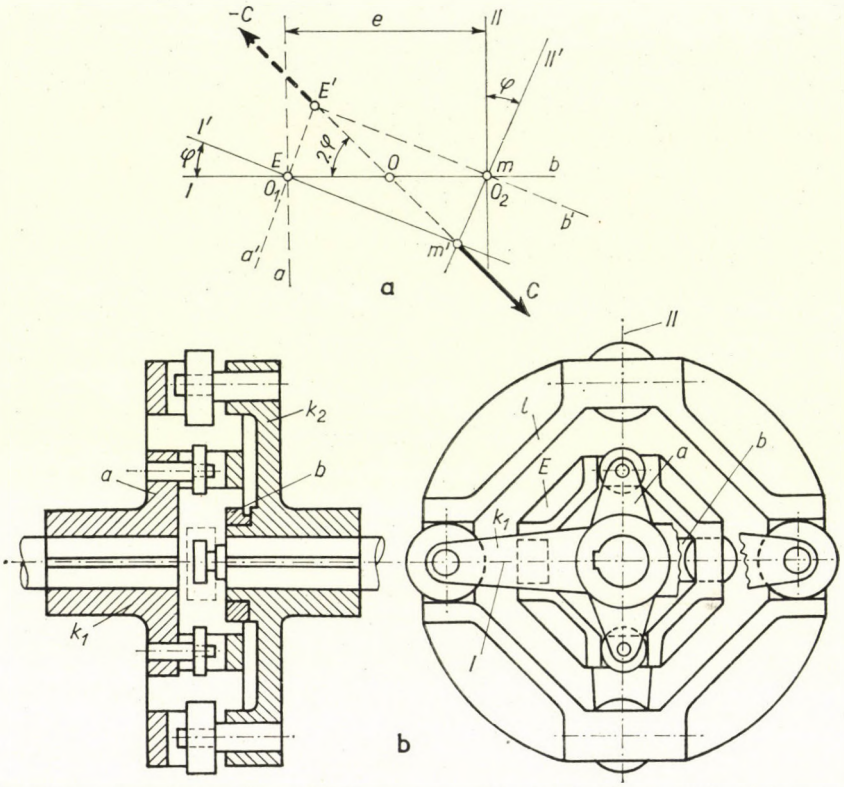


Fig. 25. Kinematic scheme of a self-balancing movable clutch: The centrifugal forces of the floating member l and of the balance-weight E are at any moment of rotation in equilibrium, e is eccentricity of parallel axes and φ the optional angle of rotation *(a)*; schematic design *(b)*

independently of shaft position. Further, since the dummy disk acting as the counterbalance moves in the very plane of the working floating disk, no tilting moment can occur.

In order to keep the centrifugal force as small as possible, the working disk is made of some material having a low specific weight, whereas the specific weight of the dummy disk — that is subject only to the centrifugal force — can be higher.

2. *Automatic balancing by means of for sliding masses*

In clutches which have to work at a higher speed it is reasonable to keep the diameters of rotating parts as small as possible; here we can rely on a method that makes the application of a more complete kinematic chain superfluous. To this end, we can choose as a basing principle the mechanical rule according to which the centre of gravity of a mass rotated by a couple always tries to find its locus in the centre of rotation.

In the two guide-way of the floating disk which intercross each other, one counterbalance is placed in each, which are thus rotated by a couple. The counterbalancing weights have the same mass but they do not hinder each other in free moving (Fig. 26a).

The mass m_1 slides in the guide-way I which rotates around the point O_1; thus, its gravity centre will try to find its locus in point O_1. In the same way,

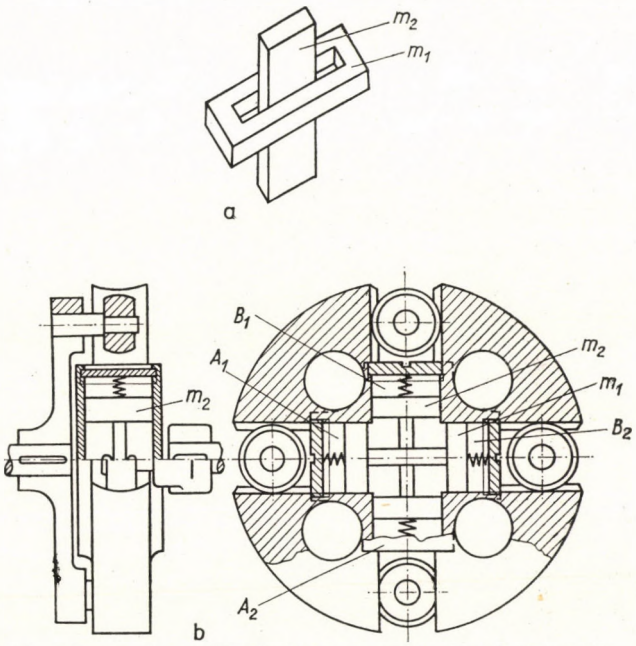


Fig. 26. Dynamic scheme of a self-balancing movable clutch: Symbolic sketch of masses free-sliding in cross direction *(a)*; schematic design *(b)*

the gravity centre of the mass m_2 will try to find its locus in point O_2, since the mass m_2 slides in the guide way II which rotates around point O_2. The schematic outline of the clutch is given in Fig. 26b.

As a matter of fact, the single points of the floating disk rotate at the same time around the points O_1 and O_2, respectively. Consequently, when calculating the action exerted by the two free sliding masses onto the disk itself not only the *relative acceleration*, but also the "*carrying*" and the "*Coriolis*"-*acceleration* are to be taken into account. Let us now consider the masses located in the rotation centres for analyzing their relative motion in the guide-way.

Let it be assumed that the position of the two free sliding masses in their respective guide ways should be regulated by means of spring supports: to be exact the floating disk, together with the free sliding masses, should be statically balanced in such a way that the gravity centre m_1 of the mass in guide-way

I and also the gravity centre m_2 of the mass in guide-way *II* should both coincide with the gravity centre *m* of the floating disk. (Of course, when the two shaft-axes are aligned, no relative displacement occurs.) The initial position of the rotating clutch according to Fig. 27a is characterized by the eccentricity

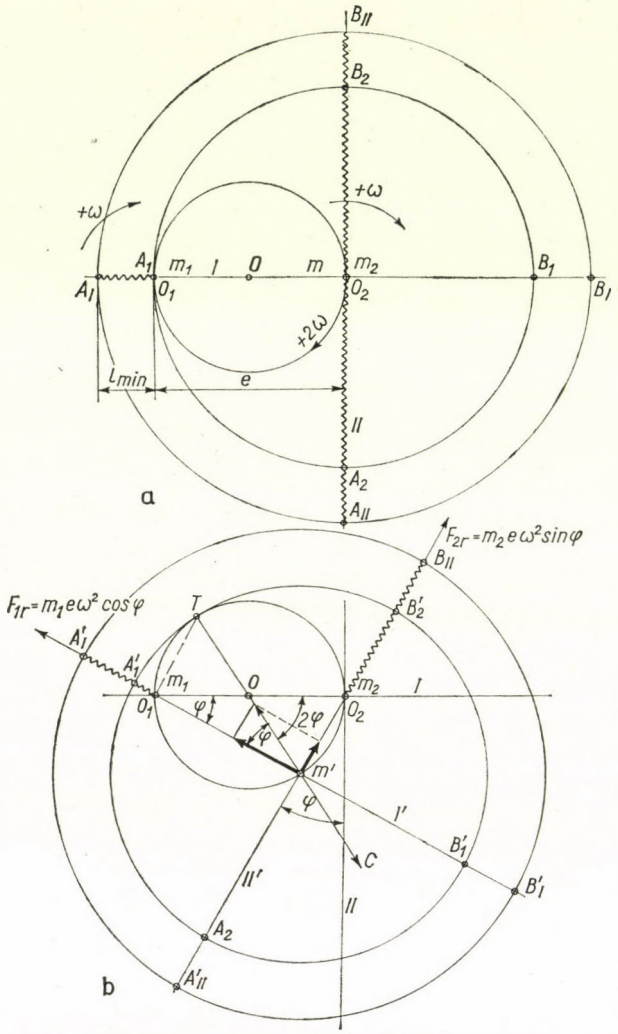


Fig. 27. Play of force of a pair of free-sliding masses in the floating member: Starting position *m* mass of the floating member, m_1 mass free-sliding in guide-way *I*, m_2 mass free-sliding in guide-way *II*, *e* eccentricity; wavy lines indicate the distances being proportional to the momentaneous length of springs *(a)*; relative displacement of the masses free-sliding in the guide-ways of the floating member *m'* denotes the mass-centre-point of the floating member in the position after being displaced, the forces F_{1r} and F_{2r} induced by the relative acceleration act in the direction of the guideway CL *(b)*

$O_1 O_2 = e$ of the two parallel shafts, and by the initial position of guide way I in the plane of the shaft-axes.

As mentioned above, the mass centre point m of the floating disk falls always onto that axis, the guide-way of which is just perpendicular to the axial plane; viz. in the initial position as delineated $m \equiv O_2$. Therefore, the length of the two spring supports of the guide-way II is the same: $\overline{A_{II}O_2} = \overline{B_{II}O_2}$. Again, in guide-way I — according to our assumption and to our motivation herebelow — the length of the supporting spring at one side is:

$$A_I O_1 = l_{\min},$$

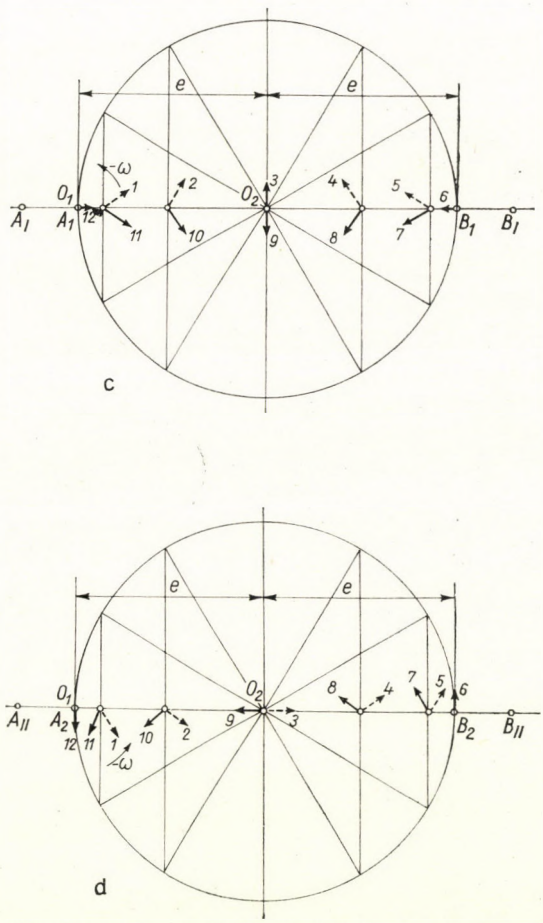


Fig. 27. According to the observation made by the imaginary traveller of the guide-way I, the arrow, that is fixed at point O_1 of space, rotates at an angular velocity ω, and performs a swinging amplitude of $2e$ *(c)*; according to the observation made by the imaginary traveller of the guideway II, the arrow that is fixed at point O_2 of space and shows in the same special direction, rotates at an angular velocity $-\omega$ at a phase difference of 90°, and performs a swinging amplitude of $2e$ *(d)*

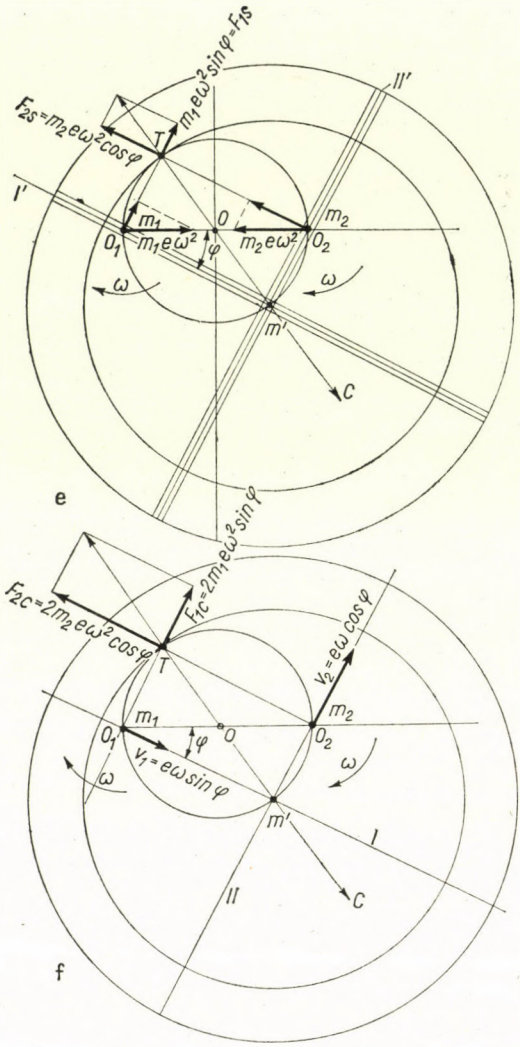


Fig. 27. Induced by the centripetal acceleration, the carrying forces F_{1s} and F_{2s} act in a perpendicular direction onto the guide-way walls *(e)*; induced by the Coriolis acceleration, the forces F_{1c} and F_{2c} act in perpendicular direction onto the guide-way walls *(f)*

and on the other side we find

$$O_1 B_I = 2e + l_{\min} = l_{\max}.$$

When the shaft made an angular displacement of the angle φ, and guide-way I comes from its initial position into the new position I' (Fig. 27b), the *relative displacement* of O_1 in the guide way is:

$$s_1 = A_1' O_1 = A_1' m' - O_1 m' = e - m' T \cos\varphi = e - e\cos\varphi = e(1 - \cos\varphi). \quad (26)$$

The relative velocity is:

$$v_1 = \frac{ds_1}{dt} = e \sin \varphi \frac{d\varphi}{dt} = e\omega \sin \varphi ; \tag{27}$$

the relative acceleration is

$$a_1 = \frac{dv_1}{dt} = e\omega \cos \varphi \frac{d\varphi}{dt} = e\omega^2 \cos \varphi . \tag{28}$$

Again, the point A'_{II} of the guide-way *II* assumes the position in A'_I after an angular displacement of 90°, thus the respective equations will be similar to the former ones, except that instead of φ we have to write $(90° - \varphi)$, namely:

$$s_2 = e\,[1 - \cos (90° - \varphi)] = e\,(1 - \sin \varphi) = B'_2 O_2 , \tag{26a}$$

$$v_2 = e\omega \cos \varphi , \tag{27a}$$

$$a_2 = e\omega^2 \sin \varphi . \tag{28a}$$

Besides, we find delineated in the diagram of Fig. 27b the direction of the relative forces

$$F_{1r} = m_1 \, a_1 = m_1 \, e\omega^2 \cos \varphi$$

and

$$F_{2r} = m_2 \, a_2 = m_2 \, e\omega^2 \sin \varphi .$$

This forces, generated by the relative accelerations, act in the direction of the centre line of the guide ways, showing away from point m'.

Let us imagine, located at point O_1 of space, a road sign arrow showing the direction towards point O_2 in space. The traveller of the guide-way *I*, who is aware of the rotating movement of his guide-way, but does not know that it performs, beside rotating, a linear shifting in relation to O_1, will observe that the arrow performs in the guide way an oscillation having the amplitude $2e$ by rotating, at the same time, at an angular speed of $-\omega$ viz. just *in the opposite sense* as compared to the shafts (Fig. 27c). Again, the other traveller in the guide way *II* will observe that the arrow in O_2 shows the same direction and will rotate at the angular speed equalling $-\omega$ and performs the alternating motion at a phases difference of 90°. (Fig. 27d).

Since the mass m_1 rotates around O_2, the respective *carrying acceleration* is equal to the centripetal acceleration $e\,\omega^2$, pointing to O_2; for the mass m_2, which rotates around O_1, the value of the carrying acceleration is the same, and its direction points to O_1. According to Fig. 27e, the carrying force acting on the wall of the guide-way *I*, is expressed:

$$F_{1s} = m_1 e\omega^2 \sin\varphi$$

and the force acting on the wall of the guide-way *II* is:

$$F_{2s} = m_2 e\omega^2 \cos\varphi .$$

As can be seen from the diagram, the direction of these component forces points away from point m'.

The *Coriolis accelerations*

$$\vec{a_{1c}} = 2\,\omega \times \vec{v_1}$$

and

$$\vec{a_{2c}} = 2\,\omega \times \vec{v_2}$$

are perpendicular to the guide-way as can be seen from the vectorial products in the above equations. These two vectors intersect each other at the tangential point T of the circles having the radii $e/2$ and e, and their respective direction points away from point m' (Fig. 27f).

Thus, the *Coriolis forces* are:

$$F_{1c} = 2\,m_1 e\omega^2 \sin\varphi ,$$

$$F_{2c} = 2\,m_2 e\omega^2 \cos\varphi .$$

According to our assumption, $m_1 = m_2 = m_0$ wherefrom we obtain the equations for the acting forces.

According to Fig. 27b, the resultant of the relative forces acting along the guide-ways *I* and *II* is:

$$F_r = \sqrt{F_{1r}^2 + F_{2r}^2} = \sqrt{m_0^2 e^2 \omega^4 (\cos^2\varphi + \sin^2\varphi)} = m_0 e\omega^2 .$$

The resultant of the carrying forces, according to Fig. 27e, is

$$F_s = \sqrt{F_{1s}^2 + F_{2s}^2} = \sqrt{m_0^2 e^2 \omega^4 (\sin^2\varphi + \cos^2\varphi)} = m_0 e\omega^2 .$$

The resultant of the *Coriolis forces* is, according to Fig. 27f:

$$F_c = \sqrt{F_{1c}^2 + F_{2c}^2} = \sqrt{4\,m_0^2 e^2\omega^4 (\sin^2\varphi + \cos^2\varphi)} = 2\,m_0 e\omega^2 .$$

The direction of all three resultants coincides with the radius OT, in the opposite sense to the centrifugal force of the floating disk; wherefrom

$$4\,m_0 e\omega^2 = m\frac{e}{2}(2\,\omega)^2 = 2\,me\omega^2 = C$$

viz.:

$$m_0 = \frac{m}{2} . \tag{29}$$

As can be derived from the above relationship, the two free sliding masses always form a counterbalance to the centrifugal force of the floating disk, independently of the variables φ, ω and e. The mass of the counterweights equals the mass of the floating disk, in the same way, as in the case discussed in Chapter 1.

According to the knowledge of the present writer, nowhere in the relevant

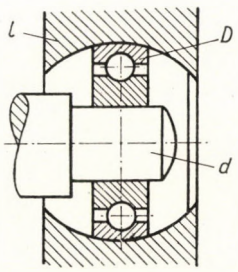


Fig. 28. Cylindrical guide-way with a rigid ball bearing located in a race-way having a spherical fitting profile

literature was a mechanical clutch type ever mentioned that is suitable not only to damp, but also to isolate the shaft vibrations in the case of coaxial or parallel roll shafts.*

3. Roller type movable clutch with cylindrical guide ways

As the simplest solution, a cylindrical guide-way presents itself combined with rollers having spherical surfaces, especially when the radius of the sphere equals the radius of the cylinder. This kind of contrivance can be made with simple rollers, or with ball bearings fitted with a spherical rigid nest (Fig. 28). Nevertheless, self-aligning roller bearings are far more preferred by the author for the following reasons: this type of bearings provide for our purposes of dimensioning a dependable basis, and because the movement around the pin and the rotation around the centre line of the guide way takes place on the surfaces of the bearing itself. In spite of the fact that the number of degrees of freedom is one more than the normal, it is easy to ensure that the plane

*This investion of the author has been classified by the *National Council of Trade Unions*, and by the *National Office of Investions* among the 10 bests investions that emerged Hungary during the last years.

of the outer ring should be kept perpendicular to the pin axis, namely, by an adequate guiding of the outer ring.

Again, there we know a special feature of the cylindric guide way, namely, when we intend to multiply the linear contact line between the roller and the guide way in the case of higher loadings (Fig. 29a and 29b).

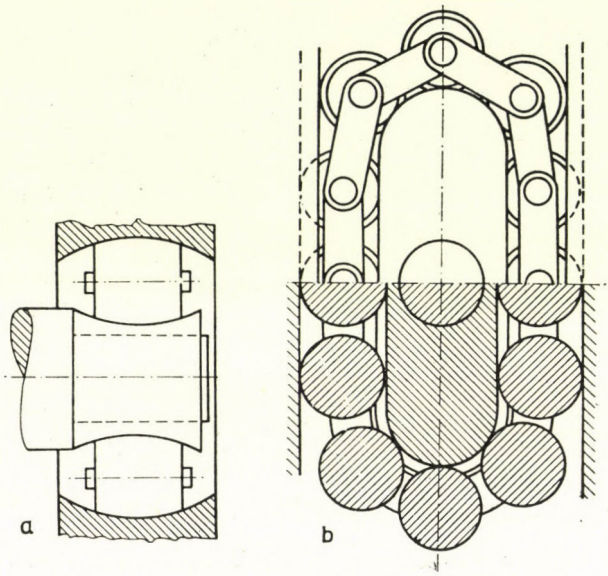


Fig. 29. Roller chain type movable clutch: Barrel type roller *(a)*; roller chain *(b)*

To this end, a closed chain is composed of barrel type rollers mounted in a guide way according to a design in which the cylinder-radius of the guide way equals the profile-radius of the rollers. The inner race-way of the chain of barrel type rollers can rotate around the pin, by which all the rollers within the chain are loaded that happened to run on the cylindric surface of the inner race-way.

REFERENCES

1. Szőke, B.: Beitrag zur Entstehung von technischen Irrtümern. *Acta Techn. Hung.* **39** (1962), 3—27.
2. Demidov, A. R.: Malomipari hengerszékek (Roller Mills). Élelmezési minisztérium Könyv és Lapkiadó Vállalata, Budapest 1951.
3. Szőke, B.: The Dynamic of Roller Mills. *Act. Techn. Hung.* **33** (1961), 65—85.
4. Szőke, B.: Malomhengerszékek munkarés beállítójának egyik szokásos szerkezeti hibája (A common error in the design of the working gap adjusting device of roll mills). *Malomipar* (1957), 9—13.
5. Szőke, B.: Az őrlőhenger hajtásának néhány hibaforrása (Some sources of errors occurring in the design of grinding roll drive). *Malomipar* (1958), 17—19.
6. Moog, O.: Der Antrieb der Walzenstühle und seine Berechnung. Hugo Matthaes Verlag, Stuttgart 1952.

7. KOLOSTORI, J.: Energiaviszonyok a hengerszéknél, az energiavisszatáplálás különböző megoldásai (Energy conditions on roll mills, various methods to solve the problem of energy feed-back). *Malomipar és Terményforgalom* (1963), 382—393.
8. PRATIQUE, J.: Contribution à l'energétique des machines à cylindres. *Bulletin des anciens élèves de l'école française de moulinerie* (1964), 169—175.
9. SZŐKE, B.: Should Grinding Rolls of Different Diameters be Operated Together? *Acta Techn. Hung.* **24** (1959), 3—44.
10. SZŐKE, B.: Járassunk-e együtt különböző átmérőjű őrlőhengereket (Should Grinding Rolls of Different Diameters be Operated together)? Mérnöki Továbbképző Intézet, Budapest 1959, No 9—501.
11. SZŐKE, B.: Behúzási viszonyok hengerjáratoknál: kúpszeletek őrlés-technológiai értelmezése (Drawing-in conditions in roll mills conical sections interpreted in the special scope of grinding technique). *Malomipar és Terményforgalom* (1963), 294—300.
12. SZŐKE, B.: Bevezetés a forgácsoló szerszámok geometriájába. A gépgyártás geometriai szemléletének alapjai (Introduction into the geometry of metal cutting tools. Basic geometric principles for machine-production engineering). Műszaki Könyvkiadó, Budapest 1966.
13. SZŐKE, B.: Self-locking of the Sliding Surface-Bearings. *Acta Techn. Hung.* **22** (1958), 329—347.
14. SZŐKE, B.: Csúszócsapágy önzáródása (Self-locking of sliding surface-bearings). *Gép* **8** (1956), 113—119.
15. HOPF, L.: Mühlentechnisches Praktikum (Band II, Mühlenbau). Hugo Mathaes Verlag, Stuttgart 1952.
16. MAROS, D.: Meddő teljesítmény a fogaskerék hajtóművekben (Idle output in toothed gears). *Gép* **12** (1960), 265—270.
17. SZŐKE, B.: One Way to Cut down Power Requirement for Flour Milling. *Acta Techn. Hung.* **31** (1960), 13—17.
18. A *Gépipari Tudományos Egyesület* megbízásából szerkesztett Műszaki Nagyjaink életrajzsorozat ("Our Great Men in Technics" A reference of biographic monographies): PÉNZES I.: MECHWART A. In manuscript 1965.
19. SZŐKE, B.—KOLOSTORI, I.: Őrlőberendezés hengersorokkal (Grinding equipment of corresponding roll mills. Employee's invention). *Malomipar és Terményforgalom* (1965), 52—54.
20. KOLOSTORI, J. and SZŐKE, B.: Őrlőberendezés hengersorokkal (Grinding equipment of compound roll mills. Employee's invention). *Malomipar és Terményforgalom* (1965), 49—52.
21. VÖRÖS, I.: Gépelemek, I. kötet (Machine Parts I, Textbook). Tankönyvkiadó, Budapest 1962.
22. SZŐKE, B.: Lamellástengelykapcsoló (Laminated Coupling) Patent Specification 137—537, 1946.
23. SZŐKE, B.: A mozgékony tengelykapcsoló néhány problémája (Some design problems on the movable clutch); *In manuscript.*
24. SZŐKE, B.: A mozgékony tengelykapcsolás áttekintése (General report on the application of movable clutches). *Malomipar és Terményforgalom* (1963), 128—135.
25. SZŐKE, B.—RITTER, E.: Önbeálló tengelykapcsoló (Self aligning clutch). Patent Specification 143055, 1954.
26. SZŐKE, B.: Önbeálló tengelykapcsoló (Self aligning clutch). Patent Specification SO-449, 1955.
27. *SKF* Golyós és görgőscsapágyak 6400 sz. főkatalógusa (Main Catalogue of Rolling Bearings SKF). Budapest 1964 and 1960.
28. *Die Kugellager-Zeitschrift Budapest* (1960); No. 143, 1—7.
29. SZŐKE, B.: A hengerszékhajtás szerepe az őrlési energiagazdálkodásban (Importance of the drive of roll mills as regards energy-economy in the mill trade). *Malomipar* (1960), 48—53.
30. SZŐKE, B.: A kardáncsuklós hajtás működésének gyakorlati elemzése (Practical Analysis of the functioning of cardan joint drive), *Malomipar és Terményforgalom* (1963), 447—448.
31. SZŐKE, B.—KOLOSTORI, I.: Görgős mozgékony tengelykapcsoló (Roller type movable clutch). Employee's invention 152100). *Malomipar* (1965), 175—182.
32. TERPLÁN, Z.: Mechanizmusok (Engineering of Mechanisms, Textbook). Tankönyvkiadó, Budapest 1962.
33. KOLOSTORI, J.—HAMZA, B.: Bestimmung und Bedeutung des Betriebs-Mahlspaltes von Walzenpaaren. *Die Mühle* **103** (1966); 103—105, 119—120.

Zweckmäßige Leitung des Energieflusses; die bewegliche, selbstausgleichende Wellenkupplung. Auf Grund theoretischer Überlegungen gibt hier Verfasser die Ableitung einer Formel zur Erklärung der Erscheinung des Energie-Rückflusses in Walzmühlen, deren Erforschung bis nun in der Fachliteratur nicht zu finden war. Laut dieser Erklärung ist es verständlich, weshalb der Rückfluß verschwindet, falls die Übersetzung die kritischen Werte 1 : 1 oder ∞ : 1 annimmt. Es wird gezeigt, daß bei richtiger Anordnung der Reihenkupplung von Walzen verschiedenen Durchmessers das Übersetzungsgetriebe durch den Energierückfluß nicht mehr überlastet wird, da der Rückfluß in geschlossener Weise aus dem System der Arbeitsmaschinen nicht heraustritt. Hiedurch bleiben die Antriebselemente von der Überbelastung verschont (Zahnräder, Wellen, Kupplungen). Die Weiterentwicklung betrifft die selbstbalancierende bewegliche Wellenkupplung. Aus den theoretischen Auslegungen geht es hervor, daß die neue Wellenkupplung als ein neuartiges Maschinenelement zu bewerten ist, da durch sie die Schwingungen nicht gedämpft, sondern isoliert werden.

Conduite efficace du flux d'énergie et accouplement auto-équilibrant mobile. A la base de considérations purement théoriques, l'auteur déduit une formule de la puissance retransmise dans un moulin à cylindres, pouvant expliquer pourquoi cette puissance disparaît quand le rapport de transmission du couplage positif s'approche de 1 : 1 ou ∞ : 1. L'auteur présente un groupement du couplage en série des cylindres de diamètres différents, où le flux d'énergie retransmis ne charge pas la transmission mais circule à l'intérieur du système, ce qui diminue considérablement la charge de tous les éléments de commande (roues dentées et accouplements). L'auteur s'occupe en outre du problème de l'accouplement auto-équilibrant mobile. Des considérations théoriques l'amènent à démontrer que cet accouplement intervient non comme un amortisseur des mouvements entre l'unité entraînante et l'unité entraînée, mais comme un nouvel élément de machine isolant les vibrations.

Целесообразное проведение потока энергии и самоуравновешивающаяся подвижная муфта сцепления *(Б. Сёке)*. Автор на основе чисто теоретических соображений выводит такие формулы потока энергии вальцовой мельничной машины по возвратной энергии, исследование которых в технической литературе до сих пор было безуспешным и которые дают ответ на вопрос, почему исчезает мощность в том случае, когда передачи принудительного сцепления приближаются к отношению 1:1 или ∞:1. Автор показывает такую группировку последовательного соединения вальцов различного диаметра, при которой возвратный энергопоток нагружает не модифицирующий привод, а циркулирует внутри самой системы и вследствие этого все элементы привода (зубчатые колеса, муфты сцепления) работают при значительно меньшей нагрузке. Далее рассматривается вопрос самоуравнивающейся подвижной муфты сцепления и автор, посредством теоретических выкладок показывает, что данная муфта сцепления между ведомыми и ведущими блоками машин работает не в качестве демпфера движения, а в качестве нового элемента машины — изолятора вибраций.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (3—4), pp. 351—370 (1968)

DISTRIBUTION OF STRESSES ACTING ON THE ROLL SURFACE IN STRIP ROLLING

PUBLICATION NO. 48 OF THE METALLURGICAL WORKING GROUP OF THE HUNGARIAN ACADEMY OF SCIENCES

L. TÓTH

UNIVERSITY OF HEAVY INDUSTRY, MISKOLC (HUNGARY)

[Manuscript received April 13, 1967]

The first part of this paper contains the description of the basic equations of rolling; the second part presents the solution of these basic equations. The approximate solution is based on knowledge of the enlargement of the pieces during rolling. The solution contains the flow diagram and the determination of the compression stresses acting on the surface of the roll. The correctness of the results is proved by comparing them with values determined by experiment.

List of symbols

b	[mm]	width of the rolled material;
b_s	[mm]	width in the neutral plane of the material in the rolling gap;
F_2	[mm²]	area of the final section of the rolled piece;
h	[mm]	thickness of the material in the rolling gap;
h_1	[mm]	thickness of the piece before being rolled;
h_2	[mm]	thickness of the ready rolled piece;
h	[mm]	width in the neutral plane of the piece;
k_f		yield stress of the rolled material;
l	[mm]	height of the upsetting specimen during the test;
l_1	[mm]	original height of the upsetting specimen;
l_d	[mm]	arc of contact;
p	[kp/mm²]	radial pressure stress acting on the roll surface;
R	[mm]	the radius of roll;
v_x, v_y, v_z	[mm/sec]	velocity of the elementary volumen $dx \cdot dy \cdot h$ during rolling;
v_{rx}	[mm/sec]	relative velocity in the x direction of the elementary volume $dx \cdot dy \cdot h$ in relation to the roll;
x_k	[mm/sec]	surface velocity of the roll;
α, y, z	[mm]	coordinates;
γ		angle between the direction of sliding and the x axis;
v		angle by which the extension of the elastic zone near the neutral section is determined;
ε_l		stress;
$\varepsilon_{r,d}$		strain of upsetting;
$\varepsilon_{r,h}$		strain of rolling;
ε_x, ε_y, ε_z		strain of the elementary volume $dx \cdot dy \cdot h$;
$\dot{\varepsilon}_x$, $\dot{\varepsilon}_y$, $\dot{\varepsilon}_z$		strain rate values of the elementary volume $dx \cdot dy \cdot h$;
φ		angle of turn of the roll;
φ_s		angle of turn of the roll up to the neutral plane;
λ		proportion factor expressed as the quotient of stress per stress rate;
μ		coefficient of friction;
σ_x, τ_{zy}	[kp/mm²]	stresses acting on the tested part of the space in the direction of the
σ_y, τ_{xz}		coordinate axes;
σ_z, τ_{xy}		
$\sigma_{x\,el}$. $\sigma_{y\,el}$.	[kp/mm²]	elastic stresses acting onto the elementary volume $dx \cdot dy \cdot h$;
τ	[kp/mm²]	shear stress acting onto the roll surface;
τ_{max}	[kp/mm²]	yield shear stress;
ξ, ψ		proportion factors, expressed as the quotient yield stress per stress.

I. Introduction

The relationship between the influencing factors and the attendant phenomena of rolling can be observed by experimental, or determined by theoretical, methods. Experiments are suitable for a direct specification of the mentioned relationship. By theoretical methods it is not possible to find a direct determination.

Knowledge of the magnitude and distribution of the stresses acting onto the surface of roll is the most important to these investigations. Therefore, our first step is to determine the stresses acting on the roll surface as functions of the influencing factors. On the basis of the stress values, most of the attendant phenomena, e.g. rolling force, rolling torque, advance of the rolled piece etc., can be calculated as functions of the influencing factors.

In the dissertation [2], the contents of which are made known in this paper, the main task was to assess the stresses acting on the roll surface.

II. Basic equations of rolling

The stresses acting on the roll surface can be assessed on the basis of the fundamental relationship expressed by means of the mathematical theory of plasticity. Since the factors by which the rolling process is influenced are not always the same, as the theoretical factors known from plasticity theorems, the relationship between these two groups of factors should be elucidated by means of metallographic knowledge (for the determination of the yield stress) and by the mechanical theory of friction (for the determination of the factor of friction).

According to the above-mentioned recognition, we can consider the basic equations of plasticity as the basic equations of rolling, of course, with the addition of methods of the assessment of yield-stress and coefficient of friction.

In order to make calculations easy, some assumptions should be set up as follows:

a) The elementary volume of the rolled piece is expressed as $dx \cdot dy \cdot h$. But, this can be considered as an acceptable approximation only when the ratio l_d/h_{av} is large enough. Yet, as can be seen in Fig. 1, these conditions cannot be applied to the region of the elastic zone in the vicinity of the neutral plane. Therefore, the stresses acting onto the roll surface will be determined only in the zones II as marked in Fig. 1. On the basis of the known neutral section, the boundaries of zone II can be specified by means of the angle γ [2, 3]. In other words, our assumption involves a uniaxial deformation in the zones II. On this basis, the outer frictional stresses acting on the outer surface

of the piece can be considered and be dealt with as forces acting on the elementary volume $dx \cdot dy \cdot h$.

b) As far as the present investigation is concerned, the elastic deformation of the roll can be neglected.

c) Along the whole width of the rolled piece, the velocity of the portion at the entry into, and at the exit out of, the rolling gap is constant. In the rolling direction as is proved by experiments, the velocity of the material within the gap can be assumed to be constant with a high grade of approximation.

d) The velocity conditions as described under *c)* are the result of the forcing action of elastic material parts before, and after, the rolling gap.

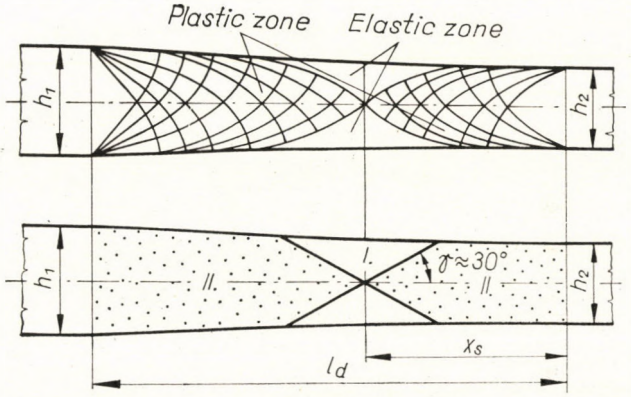


Fig. 1

Another feature of this appears as the action exerted by the elastic stresses on the material parts in the gap at the entry, and exit, sections.

On the basis of the stress values acting on the entry and exit sections, we are able to calculate the elastic stresses acting within the gap.

The distribution of stresses in the entry and exit sections as shown in Fig. 2, are values assumed on the basis of experiments carried out [2, 4]. The elastic stresses resulting in the gap can be calculated by means of elasticity mathematics.

1. The theorem of volume-constancy

On the condition, that the density of the material does not change, the law of conservation of matter can be expressed by stating that during the formation process no change of any elementary volume of the tested plastic body can occur; the respective equation is [5]:

$$\frac{\partial v_x}{\partial x} = \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} = 0. \tag{1}$$

2. *Equations of equilibrium*

Instead of the equilibrium equations generally adopted in the field of mechanics, we shall take as basis of our calculation the equilibrium of the elementary force components acting on the elementary volume as delineated in Fig. 3, in order to deduce the equilibrium equations of rolling. These are:

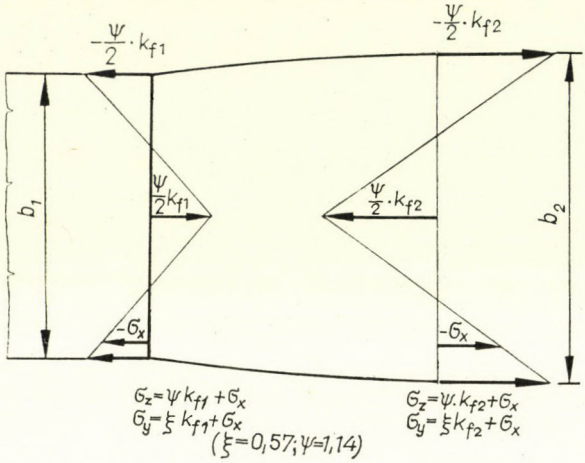


Fig. 2

$$\frac{\partial}{\partial x}(h\,\vartheta_x) = 2\,(\pm p\sin\varphi + \tau_x)\frac{1}{\cos\varphi} + \frac{\partial}{\partial x}(h\,\sigma_{x\,el.})\,, \tag{2}$$

$$\frac{\partial}{\partial y}(h\,\sigma_y) = -\,2\,\tau_y + \frac{\partial}{\partial y}(h\,\sigma_{y\,el.})\,; \tag{3}$$

$$\sigma_z = (p\cos\varphi \mp \tau_z)\frac{1}{\cos\varphi}\,. \tag{4}$$

The relationship between the stresses τ_x, τ_y and τ is expressed, according to Fig. 4, as follows:

$$\tau_x = \frac{\tau}{\sqrt{\tan^2\varphi + \dfrac{1}{\cos^2\alpha}}}\,; \quad (\tau_z = \tau_x\tan^2\varphi)\,, \tag{5}$$

$$\tau_y = \frac{\tau\cdot\tan\alpha}{\sqrt{\tan^2\varphi + \dfrac{1}{\cos^2\alpha}}}\,. \tag{6}$$

In equations (5) and (6), τ denotes the shear stress. On the basis of Coulomb's law, we can write:

$$\tau = \mu p\,. \tag{7}$$

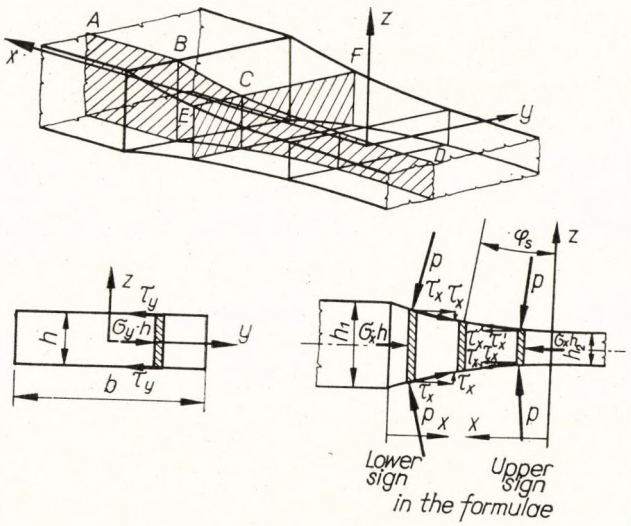


Fig. 3

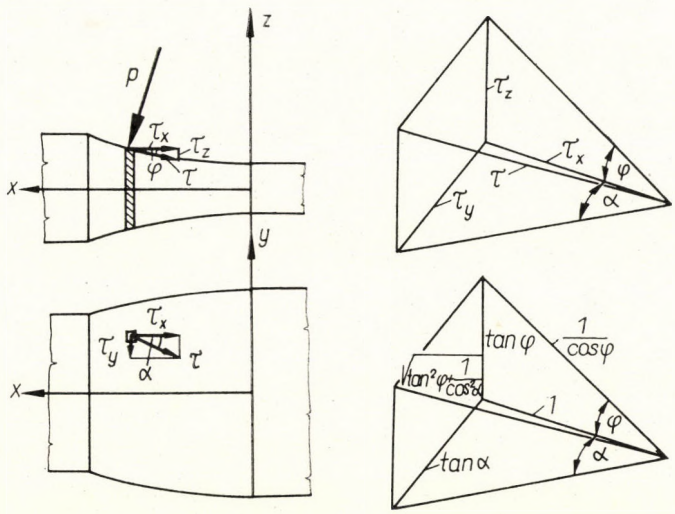


Fig. 4

When the product μp is larger than the so-called internal shear stress, sliding will occur inside the matter, and τ will assume the value:

$$\tau = \tau_{\max} = \frac{\psi k_f}{2} . \tag{8}$$

From equations (2), (3), (4), (5), (6) and (7) two equations, in which only the stress σ_x, σ_y and σ_z are contained can be deduced as expressions valid

within the sliding zone:

$$\frac{\partial \sigma_x}{\partial x} = \mp \frac{2\sigma_x}{h} \cdot \tan\varphi + \frac{2\sigma_z}{h} \cdot \frac{\pm \tan\varphi + \dfrac{\mu}{\cos\varphi \sqrt{\tan^2\varphi + 1/\cos^2\alpha}}}{1 \mp \dfrac{\mu \tan\varphi}{\cos\varphi \sqrt{\tan^2\varphi + 1/\cos^2\alpha}}} + \tag{9}$$

$$+ \frac{1}{h} \cdot \frac{\partial}{\partial x_x} (h \sigma_x \sigma_{x\ el.}),$$

$$\frac{\partial \sigma_y}{\partial y} = - \frac{2\sigma_z}{h} \cdot \frac{\dfrac{\mu \cdot \tan\alpha}{\sqrt{\tan^2\varphi + 1/\cos^2\alpha}}}{1 \mp \dfrac{\mu \cdot \tan\varphi}{\cos\varphi \sqrt{\tan^2\varphi + 1/\cos^2\alpha}}} + \frac{\partial \sigma_{y\,el.}}{\partial y} . \tag{10}$$

By taking into account that, when rolling thin strips, the angle φ is, in general, of a minute value, and the change of elastic stresses in the x direction can be neglected, the equations (9) and (10) can be written in a simplified form, viz.:

$$\frac{\partial \sigma_x}{\partial x} \cong \pm (\sigma_z - \sigma_x) \frac{2}{h} \cdot \tan\varphi + \frac{2\mu}{h} \sigma_z \cos\alpha , \tag{11}$$

where, the sign + refers to the advance zone, and the sign — refers to the lagging behind zone

$$\frac{\partial \sigma_y}{\partial y} \cong - \sigma_z \frac{2\mu}{h} \sin\alpha + \frac{\partial \sigma_{y\,el.}}{\partial y} . \tag{12}$$

From equations (2), (3), (4), (5), (6) and (8), equations for the determination of stresses can be deduced as expressions valid in the zone of adhesion:

$$\frac{\partial \sigma_x}{\partial x} = \pm (\sigma_z - \sigma_x) \frac{2 \tan\varphi}{h} +$$

$$+ \psi k_f \frac{1 + \tan^2\varphi}{2h \cos\psi \sqrt{\tan^2\varphi + \dfrac{1}{\cos^2\alpha}}} + \frac{1}{h} \frac{\partial}{\partial x_x} (h \sigma_{x\,el.}), \tag{13}$$

$$\frac{\partial \sigma_y}{\partial y} = - \frac{2}{h} \cdot \frac{\tan\alpha}{\sqrt{\tan^2\varphi + \dfrac{1}{\cos^2\alpha}}} \cdot \frac{\psi k_k}{2} + \frac{\partial \sigma_{y\,el.}}{\partial y} . \tag{14}$$

In an analogous way as above, when φ assumes a small value, equations

(13) and (14) can be simplified, as follows:

$$\frac{\partial \sigma_x}{\partial x_x} \approx \pm \frac{2 \tan \varphi}{h} (\sigma_z - \sigma_x) + \frac{\psi \cdot k_f}{2 h} \cos \alpha , \tag{15}$$

$$\frac{\partial \sigma_y}{\partial y} \approx - \frac{\psi k_f}{h} \sin \alpha + \frac{\partial \sigma_{y\,el.}}{\partial y_y} . \tag{16}$$

In equations (9), (10), (11), (12), (13), (14), (15) and (16), the angle α represents the direction of the frictional stress. Since the direction of the frictional stress is always opposite to the direction of the displacement between the two surfaces, the expression of the angle α can be written, taking into consideration of the velocity conditions as shown in Fig. 5, as follows:

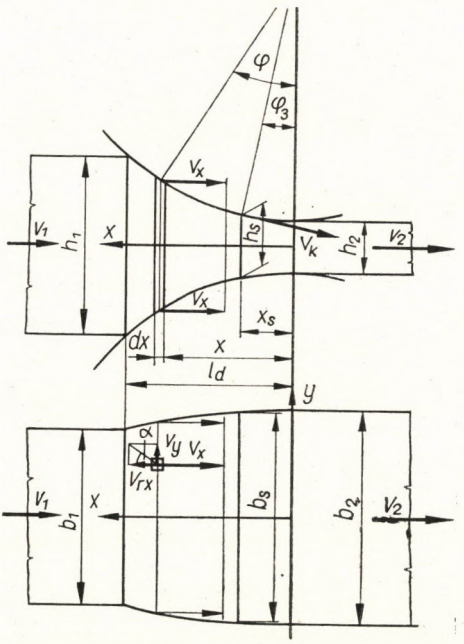


Fig. 5

$$\tan \alpha = \frac{v_y}{v_{r,x}} . \tag{17}$$

3. *Plasticity condition*

Assuming an uniaxial load of metals the plasticity condition can be expressed in the following form:

$$\sigma_1 = k_f (\varepsilon_1) . \tag{18}$$

Here σ_1 denotes the uniaxial stress, tensile or compressive, and k_f is the yield stress of the metal, the value of which is, in a cold forming process, a function only of the specific deformation ε_1. When, instead of an uniaxial load, whatever kind of load acts on any elementary volume of the metal, the condition of the plastic state occurring is expressed, in compliance with the Huber—Mises—Hencky-theory [1], by the following equation:

$$(\sigma_z - \sigma_x)^2 + (\sigma_z - \sigma_y)^2 + (\sigma_y - \sigma_x)^2 + 2(\tau_{xy}^2 + \tau_{xz}^2 + \tau_{yz}^2) = 2k_f^2(\varepsilon_r). \quad (19)$$

Now, taking into account the above condition stated under *a)* the equation (19) can be written in the following form:

$$(\sigma_z - \sigma_x)^2 + (\sigma_z - \sigma_y)^2 + (\sigma_y - \sigma_x)^2 = 2k_f^2(\varepsilon_r). \quad (20)$$

In equations (19) and (20) ε_r represents the so called reduced strain. According to Ros and Eichinger [6], the value of ε_r is:

$$\varepsilon_r^2 = \frac{2}{9}[(\varepsilon_z - \varepsilon_y)^2 + (\varepsilon_z - \varepsilon_x)^2 + (\varepsilon_y - \varepsilon_x)^2] \quad (21)$$

(namely when the condition of the constant stress rate during the formation process as construed under *a)* is strictly adhered to).

When the forming operation consists only of a simple upsetting, then

$$\varepsilon_x = \varepsilon_y = -\varepsilon_z/2;$$

taking this into account the reduced strain can be deduced, in the case of upsetting from (21):

$$\varepsilon_{r,d} = \varepsilon_z = \frac{l_1 - l}{l_1}. \quad (22)$$

Again, for rolling we have:

$$\varepsilon_x = -\varepsilon_z, \quad \varepsilon_y \cong 0.$$

Thus, from equation (21), the value of the reduced strain for the case of a rolling process can be written as:

$$\varepsilon_{r,d} = \sqrt{\frac{4}{3}}\varepsilon_z = 1{,}15\frac{h_1 - h}{h_1}. \quad (23)$$

As shown by equation (22) and (23), the difference between upsetting and rolling should be taken into account, when the yield stress which occurs

in the rolling operation should be determined. Let the decrease of height be assumed as the same for both upsetting and rolling, the value of reduced strain for rolling will be by 15% larger than it would be for pure upsetting. In other words, the yield stress is also for rolling larger than its value for upsetting.

4. Relationship between stresses and the parameters of deformation

When equation (20) is fulfilled, then the elementary volume of the piece reaches, under the influence of stresses, the state of flow.

Again, the characteristic of plastic flow is the strain rate. The relationship between the strain rate and the stresses can be expressed, according to PRANDTL and REUSS [5] as follows:

$$\frac{\partial v_x/\partial x}{\sigma_x - \sigma_{av}} = \frac{\partial v_y/\partial y}{\sigma_y - \sigma_{av}} = \frac{\partial v_z/\partial z}{\sigma_z - \sigma_{av}} = \lambda . \tag{24}$$

In this formula

$$\sigma_{av} = \frac{1}{3}(\sigma_x + \sigma_y + \sigma_z),$$

and λ assumes at any point of the plastic piece some positive value (it does not represent a matter constant). For the rolling of thin strips, equation (24) takes the following form:

$$\lambda = \frac{\partial v_x/\partial x}{\partial v_y/\partial y} = \frac{2\,\sigma_x - \sigma_y - \sigma_z}{2\,\sigma_y - \sigma_x - \sigma_z} . \tag{25}$$

This equation can be applied to cold rolling only of materials which become hard through rolling. For a clear understanding of this statement, let us investigate, in accordance with the conception of DRUCKER [7], the case of an uniaxial formation. Fig. 6 represents the function stress versus deformation, for the case of a simple tensile test. In this diagram, σ^* is the value of the yield stress coordinated to the point of deformation ε and σ belongs to the point $\varepsilon + \dot{\varepsilon} \cdot \Delta t$. The hatched area in the diagram represents the external (rolling) work which is needed to perform the elementary deformation $\dot{\varepsilon} \cdot \Delta t$; its value is but positive, when no internal work (softening) occurs:

$$\frac{1}{2}(\sigma - \sigma^*) \cdot \dot{\varepsilon} \cdot \Delta t \geqq 0 \tag{26}$$

this equation is fulfilled only when

$$\dot{\varepsilon} = \lambda\sigma \tag{27}$$

holds true.

Since the formula (27) is the basis of the Prandtl—Reuss equation, it is obvious that equation (25) — like equation (27) — holds true only for hardening materials.

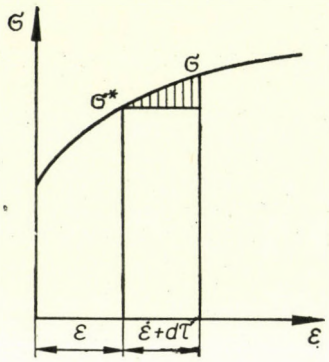


Fig. 6

III. Solution of the basic equations for rolling

The basic equations established in chapter II can be solved only on the conditions described in the same chapter. Now, by applying the so-called method of characteristics [2] to solve the partial differential equations, this method is suitable for finding the general solution by means of a numeric calculation. But, this way is very lengthy and cumbersome. Therefore, it seems to be more practicable to elaborate an approximative method on the basis of simple experiments.

1. Determination of the flow-diagram

Equation (17) serves to find the direction of the frictional stress. In relation to the surface of the roll, this direction is always opposite to the displacement direction of the tested point of the rolled piece. For the determination of the distribution of stresses, the knowledge of frictional stresses is indispensable. Again, the frictional stresses can be determined only when the angle α is known. Consequently, the stress distribution can be determined only with a known value of α.

Since the value $\tan \alpha$ is but the tangent of direction of some flow curve in the flow diagram, this latter can be delineated, too, when the function $\alpha = \alpha(x; y)$ is known.

On the basis of Fig. 5, and by taking into consideration the main conditions as laid down under *a)* and *b)*, the relative velocity $v_{r,x}$ can be deter-

mined:

$$v_{rx} = v_k \cos\varphi - v_x \tag{28}$$

where

$$v_x = v_k \frac{b_s h_s \cos\varphi_s}{bh} . \tag{29}$$

By fulfilling condition *b)*:

$$v_x = v_x(x) . \tag{30}$$

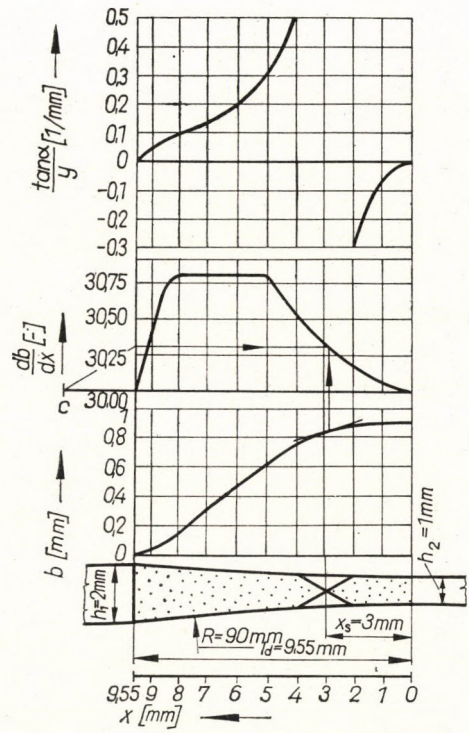


Fig. 7

By differentiating equation (30) we can state, that the strain rate $\dot{\varepsilon}_x$ is but a function of x:

$$\frac{\partial v_x}{\partial x} = \dot{\varepsilon}_x(x) . \tag{31}$$

By making use of the denotations as given in Fig. 7, and of the above given conception of the velocity of deformation $\dot{\varepsilon}_x$, this latter can be formulated as follows:

$$\frac{\partial v_z}{\partial z} = -\frac{v_z}{h/2} = \frac{v_x}{h} \cdot \frac{dh}{dx} . \tag{32}$$

On the basis of equation (1), $\dot{\varepsilon}_y$ is expressed:

$$\frac{\partial v_y}{\partial y} = -\frac{\partial v_x}{\partial x} - \frac{\partial v_z}{\partial z} . \tag{33}$$

By inserting the expressions (29) and (32) into equation (33), after performing the differentiation and the possible simplification, we obtain:

$$\frac{\partial v_y}{\partial y} = \frac{v_x}{b} \cdot \frac{db}{dx} . \tag{34}$$

As is seen from equation (34), also the strain rate $\dot{\varepsilon}_y$ is but a function of x. Thus, equation (34) can be integrated:

$$v_y = \frac{v_x}{b} \cdot \frac{db}{dx} y . \tag{35}$$

By making use of what was deduced above, and after having inserted equations (28), (29) and (35) into equation (12), and by making the possible simplification, we obtain the expression of the tangent of direction at some point of the flow curve as follows:

$$\tan \alpha = \frac{1}{\dfrac{bh \cos \varphi}{b_s h_s \cos \varphi_s} - 1} \cdot \frac{db}{dx} \cdot \frac{y}{b} . \tag{36}$$

By making use of equation (30) and when the function $b = b(x)$ is known, the value of $\tan \alpha$ can be determined at any point of the compressed surface. As a matter of principle, the equation of the flow curves can be set up by taking into account the relationship $\tan \alpha = dy/dx$. Again, we can delineate the flow diagram in a more simple way when we make use of the so-called "field of direction".

This method can be shown by observing the following example:

In Fig. 7, there are delineated, first, the function $b = b(x)$, and, second, the function db/dx obtained by differentiation. With these two, and by making use of the relationship $h = h_2 + x^2/2\,R$, the relationship $\tan \alpha/y$ can be determined on the basis of equation (36). In Fig. 8a the "field of direction" is delineated as it would be determined by means of Fig. 7; in the same figure, the system of flow curves, determined by a graphical method, is also set up, too. For comparison's sake, in Fig. 8b the flow diagram, which had been experimentally determined by SIEBEL and LUEG [8], is shown, too.

2. *Determination of the stresses acting on the compressed surface*

Once the flow diagram is delineated, the determination of stress distribution becomes simplified even when the method of the characteristic is applied. Further simplification is possible by observing the condition set up in point *b*.

The way of simplification is formed by making use of the following variables:

$$\xi = \frac{\sigma_y - \sigma_x}{k_f}, \quad \psi = \frac{\sigma_z - \sigma_x}{k_f}. \tag{37}$$

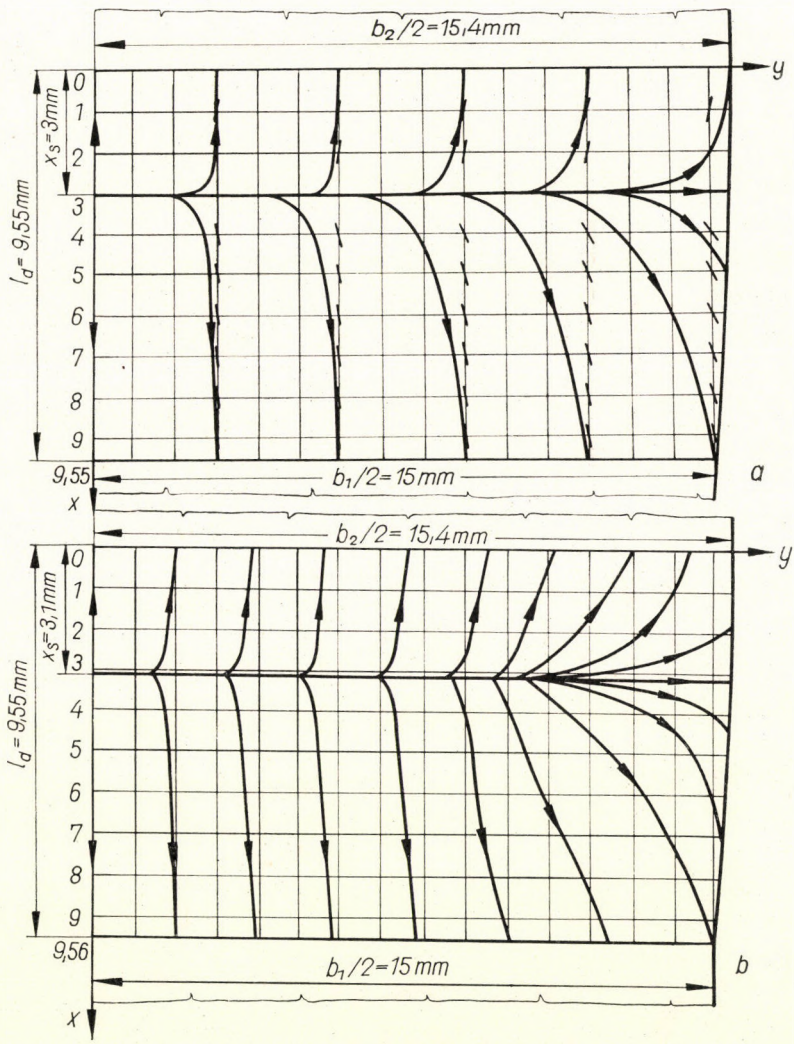


Fig. 8

With these, we obtain from equations (30) and (25)

$$\lambda = \frac{\psi - 2\xi}{\psi - \xi} \tag{38}$$

and

$$\xi^2 + \psi^2 - \xi\psi - 1 = 0\,. \tag{39}$$

From this system of (38) and (39), the variables and ψ can be determined. By substituting the given definition of the variables ξ and ψ, the system (37) can be solved:

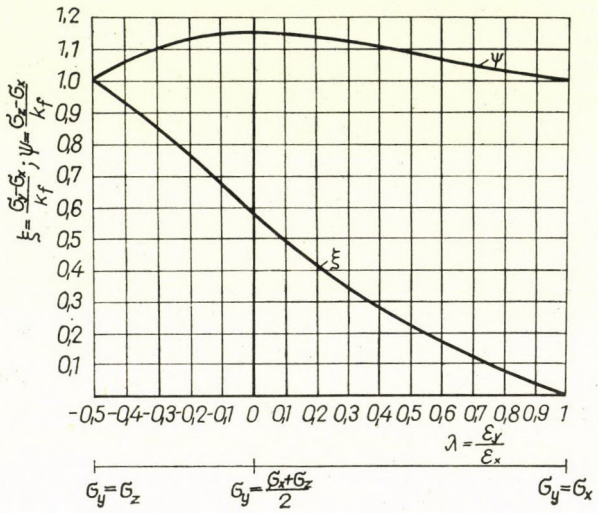


Fig. 9

$$\psi = \frac{\sigma_z - \sigma_x}{k_f} = \frac{1}{\sqrt{1 - \dfrac{1-\lambda}{2+\lambda} + \left(\dfrac{1-\lambda}{2+\lambda}\right)^2}} \tag{40}$$

and

$$\xi = \frac{\sigma_y - \sigma_x}{k_f} = \frac{\dfrac{1-\lambda}{2+\lambda}}{\sqrt{1 - \dfrac{1-\lambda}{2+\lambda} + \left(\dfrac{1-\lambda}{2+\lambda}\right)^2}}\,. \tag{41}$$

Curves representing equations (40) and (41) are shown in Fig. 9. At a given value of λ or ξ we can directly read the value of

$$\psi = (\sigma_z - \sigma_x)/k_f$$

which then serves to determine the distribution of stress.

In Fig. 10, we can see the value of λ as the quotient of the quantities $\dot{\varepsilon}_x$ and $\dot{\varepsilon}_y$ for the example given in section 1.

The value of λ once known, the values of ψ and ξ, too, can be determined by the diagram in Fig. 10. These values of ψ and ξ are likewise included in Fig. 10.

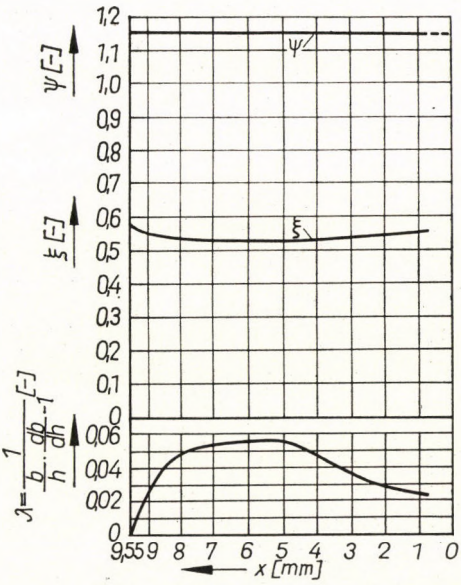


Fig. 10

Now, with the known values of the variables of ψ and ξ, as distributed on the compressed surface, the three stresses can be calculated but by finding the relevant equations. With due regard to what has been explained in chapter II, section 1., practically the equation which might be suitable for the purpose is equation (11). Namely, as regards equation (12), we are at a loss when we want to know the change of the elastic stresses. But with the chosen way, σ_z can be determined without knowing the stress σ_y. It can be directly calculated by knowing σ_x and ψ. By experimental results, we can assume that in the rolling operation under discussion, the state of deformation is approximately a plane one ($\psi = 1{,}14$).

By making use of equations (11), (36) and (37), we are able to determine the stress σ_z, when the function $k_f = k_f(x)$, and according to condition *c)*, the peripheral σ_z values are known. These σ_z values are given in Table I. Fig. 11a is delineated by making use of these tabulated values. For comparison's sake, the distribution of stress is shown in Fig. 11b as they were found by experimental measurements carried out by SIEBEL and LUEG [8] under similar conditions.

Table I

Peripheral σ_z values

x [mm]	h [mm]	$\psi \cdot k$ [kp/mm²]	$\frac{2\psi k\sigma}{h} \cdot \tan\varphi$	$\frac{2\mu}{h}$		y [mm]					
						0	3	6	9	12	15
0	1	21	0	0,28	$\cos\alpha$	1	1	1	1	1	1
					σ_x	10,5	6,3	2,1	—2,1	—6,3	10,5
					σ_z	31,5	27,3	23,1	18,9	14,7	10,5
1	1,011	20,8	0,452	0,277	$\cos\alpha$	1	0,9995	0,999	0,998	0,997	0,995
					σ_x	19,7	14,25	8,6	3,6	—1,8	—7,15
					σ_z	40,5	35,05	29,4	24,4	19	13,7
2	1,043	20,3	0,96	0,268	$\cos\alpha$	1	0,99	0,98	0,96	0,94	0,92
					σ_x	31,5	24	17,2	10,7	3,3	—2,9
					σ_z	51,8	44,7	37,5	31	23,6	17,4
3	1,095	19,9	1,2	0,256	$\cos\alpha$	—	—	—	—	—	—
					σ_x	—	—	—	—	—	—
					σ_z	—	—	—	—	—	—
4	1,18	19,4	1,45	0,237	$\cos\alpha$	1	0,98	0,96	0,91	0,98	0,79
					σ_x	24,9	19,6	13,8	9,2	2,5	—2,55
					σ_z	44,3	39	33,2	28,5	21,9	16,9
5	1,28	18,7	1,61	0,218	$\cos\alpha$	1	0,995	0,97	0,96	0,93	0,9
					σ_x	17,7	13,4	8,9	5,3	—0,14	—4,2
					σ_z	36,4	32,1	27,6	24	18,6	14,5

6	1,4	18	1,72	0,2	$\cos \alpha$	1	0,997	0,99	0,98	0,97	0,96
					σ_x	12,6	9,1	5,4	2,8	−1,82	−5,25
					σ_z	30,6	27,1	24,4	19,8	16,2	12,75
7	1,54	16,8	1,7	0,182	$\cos \alpha$	1	0,998	0,995	0,99	0,985	0,98
					σ_x	9,1	6,1	3,2	0,15	−2,8	−5,7
					σ_z	25,9	22,9	20	17	14	11,1
8	1,7	15	1,57	0,165	$\cos \alpha$	1	0,999	0,997	0,995	0,99	0,985
					σ_x	6,8	4,3	1,8	−0,7	−3,2	−5,7
					σ_z	21,8	19,3	16,8	14,3	11,8	9,3
9	1,9	12,3	1,3	0,147	$\cos \alpha$	1	0,9995	0,9993	0,999	0,998	0,997
					σ_x	5,45	3,3	1,2	−1	−3,1	−5,3
					σ_z	17,75	15,6	13,5	11,3	9,1	7
9,55	2	10	1,06	0,14	$\cos \alpha$	1	1	1	1	1	1
					σ_x	5	3	1	−1	−3	−5
					σ_z	15	13	11	9	7	5

As a matter of principle, the determination of the spatial distribution of stress, and of the flow diagram should be considered as rolling problems which belong to the scope of theoretic investigations. The solution as has been deduced in this paper is not suitable for the practizing expert. Anyhow they may serve the purpose of facilitating some further findings of a better approximation to real circumstances on the basis of a detailed analysis.

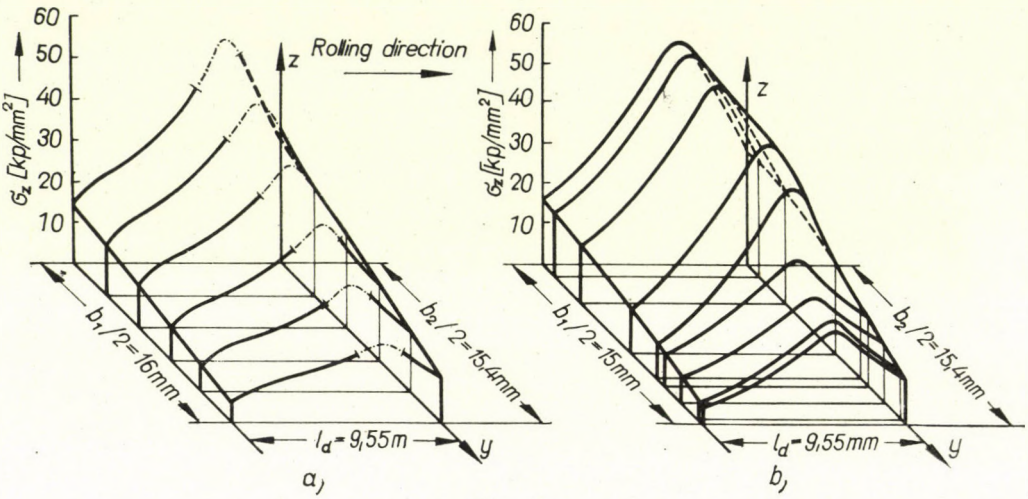


Fig. 11

REFERENCES

1. GELEJI, A.: Bildsame Formung der Metalle in Rechnung und Versuch; Berlin 1960.
2. TÓTH, L.: A hengerlés kísérő jelenségei és befolyásoló tényezői közötti összefüggések vizsgálata (Investigation of the relationship between phenomena of rolling and the influencing factors). Egyetemi doktori értekezés; Nehézipari Műszaki Egyetem, Miskolc 1964.
3. TÓTH, L.: Die Berechnung des Verformungs-Widerstandes prismatischer Körper beim Zusammendrücken ihrer parallelen Flächen. *Acta Techn. Hung.* **51** (1965), 105 — 120.
4. Теория прокатки; Металлургиздат, Москва 1962.
5. HILL, R.: The Mathematical Theory of Plasticity; Oxford 1950.
6. JOHNSON, W.—MELLOR, P. B.: Plasticity for Mechanical Engineers; London 1962.
7. DRUCKER, D. C.: A more fundamental approach to plastic stress—strain relation. *St. U. S. Nat. Cong. Appl. Mech. Chicago* (1951), Proc. I.
8. SIEBEL, E.—LUEG, W.: Untersuchungen über die Spannungsverteilung im Walzspalt. *Mitt. d. Kaiser Wilhelm Inst. für Eisenforschung* (1933).

Verteilung der Druckspannung an der Walzenoberfläche beim Auswalzen von dünnem Stahlblech. Im ersten Teil der Arbeit sind die Grundgleichungen des Walzens beschrieben. Im zweiten Teil sind die Lösungen hierfür abgeleitet. Diese Annäherungsformeln lassen sich in Kenntnis der im Laufe der Walzoperation auftretenden Deformation in der Breite des Walzgutes aufstellen. Gleichzeitig werden das Fließdiagram und die Verteilung der Druckspannung an der Walzenoberfläche mitbestimmt. Über die Richtigkeit der Lösung kann man sich durch Vergleich mit den experimentell erhaltenen Meßresultaten überzeugen.

Répartition des pressions à la surface d'un rouleau au laminage des feuillards. Dans la première partie de l'étude, on trouve la description des équations fondamentales du laminage; la deuxième partie contient la solution de ces équations. La solution approchée est basée sur la

connaissance de l'élargissement de la pièce durant le laminage. La solution contient le diagramme de fluage et la détermination des tensions de compression agissant sur la surface du rouleau. La comparaison avec des valeurs expérimentales mesurées prouve que cette solution est correcte.

Распределение напряжений, действующих на поверхность валков, при прокатке лент *(Л. Тот)*. В первой части работы дается описание основных уравнений прокатки. Во второй части работы излагается решение основных уравнений прокатки. Приближенный метод решения основан на знании расширения заготовки, возникающего при прокатке. Решение содержит определение диаграммы истечения и давящих напряжений, действующих на поверхность валков. О правильности решения можно убедиться путем сравнения с величинами, определенными экспериментальными измерениями.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (3—4), pp. 371—400 (1968)

GEOMETRIE DER SCHNECKENWÄLZFRÄSER MIT GRUNDSCHNECKEN SCHLEIFENEVOLVENTISCHER AUSFÜHRUNG*

I. LIPKA

ENTWICKLUNGSINSTITUT DER WERKZEUGMASCHINENWERKE, HALÁSZTELEK (UNGARN)

[Eingegangen am 18. November, 1965]

Der Verfasser untersucht Zahnradprofile, welche im Wälzverfahren mit Schneckenfräsern hergestellt wurden, wobei die Grundschnecke der Fräser als eine verschlungene Evolventenschnecke auf der Drehbank mit schräg gestelltem Stahl erzeugt werden kann. Es wird nachgewiesen, daß ein Zahnprofil, welches mit dem oben genannten Schneckenfräser hergestellt wird, um so mehr einer gewöhnlichen Kreisevolvente nahekommt, je kleiner der Steigungswinkel des Fräsers ist. Es wird ferner ein analytisches Kriterium vom Verfasser bei dem Zahnradprofil angewandt, welches mit einem Schneckenfräser, ausgeführt als eine verschlungene Evolventenschnecke, hergestellt wird. Dieses Kriterium dient zur Entscheidung über den Evolventencharakter des abgewälzten Zahnprofils.

I. Einleitung

In der vorigen Arbeit [1] wurden Schneckenfräser für das Wälzfräsen von Zahnrädern untersucht, bei denen die Grundschnecke des Fräsers eine Evolventenschnecke ist. In der vorliegenden Arbeit werden Zahnprofile untersucht, welche im Wälzverfahren mit Schneckenfräsern hergestellt werden, deren Grundschnecke als eine verschlungene Evolventenschnecke (Schleifenevolventenschnecke) ausgeführt ist. Auch von diesen, als verschlungene Evolventenschnecke ausgeführten Fräsern wird lediglich diejenige Schneckenart behandelt, welche auf der Drehbank mit schräg gestelltem Stahl hergestellt wird. Die Fräsfläche eines Schneckenfräsers, ausgeführt als verschlungene Evolventenschnecke, bedeutet eine geradlinige Schraubenfläche, deren Stirnschnitt (ebener Schnitt senkrecht zur Achse) eine verschlungene Kreisevolvente (Schleifenevolvente) bildet [1]. Die Charakteristik des mit schräg gestelltem Stahl hergestellten Schneckenfräsers in schleifenevolventischer Ausführung besteht darin, daß sein ebener Normalschnitt gerade oder genauer gesagt, der ebene Normalschnitt der Zahnlücke ein gerades Profil ist [2; S. 60]. Dies bedeutet, daß, wenn man die Schraubenlinie betrachtet, welche am Werkzeug-Teilzylinder liegt und durch den Halbierungspunkt der Zahnlücke geht, so schneidet die Ebene, welche zur Tangente dieses Punktes der Schraubenlinie

* Die Redaktion erwartet Stellungnahmen zu den Fragen, die in vorliegender Arbeit behandelt werden, und noch Gegenstand einer Diskussion bilden.

senkrecht steht und durch diesen Punkt geht, die Fräsfläche (Schraubenfläche) in einer Geraden. Es wird nachgewiesen, daß die Erzeugung des Zahnprofils, welches mit einer derartigen Fräs(„Schnecke, d. h. mit einer schleifenevolventischen Frässchnecke, welche mit schräg gestelltem Stahl hergestellt wird, um so mehr der Erzeugung mit Zahnstange nahekommt, je kleiner der Steigungswinkel des Fräsers ist.

Um die Zahnprofilkurve, welche durch den Fräser in der Stirnebene des Zahnrades hergestellt wird, untersuchen zu können, muß man das Gleichungssystem der Kurvenschar in Parameterform kennen, welche die Fräsfläche erzeugt. Mit seiner Hilfe kann man den Erzeugungsvorgang des Zahnprofils untersuchen, der sich in der Stirnebene des zu verzahnenden Rades infolge der Drehung des Fräsers und des Radkörpers zustandekommt — wobei jetzt die Stirnebene mit der normalen Schnittebene der Fräsfläche zusammenfällt.

II. Ableitung des Gleichungssystems in Parameterform für die verschlungene Evolventenschneckenfläche

Es wird das in Parameterform ausgedrückte Gleichungssystem für die spezielle verschlungene Evolventenschneckenfläche abgeleitet, die mit schräg gestelltem Stahl hergestellt ist. Die Fläche dieser Schnecke ist eine Schraubenfläche, die von Geraden gebildet wird, d. h., sie ist eine geradlinige Schraubenfläche, und diese Erzeugergeraden stehen senkrecht auf einer Schraubenlinie, welche auf dem Teilzylinder mit dem Radius R der Schnecke liegt. Eine derartige, spezielle Schraubenfläche entspricht einer verschlungenen Evolventenschnecke, welche auf der Drehbank mit schräg gestelltem Stahl in der Weise hergestellt wird, daß die Ebene, welche die beiden Schneiden des Stahles enthält, zur Schneckenachse um den Winkel γ geneigt ist. Der Winkel γ bedeutet hierbei den Steigungswinkel der Schraubenfläche auf dem Teilzylinder mit dem Radius R. Die Mittellinie des Stahles bildet mit den in derselben Ebene liegenden zwei Schneiden des Stahles den Eingriffswinkel α; die Mittellinie steht ferner senkrecht auf der Schneckenachse und schneidet sie. Daraus folgt, daß die Schneiden des Stahles im Verhältnis zur Schneckenachse windschiefe Geraden darstellen und von ihr in irgendeiner Entfernung k liegen. Diese Entfernung k, welche als Normaltransversale aus der Geometrie bekannt ist, ist gleich dem Radius vom Kehlzylinder der Schnecke.

Man betrachte ein räumliches, rechtwinkliges $(x; y; z)$ Koordinatensystem, bei dem die z-Achse identisch ist mit der Schneckenachse, und die x-Achse mit der Mittellinie des Drehstahles in seiner Ausgangslage zusammenfällt (Bild 1).

Da die genormten Abmessungen des Stahles (Bild 2) die folgenden sind:

$$v = \frac{m\pi}{2} \cos \gamma^*$$

und

$$l = 1{,}2 \text{ m}$$

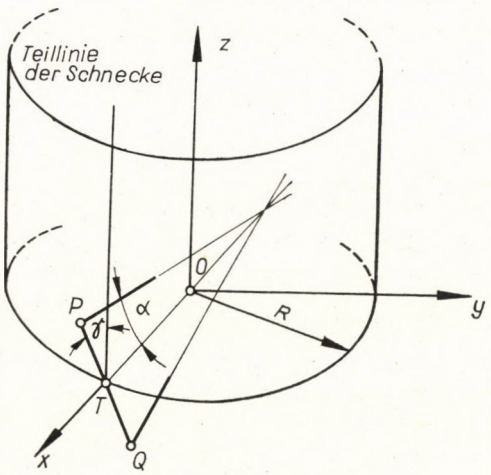


Bild 1

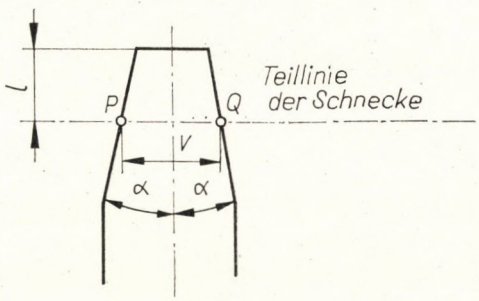


Bild 2

(m ist das Normmodul), ferner die Strecke PQ der Schneidenebene des Stahles vor der Drehung der Schneidenebene um den Winkel γ auf die Teillinie der Schnecke zu liegen kam, so werden die Koordinaten des Punktes P nach dem Bild 1 (für den Modul $m = 1$):

$$x = R;\ y = -\frac{\pi}{4} \cos \gamma \cdot \sin \gamma;\ z = \frac{\pi}{4} \cos^2 \gamma.$$

* Wenn der Normalmodul dem genormten Modul entspricht, dann schreibe man in die Formel von v und auch in die weiteren Formeln überall an die Stelle von π den Ausdruck $\pi/\cos \gamma$.

Die von dem Punkt P ausgehende Schneide bzw. ihre Verlängerung schneide die x-Achse im Punkt A (Bild 3).

Da $OT = R$; ferner aus dem rechtwinkligen Dreieck APT gemäß $PT = (\pi/4)\cos\gamma$ wird:

$$AT = \frac{\pi}{4}\cos\gamma \cdot \frac{1}{\tan\alpha},$$

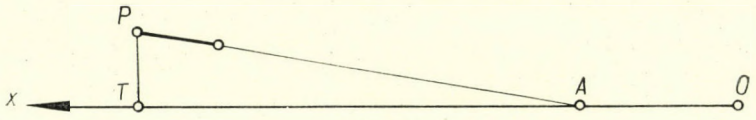


Bild 3

so sind die Koordinaten des Punktes A:

$$x = R - \frac{\pi}{4}\frac{\cos\gamma}{\tan\alpha};\ y = 0;\ z = 0\,.$$

In Kenntnis der Koordinaten der Punkte P und A können die drei Richtungskosinuswerte der Stahlschneide AP berechnet werden. Da aus dem rechtwinkligen Dreieck APT:

$$AP = \frac{\pi}{4}\frac{\cos\gamma}{\sin\alpha}$$

ist, nehmen die Richtungskosinusse folgende Werte an:

$$\frac{R - \left(R - \frac{\pi}{4}\frac{\cos\gamma}{\tan\alpha}\right)}{\frac{\pi}{4}\cdot\frac{\cos\gamma}{\sin\alpha}} = \cos\alpha;$$

$$\frac{-\frac{\pi}{4}\cos\gamma\sin\gamma}{\frac{\pi}{4}\cdot\frac{\cos\gamma}{\sin\alpha}} = -\sin\gamma\cdot\sin\alpha;$$

$$\frac{\frac{\pi}{4}\cos^2\gamma}{\frac{\pi}{4}\cdot\frac{\cos\gamma}{\sin\alpha}} = \cos\gamma\cdot\sin\alpha.$$

Mit diesen Richtungskosinussen wird das Gleichungssystem der Stahl-

schneide AP, als einer von dem Punkt A ausgehenden räumlihen Geraden:

$$x = \left(R - \frac{\pi}{4}\,\frac{\cos\gamma}{\tan\alpha}\right) + t\cos\alpha;$$
$$y = -t\sin\gamma\cdot\sin\alpha;$$
$$z = \quad t\cos\gamma\cdot\sin\alpha,$$

wo t den Entfernungsparameter bedeutet.

Die Gerade beschreibe eine Schraubenfläche, d. h. sie führe eine Schraubenbewegung um die z-Achse aus. Dreht sich die Gerade bei dieser Bewegung um die z-Achse um den Winkel τ, so wird sie gleichzeitig eine Verschiebung von der Größe $r\cdot\tau$ in Richtung der z-Achse erfahren, wobei:

$$r = \frac{S}{2\pi}.$$

Hier bedeutet S die Steigung der Schraubenfläche, d. h. die Größe der Verschiebung in Achsrichtung, welche während einer vollen Umdrehung von (2π) stattfindet. Demnach nimmt das Gleichungssystem für die mit einem schräg gestellten Stahl hergestellte Schraubenfläche folgende Form an:

$$x = \left(R - \frac{\pi}{4}\,\frac{\cos\gamma}{\tan\alpha}\right)\cos\tau + t\,(\cos\alpha\cdot\cos\tau + \sin\gamma\cdot\sin\alpha\cdot\sin\tau);$$
$$y = \left(R - \frac{\pi}{4}\,\frac{\cos\gamma}{\tan\alpha}\right)\sin\tau + t\,(\cos\alpha\cdot\sin\tau - \sin\gamma\cdot\sin\alpha\cdot\cos\tau); \qquad (1)$$
$$z = r\tau + t\cos\gamma\cdot\sin\alpha,$$

wo τ und t die Parameter sind.

Aus der dritten Gleichung des Gleichungssystems (1) ist zu ersehen, daß der auf die z-Achse bezogene Richtungskosinus der geraden Erzeugenden der Fläche $\cos\gamma\cdot\sin\alpha$ ist.

Aus dem Obigen folgt, daß wenn man den Winkel, der von der Erzeugenden und einer Ebene senkrecht zur z-Achse, d. h. zur Achse der Schraubenfläche, gebildet wird, mit ψ bezeichnet, so gilt für diesen Winkel ψ der Erzeugenden:

$$\sin\psi = \cos\gamma\cdot\sin\alpha. \qquad (1')$$

Aus dieser Beziehung folgt, da $\cos\gamma < 1$, daß

$$\psi < \alpha.$$

Der Winkel ψ der Erzeugenden der mit einem schräg gestellten Stahl hergestellten verschlungenen Evolventenschneckenfläche ist daher immer kleiner

als der Eingriffswinkel a α, d. h.als der Profilwinkel des schräg gestellten Stahles. L. SZENICZEI kommt in seinem Buch »Schneckengetriebe« [2] zu der irrtümlichen Feststellung, daß zwischen dem Erzeugungswinkel ψ und dem Eingriffswinkel α immer die Ungleichheit

$$\psi > \alpha$$

besteht, ungeachtet dessen, um welche Schnecke es sich handelt. Diese irrtümliche Feststellung rührt aus der Annahme her, daß für den im Normalschnitt der Schnecke auftretenden Eingriffswinkel α und den axialen Eingriffswinkel α_{ax} immer die Beziehung

$$\tan \alpha_{ax} = \frac{\tan \alpha}{\cos \gamma}$$

gültig ist, wobei γ den Steigungswinkel der Schnecke auf dem Teilzylinder bedeutet. Die zuletzt angeschriebene Beziehung gilt jedoch nur für die Evolventenschnecke, wogegen bei einem axialen Eingriffswinkel einer verschlungenen Evolventenschnecke, welche mittels Schrägstahls hergestellt ist, folgende Beziehung besteht:

$$\tan \alpha_{ax} = \frac{\tan \alpha}{\cos \gamma} \cdot \frac{1 - \frac{\pi}{2} \sin^3 \gamma \cdot \cot \alpha \cdot \cos \gamma}{\sqrt{1 + \pi \tan \alpha \cdot \sin^3 \gamma \cdot \cos \gamma - \left(\frac{\pi}{2}\right)^2 \sin^4 \gamma \cdot \cos^2 \gamma}} .$$

In das oben abgeleitete Gleichungssystem (1) soll jetzt der Radius k des Kehlzylinders eingeführt werden. Zu diesem Zweck betrachte man die zur Koordinatenebene $(x; y)$ senkrechte Projektion der geraden Erzeugenden der Schraubenfläche, da die Entfernung dieser Projektion vom Ausgangspunkt O dem Radius k gleich ist. Das in Parameterform ausgedrückte Gleichungssystem der in der Projektion entstehenden Geraden wird durch die ersten beiden Gleichungen des unter (1) angeschriebenen Gleichungssystems gebildet, d.h.

$$x = \left(R - \frac{\pi}{4} \frac{\cos \gamma}{\tan \alpha}\right) \cos \tau + t (\cos \alpha \cdot \cos \tau + \sin \gamma \cdot \sin \alpha \cdot \sin \tau) ;$$

$$y = \left(R - \frac{\pi}{4} \frac{\cos \gamma}{\tan \alpha}\right) \sin \tau + t (\cos \alpha \cdot \sin \tau - \sin \gamma \cdot \sin \alpha \cdot \cos \tau) .$$

Zur Eliminierung des Parameters t multipliziere man beide Seiten der ersten Gleichung mit dem Ausdruck

$$(\sin \gamma \cdot \sin \alpha \cdot \cos \tau - \cos \alpha \cdot \sin \tau) ,$$

ferner beide Seiten der zweiten Gleichung mit dem Ausdruck

$$(\cos\alpha\cdot\cos\tau+\sin\gamma\cdot\sin\alpha\cdot\sin\tau),$$

dann addiere man die entsprechenden Seiten, wobei man für die Projektion der Erzeugenden folgende Gleichung gewinnt:

$$\begin{aligned}x(\sin\gamma\cdot\sin\alpha\cdot\cos\tau-\cos\alpha\cdot\sin\tau)+\\+y(\cos\alpha\cdot\cos\tau+\sin\gamma\cdot\sin\alpha\cdot\sin\tau)=\\=\left(R-\frac{\pi}{4}\,\frac{\cos\gamma}{\tan\alpha}\right)\sin\gamma\sin\alpha.\end{aligned}$$

Die Hessesche Normalform dieser Gleichung wird, da die Summe der Quadrate der Koeffizienten von x und y

$$\begin{aligned}(\sin\gamma\cdot\sin\alpha\cdot\cos\tau-\cos\alpha\cdot\sin\tau)^2+(\cos\alpha\cdot\cos\tau+\sin\gamma\cdot\sin\alpha\cdot\sin\tau)^2=\\=\cos^2\alpha+\sin^2\gamma\cdot\sin^2\alpha=1-\cos^2\gamma\cdot\sin\alpha\end{aligned}$$

ist, folgende Gestalt annehmen:

$$\begin{aligned}x\,\frac{\sin\gamma\cdot\sin\alpha\cdot\cos\tau-\cos\alpha\cdot\sin\tau}{\sqrt{1-\cos^2\gamma\cdot\sin^2\alpha}}+\\+y\,\frac{\cos\alpha\cdot\cos\tau+\sin\gamma\cdot\sin\alpha\cdot\sin\tau}{\sqrt{1-\cos^2\gamma\cdot\sin^2\alpha}}=\\=\left(R-\frac{\pi}{4}\,\frac{\cos\gamma}{\tan\alpha}\right)\frac{\sin\gamma\cdot\sin\alpha}{\sqrt{1-\cos^2\gamma\cdot\sin^2\alpha}}.\end{aligned}$$

Da das konstante Glied der Normalform der Entfernung des Ausgangspunktes O von der Geraden gleich ist, so liefert das konstante Glied den Radius k des Kehlzylinders, und daher wird:

$$k=\left(R-\frac{\pi}{4}\,\frac{\cos\gamma}{\tan\alpha}\right)\frac{\sin\gamma\cdot\sin\alpha}{\sqrt{1-\cos^2\gamma\cdot\sin^2\alpha}}.$$

Nun wird der Radius k des Kehlzylinders in das Gleichungssystem (1) der Schraubenfläche in der Weise eingeführt, daß man als Ausgangspunkt den Punkt auf der geraden Erzeugenden der Fläche nimmt — anstelle des jetzigen Punktes A —, der auf dem Kehlzylinder liegt. Die Koordinaten $(x; y)$ dieses Punktes erhält man, wenn man die Koeffizienten von x bzw. y in der vorher angeschriebenen Hesseschen Normalgleichung der Projektion der geraden

Erzeugenden mit Radius k des Kehlzylinders multipliziert. Demnach sind die Koordinaten x und y des Berührungspunktes der zum Parameter τ gehörenden geraden Erzeugenden und des Kehlzylinders:

$$x_0 = k\frac{\sin\gamma\cdot\sin\alpha\cdot\cos\tau-\cos\alpha\cdot\sin\tau}{\sqrt{1-\cos^2\gamma\cdot\sin^2\alpha}}\,,\quad y_0 = k\frac{\cos\alpha\cdot\cos\tau+\sin\gamma\cdot\sin\alpha\cdot\sin\tau}{\sqrt{1-\cos^2\gamma\cdot\sin^2\alpha}}\,.$$

Damit man auch die Koordinate z_0 dieses Punktes erhält, multipliziere man die unter (1) angegebene erste Gleichung mit $\sin\tau$, die zweite Gleichung dagegen mit $(-\cos\tau)$, dann addiere man die entsprechenden Seiten; in diesem Fall ist:

$$x\sin\tau - y\cos\tau = t\sin\gamma\cdot\sin\alpha\,.$$

Setzt man hier die obigen Werte von x_0 und y_0 ein, so ergibt sich

$$-\frac{k\cos\alpha}{\sqrt{1-\cos^2\gamma\cdot\sin^2\alpha}} = t\sin\gamma\cdot\sin\alpha\,,$$

woraus

$$t = -\frac{k\cos\alpha}{\sin\gamma\cdot\sin\alpha\sqrt{1-\cos^2\gamma\cdot\sin^2\alpha}}\,,$$

wo $\gamma\neq 0$ ist. Setzt man diesen Wert von t in den unter (1) angegebenen Ausdruck von z ein, so wird

$$z_0 = r\cdot\tau - \frac{k\cos\alpha\cdot\cos\gamma}{\sin\gamma\sqrt{1-\cos^2\gamma\cdot\sin^2\alpha}} = r\tau - \frac{k\cos\alpha}{\tan\gamma\sqrt{1-\cos^2\gamma\cdot\sin^2\alpha}}\,.$$

Nimmt man nun in dem Gleichungssystem (1) — welches das Gleichungssystem der Geraden bei konstantem τ und veränderlichem t der geraden Erzeugenden der Schraubenfläche ist — als Ausgangspunkt den Punkt $(x_0;\ y_0;\ z_0)$ der Geraden, dann erhält das Gleichungssystem der geraden Erzeugenden folgende Form:

$$\begin{aligned}
x &= k\frac{\sin\gamma\cdot\sin\alpha\cdot\cos\tau-\cos\alpha\cdot\sin\tau}{\sqrt{1-\cos^2\gamma\cdot\sin^2\alpha}} + \\
&\quad + t(\cos\alpha\cdot\cos\tau+\sin\gamma\cdot\sin\alpha\cdot\sin\tau)\,; \\
y &= k\frac{\sin\gamma\cdot\sin\alpha\cdot\sin\tau+\cos\alpha\cdot\cos\tau}{\sqrt{1-\cos^2\gamma\cdot\sin^2\alpha}} + \\
&\quad + t(\cos\alpha\cdot\sin\tau-\sin\gamma\cdot\sin\alpha\cdot\cos\tau)\,; \\
z &= r\tau - \frac{k\cos\alpha}{\tan\gamma\sqrt{1-\cos^2\alpha\cdot\sin^2\gamma}} + t\cos\gamma\cdot\sin\alpha\,.
\end{aligned} \tag{1''}$$

Damit wurde das Gleichungssystem der Schraubenfläche in einer Form dargestellt, bei der die geraden Erzeugenden der Fläche von den auf dem Kehlzylinder liegenden Punkten ausgehen. Das unter (1″) angegebene Gleichungssystem der Schraubenfläche wird noch in folgender Weise umgeformt: Man definiere den Winkel χ in nachfolgender Weise:

$$\cos\chi = \frac{\sin\gamma\cdot\sin\alpha}{\sqrt{1-\sin^2\alpha\cdot\cos^2\gamma}}\,,\quad \sin\chi = \frac{\cos\alpha}{\sqrt{1-\sin^2\alpha\cdot\cos^2\gamma}}\,. \tag{2}$$

Wenn man die hieraus sich ergebenden Werte von $\sin\gamma$. $\sin\alpha$ und $\cos\alpha$ in (1″) einsetzt, und wenn man ferner den Winkel ψ einführt, den die Erzeugenden mit der auf der z-Achse senkrecht stehenden Ebene einschließen, wobei der Winkel ψ den Ergänzungswinkel für den Winkel bedeutet, der von den Erzeugenden und der z-Achse gebildet wird, und weiterhin nach (1′):

$$\cos\psi = \sqrt{1-\sin^2\alpha\cdot\cos^2\gamma} \tag{3}$$

ist (nach der dritten Gleichung aus (1) ist nämlich der Richtungskosinus der Erzeugenden auf die z-Achse bezogen: $\cos\gamma\sin\alpha$); so wird das Gleichungssystem unter (1″) folgende Form annehmen:

$$\begin{aligned} x &= k\cos(\tau+\chi) + \cos\psi\cdot\sin(\tau+\chi)\cdot t\,;\\ y &= k\sin(\tau+\chi) - \cos\psi\cdot\cos(\tau+\chi)\cdot t\,;\\ z &= r\cdot\tau - \frac{k\cos\alpha}{\tan\gamma\cos\psi} + t\cdot\sin\psi\,,\ (\gamma\neq 0)\,. \end{aligned} \tag{1*}$$

Dies bedeutet das in Parameterform dargestellte Gleichungssystem der Fläche der verschlungenen Evolventenschnecke, die mittels Schrägstahl, d. h. um den Winkel γ schräg gestellten Stahl hergestellt worden ist. τ und t sind die Parameter; χ ist der Winkel, dessen Definition unter (2) gegeben worden ist; α ist der Eingriffswinkel des Werkzeugs (eines Stahles mit gerader Schneide), angegeben im Bild 2; k ist ferner der Radius des Kehlzylinders.*

III. Bestimmung des ebenen Normalschnittes der Schleifenevolventenschnecke

Die normale Schnittebene steht senkrecht auf der Schraubenlinie, die auf dem Teilzylinder mit dem Radius R liegt, wobei die Schraubenlinie einen Steigungswinkel γ besitzt. Das Gleichungssystem in Parameterform der

* Der Umstand, daß die im Gleichungssystem unter (1*) gegebene Fläche von den drei Schneckenarten in der Tat der Schleifenevolventenschnecke entspricht, kann durch die Bestimmung der Stirnschnittkurve der Fläche entschieden werden. Ihre Untersuchung wird im Anhang vorgenommen.

Schraubenlinie, die auf dem Zylinder mit dem Radius R liegt und den Steigungswinkel γ besitzt, ist von folgender Form:

$$\begin{aligned} x &= R\cos\tau; \\ y &= R\sin\tau; \\ z &= r\cdot\tau, \end{aligned} \tag{4}$$

wobei

$$r = R\tan\gamma \tag{4'}$$

ist [2; S. 42].

In dem zum Parameterwert τ gehörenden Punkt der Schraubenlinie sind die Richtungskosinusse der Tangente proportional folgenden Größen:

$$-R\sin\tau;\ R\cos\tau;\ r.$$

Es ist leicht nachzuweisen, daß auf dieser Tangente die zu τ gehörende und unter (1*) erzeugte gerade Erzeugende der Schraubenfläche senkrecht steht. Komponiere man die Richtungskosinusse der Erzeugenden mit den oben angeschriebenen Größen, welche proportional den Richtungskosinussen der Schraubenlinientangente sind:

$$\begin{aligned} -\cos\psi\cdot\sin(\tau+\chi)\cdot R\sin\tau - \cos\psi\cdot\cos(\tau+\chi)\cdot R\cos\tau + \sin\psi\cdot r = \\ = -\cos\psi\cdot R\cdot\cos\chi + \sin\psi\cdot r. \end{aligned}$$

Setzt man die Werte von $\cos\chi$ aus (2) und von r aus (4') ein, wobei man berücksichtigt, daß nach (1') $\sin\psi = \cos\gamma\ \sin\alpha$ ist, so wird

$$-\cos\psi\cdot R\cdot\cos\chi + \sin\psi\cdot r = -R\sin\gamma\sin\alpha + R\sin\gamma\cdot\sin\alpha = 0.$$

Da der Wert dieser Komposition 0 ist, so steht die Erzeugende in der Tat senkrecht auf der entsprechenden Tangente der Schraubenlinie.

Nun bestimme man die Gleichung der Ebene, welche durch den Punkt der Schraubenlinie (4) geht, der zum Parameterwert τ gehört und senkrecht zur Tangente steht, die zu diesem Punkt der Schraubenlinie gehört. Die Richtungskosinusse der Normalen dieser Ebene stimmen offensichtlich mit den Richtungskosinussen der Schraubenlinientangente überein, da die Ebene senkrecht auf der Tangente der Schraubenlinie steht. Demnach lautet die Gleichung der Ebene, welche den Normalschnitt der Schraubenfläche liefert:

$$(x - R\cos\tau)(-R\sin\tau) + (y - R\sin\tau)R\cos\tau + (z - r\,\tau)r = 0$$

oder in entwickelter Form:

$$-x\cdot R\sin\tau + y\,R\cos\tau + (z - r\,\tau)r = 0. \tag{5}$$

Diese Ebene schneidet die Schraubenfläche in der Geraden, welche zum gegebenen Wert τ gehört und unter (1*) erzeugt worden ist, da sofern man die unter (1*) angegebenen Ausdrücke in die Gleichung der Ebene von (5) einsetzt, die Gleichung erfüllt wird. Wenn man nämlich die Ausdrücke von (1*) in (5) einsetzt und dabei berücksichtigt, daß $\sin\psi = \cos\gamma \sin\alpha$ ist, so wird:

$$
\begin{aligned}
&-\left[k\cos(\tau+\chi)+\cos\psi\cdot\sin(\tau+\chi)\cdot t\right]R\sin\tau+\\
&+\left[k\sin(\tau+\chi)-\cos\psi\cdot\cos(\tau+\chi)\cdot t\right]R\cos\tau+\\
&+\left(r\tau-\frac{k\cos\alpha}{\tan\gamma\cdot\cos\psi}+t\cos\gamma\cdot\sin\alpha-r\tau\right)r=\\
&=\left[-kR\cos(\tau+\chi)\cdot\sin\tau+kR\sin(\tau+\chi)\cdot\cos\tau-\frac{r\cdot k\cdot\cos\alpha}{\tan\gamma\cos\psi}\right]+\\
&+t\left(-R\cos\psi\cdot\sin\tau\cdot\sin(\tau+\chi)-\right.\\
&\left.-R\cos\psi\cdot\cos\tau\cdot\cos(\tau+\chi)+r\cos\gamma\cdot\sin\alpha\right).
\end{aligned}
$$

Hier wird der Ausdruck in der eckigen Klammer, da nach (4′) $r = R\tan\gamma$ ist,

$$
\left[kR\left(\sin(\tau+\chi)\cdot\cos\tau-\cos(\tau+\chi)\cdot\sin\tau\right)-\frac{k\cos\alpha}{\tan\gamma\cos\psi}r\right]=
$$

$$
=kR\sin\chi-k\frac{\cos\alpha}{\tan\gamma\cos\psi}R\tan\gamma=kR\sin\chi-k\frac{\cos\alpha}{\cos\psi}\cdot R,
$$

setzt man hier den Wert $\sin\chi = \cos\alpha/\cos\psi$ aus (2) ein, unter Berücksichtigung von (3), so ist:

$$
kR\frac{\cos\alpha}{\cos\psi}-k\frac{\cos\alpha}{\cos\psi}R=0.
$$

Der in der runden Klammer stehende Ausdruck: der Koeffizient von t, verschwindet auch identisch, da

$$
\begin{aligned}
&R\cos\psi\left[-\sin\tau\cdot\sin(\tau+\chi)-\cos\tau\cdot\cos(\tau+\chi)\right]+r\cos\gamma\cdot\sin\alpha=\\
&=-R\cos\psi\cdot\cos\chi+r\cos\gamma\cdot\sin\alpha
\end{aligned}
$$

und wenn man hier den folgenden Wert aus (2)

$$
\cos\chi=\frac{\sin\gamma\cdot\sin\alpha}{\cos\psi}
$$

und den Wert $r = R \tan \gamma$ einsetzt, so ist:

$$- R \cos \psi \frac{\sin \gamma \cdot \sin \alpha}{\cos \psi} + R \sin \gamma \cdot \sin \alpha = 0.$$

Die Ebene (5), welche auf der Schraubenlinie (4) bzw. auf ihrer Tangente senkrecht steht, schneidet die Schraubenfläche in der zu einem gegebenen Wert von τ gehörenden, unter (1*) angeführten Geraden, in der geraden Erzeugenden der Fläche, und in dieser Weise ist der Normalschnitt der Schraubenfläche tatsächlich eine Gerade.

IV. Zahnprofilwälzung mittels Schneckenfräser mit Grundschnecken schleifenevolventischer Ausführung

In diesem Abschnitt wird folgender Satz bewiesen:

Mit Hilfe eines Schneckenfräsers in schleifenevolventischer Ausführung, wobei der Fräser mit schräg gestelltem Stahl hergestellt worden ist, kann, sofern die Fräserachse dem Steigungswinkel entsprechend schräg gestellt wird, im Abwälzverfahren ein Zahnprofil hergestellt werden, das um so mehr einer Kreisevolvente angenähert ist, je kleiner der Steigungswinkel des Fräsers ist.

Man drehe die Schraubenfläche, welche durch die Gleichungen von (1*) bestimmt ist, um die x-Achse in der Weise, daß die T-Achse der Schraubenfläche, nach Bild 4, eine Lage einnehme, in der sie mit der y-Achse einen Winkel γ bildet.

Für diese Lage mit der schräg gestellten Achse erhält man das Gleichungssystem der Schraubenfläche in der Weise, daß man für die Ausdrücke y und z aus (1*) die Rotation

$$y' = y \cos \varphi - z \sin \varphi;$$
$$z' = y \sin \varphi + z \cos \varphi$$

anwendet, wobei $\varphi = -(90° + \gamma)$ ist. In diesem Fall bleibt x unverändert, und wenn man die Ausdrücke y', z' erneut nur mit y und z bezeichnet, so nimmt das Gleichungssystem der Schraubenfläche mit der schräg gestellten Achse folgende Form an:

$$\begin{aligned} x &= k \cos (\tau + \chi) + \cos \psi \cdot \sin (\tau + \chi) \cdot t; \\ y &= -(k \sin (\tau + \chi) - \cos \psi \cdot \cos (\tau + \chi) \cdot t) \sin \gamma + \\ &\quad + [r \cdot \tau - \tau_0 + t \sin \psi] \cos \gamma; \\ z &= -(k \sin (\tau + \chi) - \cos \psi \cdot \cos (\tau + \chi) \cdot t) \cos \gamma - \\ &\quad - [r \tau - \tau_0 + t \sin \psi] \sin \gamma, \end{aligned} \tag{6}$$

wobei:

$$\tau_0 = \frac{k \cos \alpha}{\tan \gamma \cos \psi} \tag{7}$$

ist.

Man drehe die Fläche um ihre T-Achse, und dies entspricht der Drehung des Schneckenfräsers um seine T-Achse, einer Komponente der Wälzbewegung. Dreht man die Schraubenfläche um die T-Achse um irgendeinen Winkel Φ, so ist dies gleichbedeutend damit, daß im Gleichungssystem der Schraubenfläche in (1*) in die Ausdrücke von x und y anstelle des Winkels τ der Winkel

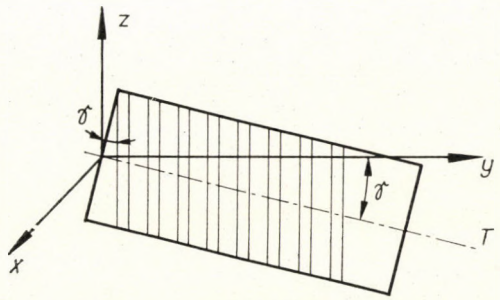


Bild 4

$(\tau + \Phi)$ kommt. Dementsprechend erhält man das Gleichungssystem der Fläche (6), welche um die T-Achse um den Winkel Φ gedreht worden ist, indem in den Ausdruck von x im System (6) anstelle von τ der Wert $(\tau + \Phi)$ gesetzt wird; in der Beziehung für y, ebenfalls in (6), setzt man in den Ausdruck, der sich in der runden Klammer neben $\sin \gamma$ befindet, anstelle von τ den Wert $(\tau + \Phi)$; im Ausdruck von z schreibt man ferner in dem Faktor, der sich in der runden Klammer neben $\cos \gamma$ befindet, anstelle von τ den Wert $(\tau + \Phi)$. In diesem Fall wird das Gleichungssystem der um die T-Achse um den Winkel Φ gedrehten Fläche:

$$\begin{aligned} x &= k \cos(\tau + \Phi + \chi) + \cos\psi \cdot \sin(\tau + \Phi + \chi) \cdot t; \\ y &= -[k \sin(\tau + \Phi + \chi) - \cos\psi \cdot \cos(\tau + \Phi + \chi) \cdot t] \sin\gamma + \\ &\quad + [r\tau - \tau_0 + t \sin\psi] \cos\gamma; \\ z &= -[k \sin(\tau + \Phi + \chi) - \cos\psi \cdot \cos(\tau + \Phi + \chi) \cdot t] \cos\gamma - \\ &\quad -[r\tau - \tau_0 + t \sin\psi] \sin\gamma. \end{aligned} \tag{6'}$$

Das Gleichungssystem unter (6) bedeutet die Darstellung in Parameterform der Schneckenfläche, die um die T-Achse gedreht wird. Drehe man die Erzeugende dieser Fläche, welche zum Parameterwert τ gehört, um den Winkel

$$\Phi = 2\pi - \tau.$$

Nach der Formel (6′) wird, wenn man bei ihr den Ausdruck $\Phi + \tau = 2\pi$ einführt, die zu τ gehörende Erzeugende bei der Drehung in eine Gerade übergehen, deren Gleichungssystem wie folgt aussieht:

$$\begin{aligned} x &= k\cos\chi + \cos\psi \cdot \sin\chi \cdot t; \\ y &= -(k\sin\chi - \cos\psi \cdot \cos\chi \cdot t)\sin\gamma + [r\tau - \tau_0 + t\sin\psi]\cos\gamma; \\ z &= -(k\sin\chi - \cos\psi \cdot \cos\chi \cdot t)\cos\gamma - [r\tau - \tau_0 + t\sin\psi]\sin\gamma, \end{aligned}$$

oder nach dem Parameter t geordnet:

$$\begin{aligned} x &= k\cos\chi + \cos\psi \cdot \sin\chi \cdot t; \\ y &= -k\sin\chi \cdot \sin\gamma + (r\tau - \tau_0)\cos\gamma + \\ &\quad + [\cos\psi \cdot \cos\chi \cdot \sin\gamma + \sin\psi \cdot \cos\gamma] \cdot t; \\ z &= -k\sin\chi \cdot \cos\gamma - (r\tau - \tau_0)\sin\gamma + \\ &\quad + [\cos\psi \cdot \cos\chi \cdot \cos\gamma - \sin\psi \cdot \sin\gamma] \cdot t. \end{aligned}$$

Dieses Gleichungssystem kann noch in folgender Weise umgeformt und vereinfacht werden: da nach (2) und (3)

$$\cos\psi \cdot \cos\chi = \sin\gamma \cdot \sin\alpha,$$

ferner nach (1′)

$$\sin\psi = \sin\alpha \cdot \cos\gamma$$

ist, so ist im Ausdruck von z der Koeffizient von t:

$$\begin{aligned} \cos\psi \cdot \cos\chi \cdot \cos\gamma - \sin\psi \cdot \sin\gamma &= \sin\gamma \cdot \sin\alpha \cdot \cos\gamma - \\ - \sin\alpha \cdot \cos\gamma \cdot \sin\gamma &= 0; \end{aligned}$$

und im Ausdruck von y der Koeffizient von t:

$$\cos\psi \cdot \cos\chi \cdot \sin\gamma + \sin\psi \cdot \cos\gamma = \sin\alpha \cdot \sin^2\gamma + \sin\alpha \cdot \cos^2\gamma = \sin\alpha.$$

Unter Berücksichtigung dieser Vereinfachungen wird das Gleichungssystem der Geraden, in welche die gerade Erzeugende, die zum Parameterwert τ gehört, beim Drehen der Schnecke um den Winkel $\Phi = 2\pi - \tau$ übergeht, folgende Form annehmen:

$$\begin{aligned} x &= k\cos\chi + \cos\psi \cdot \sin\chi \cdot t; \\ y &= -k\sin\chi \cdot \sin\gamma + (r \cdot \tau - \tau_0)\cos\gamma + t \cdot \sin\alpha; \\ z &= -k\sin\chi \cdot \cos\gamma - (r\tau - \tau_0)\sin\gamma. \end{aligned}$$

Führt man hier noch die Werte von $\cos \chi$ und $\sin \chi$ aus (2) ein, und berücksichtigt, daß nach (7)

$$\tau_0 = \frac{k \cos \alpha}{\tan \gamma \cos \psi}$$

gilt, so gewinnt man für das Gleichungssystem der Geraden:

$$\begin{aligned} x &= k \frac{\sin \gamma \cdot \sin \alpha}{\cos \psi} + \cos \alpha \cdot t ; \\ y &= - \frac{k \cos \alpha}{\sin \gamma \cdot \cos \psi} + r \cos \gamma \cdot \tau + t \cdot \sin \alpha ; \\ z &= - r \sin \gamma \cdot \tau . \end{aligned} \tag{8}$$

Die Darstellung unter (8) zeigt, daß beim Drehen der Fläche um den Winkel $\Phi = (2\pi - \tau)$ die zum Parameterwert τ gehörende Erzeugende in eine Gerade übergeht, die parallel mit der Ebene $(x; y)$ ist. Der Wert der Koordinate $z = -r \sin \gamma \cdot \tau$ ist nämlich unabhängig vom Parameter t, und dieser Wert z ist gerade gleich der Entfernung der Geraden von der Ebene $(x; y)$. Der Wert dieser Entfernung ändert sich, wenn τ sich von 0 bis 2π verändert, d. h. wenn die geraden Erzeugenden der Schraubenfläche infolge der Drehung nacheinander in eine (horizontale) Lage, parallel mit der Ebene $(x; y)$ gelangen. Die Änderung der Entfernung z wird jedoch nur von kleinem Ausmaß sein, da der Wert des im Ausdruck von z auftretenden Faktors $r \cdot \sin \gamma$ klein ist, da γ klein und $r = 0{,}5$ ist. Wenn der Winkel der Drehung, $\Phi = (2\pi - \tau)$, wächst, so vermindert sich τ, und das bedeutet soviel, daß die Entfernung der Erzeugenden, die nacheinander in eine mit der $(x; y)$-Ebene parallele Lage geraten, von der parallelen Ebene $(x; y)$ sich dauernd vermindert. Demnach ist es offensichtlich, daß im Fall eines sehr kleinen Steigungswinkels γ, wenn der Wert von z gemäß (8) ebenfalls sehr klein ist, die geraden Erzeugenden der Schraubenfläche bei der Drehung des Fräsers nacheinander in eine angenähert horizontale Ebene, in eine Stirnebene des zu verzahnenden Rades gelangen, wo eine Schar von Geraden entsteht. Demnach kann man sagen, daß in der horizontalen Ebene $(x; y)$ eine Geradenschar entsteht, die von den ersten zwei Gleichungen des Gleichungssystems unter (8) erzeugt wird, und welches etwas gekürzt in folgender Form angeschrieben wird:

$$\begin{aligned} x &= a + t \cdot \cos \alpha ; \\ y &= b + r \cos \gamma \cdot \tau + t \cdot \sin \alpha , \end{aligned} \tag{8'}$$

wobei

$$a = k \frac{\sin \gamma \cdot \sin \alpha}{\cos \psi} ; \quad b = - \frac{k \cos \alpha}{\sin \gamma \cdot \cos \psi}$$

ist.

Dieses Gleichungssystem mit zwei Parametern, wo τ und t die Parameter sind, erzeugt eine Schar der parallelen Geraden. Diese Geradenschar, betrachtet aus der Ebene, welche auf der Achse des durch Abwälzung zu erzeugenden Zahnrades senkrecht steht und bei der Wälzbewegung sich gemeinsam mit dem Zahnradkörper dreht, wird eine Schar aus nicht mehr parallelen Geraden bilden. Die Umhüllende dieser Schar liefert dann das durch den Fräser erzeugte Zahnprofil im Stirnschnitt. Den ganzen Erzeugungsvorgang des Zahnprofils kann man sich so vorstellen, daß das Zahnrad (das Arbeitsstück) unbeweglich steht (sich also nicht um die Achse dreht), wogegen der Fräser, der sich um seine eigene Achse dreht, auch noch um den Mittelpunkt des Zahnrades eine Drehbewegung ausführt. Dieser Betrachtungsweise entsprechend wird zur Winkeldrehung von der Größe $(2\,\pi - \tau)$ um die T-Achse des Schneckenfräsers, oder was dasselbe bedeutet, zur Winkeldrehung des Schneckenfräsers um die T-Achse von der Größe $(-\tau)$ eine Winkeldrehung des Fräsers um den Mittelpunkt des Zahnrades um irgendeine Größe von $(-\varkappa\tau)$ gehören (wo $\varkappa$ einen geeigneten Faktor bedeutet). Das Koordinatensystem $(x; y)$ wird parallel zu sich selbst verschoben in der Weise, daß der Ursprung 0 des Koordinatensystems mit dem Mittelpunkt des Zahnrades zusammenfalle. In diesem Fall verändern sich die konstanten Glieder a und b in den Gleichungen unter (8′). Danach wird das Koordinatensystem $(x; y)$ in Richtung der y-Achse, mit dem Zahnradkörper zusammen, parallel zu sich selbst so verschoben, daß in dem Ausdruck von y aus (8′) das konstante Glied verschwindet, d. h. $b = 0$ wird. Vollzieht der Fräser um seine T-Achse eine Winkeldrehung von der Größe $(-\tau)$ und um den Mittelpunkt des Zahnrades eine (relative) Winkeldrehung von der Größe $(-\varkappa\tau)$, so wird nach den obigen Ausführungen das in Parameterform ausgedrückte Gleichungssystem der im Stirnschnitt entstehenden Geradenschar folgende Form annehmen:

$$x = (a + t\cos\alpha)\cos\varkappa\,\tau + (r\cos\gamma\cdot\tau + t\sin\alpha)\sin\varkappa\,\tau;$$
$$y = -(a + t\cos\alpha)\sin\varkappa\,\tau + (r\cos\gamma\cdot\tau + t\sin\alpha)\cos\varkappa\,\tau.$$

Das letztere Gleichungssystem kann noch in folgender Weise angeschrieben werden:

$$x = a\cos\varkappa\,\tau + \tau\,r\cos\gamma\cdot\sin\varkappa\,\tau + t\cdot\cos(\alpha - \varkappa\tau);$$
$$y = -a\sin\varkappa\,\tau + \tau\,r\cos\gamma\cdot\cos\varkappa\,\tau + t\cdot\sin(\alpha - \varkappa\tau).$$

Zur Eliminierung des Parameters t multipliziere man die erste Gleichung mit $\sin(\alpha - \varkappa\tau)$, die zweite mit $-\cos(\alpha - \varkappa\tau)$, und dann addiere man die entsprechenden Seiten:

$$x\sin(\alpha - \varkappa\tau) - y\cos(\alpha - \varkappa\tau) = a\sin\alpha - \tau\,r\cos\gamma\cdot\cos\alpha.$$

Diese lineare Gleichung bedeutet die Gleichung der Geradenschar, deren Hüllkurve das Zahnprofil bildet. Hierbei bedeutet τ den Parameter der Schar, und wenn man diesen aus der Gleichung eliminiert, so gewinnt man die Gleichung der Hüllkurve. Schreibe man die obige Gleichung der Geradenschar in einer auf 0 reduzierten Form auf:

$$x \sin(\alpha - \varkappa\tau) - y \cos(\alpha - \varkappa\tau) + \tau r \cos\gamma \cdot \cos\alpha - a \sin\alpha = 0;$$

und bezeichne man die linke Seite der Gleichung mit:

$$F(x;\ y;\ \tau) = 0.$$

Zur Eliminierung des Parameters τ muß der folgende partielle Differentialquotient gebildet werden:

$$\frac{\partial F(x;y;\tau)}{\partial\tau} = -x\varkappa \cos(\alpha - \varkappa\tau) - y\varkappa \sin(\alpha - \varkappa\tau) + r \cos\gamma \cdot \cos\alpha.$$

Wenn nämlich aus der Gleichung

$$\frac{\partial F(x;y;\tau)}{\partial\tau} = 0$$

der Parameter τ ausgedrückt und dieser Wert in die Gleichung

$$F(x;\ y;\ \tau) = 0$$

eingesetzt wird, so erhält man die gesuchte Gleichung der Hüllkurve. Die Eliminierung wird in folgender Weise durchgeführt.

Man betrachte folgendes, aus zwei Gleichungen bestehendes System:

$$F(x;y;\tau) = 0\,;\quad \frac{\partial F(x;y;\tau)}{\partial\tau} = 0\,,$$

d.h. folgende zwei Gleichungen:

$$x \sin(\alpha - \varkappa\tau) - y \cos(\alpha - \varkappa\tau) = a \sin\alpha - \tau r \cos\gamma \cdot \cos\alpha;$$
$$-x\varkappa \cos(\alpha - \varkappa\tau) - y\varkappa \sin(\alpha - \varkappa\tau) = -r \cos\gamma \cdot \cos\alpha.$$

Wenn man dieses letztere Gleichungssystem nach x und y auflöst, so gewinnt man bereits das Gleichungssystem in Parameterform der Hüllkurve, d. h. die

folgende Darstellung der Koordinaten x und y mit dem Parameter τ:

$$x = \frac{1}{\varkappa}\left\{\varkappa\,(a\sin\alpha - \tau\, r\cos\gamma\cdot\cos\alpha)\sin(\alpha-\varkappa\tau) + \right.$$
$$\left. + r\cos\gamma\cdot\cos\alpha\cdot\cos(\alpha-\varkappa\tau)\right\};$$

$$y = \frac{1}{\varkappa}\left\{r\cos\gamma\cos\alpha\cdot\sin(\alpha-\varkappa\tau) + \right.$$
$$\left. + \varkappa\,(\tau\, r\cos\gamma\cdot\cos\alpha - a\sin\alpha)\cos(\alpha-\varkappa\tau)\right\}.$$

Schreibt man in die Klammerausdrücke anstelle von $r\cos\gamma$ den aus der Beziehung [2; S. 42]

$$r = R\tan\gamma$$

sich ergebenden Wert von

$$r\cos\gamma = R\sin\gamma,$$

ferner unter Einführung der Abkürzung

$$R\cos\alpha = g,$$

kann das obige Gleichungssystem auch in folgender Form angeschrieben werden:

$$x = (a\sin\alpha - \tau\sin\gamma\cdot g)\sin(\alpha-\varkappa\tau) + \frac{g\sin\gamma}{\varkappa}\cos(\alpha-\varkappa\tau);$$

$$y = (\tau g\sin\gamma - a\sin\alpha)\cos(\alpha-\varkappa\tau) + \frac{g\sin\gamma}{\varkappa}\sin(\alpha-\varkappa\tau).$$

Aus den ersten Gliedern dieser Ausdrücke klammere man den Bruch

$$\frac{g\sin\gamma}{\varkappa}$$

aus, und so wird:

$$x = \frac{g\sin\gamma}{\varkappa}\left(\frac{\varkappa a\sin\alpha}{g\sin\gamma} - \varkappa\tau\right)\sin(\alpha-\varkappa\tau) + \frac{g\sin\gamma}{\varkappa}\cos(\alpha-\varkappa\tau);$$

$$y = -\frac{g\sin\gamma}{\varkappa}\left(\frac{\varkappa a\sin\alpha}{g\sin\gamma} - \varkappa\tau\right)\cos(\alpha-\varkappa\tau) + \frac{g\sin\gamma}{\varkappa}\sin(\alpha-\varkappa\tau).$$

Drehe man diese Kurve, welche mit Hilfe des in Parameterform ausgedrückten Gleichungssystems hergestellt worden ist, um den Ursprung des Koordinatensystems um den Winkel

$$\left(\frac{\varkappa\cdot a\cdot\sin\alpha}{g\sin\gamma} - \alpha\right).$$

Dann erhält man für x und y folgende Ausdrücke:

$$x = \frac{g \sin \gamma}{\varkappa} \left(\frac{\varkappa a \sin \alpha}{g \sin \gamma} - \varkappa \tau \right) \cdot \sin \left(\frac{\varkappa a \sin \alpha}{g \sin \alpha} - \varkappa \tau \right) + \\ + \frac{g \sin \gamma}{\varkappa} \cos \left(\frac{\varkappa a \sin \alpha}{g \sin \gamma} - \varkappa \tau \right);$$

$$y = - \frac{g \sin \gamma}{\varkappa} \left(\frac{\varkappa a \sin \alpha}{g \sin \gamma} - \varkappa \tau \right) \cos \left(\frac{\varkappa a \sin \alpha}{g \sin \gamma} - \varkappa \tau \right) + \\ + \frac{g \sin \gamma}{\varkappa} \sin \left(\frac{\varkappa a \sin \alpha}{g \sin \gamma} - \varkappa \tau \right).$$

Führt man hier der Kürze halber den Winkel

$$\varphi = \frac{\varkappa a \sin \alpha}{g \sin \gamma} - \varkappa \tau$$

als einen neuen Parameter ein, dann nimmt das Gleichungssystem folgende Gestalt an:

$$x = \frac{g \sin \gamma}{\varkappa} \varphi \cdot \sin \varphi + \frac{g \sin \gamma}{\varkappa} \cos \varphi,$$

$$y = - \frac{g \sin \gamma}{\varkappa} \varphi \cdot \sin \varphi + \frac{g \sin \gamma}{\varkappa} \sin \varphi.$$

Das in dieser Weise erhaltene Gleichungssystem ist offensichtlich das in Parameterform hergestellte Gleichungssystem einer gewöhnlichen Kreisevolvente, und zwar einer Kreisevolvente, die von einem Grundkreis mit dem Radius $g \sin \gamma / \varkappa$ abgewickelt wurde. Nun bestimme man den Wert von diesem Grundkreisradius. Wenn die Zähnezahl des zu verzahnenden Rades z ist, so wird $\varkappa = 1/z$, da bei einer einmaligen vollen Umdrehung des Fräsers um seine Achse das zu verzahnende Rad lediglich einen $1/z$ Teil einer vollen Umdrehung ausführt.

In dem oben abgeleiteten Wert für den Grundkreisradius

$$\frac{g \sin \gamma}{\varkappa}$$

steht im Zähler der Wert g, welcher seiner Definition entsprechend

$$g = R \cos \alpha$$

ist. Oder anstelle von R den Wert

$$R = \frac{r}{\tan \gamma}$$

schreibend, wird

$$g = \frac{r \cos \alpha}{\tan \gamma} .$$

In diesem zuletzt angeschriebenen Bruch ist r die Konstante der Schraubenfläche (Fräserfläche), und nach der Definition ist

$$r = \frac{S}{2\pi},$$

wobei die Steigung der Schraubenfläche:

$$S = \pi m_{ax},$$

$$m_{ax} = \frac{m_n}{\cos \gamma}$$

ist (m_{ax} = axialer Modul, m_n = Normalmodul). Da jedoch beim zylindrischen Zahnrad-Wälzfräser der Normalmodul m_n die Norm bedeutet [3; S. 7], so ist für $m_n = 1$:

$$S = \frac{\pi}{\cos \gamma}$$

und in dieser Weise

$$r = \frac{1}{2 \cos \gamma} .$$

Schreibt man diesen Wert in den oben angegebenen Ausdruck von g:

$$g = \frac{\cos \alpha}{2 \sin \gamma} .$$

Wird der so gewonnene Wert für g und der von $\varkappa = 1/z$ in die vorher angegebene Beziehung des Grundkreisradius eingesetzt, so erhält man für den Grundkreisradius:

$$\frac{g \sin \gamma}{\varkappa} = \frac{\cos \alpha}{2 \sin \gamma} \sin \gamma \cdot z = \frac{z}{2} \cos \alpha .$$

Der Grundkreisradius stimmt daher mit dem Radius

$$r_a = \frac{z}{2} \cos \alpha , \quad (m = 1)$$

des Zahnrades mit der Zähnezahl z überein.

Damit wurde nachgewiesen, daß wenn man den Steigungswinkel γ des Schneckenfräsers mit einer schleifenevolventischen Grundschnecke genügend klein wählt, so nähert sich das Zahnprofil, welches durch Wälzfräser erzeugt im Stirnschnitt des Zahnrades entsteht, mit beliebiger Genauigkeit der Kreisevolvente an. Liegt nämlich ein Fräser mit kleinem Steigungswinkel vor, so entspricht die Drehung des Fräsers und des zu verzahnenden Rades zusammen einer Abwälzbewegung, in guter Annäherung der Evolventenerzeugung mittels Zahnstange. In den obigen Ausführungen wurde gezeigt, daß die geraden Erzeugenden (die Schneiden) der Fräsfläche bei der Drehung des Fräsers nacheinander in eine parallele Lage zur Stirnebene des Zahnrades, zur horizontalen Ebene $(x; y)$ gelangen, wobei in diesen Lagen die Entfernung von der horizontalen Ebene $(x; y)$ sich dauernd vermindert, wenn der Drehwinkel

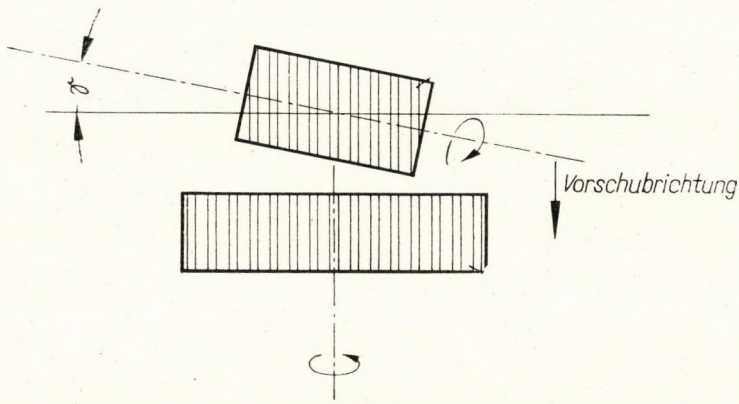


Bild 5

des Fräsers wächst, und damit gelangen diese Lagen der Schneiden um so näher zueinander und zur $(x; y)$ Ebene, je kleiner der Winkel γ ist. Betrachtet man nun genauer den bei der Abwälzung auftretenden Bewegungsvorgang, so sieht man, wie es auch aus dem Bild 5 hervorgeht, daß während der Fräser und das Zahnrad sich im Sinn der angegebenen Pfeilrichtung drehen, der Fräser gleichzeitig parallel mit der Achse des Arbeitsstückes, in Pfeilrichtung, eine Vorschubbewegung ausführt, mit dem Zweck, daß das Rad in seiner vollen Breite bearbeitet wird. Diese Vorschubbewegung bewirkt, daß die bei der Fräserdrehung entstehenden und mit der horizontalen $(x; y)$-Ebene parallelen Lagen der Erzeugenden sich in vertikaler Richtung nach unten verschieben. Die nacheinander entstehenden horizontalen Lagen der Erzeugenden beim Drehen des Fräsers nähern sich jedoch der horizontalen Ebene $(x; y)$ von unten — nach (8) ist nämlich $z < 0$ —, bei dieser Drehung verschieben sie sich also nach oben. Die entgegengesetzten Verschiebungen der horizontalen Lagen der Erzeugenden werden sich gegenseitig kompensieren. Daraus folgt, daß wenn man eine feste Stirnebene des Zahnradkörpers betrachtet, so werden die

geraden Erzeugenden (Schneiden) der Fräserfläche, im Fall einer entsprechenden Vorschubgeschwindigkeit des Fräsers, nacheinander alle in die betrachtete fixe Stirnebene gelangen, wodurch dieselbe Situation erreicht wird wie im Fall der Evolventenerzeugung mittels Zahnstange.

V. Anwendung des Evolventen-Kriteriums beim Zahnradprofil, das mittels Schneckenfräser mit schleifenevolventischer Grundschnecke erzeugt wurde

Für den exakten Nachweis der Folgerung, die am Ende des vorigen Kapitels gezogen wurde, sei das analytische Kriterium, welches in einer vorigen Arbeit [4] abgeleitet wurde, angewandt. Mit Hilfe dieses Kriteriums kann entschieden werden, ob ein Zahnradprofil, welches mit irgendeinem Wälzwerkzeug erzeugt worden ist, eine gewöhnliche Kreisevolvente sein kann, oder eine hiervon abweichende Kurve sein muß. Dieses Kriterium wird jetzt bei einem Zahnprofil angewandt, das mit einem Schneckenfräser in Schleifenevolventenform erzeugt worden ist, wobei die Herstellung des Fräsers mittels schräg gestellten Stahls erfolgte. Die Anwendung erfolgt in der Weise, daß außer der Drehbewegung des Fräsers auch seine geradlinige Vorschubbewegung berücksichtigt wird, d. h. die Wälzbewegung des Fräsers wird in ihrer Gesamtheit betrachtet, wobei noch die Voraussetzung besteht, daß die Zahl der Fräserschneiden unendlich ist.

Zur Anwendung des Evolventen-Kriteriums muß man das zwei Parameter enthaltende Gleichungssystem der Kurvenschar kennen, welche in der Stirnebene des Zahradkörpers durch die Wälzbewegung des Fräsers, durch seine Drehung und Vorschubbewegung entsteht. In einem Koordinatensystem $(x; y)$, welches in der Stirnebene eines Zahnrades liegt, und wobei den Koordinatenurpsrung der Mittelpunkt des Zahnrades bildet, sei das in Parameterform ausgedrückte Gleichungssystem dieser Kurvenschar:

$$x = X(t; \Phi) + a;$$
$$y = Y(t; \Phi),$$

wo a die Abszisse des Fräsermittelpunktes bedeutet. Um die Funktionen $X(t; \Phi)$, $Y(t; \Phi)$ mit zwei Veränderlichen zu erhalten, muß man die Durchdringungskurven der um den Winkel Φ gedrehten Schraubenfläche und der Stirnebene $z = 0$ bestimmen; hierbei wird noch berücksichtigt, daß die Schraubenfläche (6′) jetzt außer der Drehung noch eine Vorschubbewegung parallel mit der Achse des Werkstückes ausführt. Die Berücksichtigung dieser Vorschubbewegung bedeutet soviel, daß im Gleichungssystem unter (6′) der um den Winkel Φ gedrehten Schraubenfläche im Ausdruck von z noch das

Glied $c \cdot \Phi$ (c = konstant) auftritt, da die Vorschubbewegung des Fräsers parallel mit der z-Achse ist. Dabei erscheinen die unter (6′) angeführten Gleichungen, nach dem Parameter t geordnet, in folgender Form:

$$\begin{aligned} x &= k \cos(\tau + \Phi + \chi) + \cos\psi \cdot \sin(\tau + \Phi + \chi) \cdot t; \\ y &= -k \sin(\tau + \Phi + \chi) \cdot \sin\gamma + (r\,\tau - \tau_0)\cos\gamma + \\ &\quad + [\cos\psi \cdot \cos(\tau + \Phi + \chi) \cdot \sin\gamma + \sin\psi \cdot \cos\gamma]t; \\ z &= -k \sin(\tau + \Phi + \chi) \cdot \cos\gamma - (r\,\tau - \tau_0)\sin\gamma + \\ &\quad + [\cos\psi \cdot \cos(\tau + \Phi + \chi) \cdot \cos\gamma - \sin\psi \cdot \sin\gamma]t - c \cdot \Phi. \end{aligned} \tag{6″}$$

Mit dem Gleichungssystem von drei Parametern, wobei τ, Φ und t die Parameter sind, wurde eine Flächenschar hergestellt; ihre Durchdringung mit der Stirnebene $z = 0$ läßt eine Kurvenschar in der Stirnebene entstehen. Diese Kurvenschar kann durch ein Gleichungssystem mit zwei Parametern hergestellt werden. Wenn die Veränderlichen Φ und t — wie das bereits in dem vorher angeschriebenen Gleichungssystem der Fall war — als die beiden Parameter gewählt werden, so muß, da $z = 0$ ist, entsprechend der dritten Gleichung aus (6″), zwischen dem Parameter τ und dem Parameterpaar (Φ; t) eine Beziehung bestehen. Entsprechend der Annahme $z = 0$ muß die dritte Gleichung aus (6″) für jeden Wert von t verschwinden. Diese Gleichung

$$\begin{aligned} &-k \sin(\tau + \Phi + \chi)\cos\gamma - (r\,\tau - \tau_0)\sin\gamma + \\ &\quad + [\cos\psi \cdot \cos(\tau + \Phi + \chi) \cdot \cos\gamma - \sin\psi \cdot \sin\gamma]\,t - c\,\Phi = 0 \end{aligned}$$

verschwindet dann bei jedem Wert von t, wenn

$$\cos\psi \cdot \cos(\tau + \Phi + \chi) \cdot \cos\gamma - \sin\psi \cdot \sin\gamma = 0 \tag{9}$$

und außerdem wenn

$$k \sin(\tau + \Phi + \chi) \cdot \cos\gamma + (r\,\tau - \tau_0)\sin\gamma + c\,\Phi = 0 \tag{9′}$$

ist.

Aus (9) ergibt sich unmittelbar, daß

$$\cos(\tau + \Phi + \chi) = \tan\psi \cdot \tan\gamma \tag{9″}$$

und aus (9′), daß

$$\sin(\tau + \Phi + \chi) = \frac{(\tau_0 - r\tau)\sin\gamma - c\Phi}{k\cos\gamma}$$

ist.

Aus dem jetzt gewonnenen Wert für $\cos(\tau + \Phi + \chi)$ ist

$$\tau + \Phi + \chi = \arccos(\tan\psi \cdot \tan\gamma),$$

woraus

$$\tau = \arccos(\tan\psi \cdot \tan\gamma) - \chi - \Phi$$

oder kürzer ausgedrückt:

$$\tau = a - \Phi$$

ist, wobei

$$a = \arccos(\tan\psi \cdot \tan\gamma) - \chi.$$

Nach den Beziehungen (2) und (3) ist

$$\cos\chi = \frac{\sin\gamma \cdot \sin\alpha}{\cos\psi}, \tag{10}$$

ferner, da gemäß (3) $\sin\psi = \sin\alpha \cdot \cos\gamma$ gilt, so ist

$$\sin\alpha = \frac{\sin\psi}{\cos\gamma},$$

setzt man dies nun in den vorherigen Ausdruck von $\cos\chi$ ein, so wird:

$$\cos\chi = \frac{\sin\gamma}{\cos\psi} \cdot \frac{\sin\psi}{\cos\gamma} = \tan\psi \cdot \tan\gamma.$$

Vergleicht man die letzte Gleichheit mit der Beziehung unter (9″):

$$\cos(\tau + \Phi + \chi) = \tan\psi \cdot \tan\gamma,$$

so kann man wegen der Gleichheit der rechten Seiten folgendes anschreiben:

$$\chi = \tau + \Phi + \chi,$$

d.h.

$$\tau + \Phi = 0$$

und in dieser Weise ist

$$\tau = -\Phi.$$

Man setze diesen Wert von τ in die Ausdrücke von x und y der Beziehungen unter (6″) ein. In diesem Fall, da nach (10), (2) und (3)

$$\cos\chi = \frac{\sin\gamma \cdot \sin\alpha}{\cos\psi}, \qquad \sin\chi = \frac{\cos\alpha}{\cos\psi}, \qquad \sin\psi = \sin\alpha \cdot \cos\gamma,$$

erhält man für x und y:

$$\begin{aligned} x &= k\cos\chi + \cos\psi\cdot\sin\chi\cdot t = k\,\frac{\sin\gamma\cdot\sin\alpha}{\cos\psi} + t\cos\alpha\,; \\ y &= -k\sin\chi\cdot\sin\gamma + (-r\Phi - \tau_0)\cos\gamma + \\ &\quad + (\cos\psi\cdot\cos\chi\cdot\sin\gamma + \sin\psi\cdot\cos\gamma)\,t = \\ &= -k\frac{\cos\alpha}{\cos\psi}\sin\gamma + (-r\Phi - \tau_0)\cos\gamma + t\cdot\sin\alpha\,. \end{aligned} \tag{11}$$

Die Beziehung unter (9′), sofern $\tau = -\Phi$ ist, wird nur dann bestehen, wenn für jeden Wert von Φ identisch gilt, daß

$$k\sin\chi\cdot\cos\gamma - \tau_0\sin\gamma - \Phi(r\sin\gamma - c) = 0\,.$$

Da nach (7)

$$\tau_0 = \frac{k\cos\alpha}{\tan\gamma\cdot\cos\psi}$$

ferner nach (2) und (3)

$$\sin\chi = \frac{\cos\alpha}{\cos\psi}$$

ist, so gilt

$$k\sin\chi\cdot\cos\gamma - \tau_0\sin\gamma = k\,\frac{\cos\alpha}{\cos\psi}\cos\gamma - \frac{k\cos\alpha}{\tan\gamma\cdot\cos\psi}\sin\gamma = 0\,.$$

Daher besteht die Beziehung (9′) für jeden Wert von Φ in dem Fall, wenn in der vorigen Gleichung der Koeffizient von Φ gleich 0 ist, d.h.

$$r\sin\gamma - c = 0\,.$$

Dies bedeutet soviel, daß der Wert von c, d.h. der Wert des Fräservorschubs, bei einer Winkeldrehung vom Wert 1:

$$c = r\sin\gamma\,.$$

Die Gleichungen unter (11) liefern jetzt das in Parameterform ausgedrückte Gleichungssystem der Kurvenschar, welche infolge der Durchdringung der Fräserfläche und der Stirnebene $z = 0$ entsteht, wobei die Fräserfläche eine Drehbewegung um ihre Achse und außerdem eine geradlinige Vorschubbewegung ausführt. Dieses Gleichungssystem kann unter Berücksichtigung, daß $r\cos\gamma = 1/2$, ferner daß nach (7)

$$\tau_0 = \frac{k\cos\alpha}{\tan\gamma\cdot\cos\psi}$$

ist, auch in folgender Form angeschrieben werden:

$$x = \frac{k \sin\gamma \cdot \sin\alpha}{\cos\psi} + t\cos\alpha\,;$$
$$y = -\frac{k\cos\alpha}{\cos\psi \cdot \sin\gamma} - \frac{1}{2}\Phi + t\sin\alpha\,. \tag{11'}$$

Bei diesen Funktionen x und y mit zwei Veränderlichen sei das Evolventen-Kriterium angewandt, damit über den Evolventencharakter des durch Wälzverfahren hergestellten Zahnprofils entschieden wird. Da die partiellen Ableitungen nach t und Φ

$$x_t = \cos\alpha\,,\quad y_t = \sin\alpha\,;$$
$$x_\Phi = 0\,,\qquad y_\Phi = -\frac{1}{2}$$

sind, ist daher das Quadrat der im Kriterium auftretenden Funktionaldeterminanten:

$$\begin{vmatrix} x_\Phi & y_\Phi \\ x_t & y_t \end{vmatrix}^2 = \begin{vmatrix} 0 & -\frac{1}{2} \\ \cos\alpha & \sin\alpha \end{vmatrix}^2 = \frac{\cos^2\alpha}{4}\,;$$

ferner ist die ebenfalls auftretende Quadratsumme:

$$x_t^2 + y_t^2 = 1\,.$$

Mit den obigen Ausdrücken wird die Gleichung, welche die Erfüllung des Kriteriums ausdrückt:

$$\begin{vmatrix} x_\Phi & y_\Phi \\ x_t & y_t \end{vmatrix}^2 - \varkappa^2 r_a^2 (x_t^2 + y_t^2) = \frac{\cos^2\alpha}{4} - \varkappa^2 r_a^2 = 0\,.$$

Sie wird tatsächlich erfüllt, da

$$r_a = \frac{z}{2}\cos\alpha$$

und

$$\varkappa = \frac{1}{z}$$

(z = Zähnezahl) sind.

Nach diesen Ausführungen kann leicht nachgewiesen werden, daß das abgewälzte Zahnprofil in der Tat eine Kreisevolvente ist. Dazu muß gezeigt

werden, daß in dem Gleichungssystem unter (8a) der Arbeit [4] jedem Wert von Φ ein Wert von t zugeordnet werden kann, bei dem die beiden Gleichungen unter (8a) gleichzeitig erfüllt werden. Die Erfüllung des Gleichungssystem (8a) in der Arbeit [4] ist nicht nur die notwendige, sondern gleichzeitig auch die *hinreichende* Bedingung dafür, daß das durch Abwälzung hergestellte Zahnprofil eine gewöhnliche Kreisevolvente sei.

Betrachte man die zweite Gleichung unter (8a), die in der Arbeit [4] abgeleitet ist, und die mit den jetztigen Bezeichnungen folgende Form hat:

$$\begin{vmatrix} x_t & x_\Phi \\ y_t & y_\Phi \end{vmatrix} + \varkappa \{x x_t + y y_t\} = 0 .$$

Wenn man hier die Ausdrücke für x und y aus (11′) und die Werte der partiellen Ableitungen einsetzt, so gewinnt man:

$$\begin{vmatrix} x_t & x_\Phi \\ y_t & y_\Phi \end{vmatrix} + \varkappa \{x x_t + y y_t\} = -\frac{\cos\alpha}{2} +$$

$$+ \frac{1}{z}\left\{ -\frac{k \sin\alpha \cdot \cos\alpha \cdot \cos^2\gamma}{\cos\psi \sin\gamma} - \frac{\sin\alpha}{2}\Phi + t \right\} = 0 ,$$

woraus

$$t = \frac{\sin\alpha}{2} \cdot \Phi + k \frac{\sin\alpha \cdot \cos\alpha \cdot \cos^2\gamma}{\cos\psi \cdot \sin\gamma} + \frac{z}{2}\cos\alpha$$

ist.

Unter dessen Berücksichtigung wird die linke Seite der ersten Gleichung unter (8a):

$$x x_\Phi + y y_\Phi = \frac{1}{2}\left(\frac{k\cos\alpha}{\cos\psi \cdot \sin\gamma} + \frac{1}{2}\Phi - t\sin\alpha \right) =$$

$$= \frac{\cos\alpha}{2}\left(\frac{k}{\cos\psi \cdot \sin\gamma} + \frac{1}{2}\Phi\cos\alpha - \frac{k\sin^2\alpha \cdot \cos^2\gamma}{\cos\psi \cdot \sin\gamma} - \frac{z}{2}\sin\alpha \right)$$

und die rechte Seite

$$r_a \sqrt{(x_\Phi - \varkappa y)^2 + (\varkappa x + y_\Phi)^2} =$$

$$= \frac{\cos\alpha}{2} \sqrt{\left(\frac{k}{\cos\psi \cdot \sin\gamma} + \frac{1}{2}\Phi\cos\alpha - \frac{k\sin^2\alpha \cdot \cos^2\gamma}{\cos\psi \cdot \sin\gamma} - \frac{z}{2}\sin\alpha \right)^2} .$$

So wird die erste Gleichung von (8a) ebenfalls erfüllt, was zu beweisen war.

ANHANG

Es wird nachgewiesen, daß der Stirnschnitt (der zur z-Achse senkrechte Schnitt) der Schraubenfläche unter (1*) eine Schleifen-Kreisevolvente ist, und in dieser Weise das Gleichungssystem unter (1*) die Fläche einer verschlungenen Evolventenschnecke erzeugt.

Aus der Gleichung $z = 0$ von (1*) ergibt sich der Wert des Parameters t zu:

$$t = \frac{k \cos \alpha}{\tan \gamma \cdot \cos \psi \cdot \sin \psi} - \frac{r}{\sin \psi} \tau .$$

Schreibt man diesen Wert von t in die Ausdrücke x und y von (1*), so erhält man das Gleichungssystem in Parameterform der im Stirnschnitt entstehenden Kurve:

$$x = k \cos (\tau + \chi) + \frac{k \cos \alpha}{\tan \gamma \sin \psi} \sin (\tau + \chi) - r \cot \psi \cdot \sin (\tau + \chi) \cdot \tau ;$$

$$y = k \sin (\tau + \chi) - \frac{k \cos \alpha}{\tan \gamma \cdot \sin \psi} \cos (\tau + \chi) + r \cot \psi \cdot \cos (\tau + \chi) \cdot \tau .$$

Wird aus dem zweiten und dritten Glied dieser Gleichungen der Faktor $r \cot \psi$ ausgeklammert, so wird

$$x = k \cos (\tau + \chi) - r \cot \psi \left(\tau - \frac{k \cos \alpha}{r \tan \gamma \cdot \cos \psi} \right) \sin (\tau + \chi) ;$$

$$y = k \sin (\tau + \chi) + r \cot \psi \left(\tau - \frac{k \cos \alpha}{r \tan \gamma \cdot \cos \psi} \right) \cos (\tau + \chi) .$$

Man drehe die in dieser Weise hergestellte Kurve um den Ursprung 0 um den Winkel

$$- \left(\chi + \frac{k \cos \alpha}{r \tan \gamma \cdot \cos \psi} \right) .$$

Wenn man jetzt die Koordinaten der Punkte der gedrehten Kurve erneut mit x und y bezeichnet, so erhält man folgende Ausdrücke:

$$x = k \cos \left(\tau - \frac{k \cos \alpha}{r \tan \gamma \cdot \cos \psi} \right) - r \cot \psi \left(\tau - \frac{k \cos \alpha}{r \tan \gamma \cdot \cos \psi} \right) \sin \left(\tau - \frac{k \cos \alpha}{r \tan \gamma \cdot \cos \psi} \right) ;$$

$$y = k \sin \left(\tau - \frac{k \cos \alpha}{r \tan \gamma \cdot \cos \psi} \right) + r \cot \psi \left(\tau - \frac{k \cos \alpha}{r \tan \gamma \cdot \cos \psi} \right) \cos \left(\tau - \frac{k \cos \alpha}{r \tan \gamma \cdot \cos \psi} \right) .$$

Führt man hier anstelle von τ den neuen Parameter

$$\varphi = \tau - \frac{k \cos \alpha}{r \tan \gamma \cdot \cos \psi}$$

ein, so wird:

$$x = k \cos \varphi - r \cot \psi \cdot \varphi \cdot \sin \varphi ;$$

$$y = k \sin \varphi + r \cot \psi \cdot \varphi \cdot \cos \varphi .$$

Dieses letztere Gleichungssystem erzeugt eine Kreisevolvente, welche vom Grundkreis mit dem Radius

$$r \cot \psi$$

durch Abwälzung hergestellt worden ist; der obige Radius ist gleichzeitig der Grundzylinderradius der Schnecke. Wenn für diesen Radius gilt, daß

$$k < r \cot \psi,$$

d. h. der Kehlzylinderradius der Schnecke kleiner als der Radius des Grundzylinders ist, so wird vom obigen Gleichungssystem in Parameterform eine Schleifen-Kreisevolvente hergestellt. Dazu muß man folgende Ungleichung beweisen, wobei der im Abschnitt II für den Kehlzylinderradius k abgeleitete Ausdruck verwendet wird:

$$\left(R - \frac{\pi}{4} \cdot \frac{\cos \gamma}{\tan \alpha}\right) \frac{\sin \gamma \cdot \sin \alpha}{\sqrt{1 - \cos^2 \gamma \cdot \sin^2 \alpha}} < r \cot \psi.$$

Da hier nach (3)

$$\cos \psi = \sqrt{1 - \sin^2 \alpha \cdot \cos^2 \gamma}$$

und so

$$\cot \psi = \frac{\sqrt{1 - \sin^2 \alpha \cdot \cos^2 \gamma}}{\sin \alpha \cdot \cos \gamma},$$

ferner aus (4′)

$$R = \frac{r}{\tan \gamma}$$

ist, ist die zu beweisende Ungleichung gleichwertig mit folgendem Ausdruck:

$$\left(\frac{r}{\tan \gamma} - \frac{\pi}{4} \cdot \frac{\cos \gamma}{\tan \alpha}\right) \sin^2 \alpha \cdot \sin \gamma \cdot \cos \gamma < r(1 - \sin^2 \alpha \cdot \cos^2 \gamma).$$

Führt man die Multiplikation auf der linken Seite der Ungleichung durch, so wird:

$$r \sin^2 \alpha \cdot \cos^2 \gamma - \frac{\pi}{4} \sin \alpha \cdot \cos \alpha \cdot \sin \gamma \cdot \cos^2 \gamma < r(1 - \sin^2 \alpha \cdot \cos^2 \gamma).$$

Diese Ungleichung besteht offensichtlich, wenn die rechte Seite größer ist als das erste Glied der linken Seite, d. h. wenn folgende Beziehung besteht:

$$\cos^2 \gamma \cdot \sin^2 \alpha < 1 - \sin^2 \alpha \cdot \cos^2 \gamma.$$

Aber

$$\sin^2 \alpha < 1 - \sin^2 \alpha$$

wenn

$$\alpha < 45°$$

ist.

In dieser Weise besteht für $\alpha = 20°$, da $\cos \gamma \leqq 1$ ist, offensichtlich die obige Ungleichung

$$\cos^2 \gamma \cdot \sin^2 \alpha < 1 - \sin^2 \alpha \cdot \cos^2 \gamma.$$

Damit wurde nachgewiesen, daß

$$k < r \cot \psi$$

und somit bedeuten die Beziehungen unter (1*) das in Parameterform ausgedrückte Gleichungssystem der Fläche der verschlungenen Evolventenschnecke.

SCHRIFTTUM

1. Lipka, I.: Geometrie der Abwälzfräser mit Grundschneckenevolventenförmiger Ausführung. *Acta Techn. Hung.* **61** (1968), 3—4.
2. Szeniczei, L.: Csigahajtóművek (Schneckengetriebe); Budapest 1957.

3. SZENICZEI, L.: Fogaskerék lefejtő marók geometriája (Geometrie der Zahnrad-Wälzfräser); Budapest 1955.
4. LIPKA, I.: Wältheorie von Zahnrädern mit einem Evolventenprofil. *Acta Techn. Hung.* **60** (1968), 1—2.

The Geometry of Generating Cutters with Convolute Basic Worm. The author investigates gear profiles generated by a hob the basic worm of which is a convolute worm made on a lathe with a dipped tool. He shows that a tooth profile generated with such a hob approaches the ordinary involute of a circle the better, the smaller the helix angle of the cutter is. The author finally applies the analytical criterion for deciding the involute character of the generated tooth profile to the gear profile made with a convolute hob.

Géométrie des fraises vis-mère à vis de base convolutée. L'auteur examine des profils de dentures engendrées par des fraises vis-mère, dont la vis de base est une vis convolutée taillée au tour par un couteau incliné. Il montre qu'un profil de dent engendré par une telle fraise approche d'autant plus la développante circulaire ordinaire que le pas d'hélice de la fraise est plus petit. Le critère analytique, servant à déterminer le caractère de développante du profil engendré est enfin appliqué au profil de dent produit par une fraise vis-mère à vis de base convolutée.

Геометрия обкаточных фрез с конволютным червяком *(И. Липка)*. Автор исследует профили зубчатых колес, изготовленных обкаткой при помощи таких червячных фрез, червяк которых представляет собою конволютный червяк, изготовленный на токарном станке при помощи установленного под углом резца. Показано, что изготовленный при помощи такой червячной фрезы методом обкатки профиль зуба приближается к обычной круговой эвольвенте тем лучше, чем меньше угол шага фрезы. Наконец, аналитический дается критерий, служащий для определения эвольвентного характера обкатанного профиля зуба, для профиля зубчатого колеса, изготовленного при помощи конволютной червячной фрезы.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (3—4), pp 401—412 (1968)

ZUR FRAGE DER KONZENTRATIONSABHÄNGIGKEIT DER POISSON-ZAHL*

H. WAWRA
BERGAKADEMIE FREIBERG (SACHSEN)

[Eingegangen am 27 April, 1967]

Die Konzentrationsabhängigkeit der Poisson-Zahl nach Indirektmessung an Kupfer- und Messingzylindern wird detailliert dargestellt und diskutiert. Man unterscheidet Typen μI als Konzentrationsdiagramme deformierter Werkstoffe, Typen μII weichgeglühter Proben und Mischtypen, die alle durch charakteristische Extremwerte ausgezeichnet sind, vorwiegend als Ergebnis der Abweichungen von statistischer Verteilung der Kristallite des Vielkristallhaufwerkes.

I. Einleitung

Werte der Poisson-Zahl unter 0,25 und über 0,5 wurden bei metallischen Werkstoffen mit größerer Skepsis [1] betrachtet. Einflüsse des Prüfverfahrens sollten dabei eine ausschlaggebende Rolle spielen (z. B. Anomalien durch Anwendung indirekter Bestimmungsverfahren), so daß der Vorschlag gemacht wird, auf die Angabe derartiger Werte zu verzichten, da begründete Rückschlüsse auf die Werkstoffstruktur offenbar nicht möglich sind. Die seinerzeit berechtigte Forderung darf nun aber fallen gelassen werden, da einmal auch mit direkten Bestimmungsmethoden u. a. die obere Grenze von 0,5 überschritten wurde [2], zum andern vom Verfasser darauf aufmerksam gemacht wird, daß der Wert, der durch indirekte Bestimmungsmethoden nach für Isotropie gültigen Formeln berechnet ist, anomale Poisson-Zahlen ergibt, die durchaus gute Rückschlüsse auf die Strukturcharakteristik des Werkstoffes ermöglichen. LARSON hat an 11 verschiedenen Titanlegierungen sporadische Tests im Hinblick auf Konzentrationsänderung, Strukturwechsel und Einflüsse auf die Poisson-Zahl durchgeführt [3] und dabei (mit Hilfe einer direkten Bestimmungsmethode) sehr schöne Werte vorlegen können, von denen weit über 50 einen Betrag über 0,5 ergaben, wobei neuerdings auch Poisson-Zahlen mit negativem Vorzeichen [4, 5] als mögliche Indikation bisher noch unbekannter Strukturveränderungen angesehen werden.

Inwieweit die Poisson-Zahl in Abhängigkeit von den Konzentrationsverhältnissen einer α-Phase bzw. auch unterschiedlichen Entfestigungsgraden

* Eine Weiterführung der Dr.-Ing. Dissertation.

eine Indikation von Strukturänderungen im Werkstoff ist, wird im folgenden in spezifischer Form detailliert und umfassend am Beispiel des Messings experimentell untersucht.

II. Experimentelle Methodik

Es wurden die Messinglegierungen Ms 99, Ms 94, Ms 90, Ms 80 und Ms 72 (die Zahl hinter dem Kurzzeichen gibt den Kupfergehalt an) sowie Elektrolytkupferbolzen mit Schale stranggepreßt und anschließend in drei Zügen insgesamt 76% kaltgezogen. An davon abgeschnittenen zylindrischen Probekörpern erfolgte eine Anlaßbehandlung bei $T = 160$, 200, 240, 280, 320, 360, 400 und 500 °C, sowie $t = 10$, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 140 und 180 Minuten im Hochtemperatur-Trockenschrank. Diese thermische Einwirkung ist bei allen 6 Werkstoffen völlig analog erfolgt, wobei auch für jede Anlaßtemperatur und Anlaßzeit eine neue Probe vorlag, die ohne besonderes Auswahlprinzip den etwa 7 m langen gezogenen Stangen entnommen wurden.

Zur Berechnung des Youngschen Moduls (E-Moduls) aus der Longitudinalschwingung der zylindrischen Probekörper wurden folgende Grundformeln herangezogen:

$$v_{\text{long.}} = \sqrt{\frac{E}{\varrho}} \tag{1}$$

$$f_p = \frac{p v_{\text{long.}}}{2l}, \tag{2}$$

$$E = \frac{4l^2\varrho}{p^2} (f_p^{\text{long.}})^2; \tag{3}$$

dabei bedeuten: $v_{\text{long.}}$ die Schallgeschwindigkeit in longitudinaler Richtung, E ist der Youngsche Modul, ϱ ist die Dichte, l die Stablänge, f die Frequenz und p die Ordnungszahl der Schwingung. Wird dabei [6] Formel (1) in Formel (2) eingesetzt und nach dem Youngschen Modul umgeformt (3), so erhält man diesen Wert E proportional abhängig von der Frequenz. Für die Grundwelle der Longitudinalschwingung unter Berücksichtigung der Beziehung

$$\text{Massendichte} = \text{Spezifisches Gewicht}/g \tag{4}$$

erhält man für den Youngschen Modul eines runden Stabes:

$$E = 4{,}0775 \cdot 10^{-8} \gamma l^2 (f_1^{\text{long.}})^2. \tag{5}$$

Hier erscheint der Youngsche Modul in kp/mm², wenn l in cm und das spezifische Gewicht γ in p/cm³ eingesetzt wird. Statt des spezifischen Gewichtes kann auch nach der Beziehung

$$\text{Spezifisches Gewicht} = P \text{ (Gewicht)}/d^2 \cdot l \text{ (Volumen)} \tag{6}$$

die Formel (5) umgeformt werden zu

$$E = 5{,}1916 \cdot 10^{-8} \cdot P \frac{l}{d^2} (f_1^{\text{long.}})^2, \tag{7}$$

wenn P das Probengewicht in pond, die Länge l und der Probendurchmesser d in cm, und die longitudinale Grundfrequenz f_1^{long} in Hz gemessen werden.

Die Formel zur Berechnung des Torsionsmoduls G ergibt sich in völlig analoger Form, wenn für die Schallgeschwindigkeit der Longitudinalwelle die der Torsionswelle, für den Youngschen Modul E der Torsionsmodul G eingesetzt wird. Für einen zylindrischen Stab erhält man also über

$$v_{\text{tors.}} = \sqrt{\frac{G}{\varrho}}, \tag{8}$$

sowie

$$f_p = \frac{p v_{\text{tors.}}}{2 \cdot l} \tag{9}$$

die Beziehung:

$$G = \frac{4 l^2 \varrho}{p^2} (f_p^{\text{tors.}})^2; \tag{10}$$

und für die Grundwelle der Torsionsschwingung runder Stäbe unter Einbeziehung von Formel (6) wiederum

$$G = 5{,}1916 \cdot 10^{-8} \cdot P \frac{l}{d^2} (f_1^{\text{tors.}})^2. \tag{11}$$

Da sich die bloße Frequenzangabe bei derartigen Versuchen mehr und mehr einbürgert [7], wurde auch analog [8] auf besondere Korrekturen dieser E- und G-Werte verzichtet. Die Poisson-Zahl ist dann über

$$\mu = \frac{E}{2G} - 1 \tag{12}$$

bestimmt worden, wobei der Fehler von E und G etwa 0,1%, der von μ etwa 0,3 ÷ 0,7% beträgt. Die Zeit zwischen Kaltverformung und Durchführung der Messungen in dem Frequenzmeßgerät Elastomat [6] betrug etwa 3 Monate.

III. Experimentelle Ergebnisse

Alle Bilder zeigen die Abhängigkeit der μ-, E-, G-Werte bzw. Vickershärtewerte (HV) von der Zinkkonzentration von α-Messing bei einer jeweils konstanten Anlaßtemperatur T und variierter Anlaßzeit. So veranschaulichen die Bilder 1, 2 und 3 die μ-, E-, und G-Werte für eine Anlaßtemperatur von $T = 200$ °C. Bild 4 bringt den Verlauf der Poisson-Zahl für $T = 240$ °C. Die Bilder 5, 6, 7 und 8 zeigen den Verlauf der μ-, E-, G- und HV-Werte bei $T =$

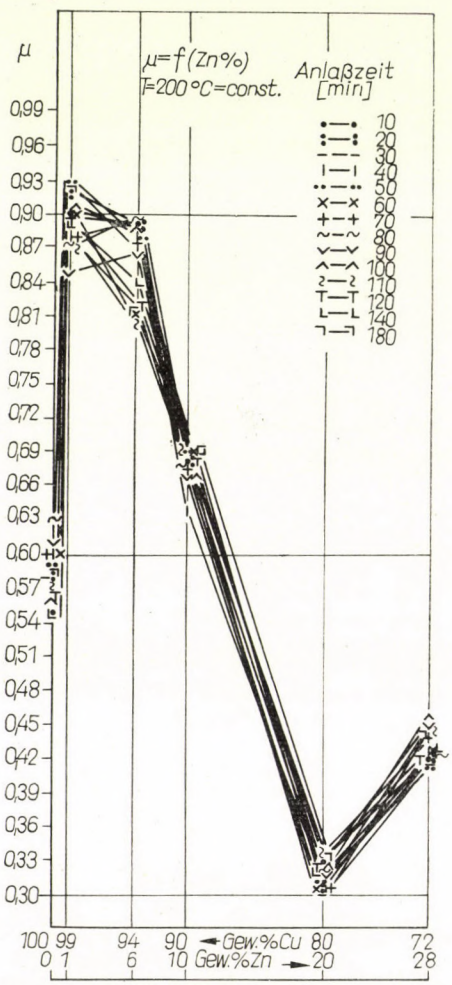


Bild 1. Konzentrationsabhängigkeit der Poisson-Zahl von verformten α-Messingstangen (Anlaßtemperatur $T = 200$ °C)

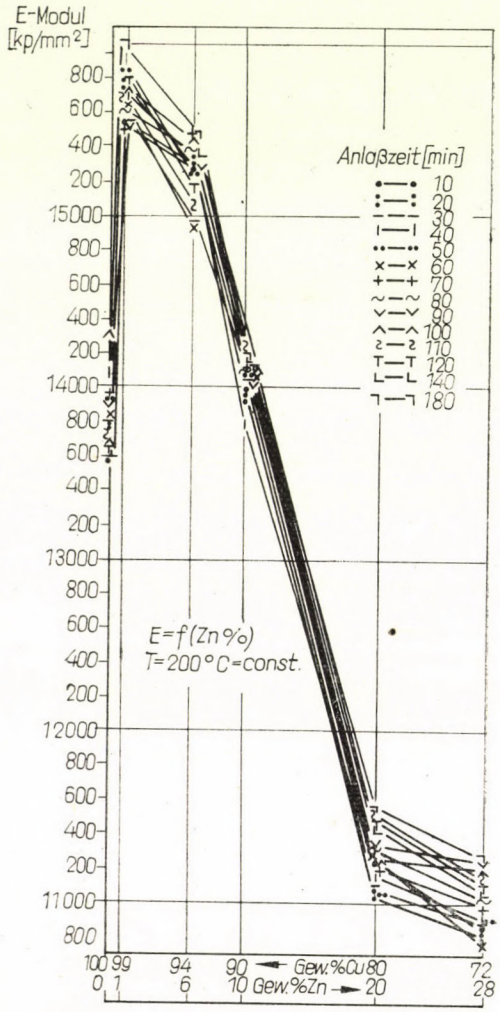


Bild 2. Konzentrationsabhängigkeit des Youngschen Moduls von verformten α-Messingstangen (Anlaßtemperatur $T = 200$ °C)

$= 320$ bzw. 500; Bild 9 den μ-Wert bei $T = 360$ °C; und Bild 13 den μ-Wert bei $T = 500$ °C. Die Abhängigkeiten für $T = 400$ °C bringen die Bilder 10, 11 und 12. Eine Beschreibung dieser Darstellungen kann zweckmäßig in Verbindung mit ihrer Diskussion erfolgen.

IV. Diskussion

Bild 1 zeigt den Verlauf der Poisson-Zahl im schwach angelassenen Zustand. Bedeutende Veränderungen der Gesamtcharakteristik gegenüber

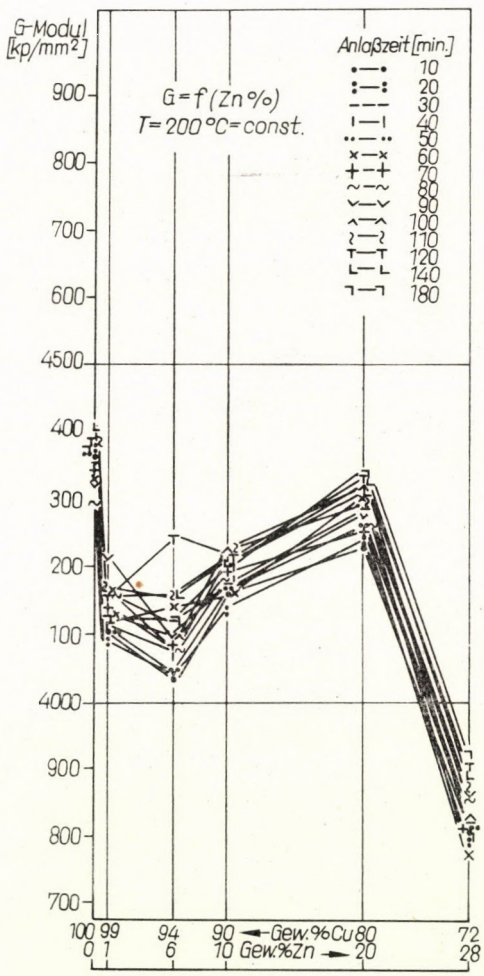


Bild 3. Konzentrationsabhängigkeit des Torsionsmoduls von verformten α-Messingstangen (Anlaßtemperatur $T = 200$ °C)

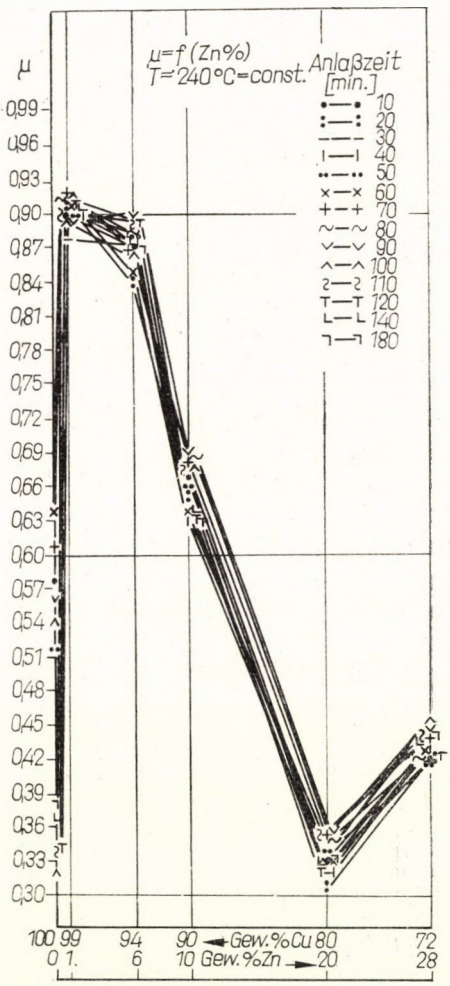


Bild 4. Konzentrationsabhängigkeit der Poisson-Zahl von verformten α-Messingstangen (Anlaßtemperatur $T = 240$ °C)

nicht angelassenen Proben sind nicht festzustellen, so daß dieser Verlauf typisch für unterhalb Rekristallisationstemperatur angelassene Werkstoffe ist. Dabei wird im folgenden nicht mehr besonders darauf hingewiesen, daß es sich bei diesen Ergebnissen um die beschriebenen Kupfer- und Messingwerkstoffe handelt, und die entzelnen Werte durch bewußtes Anwenden einer nur für Isotropie gültigen Formel entstanden, also hier mit anomalen Zahlen im allgemenien gearbeitet wird. Diese sind in Bild 1, wie es auch durchgeführte Texturaufnahmen belegen, das Ergebnis der Konzentrationsabhängigkeit des Quotienten $Q_{[111]/[100]}$ der vorliegenden Doppelfaser-Deformationstextur [111] +

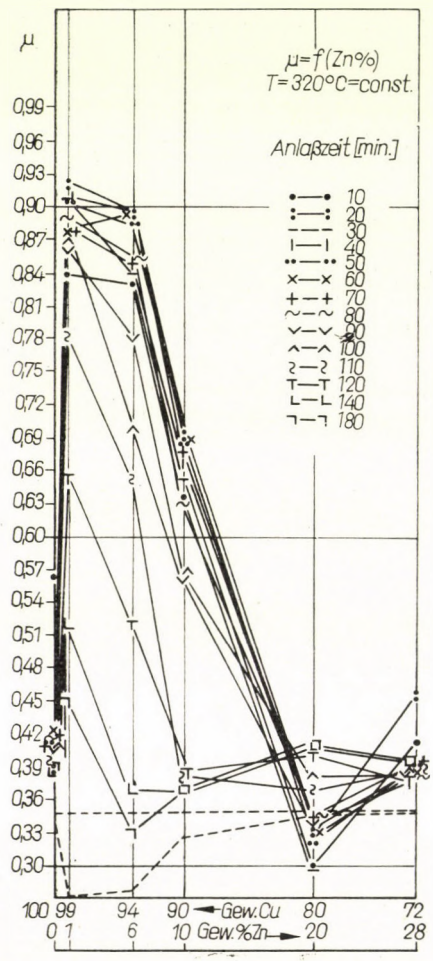


Bild 5. Konzentrationsabhängigkeit der Poisson-Zahl von verformten α-Messingstangen (Anlaßtemperatur $T = 320$ °C)

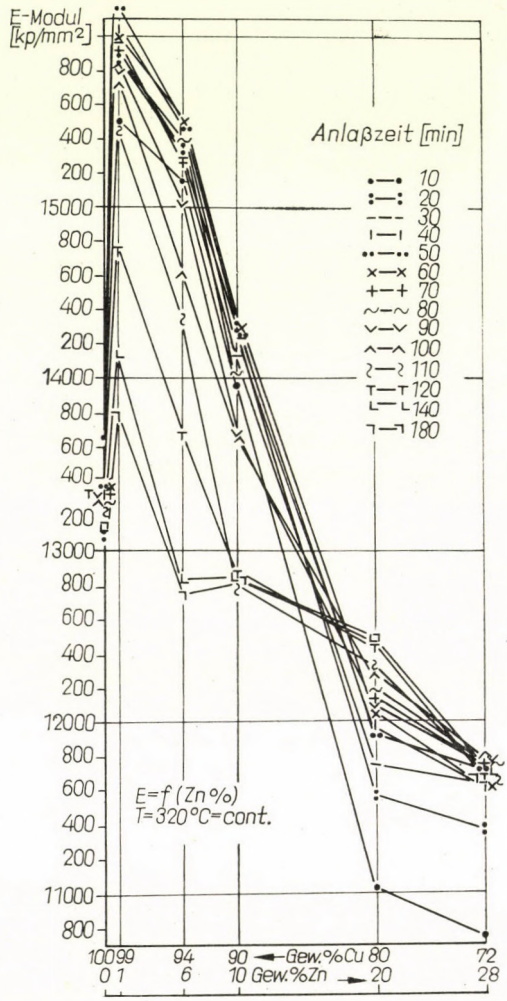


Bild 6. Konzentrationsabhängigkeit des Youngschen Moduls von verformten α-Messingstangen (Anlaßtemperatur $T = 320$ °C)

+ [100]. Daneben sind die Konzentrationsabhängigkeit von Deformationsspannungen, Änderungen des den Werkstoff durchsetzenden Netzwerkes (z. B. Korngrenzen), Rißstrukturen u. a. m. zu beachten. Wir wollen uns vorwiegend mit dem Texturfaktor befassen, der hier oft bei der Deutung der Ergebnisse dominierende Bedeutung besitzt.

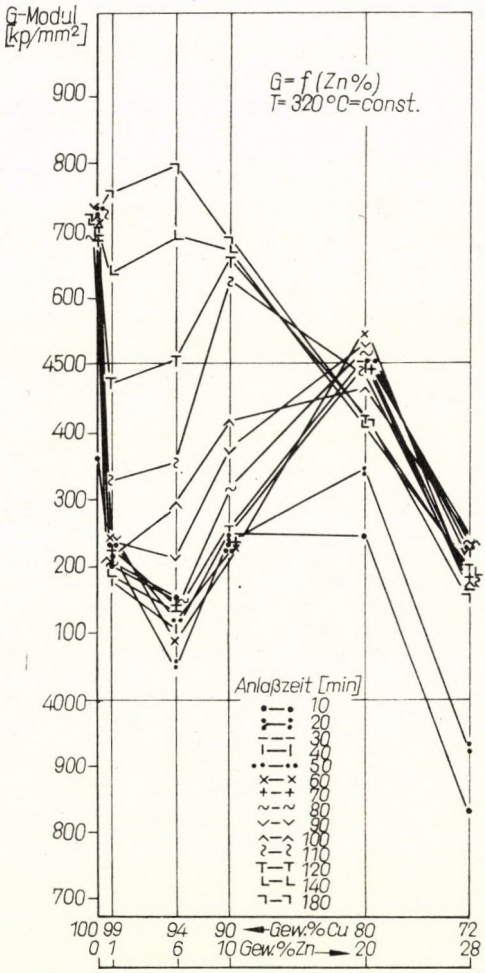


Bild 7. Konzentrationsabhängigkeit des Torsionsmoduls von verformten α-Messingstangen (Anlaßtemperatur $T = 320$ °C)

Die Konzentrationsabhängigkeit von μ unterhalb Rekristallisationstemperatur (Typ μ *I)* ist mit ihrem für Ms 99 und Ms 94 ausgeprägten Maximum ein Ergebnis der in diesem Konzentrationsbereich vorliegenden Q_{max}-Werte. Und zwar sinkt, von Ms 99 ausgehend, der Anteil der [111]-Komponenten der Doppelfasertextur und die [100]-Komponenten nehmen zu. Da in [111]-

Richtung sowohl bei Kupfer als auch bei den Messingen maximale Youngsche Moduln aber minimale Torsionsmoduln, dagegen in [100]-Richtung minimale Youngsche aber maximale Torsionsmoduln im Einkristall vorliegen [9], verhält sich ein texturiertes Vielkristallhaufwerk ähnlich. Der *E*-Modul (Bild 2) besitzt für Ms 99 ein Maximum und auch das Minimum des Torsionsmoduls bildet sich im Bereich niedriger Zinkkonzentrationen aus. Die mit zunehmendem Zinkgehalt zunehmenden Deformationsspannungen bewirken aber, daß das Minimum zu Ms 94 hin verlagert wird, obwohl (nur unter Berücksichtigung des Texturfaktors) dieses bei Ms 99 erwartet worden wäre.

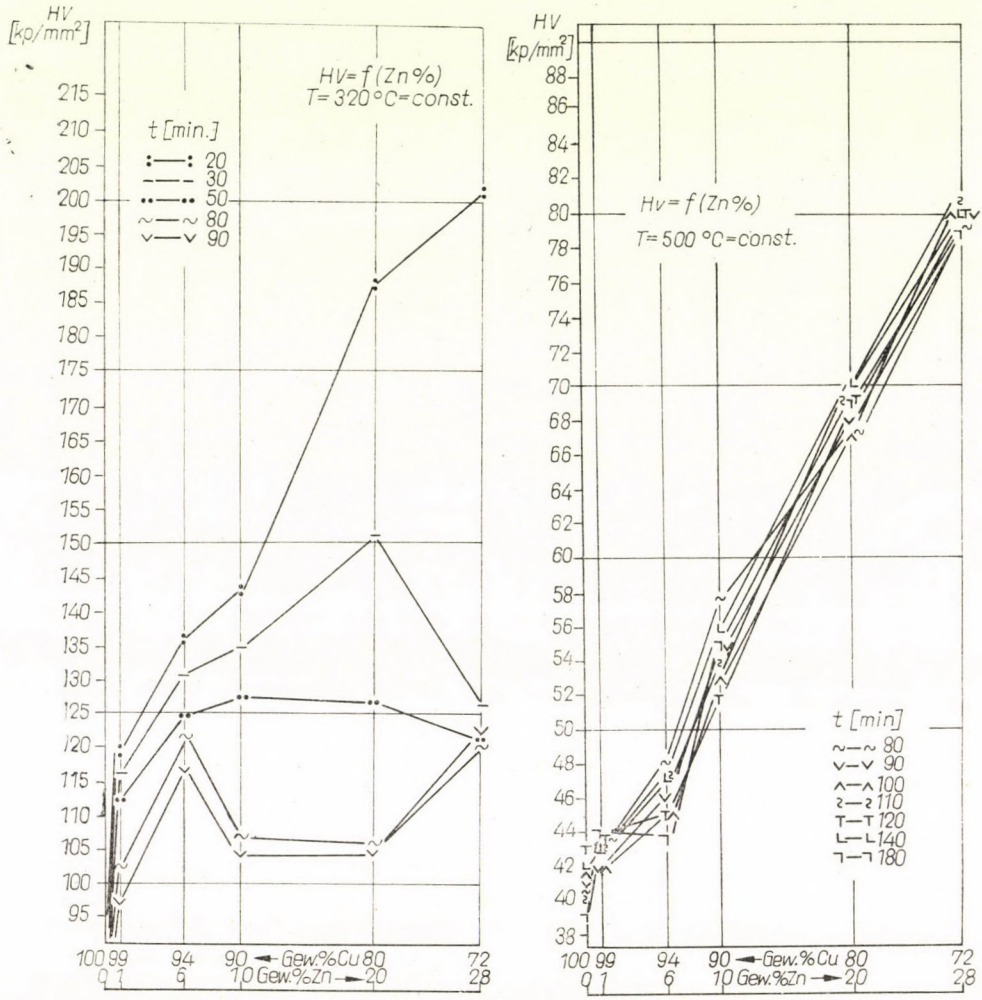


Bild 8. a) Konzentrationsabhängigkeit der Vickershärte von verformten α-Messingstangen (Anlaßtemperatur $T = 320$ °C); *b)* Konzentrationsabhängigkeit der Vickershärte von verformten α-Messingstangen (Anlaßtemperatur $T = 500$ °C)

Dieses Verhalten des Torsionsmoduls (Bild 3), das hier den Einfluß des Spannungsfaktors durch Modulverminderung gut erkennen läßt, ist nochmals zwischen Ms 80 und Ms 72 anschaulich nachweisbar. Die [100]-Komponente ändert sich in diesem Bereich nur noch ganz schwach, die Spannungen nehmen aber sehr stark zu, so daß der Torsionsmodul G stark abnimmt (Bild 3). Eine schwache [100] — Abnahme zu Ms 72 hin würde gleichfalls einen solchen Verlauf erklären, wodurch μ = Minima um Ms 80 entstehen.

Ein Typ μIb bildet sich für $T = 240$ °C aus (Bild 4), wobei das Kupfer bei höheren Anlaßzeiten rekristallisiert und sich so genähert regellose Kristallitenverteilung ausbildet. Wir haben damit einen Mischtyp von Konzentra-

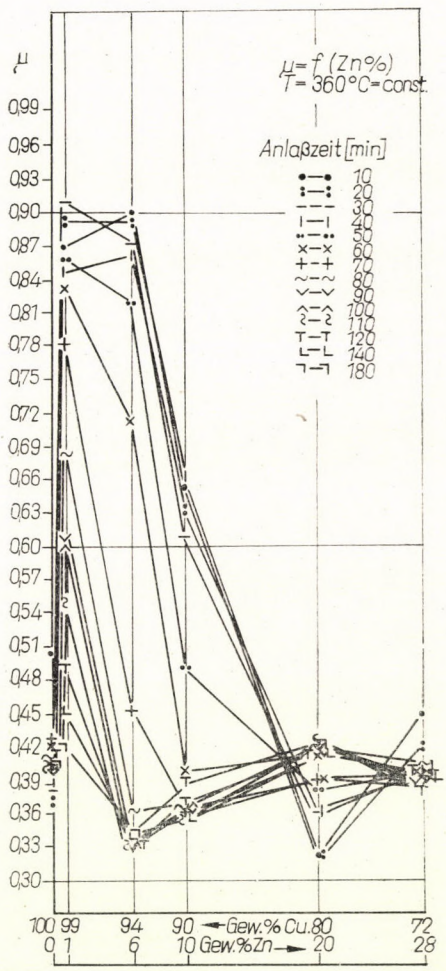


Bild 9. Konzentrationsabhängigkeit der Poisson-Zahl von verformten α-Messingstangen (Anlaßtemperatur $T = 360$ °C)

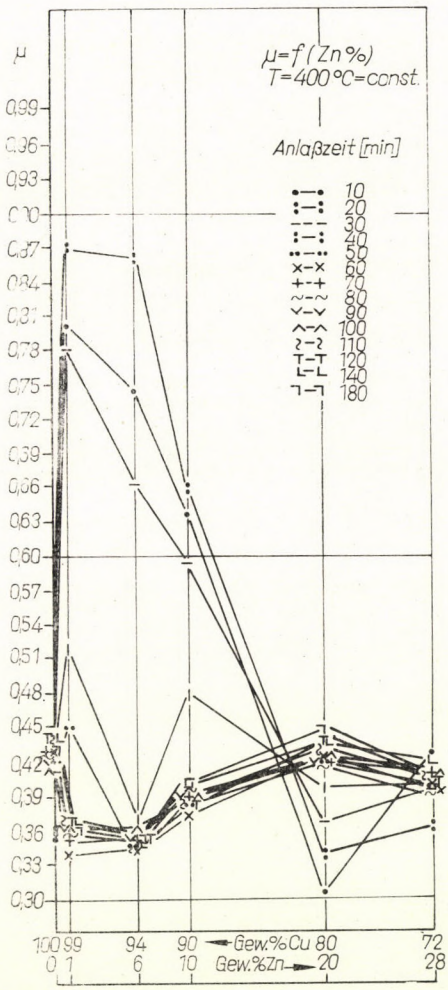


Bild 10. Konzentrationsabhängigkeit der Poisson-Zahl von verformten α-Messingstangen (Anlaßtemperatur $T = 400$ °C)

tionsdiagrammen der Poisson-Zahl, da Werkstoffe mit stark unterschiedlichen Entfestigungsgraden bzw. Anisotropiegraden im genannten Konzentrationsintervall vorliegen. Der Bereich für Quasiisotropiewerte der Poisson-Zahl des Kupfers und der α-Messinge liegt um 0,35 (vgl. z. B. [10]).

Mit Bild 5 wird in detaillierter Form die Veränderung der Poisson-Zahl während Entfestigung (Vickershärte Bild 8) erfaßt, wobei Bild 6 und Bild 7 die Rekristallisationskinetik durch Bestimmung des Youngschen und Torsionsmoduls fixiert. Diese Veränderungen sind begründet durch die Auflösung der Doppelfaser-Deformationstextur, den Abbau von Spannungen usw., wobei

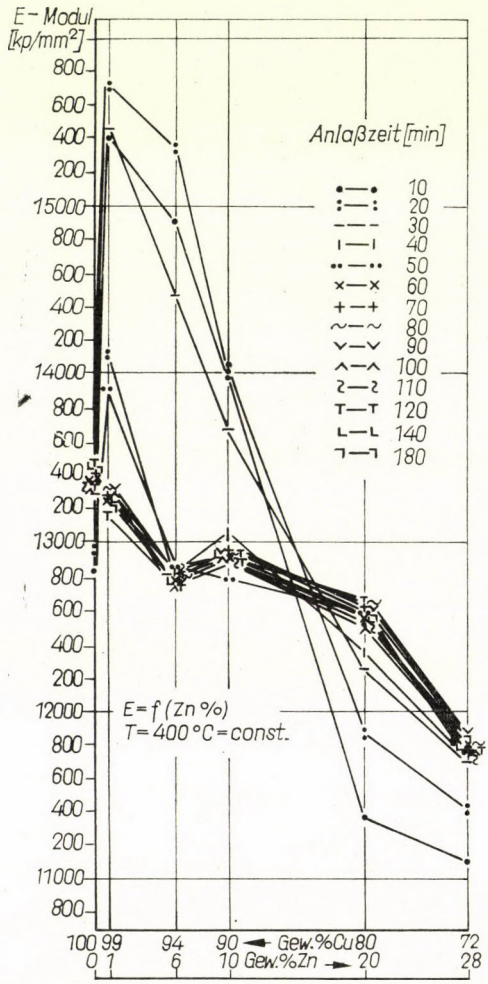


Bild 11. Konzentrationsabhängigkeit des Youngschen Moduls von verformten α-Messingstangen (Anlaßtemperatur $T = 400$ °C)

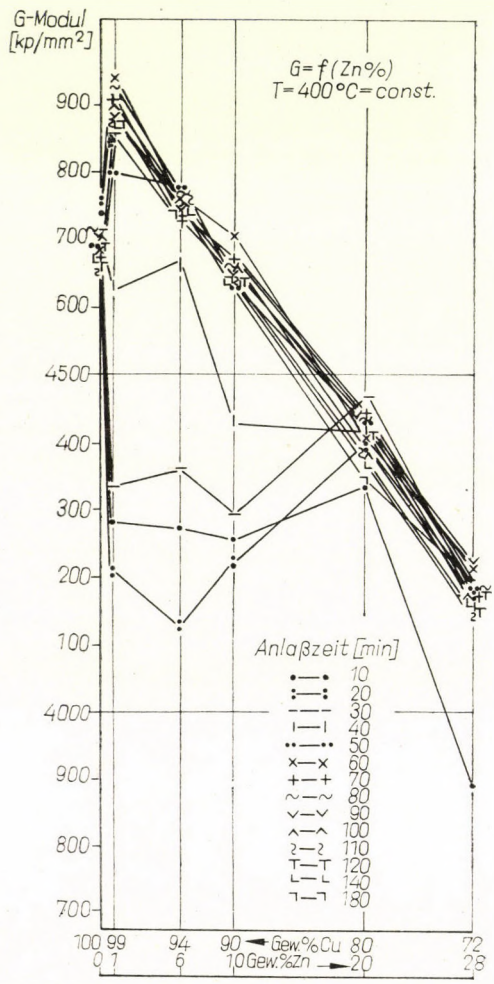


Bild 12. Konzentrationsabhängigkeit des Torsionsmoduls von verformten α-Messingstangen (Anlaßtemperatur $T = 400$ °C)

sich in allen Fällen genähert regellose Kristallitenorientierung mit μ-Werten um 0,35 ausbildet. Die Gegenläufigkeit der μ-Wert-Veränderungen des Ms 94 oder Ms 80 beispielsweise ist dadurch zu erklären. Diese Quasiisotropie-Bereiche sind sehr schnell durchlaufen, so daß nicht in jedem Fall bei der vorgegebenen

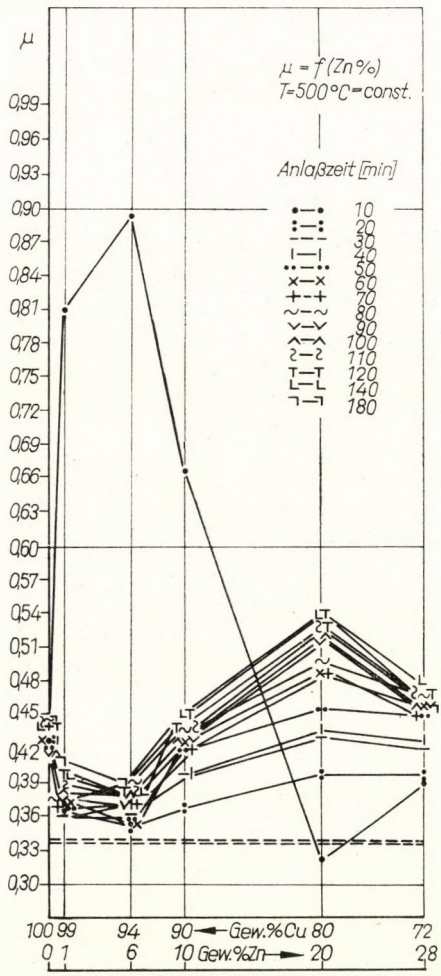


Bild 13. Konzentrationsabhängigkeit der Poisson-Zahl von verformten α-Messingstangen (Anlaßtemperatur $T = 500$ °C)

Wahl der Anlaßzeiten und -temperaturen auch immer eine Probe mit regelloser Orientierung vorlag. Bild 9 zeigt die μ-Werte in einem schon fortgeschrittenerem Stadium der Entfestigung, wobei sich zwischen Ms 94 und Ms 72 schon der Typ μ *II* herausbildet als Konzentrationsdiagramm des bei einer bestimmten Temperatur weichgeglühten Probenzustandes. Wie durchgeführte Texturaufnahmen zeigen, ist aber der z. B. bei einer Anlaßtemperatur von 600 °C

auftretende starke Anisotropiegrad auch hier schon deutlich ausgeprägt, was durch die in vorliegender Form ermittelte Poisson-Zahl bereits zwischen $T = 360$ u. 400 °C nachgewiesen ist. In diesem Nachweis der konzentrationsabhängigen Stärke des kristallografisch bedingten Anisotropiegrades liegt eine der besonderen Bedeutung derartiger Messungen.

Bild 10 und 13 in Verbindung mit Bild 11 und 12 läßt des weiteren erkennen, daß die hochzinkhaltigen α-Messinge gegenüber den hochkupferhaltigen trotz gleicher Anlaßbehandlung sich schon mit ihren μ-Werten bedeutend weiter vom Isotropiewert entfernt haben, damit einen höheren Grad von Texturausbildung anzeigen als Ms 99 oder Ms 94. Es muß hiermit festgestellt werden, daß die Konstanz der Probenvorbehandlung bzw. Konstanz der Parameter der thermischen Einwirkung durchaus nicht auch eine Gewähr für Regellosigkeit oder konstanten Anisotropiegrad der Versuchsproben für Konzentrationsdiagramme ist.

SCHRIFTTUM

1. Franz, H.: *Z. f. Metallk.* **49** (1958), 510.
2. Köster, W.—Scherb, J.: *Z. f. Metallk.* **49** (1958), 501.
3. Larson, F.: *Trans. A. S. M.* **57** (1964), 620.
4. Pursey, H.: *J. Inst. Met.* **86** (1957/58), 362.
5. Robertson, J.—Lipsitt, H.: *J. appl. Phys.* **36** (1965), 2843.
6. Förster, F.: Bedienungsanleitg. Elastomat; Reutlingen 1963.
7. Izumi, O.: *J. Jap. Inst. Met. (Sendai)* **23** (1959), 209.
8. Köster, W.: *Z. f. Metallk.* **39** (1948), 1.
9. Hearmon, R.: *Rev. mod. Physics* **18** (1946), 409.
10. Köster, W.—Franz, H.: *Met. Rev.* **6** (1961), 1.

The Problem of Poisson's Ratio as a Function of Concentration. Poisson's ratio as a function of concentration according to indirect measurements on copper- and brass-cylinders is described and discussed in detail. We distinguish types μ *I* as concentration diagrams of deformed materials, types μ *II* of soft-annealed specimens, and mixed types, which are all distinguished by characteristic extreme values, chiefly as a result of the deviations of statistical distribution of the crystallites of the polycristallite aggregation.

Le problème de la constante de Poisson en fonction de la concentration. La constante de Poisson en fonction de la concentration est décrite et discutée en détail sur la base de mesures indirectes faites sur des cylindres de cuivre et d'étain. On distingue les types de courbes de concentration μ *I* se rapportant aux matériaux déformés des types μ *II* d'échantillons recuits et des types mixtes, qui se reconnaissent tous à leurs valeurs extrêmes caractéristiques et sont surtout le résultat des déviations produites dans l'agrégation polycristalline, par rapport à la distribution statistique des cristallites.

К вопросу зависимости от концентрации числа Пуассона *(Х. Вавра)*. Подробное изложение и обсуждение зависимости числа Пуассона от концентрации, на основе косвенных измерений, проведенных на медных и латунных цилиндрах. Различаются типы μI — как диаграмма концентрации деформированных материалов, μII — как диаграмма концентрации отожженных проб и смешанные типы. Все типы отличаются характерными критическими значениями, и прежде всего являются результатами отклонений от статистики распределения кристаллитов и поликристаллических скоплений.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (3—4), pp. 413—428 (1968)

THE DETERMINATION OF EQUILIBRIUM AND KINETICS IN THE PROCESSES OCCURRING IN OXIDATION-SMELTING

Z. HORVÁTH
DOCTOR OF TECHN. SC.
UNIVERSITY OF HEAVY INDUSTRY, MISKOLC (HUNGARY)

[Manuscript received June 5, 1967]

The work describes the processes which take place during oxidation-smelting and then analyses the possibilities and the most suitable working methods of eliminating the impurity by the same process for different cases. After accounting for the equilibrium conditions, there is a treatment of the kinetics of the reactions taking place in individual equipment for different cases.

I. Introduction

Oxidation-smelting is used in many instances in metallurgy for refining of metals. So for example C, Mn, Si and P is separated from iron by smelting in this way during the manufacture of steel; the impurities Fe, Zn, S of copper in the metallurgy of copper; Sn, As, Sb, Zn in the metallurgy of lead; Fe, Zn, S in the metallurgy of nickel; Pb from noble metals in the metallurgy of noble metals; As, S in the metallurgy of antimony etc. The applicability of the process is based on the principle — as Fig. 1 also shows — that the chemical affinity to oxygen of the above-mentioned elements constituting the impurities is greater than that of the base metal. That is, the elements present in the form of impurity oxidize under a more negative change of thermodynamical normal potential than the base metal. The oxidizing agent is generally oxygen, the oxide of a base metal, $NaNO_3$ or NaOH.

II. The determination of equilibrium

Assuming the valency of the metal impurity (Sz) and that of the base metal (Me) to be 2 in each case, the processes which take place during the oxidation with O_2 can be written in the following general form:

$$2[Sz] + O_2 = 2(SzO), \qquad (A)$$

$$2[Me] + O_2 = 2(MeO). \qquad (B)$$

In the case of oxidation for conditions differing from the normal the affinity

for oxygen with the thermodynamical potentials can be expressed in the following manner:

$$\Delta G_A = \Delta G_A^\circ + RT \ln \frac{A_{SzO}^2}{A_{Sz}^2 \cdot P_{O_2}} = \Delta G_A^\circ + RT \ln I_A \qquad (1)$$

and

$$\Delta G_B = \Delta G_B^\circ + RT \ln \frac{A_{MeO}^2}{A_{Me}^2 \cdot P_{O_2}} = \Delta G_B^\circ + RT \ln I_B. \qquad (2)$$

In (1) and in (2) ΔG° can be determined from the thermodynamical normal potential of the oxide formation (possibly oxide formation and solution),

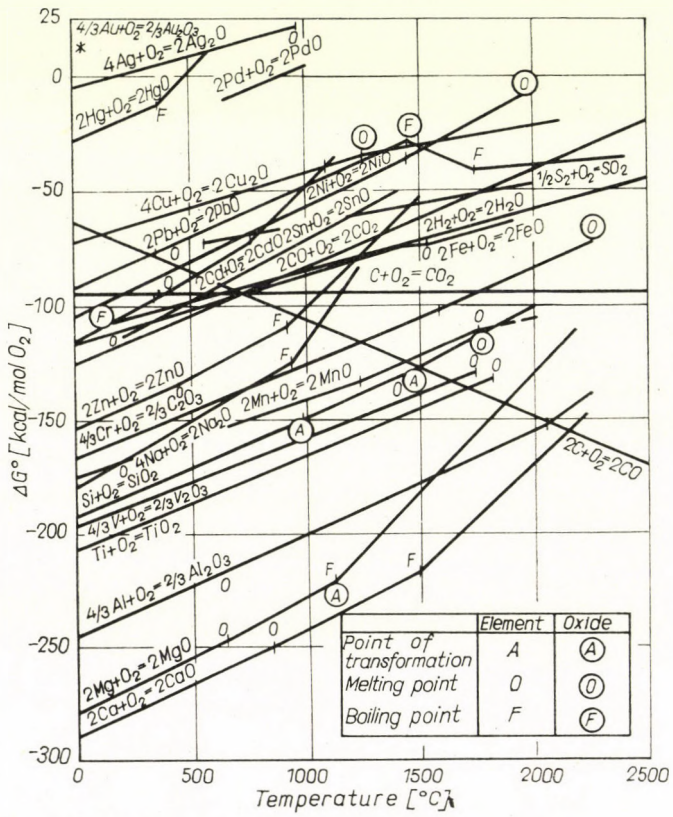


Fig. 1

A occurring in the fraction which comes after ln means the activity corresponding to the given composition of the mixture, while P denotes the partial pressure.

If for a constant temperature we express ΔG as a function of ln I, then — as Fig. 2 also shows — we obtain a straight line, which intersects the ordinate at a distance ΔG° and whose slope is RT.

At the beginning of the operation the amount of element consituting the impurity is small and that of the base metal is usually very high (exception is the cupellation). For this reason point *1* in the straight line ΔG_A, corresponding to the condition under reference, being greater than the value of I_A, is further from the ordinate than point *2*, relating to the starting condition and lying in the straight line ΔG_B. If $\Delta G_{A_1} < \Delta G_{B_2}$, then the element constituting the impurity oxidizes first. With the beginning of the oxidation the concentration of the impurity decreases, while that of the base metal increases.

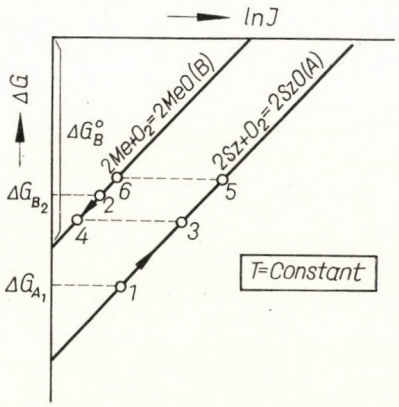


Fig. 2

For this reason point *1* moves upwards while the point *2* moves downwards in the straight lines. The impurity alone is oxidized up till $\Delta G_A < \Delta G_B$. In point *3* and in point *4* $\Delta G_{A_3} = \Delta G_{B_4}$, that is the affinity for oxygen for the elements Sz and Me are equal to each other. Starting from this point the concentration of the impurity can only be reduced at the cost of the oxidation of a part of the base metal.

A mathematical idea can also be obtained of what has been said in the preceeding. Let there be in a kg of base metal b kg of impurity, and let x kg of the latter oxidize away before both begin to oxidize at the same time, and z kg at the termination of the process. Let y kg of base metal be converted into oxide before it reaches the purity specified by the standard, and let the oxide of the base metal possess unit activity. Then knowing the corresponding mol-gram (M_{Sz}, M_{Me} respectively), the mol-number (n) at the start, corresponding to the simultaneous oxidation and at the termination can be given in the manner shown in *Table I.*

1.

With the knowledge of these the concentrations corresponding to the *simultaneous oxidation* can be calculated from the equation

$$\Delta G_{A_3} = \Delta G_{B_4}.$$

In these relationships, in place of the instantaneous activities the product of the coefficients of activity (γ), the mol-number (n) and the total mol-number (Σn) must be written. According to these

$$\Delta G_B^\circ + RT \ln \frac{(\Sigma n_{metal})^2}{\gamma_{Me}^2 (a/M_{Me})^2 \cdot P_{O_2}} = \Delta G_A^\circ + RT \ln \frac{A_{SzO}^2 \cdot (\Sigma n_{metal})^2}{\gamma_{Sz}^2 [(b-x)/M_{sz}]^2 P_{O_2}} .$$

Table I

Mol-number, n	Start (State *1* and *2*)	Simultaneous Oxidation (State *3* and *4*)	Finish (State *5* and *6*)
n_{Me}	$\frac{a}{M_{Me}}$	$\frac{a}{M_{Me}}$	$\frac{a-y}{M_{Me}}$
n_{Sz}	$\frac{b}{M_{Sz}}$	$\frac{b-x}{M_{Sz}}$	$\frac{b-z}{M_{Sz}}$
n_{MeO}	—	—	$\frac{y}{M_{Me}}$
n_{SzO}	—	$\frac{x}{M_{Sz}}$	$\frac{z}{M_{Sz}}$

From this relationship

$$\frac{\gamma_{Sz}}{\gamma_{Me}} \frac{b-x}{a} \frac{M_{Me}}{M_{Sz}} \frac{1}{A_{SzO}} = \exp \frac{\Delta G_A^\circ - \Delta G_B^\circ}{2RT}, \tag{3}$$

and

$$x = b - a \frac{\gamma_{Me}}{\gamma_{Sz}} \frac{M_{Sz}}{M_{Me}} A_{SzO} \exp \frac{\Delta G_A^\circ - \Delta G_B^\circ}{2RT} . \tag{4}$$

a) If the oxide of the element constituting the impurity is in *a separately condensed phase* and *other oxides are not present* in a dissolved state beside it, then $A_{SzO} = 1$ and from (4) the amount of the oxidized elements forming the impurity till the onset of simultaneous oxidation is:

$$x = b - a \frac{\gamma_{Me}}{\gamma_{Sz}} \cdot \frac{M_{Sz}}{M_{Me}} \exp \frac{\Delta G_A^\circ - \Delta G_B^\circ}{2RT} . \tag{5}$$

In this case the coefficient of activity of the base metal and the impurity (impure element) primarily influences the value of x.

b) If the *oxide of the impurity* is in solution in an oxide phase *(in slag)*

prepared by the addition of foreign additives, then

$$A_{SzO} = \gamma_{SzO} \frac{(x/M_{Sz})}{\Sigma n_{oxide}}$$

and from (3)

$$\frac{\gamma_{Sz}}{\gamma_{Me}} \frac{1}{\gamma_{SzO}} \frac{b-x}{a} \frac{M_{Me} \cdot \Sigma n_{oxide}}{x} = \exp \frac{\Delta G_A^\circ - \Delta G_B^\circ}{2\,RT},$$

and

$$\frac{b-x}{ax} = \frac{\gamma_{Me}}{\gamma_{Sz}} \frac{\gamma_{SzO}}{M_{Me}\,\Sigma n_{oxide}} \exp \frac{\Delta G_A^\circ - \Delta G_B^\circ}{2\,RT}. \tag{6}$$

It follows from (6), that the greater the fractional value resulting from the coefficients of activity and the lesser is the oxide molecules in the slag, the greater is $(b - x)$.

c) If the resulting oxide *dissolves* in the base metal, then

$$A_{SzO} = \gamma_{SzO} \cdot \frac{x/M_{Sz}}{\Sigma n_{metal}}$$

and from (3)

$$\frac{\gamma_{Sz}}{\gamma_{Me}} \frac{1}{\gamma_{SzO}} \frac{b-x}{a} \frac{M_{Me}\,\Sigma n_{metal}}{x} = \exp \frac{\Delta G_A^\circ - \Delta G_B^\circ}{2\,RT},$$

and

$$\frac{b-x}{ax} = \frac{\gamma_{Me}}{\gamma_{Sz}} \frac{\gamma_{SzO}}{M_{Me}\,\Sigma n_{metal}} \exp \frac{\Delta G_A^\circ - \Delta G_B^\circ}{2\,RT}. \tag{7}$$

If there is only an impure element in the base metal, then

$$\Sigma n_{metal} = \frac{a}{M_{Me}} + \frac{b}{M_{Sz}}$$

and

$$\frac{b-x}{x} = \frac{a}{M_{Me}} \frac{M_{Me} \cdot M_{Sz}}{aM_{Sz} + bM_{Me}} \frac{\gamma_{Me} \cdot \gamma_{SzO}}{\gamma_{Sz}} \exp \frac{\Delta G_A^\circ - \Delta G_B^\circ}{2\,RT}.$$

d) If the resultant oxide should be *gaseous*, then $A_{SzO} = P_{SzO}$ and according to (4)

$$x = b - a \frac{\gamma_{Me}}{\gamma_{Sz}} \frac{M_{Sz}}{M_{Me}} P_{SzO} \exp \frac{\Delta G_A^\circ - \Delta G_B^\circ}{2\,RT}, \tag{8}$$

thus the reduction in P_{SzO} assists in the elimination of the impurity.

e) If the oxide of the impurity *reacts with one of the constituents present in the slag*, then at the initial concentration indicated in Fig. 3 the impurity

undergoes further oxidation alone. This happens because the impurity reacts in the presence of the separately given RO according to the equation

$$2[\mathrm{Sz}] + \mathrm{O_2} + 2(\mathrm{RO}) = 2(\mathrm{SzO} \cdot \mathrm{RO}) \qquad \text{(C)}$$

during a change of the thermodynamical potential which is more negative than when in the absence of RO.

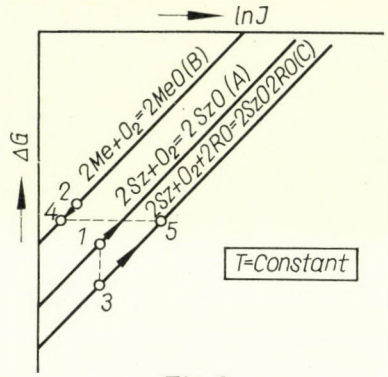


Fig. 3

In this case the quantity of the impurity oxidized till the onset of simultaneous oxidation and dissolving in the slag in the complex form can be calculated in the following manner from the equation

$$\Delta G_{\mathrm{B}_4} = \Delta G_{\mathrm{C}_5}:$$

$$\Delta G^{\circ}_{\mathrm{B}} + RT \ln \frac{(\Sigma n_{\mathrm{metal}})^2}{\gamma^2_{\mathrm{Me}} (a/M_{\mathrm{Me}})^2 \cdot P_{\mathrm{O_2}}} =$$

$$= \Delta G^{\circ}_{\mathrm{C}} + RT \ln \frac{\gamma_{2(\mathrm{SzO}\cdot\mathrm{RO})} \cdot \dfrac{x}{2\,M_{\mathrm{Sz}}} (\Sigma n_{\mathrm{metal}})^2 \cdot (\Sigma n_{\mathrm{oxide}})^2}{\Sigma n_{\mathrm{oxide}} \gamma^2_{\mathrm{Sz}} \left(\dfrac{b-x}{M_{\mathrm{Sz}}} \right)^2 \cdot P_{\mathrm{O_2}} \cdot \gamma^2_{\mathrm{RO}} \left(\dfrac{cM_{\mathrm{Sz}} - xM_{\mathrm{RO}}}{M_{\mathrm{Sz}} \cdot M_{\mathrm{RO}}} \right)^2} .$$

From this relationship

$$\frac{\gamma_{\mathrm{Sz}} \cdot \gamma_{\mathrm{RO}}}{\gamma_{\mathrm{Me}} \sqrt{\gamma_{2(\mathrm{SzO}\cdot\mathrm{RO})}}} \frac{\sqrt{2}\, M_{\mathrm{Me}}}{M_{\mathrm{RO}} \cdot \sqrt{M^3_{\mathrm{Sz}}}} \frac{1}{\sqrt{\Sigma n_{\mathrm{oxide}}}} \frac{(b-x)(cM_{\mathrm{Sz}} - xM_{\mathrm{RO}})}{ax} =$$

$$= \exp \frac{\Delta G^{\circ}_{\mathrm{C}} - \Delta G^{\circ}_{\mathrm{B}}}{2\,RT}$$

and

$$\frac{(b-x)(cM_{Sz}-xM_{RO})}{ax} = \\ = \frac{\gamma_{Me}\sqrt{\gamma_{2(SzO\cdot RO)}}}{\gamma_{Sz}\cdot\gamma_{RO}}\frac{M_{RO}\sqrt{M_{Sz}^3}}{M_{Me}\sqrt{2}}\sqrt{\Sigma n_{oxide}}\exp\frac{\Delta G_C^\circ-\Delta G_B^\circ}{2RT}. \tag{9}$$

In (9) c means the quantity of RO present in the slag at the start.

It may be concluded from Fig. 3 and from (9) that the presence of an oxide in the slag reacting with the oxide of the impurity is advantageous, since in such a case the impurity alone can be oxidised to a much greater extent.

2

At the concluding stage of the smelting by oxidation it is possible to calculate the measurable values of concentration on the basis of Fig. 2 from the equation $\Delta G_{A_5}^\circ = \Delta G_{B_6}^\circ$. Corresponding to this at the end of the operation

$$\Delta G_B^\circ + RT\ln\frac{A_{MeO}^2(\Sigma n_{metal})^2}{\gamma_{Me}^2\left(\frac{a-y}{M_{Me}}\right)^2\cdot P_{O_2}} = \Delta G_A^\circ + RT\ln\frac{A_{SzO}^2(\Sigma n_{metal})^2}{\gamma_{Sz}^2\left(\frac{b-z}{M_{Sz}}\right)^2\cdot P_{O_2}}.$$

From this relationship

$$\frac{\gamma_{Sz}}{\gamma_{Me}}\frac{b-z}{a-y}\frac{M_{Me}}{M_{Sz}}\frac{A_{MeO}}{A_{SzO}} = \exp\frac{\Delta G_A^\circ-\Delta G_B^\circ}{2RT}, \tag{10}$$

and

$$\frac{b-z}{a-y} = \frac{\gamma_{Me}}{\gamma_{Sz}}\frac{M_{Sz}}{M_{Me}}\frac{A_{SzO}}{A_{MeO}}\exp\frac{\Delta G_A^\circ-\Delta G_B^\circ}{2RT}. \tag{11}$$

In order to calculate the two unknown (y and z) in (10) and (11) it is necessary to use the equation

$$B = \frac{b-z}{a-y} \tag{12}$$

from the specification laid down in the standard.

Here also we can calculate the final concentration in many ways corresponding to the given data.

a) If the impurity and the oxide of the base metal *dissolves in the oxide phase*, then

$$A_{SzO} = \gamma_{SzO}\frac{z/M_{Sz}}{\Sigma n_{oxide}}$$

and

$$A_{MeO} = \gamma_{MeO} \frac{y/M_{Me}}{\Sigma n_{oxide}} .$$

The substitution of these equations into (10) gives

$$\frac{\gamma_{Sz}}{\gamma_{Me}} \frac{\gamma_{MeO}}{\gamma_{SzO}} \frac{b-z}{a-y} \frac{y}{z} = \exp \frac{\Delta G^{\circ}_{A} - \Delta G^{\circ}_{B}}{2RT}$$

and

$$\frac{b-z}{a-y} \frac{y}{z} = \frac{\gamma_{Me}}{\gamma_{Sz}} \frac{\gamma_{SzO}}{\gamma_{MeO}} \exp \frac{\Delta G^{\circ}_{A} - \Delta G^{\circ}_{B}}{2RT} . \qquad (13)$$

Since further oxidation after simultaneous oxidation starts can only be justified if the concentration of the impurity shows a further fall, that is if

$$\frac{b-z}{a-y} < \frac{b-x}{a} ,$$

further continuation of the process can only have meaning if

$$\frac{z}{y} (13) < x (7) ,$$

that is if

$$\frac{z}{y \cdot \gamma_{MeO}} < \frac{x}{M_{Me} \Sigma n_{oxide}} .$$

Since the inequality is identical with

$$\frac{z}{x} < \frac{A_{MeO}}{M_{Me}}$$

and in case of successful operation $z/x > 1$, it follows that in this case elimination of the impurity by further oxidation can in no way be economical.

b) If *the oxide of the base metal dissolves in the base metal*, while *the oxide of the impurity passes on into the slag*

$$A_{SzO} = \gamma_{SzO} \frac{z/M_{Sz}}{\Sigma n_{oxide}}$$

and

$$A_{MeO} = \gamma_{MeO} \frac{y/M_{Me}}{\Sigma n_{metal}} .$$

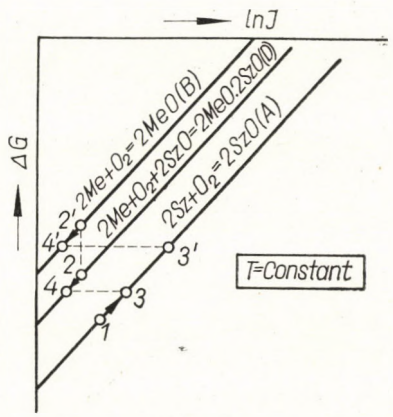
In this case on the basis of (10) we can write

$$\frac{\gamma_{Sz}}{\gamma_{Me}} \frac{\gamma_{MeO}}{\gamma_{SzO}} \frac{b-z}{a-y} \frac{y}{z} \frac{\Sigma n_{oxide}}{\Sigma n_{metal}} = \exp \frac{\Delta G_A^\circ - \Delta G_B^\circ}{2RT}$$

and

$$\frac{b-z}{a-y} \frac{y}{z} = \frac{\gamma_{Me}}{\gamma_{Sz}} \frac{\gamma_{SzO}}{\gamma_{MeO}} \frac{\Sigma n_{metal}}{\Sigma n_{oxide}} \exp \frac{\Delta G_A^\circ - \Delta G_B^\circ}{2RT} . \tag{14}$$

In such a case it is preferable to oxidize the solution of the oxide in the base metal to the maximum extent. This case occurs, for example, on the refining of Cu and Ni. Here there is no restriction, that on further oxidation the ratio of the impurity and base metal should be reduced, since metals in the forms of oxides in solution can again be reduced by means of a later process.



Fig, 4

c) If the oxide of the base metal dissolves in the base metal, and the *oxide of the impurity is gaseous* (for example during desulfurizing of copper):

$$A_{SzO} = P_{SzO} ,$$

$$A_{MeO} = \gamma_{MeO} \frac{y/M_{Me}}{\Sigma n_{metal}}$$

and

$$\frac{b-z}{a-y} y = \frac{\gamma_{Me}}{\gamma_{Sz}} \frac{1}{\gamma_{MeO}} M_{Sz} \Sigma n_{metal} P_{SzO} \exp \frac{\Delta G_A^\circ - \Delta G_B^\circ}{2RT} . \tag{15}$$

It can be seen from (15), that the greater the amount of the oxide of base metal dissolving in the base metal is, and the lesser the partial pressure of the oxide of the impurity, the greater would be the amount of impurity that can be eliminated from a given metal.

d) If the impurity and the oxide of the base metal *form a chemical compound*, then in the case of initial concentration (point 1 and 2*)* indicated in Fig. 4 the impurity oxidizes till the start of simultaneous oxidation. In such a case, however, the base metal no longer oxidizes according to (B), but according to the reaction

$$2[\mathrm{Me}] + \mathrm{O}_2 + 2(\mathrm{SzO}) = (2\mathrm{MeO} \cdot 2\,\mathrm{SzO}). \qquad (\mathrm{D})$$

Since the change of the thermodynamical potential corresponding to the process is more negative than that of the reaction (B) due to the formation

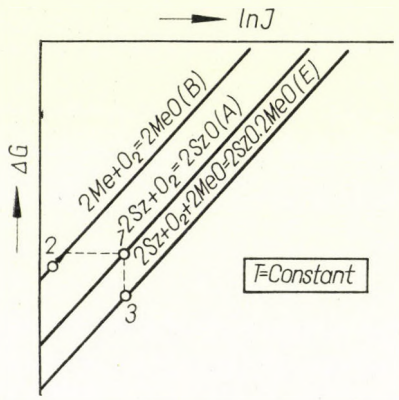


Fig. 5

of a chemical compound, the simultaneous oxidation takes place quicker. In this case the condition of oxidation of the base metal is

$$\Delta G_{\mathrm{D}_4} = \Delta G_{\mathrm{A}_3}.$$

From this equation the quantity of impurity oxidized at reaching point 3 can be calculated with the aid of an equation similar to (6):

$$\frac{b-x}{ax} = \frac{\gamma_{\mathrm{Me}}}{\gamma_{\mathrm{Sz}}} \frac{\gamma_{\mathrm{SzO}}}{M_{\mathrm{Me}} \Sigma n_{\mathrm{oxide}}} \exp \frac{\Delta G^{\circ}_{\mathrm{A}} - \Delta G^{\circ}_{\mathrm{D}}}{2RT}. \qquad (16)$$

It can be seen from Fig. 4, though it may also be concluded from (15), that the formation of the compound is not favourable for refining. If at the beginning the base metal oxidizes during a more negative change of thermodynamical potential, then — as Fig. 5 also shows — after the oxidation of a certain quantity of a base metal the impurity oxidizes according to the reaction

$$2[\mathrm{Sz}] + \mathrm{O}_2 + 2(\mathrm{MeO}) = (2\mathrm{SzO} \cdot 2\mathrm{MeO}). \qquad (\mathrm{E})$$

Summarizing the discussion so far, it may be established from the analysis of the question related to equilibrium alone, that for a given case the impurity can and considerably easily be eliminated by smelting through oxidation, if

the affinity for oxygen of the element constituting the impurity is large;

the oxide of the impurity can further react with one of the components of the slag;

in case of a gaseous oxide of the impurity its partial pressure is small;

the oxide of the impurity does not dissolve in the base metal; and

the oxide of the base metal easily dissolves in the base metal.

III. The determination of the kinetics

The preceeding is based on the assumption that the bath is homogeneous at every instant. This is not the case in actual practice. The speed of oxidation, and even the mechanism of the process taking place depend on the nature of work.

1. Oxidation of the inert smelting bath

a) If the gaseous oxidizing agent (air, air enriched with oxygen or oxygen) *comes into contact with the surface* of the bath or is blown onto the surface layer, the impurity in the surface layer can only oxidize according to the reaction stated in the previous chapter.

The oxide of the impurity either passes into the gaseous phase or in an unchanged form probably reacts with a component of the slag and passes into the slag, or is dissolved in the smelting bath. In this layer the concentration of the impurity rapidly decreases. Thus, before long the base metal oxidizes simultaneously with the impurity. The oxide of the base metal either remains unchanged, or passes into the slag after reacting with the oxide of the impurity, or becomes dissolved in the smelting bath.

If the oxide of the base metal does not dissolve in the smelting bath, then owing to the rapid decrease in the concentration of the impurity in the surface layer in contact with O_2, the simultaneous oxidation gets off to a quick start. For this reason the impurity reaching the surface oxidizes in the presence of the oxide of the base metal, and probably — though not in all cases — complex oxide is formed according to the equation

$$2[Sz] + O_2 + 2(MeO) = (2SzO \cdot 2MeO). \qquad (E)$$

In such a case, naturally not only O_2, but also MeO occurring freely or in complex state can oxidize, so that one must also take the possibility of the

following reaction into account:

$$2[Sz] + 2(MeO) = 2(SzO) + 2[Me] \quad (G) = (A) - (B),$$
$$2[Sz] + (2MeO \cdot 2Szo) = 4(SzO) + 2[Me] \quad (H) = (A) - (D).$$

In accord with the preceeding Fig. 6 shows, that the reaction (E) is the most advantageous. A decrease of a greater amount of impurity is expected in (G) than in (H), that is to say, that the complex compound can be decomposed at the cost of a larger amount of energy than that required for metal oxide. (In order to make it easily understandable, we neglected minute changes in the concentration of the base metal in Fig. 6.) The above mentioned pro-

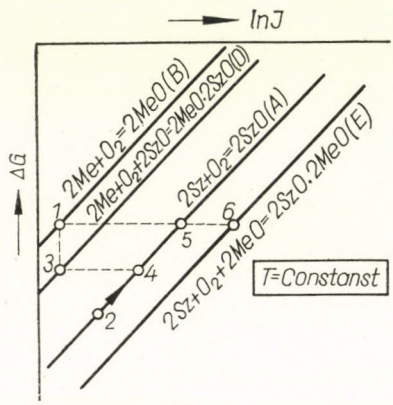


Fig. 6

cesses can only take place at the boundary of the metal and oxide phase.

Diffusion, or temperature, or the difference in density, or may be motion due to the kinetic energy of the oxidizing agent blown into the bath causes the migration of these elements as was mentioned previously.

At a given point in the field the concentration can decrease with time because of the motion of the bath, the diffusion can carry away material from a given position, and can finally make use of it in some kind of reaction. We can express this mathematically in the following manner [7; 92]:

$$-\frac{\partial c}{\partial t} = \operatorname{div}[c\bar{v}] - \operatorname{div}[D \operatorname{grad} c] + v_i v_r,$$

$$\text{local change} = \text{flow} + \text{diffusion} + \text{reaction}.$$

In this relationship is:

c the concentration [mol cm^{-3}];
t the time [sec];
$\bar{v}$ is the velocity of flow [cm sec^{-1}];

D the diffusion constant [$cm^2\ sec^{-1}$];
ν_i the mole number;
v_r is the velocity of reaction [$mol\ cm^{-3}\ sec^{-1}$].

If the quantity of the material transferred by the *diffusion* predominates ($v_{diffusion} < v_{reaction}$), we can express the process by the following equation:

$$\frac{\partial c}{\partial t} - \operatorname{div}\left[D \operatorname{grad} c\right] = 0. \tag{17}$$

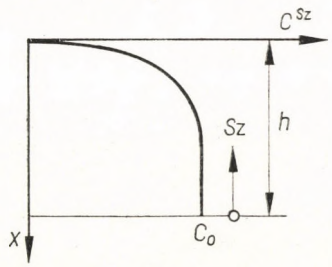


Fig. 7

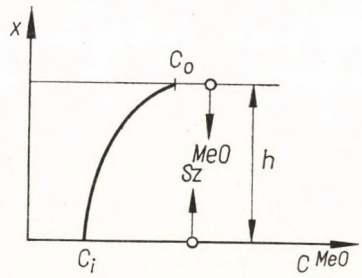


Fig. 8

This means, that at a given point the change of concentration taking place in the unit of time is equal to the quantity of material carried by diffusion. In this relationship c is the concentration of the impurity [$mol\ cm^{-3}$].

If the diffusion constant can really be regarded as constant — as Fig. 7 which illustrates the position, observable at the start of the process, also shows — it is only a case of unidirectional diffusion, and (17) takes up the following simplified form in the rectangular coordinates:

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2}. \tag{18}$$

Since

$$c(x;\ 0) = c_0,$$
$$c(0;\ t) = 0;$$

$$c(x;\ t) = c_0\ \operatorname{erf}\left(\frac{x}{2\sqrt{Dt}}\right). \tag{19}$$

If the oxide of the base metal dissolves in the smelting bath, the speed of the process considerably increases, because the impurity does not only oxidize in the surface layer, but also in the interior of the bath. Here the diffusion of the oxide of the base metal in a direction opposite to that of the impurity ensures the oxygen requirement of the reaction.

The diffusion of the oxide of the base metal could be shown to observe a certain law by making use of the experimental data of the concentration distribution corresponding to Fig. 8. It can be shown that for a certain case

the concentration varies with the distance in accordance with an experimental curve, that is

$$c = A_1 \exp\left(x\sqrt{\frac{k}{D}}\right) + A_2 \exp\left(-x\sqrt{\frac{k}{D}}\right). \tag{20}$$

If the boundary conditions corresponding to Fig. 8 are:

$$c^{\mathrm{MeO}}(x = h) = c_0 = A_1 \exp\,(h\sqrt{k/D}) + A_2 \exp\,(-h\sqrt{k/D})$$

and

$$\left(\frac{\partial c}{\partial x}\right)_{x=0} = 0 = A_1\sqrt{k/D} - A_2\sqrt{k/D}\,,$$

then

$$\begin{aligned} A_1 &= A_2 \\ c_0 &= 2A_1 \cosh\,(h\sqrt{k/D}), \\ c^{\mathrm{MeO}} &= 2A_1 \cosh\,(x\sqrt{k/D}, \\ c_i^{\mathrm{MeO}} &= 2A_1\,; \\ \frac{c^{\mathrm{MeO}}}{c_0} &= \frac{\cosh\,(x\sqrt{k/D})}{\cosh\,(h\sqrt{k/D})} \end{aligned} \tag{21}$$

and

$$\frac{c_0}{c_i^{\mathrm{MeO}}} = \cosh\,(h\sqrt{k/D})\,.$$

As a result of the two directional diffusions the refining in this case takes half the time it would if the oxide of the base metal does not dissolve in the bath.

b) If the gaseous oxidizing agent is blown *across the smelting bath* or into a part of it, as for example in the converter process or oxygen process of manifacture of steel, the work of diffusion considerably decreases. This means that the motion of the bath largely aids the equalization of the difference in concentration, and the oxygen reaches the interior of the bath even in the case when the oxide of the base metal is not soluble. For these reasons, we can, in such a case count on a quick oxidation.

2. *Oxidation in the moving smelting bath*

As a matter of fact what has already been stated in 1*b* can also be enumerated here. However, only two cases will be discussed here. The first is when the smelting bath is moved with respect to the relatively stationary oxidizing agent, and the second is when it is moved with respect to the moving oxidizing agent.

a) The oxidizing agent remains comparatively *undisturbed* for example in the Harris process. That is to say here Sn, As, Sb, or lead containing Zn is

pumped across to the salt bath containing the oxidizing agent so long as the amount of the impurity does not reduce to the desired level. During pumping the element constituting the impurity continually approaches closer to the newer and newer oxidizing agent, and can be oxidizing during a very short time.

b) Regarding the *simultaneous movement* of the smelting bath and the oxidizing agent mention can be made of the rotating drum furnace for smelting through oxidation. In this case, if the oxide of the metal does not dissolve in the bath, the oxidation will be more rapid, since there is continual renewal of surface, and since not only diffusion carries the impurity to the oxidizing position, but also a very powerful motion of the material.

If the oxide of the base metal dissolves in the bath, MeO not only diffuses into the interior of the bath from the top, but also from the bottom. New metal film continually draws in the moving lining at the bottom of the furnace which after oxidation turns to the bottom of the furnace. MeO which can be found all around the furnace, wanders at a nearly radial path towards the interior of the bath, and in this manner oxidizes the impurity which diffuses in the opposite direction to itself. What makes the work of diffusion easier here is not the shorter path of diffusion only, but also a very effective transfer of heat in the rotating drum furnace. It is because here the bath not only becomes heated from the top through direct heat transfer, gas radiation and radiation from the dome, but also from the bottom, since strongly heated lining continually turns towards the bottom of the furnace. Thus the temperature in the middle of the furnace is the least. The temperature obtained in this manner — and the difference of density, as well as the impurity of low density only increases the resulting motion due to upward flow, and brings about turbulence in flow which is already intense due to mechanical excitation.

REFERENCES

1. Horváth, Z.—Mihalik, A.: Elméleti kohászattan I (Theoretical Metallurgy I). Tankönyvkiadó, Budapest 1966.
2. Horváth, Z.: Termodinamikai alapfogalmak (Fundamentals of Thermodynamics); Miskolc 1967.
3. Horváth, Z.: Reakciókinetikai alapfogalmak (Fundamentals of Reaction Kinetics); Miskolc 1967.
4. Erdey-Grúz, T.—Schay, G.: Elméleti fizikai kémia III (The Theory of Physical Chemistry III). Tankönyvkiadó, Budapest 1962.
5. Verő, J.: Az ipari vasötvözetek metallografiája; Vaskohászati enciklopédia, IX/2 (The Metallography of Ferrous Alloys used in Industry). Akadémia kiadó, Budapest 1964.
6. Berecz, E.: Alkalmazott fizikai kémia (Applied Physical Chemistry). Tankönyvkiadó, Budapest 1964.
7. Benedek, P.—László, A.: A vegyészmérnöki tudomány alapjai (The Fundamentals of the Science of Chemical Engineering). Műszaki Könyvkiadó, Budapest 1964.
8. Horváth, A.: Kohászati fizikai-kémiai számítások (Physical and Chemical Computations in Metallurgy). Műszaki könyvkiadó, Budapest 1964.
9. Geters, F.: Zur Kinetik des Frischens unter besonderer Berücksichtigung des Sauerstoffaufblas-Verfahrens. *Archiv f. d. Eisenhüttenwesen* (1966), 209.

10. BISWAS, A. K.—BASHFORTH, G. R.: The Physical Chemistry of Metallurgical Processes. Chapman and Hall Ltd, London 1962.
11. HORVÁTH, Z.: Thermodynamics of the Reactions Taking Place during the Oxidation Period of Copper Refining. *Acta Techn. Hung.* **25** (1959), 159—186.
12. GILCHRIST, J. D.: Extraction Metallurgy. Pergamon Press, London 1967.

Kinetische und Gleichgewichtsuntersuchung der Vorgänge beim oxydierenden Schmelzen. In der Arbeit werden die beim oxydierenden Schmelzen in Betracht kommenden Vorgänge besprochen und weiterhin wird es untersucht, wie in den verschiedenen Fällen die Möglichkeiten und zweckentsprechendsten Arbeitsmethoden für die Entfernung der verunreinigenden Elemente festgestellt werden können. Nach Berechnung der Gleichgewichtsverhältnisse wird auch die Kinetik der in den einzelnen Einrichtungen stattfindenden Reaktionen besprochen.

Examen de l'équilibre et de la cinétique des phénomènes accompagnant la fusion oxydante. Le travail décrit les phénomènes pouvant éventuellement se produire lors de la fusion oxydante et examine ensuite comment, dans les différents cas, on arrive à déterminer les possibilités d'éliminer les éléments polluants par la fusion oxydante et quel est le mode de travail le mieux approprié. Après le calcul des conditions d'équilibre, l'auteur parle aussi de la cinétique des réactions se déroulant dans les différentes installations.

Равновесные и кинетические исследования процессов при окислительной плавке *(З. Хорват)*. В работе описываются процессы, которые, по всей вероятности, протекают при окислительной плавке; затем говорится о том, каким образом можно определить в различных случаях условия и технологию удаления элементов, присутствующих в качестве примеси, при окислительной плавке. После расчета условий равновесия рассматривается также вопрос кинетики реакций, протекающих в различных случаях в отдельных видах оборудования.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (3—4), pp. 429—438 (1968)

A THEORETICAL APPROACH TO THE DETERMINATION OF THE AMOUNT OF WEAR DURING SLIDING BETWEEN SURFACES

PRANAB KUMAR DE

CAND. OF TECHN. SC.

[Manuscript received June 20, 1967]

The present work takes into consideration a model of wear in which some of the asperities occurring in the surfaces of contact, tear away during sliding to give rise to loose particles. These particles alone constitute wear in this model. The tearing of particles would require a certain amount of mechanical work to be done on them. This work may be associated with the breaking work of the material which is the work necessary to break by tearing a unit volume of the material. The geometry of the asperities and the forces acting on them led to a method for finding the total work of tearing for a particular length of sliding. The total tearing work and the breaking work of the material then permits the evaluation of the volume of wear particles for a given distance of sliding.

I. Introduction

The sliding of one body over another, when the surfaces of contact are under pressure on account of a certain load, causes some of the asperities of the surfaces to deform elastically, causes some others to deform plastically and causes the rest — in which the plastic deformation exceeds a certain limit — to tear away from the parent body. The present model attributes wear to these torn particles alone. Work must be expended on the asperities for each of the three types of deformations, and the total work during sliding may be assumed to be equivalent to the sum of the work done on the asperities in order to bring about elastic deformation in some of them, plastic deformation in some others and tearing in the rest of the asperities. The frictional force and the distance of sliding would easily give the total work necessary for sliding. It is assumed that the ratio between the work done for the three kinds of deformation for a small fraction of the sliding area is equal to that of the whole area. This ratio enables the determination of the work expended in tearing away asperities over the whole surface of the sliding area from the total work of sliding. Breaking work which is the work necessary to be done in order to break by tearing unit volumes of the material, then readily enables the calculation of the volume of the torn particles which represent wear.

The main task in this problem is to find a process for determining the work expended in the three cases of deformation of the asperities. The other parameters necessary for evaluating wear do not pose any special problem in

their solution. The ranges of deformations of asperities corresponding to the three cases are, however, necessary for finding the work expended. The means for finding them occur in the following.

The range of deformations (elastic) of asperities vary from zero to deformation a_k, corresponding to the elastic limit. KRAGELSKII [3] gives the following expression for a_k:

$$a_k = 240 \frac{\sigma_F^2}{E^2} \cdot R$$

where σ_F yield point of the fully work-hardened material;
E elastic modulus of the material;
R radius of curvature at the tip of the asperity.

The range of plastic deformations of asperities vary from a_k to the deformation a_p, corresponding to the onset of tearing of asperities. Once again KRAGELSKII states the equation for a_p which is the following:

$$a_p = \left(\frac{\pi}{5{,}6} - 1{,}3 \cdot \mu \right)^2 \cdot R$$

where R is the radius of curvature of the asperity as before, and μ is the coefficient of friction of the sliding surface of the material.

The range of deformations of asperities in case of tearing vary from a_p to deformation a, corresponding to the approach between the surfaces due to the load. KRAGELSKII expresses a in the following manner:

$$\varepsilon = \left(\frac{\delta}{b} \right)^{1/\nu} + \frac{\varepsilon_k}{2}$$

where ε $= a/H$ in which H is the maximum height between asperities;
δ $= q_c/p_m$ in which q_c is the mean pressure over the contour area, and p_m is the pressure on the asperity corresponding to the onset of plasticity;
b a constant;
ν a constant;
ε_k $= a_k/H$ in which a_k is the compression of the asperity at the onset of plasticity.

In the above expression q_c becomes equal to the mean pressure over the apparent area, if waviness is neglected. For the determination of p_m, V. N. MAROCHKIN [4] obtained the following relationship:

$$p_m = (1 + \gamma + \nu)\, \sigma_s = c\sigma_s$$

where p_m is the yield pressure;
γ is the semi-vertical angle of the asperity;
ν is the coefficient of axis symmetry;
σ_s is the yield point of a work hardened specimen.

Fig. 1 shows the relationship between c and γ, according to the calculation made by V. N. MAROCHKIN.

Thus the determination of a_k, a_p and a according to the above would provide the limits whithin which each of the three types of deformation occurs. However, before commencing to deal with the question of the work expended it would be preferable to take up the aspect relating to the number of asperities involved in interaction between surfaces.

II. Number of asperities in contact

The following leads to an expression for the determination of the number of asperities of a surface suffering deformation owing to the approach of another surface pressed down on the former.

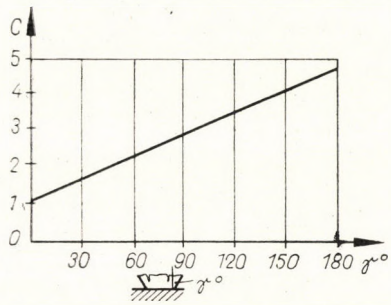


Fig. 1. The relationship between c and γ, according to the calculation made by V. N. MAROCHKIN

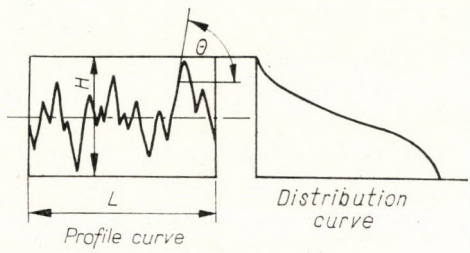


Fig. 2

G. YOSHIMOTO and T. TSUKIZOE [2] state that when the distribution curve which is obtained from the profile curve in Fig. 2 has a normal distribution, then the following expression can be written:

$$f(u) = \frac{1}{\sqrt{2\pi\delta}} \exp\left[-\frac{1}{2}\left(\frac{u}{\delta}\right)^2\right]$$

where $f(u)$ probability density; u deviation from the median of the profile curve; σ standard deviation.

Then the number of surface units between $u = \mu$ and $u = \mu + du$ (Fig. 3) is given by $Nf(u)du$, where N is the total number of surface units contained within measured length L. Hence the number of interaction of the profile curve on the parallel line $u = \mu$ will be:

$$Nf(u) = \varepsilon\sigma$$

where $\varepsilon\sigma$ is the projection of one surface unit.

When the asperities on the profile curve have an equal base angle Θ, we obtain

$$N = L \tan \Theta.$$

Thus, the number of asperities traversed by the parallel line u= 1 is

$$\frac{1}{2} N f(\mu) = \frac{1}{2} L \tan \Theta.$$

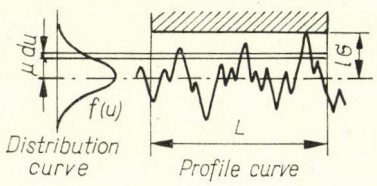


Fig. 3

When a parallel line is brought close to the profile curve, consider the contact to begin where $u = 1$. As the parallel line moves further from $u = 1$ through a distance a_m, let the number of asperities to come into contact be M. M is given by:

$$M = \frac{1}{2} L \tan \Theta [f(l\sigma - a_m) - f(l\sigma)] =$$

$$= \frac{1}{2\sqrt{2\pi}} \cdot \frac{L}{\sigma} \tan \Theta \left\{ \exp \left[-\frac{1}{2} \left(l - \frac{a_m}{\sigma} \right)^2 \right] - \exp \left(-\frac{1}{2} l^2 \right) \right\} =$$

$$= \frac{1}{2\sqrt{2\pi}} \exp \left(-\frac{1}{2} l^2 \right) \left[\exp \left(\frac{l a_m}{\sigma} - \frac{a_m^2}{2\sigma^2} \right) - 1 \right] \frac{L}{\sigma} \tan \Theta$$

where l is dependent on N and L. Hence

$$M = \frac{j \exp (l^2/2)}{2\sqrt{2\pi}} \left[\exp \left(\frac{j l a_m}{H} - \frac{j^2 a_m^2}{2H^2} \right) - 1 \right] \frac{L}{H} \tan \Theta$$

where H is the maximum height of asperities given by $H = j\sigma$ (j is dependent on the measured length L used for measuring H).

In the case of contact between two metal surfaces substituting $H = \sqrt{H_1^2 + H_2^2}$ in the previous equation we obtain:

$$M = \frac{j \exp (-l^2/2)}{2\sqrt{2\pi}} \left\{ \exp \left[\frac{j l a_m}{\sqrt{H_1^2 + H_2^2}} - \frac{j^2 a_m^2}{2(H_1^2 + H_2^2)} \right] - 1 \right\} \frac{L}{H} \tan \Theta \quad (1)$$

where $j = 5{,}0$, $L = 3$ mm for $6 < H < 25\mu$, $l = 3{,}3$. Thus, the above expression for M represents the number of asperities on the surface to undergo deformation when the approach to the surface is a_m. The subsequent step is to find the work done for the three types of deformation.

III. Work done on asperities for elastic deformation

The work W_{R_1} done on asperity is

$$W_{R_1} = \frac{1}{2} F_i \cdot a_i \,[\text{kpmm}] \tag{2}$$

where F_i is the load on the asperity;
a_i is the compression on the asperity due to the approach of contacting surfaces.

Applying Hertz's theory F_i can be expressed in the following manner:

$$F_i = \left(\frac{a_i}{0{,}83}\right)^{3/2} \cdot R^{1/2} \cdot E \cdot \frac{1}{1-\mu^2} .$$

Substituting the above value of F_i into equation (2), we have

$$W_{R_1} = C_R \cdot a_i^{5/2} \;\; [\text{kpmm}] \tag{3}$$

where

$$C_R = \frac{1}{2} \cdot \frac{R^{1/2} \cdot E}{0{,}83^{3/2}\,(1-\mu^2)} .$$

The number of asperities for which the extent of compression varies from zero and a_i is equal to the difference between the number of asperities in contact due to approach between the surfaces, a, and the number of asperities in contact due to the approach, $(a - a_i)$. This is obtained by substituting a and $a - a_i$ into the place of a_m in equation (2). Let

$$M_a - M_{(a-a_i)} = M_{a_i}$$

then after suitable substitutions and subsequent simplifications,

$$M_{a_i} = K\{\exp(pa - qa^2) - \exp[p(a-a)_i - q(a-a_i)^2]\}$$

where

$$K = \frac{j}{2\sqrt{2\pi}} \cdot \frac{L}{\sqrt{H_1^2 + H_2^2}} \tan\Theta \cdot \exp\left(-\frac{l^2}{2}\right),$$

$$p = \frac{jl}{\sqrt{H_1^2 + H_2^2}}, \quad q = \frac{j^2}{2\,(H_1^2 + H_2^2)} .$$

Similarly the number of asperities, whose deformation varies from 0 and $a_i + \delta a_i$, is the following:

$$M_{a_i+\delta a_i} = K\left\{\exp(pa - qa^2) - \exp\left[p(a - a_i - \delta a_i) - q\,(a - a_i - \delta a_i)\right]\right\}.$$

Hence the number of asperities whose deformation varies between a_i and $a_i + + \delta\, a_i$ is given by

$$\begin{aligned} M_{\delta a_i} &= M_{a_i+\delta a_i} - M_{a_i} = \\ &= K \exp\left[p(a - a_i) - q(a - a_i)^2\right]\left\{-\exp\left[-p\delta a_i + \right.\right. \\ &\left.\left. + 2\,q\delta a_i(a - a_i) - q\delta a_i^2\right] + 1\right\}. \end{aligned}$$

Thus, $M_{\delta a_i}$ for δa_i approaching zero to the limit is equivalent to the number of those asperities only which have the same deformation a_i. It follows, therefore, that

$$\begin{aligned} dM = \lim_{\delta a_i \to 0} M_{\delta a_i} &= K\left[p - 2\,q(a - a_i)\right] \exp\left[p(a - a_i) - \right. \\ &\left. - q(a - a_i)^2\right] da_i. \end{aligned} \tag{4}$$

From equations (3) and (4) the work done dW_R on all the asperities undergoing the same compressive deformation, a_i, is the following:

$$dW_R = C_R\, K a_i^{5/2}\left[p - 2\,q(a - a_i)\right] \exp\left[p(a - a_i) - q(a - a_i)^2\right] da_i.$$

Consequently, the total work W_R when the compressive deformation of the asperities vary from 0 to a_k, which is till the onset of plasticity, is the following:

$$W_R = \int_0^{a_k} C_R\, K a_i^{5/2}\left[p - 2a\,(a - a_i)\right] \exp\left[p\,(a - a_i) - q\,(a - a_i)^2\right] da_i. \tag{5}$$

Recourse to numerical method, e.g. Simpson's Rule, is necessary for the evaluation of the above.

IV. Work done on the asperities for plastic deformation

Let the mean pressure on asperity be p_m. This is assumed to be constant for all compressions in the range of plastic deformation. The contact area A_i for a compression a_i according to KRAGELSKII is $\pi R a_i$. Therefore, the load on the asperity is $\pi R a_i q_m$.

The work, W_{P1}, done on an asperity is the following:

$$W_{P_1} = \frac{1}{2} \pi R a_i p_m a_k + \pi R a_i p_m (a_i - a_k) = C_P a_i \left(a_i - \frac{a_k}{2} \right) \qquad (6)$$

where $C_P = \pi R p_m$.

The above takes into account both elastic and plastic deformation. Equation (4) once again gives the number of asperities corresponding to the compressive deformation a_i.

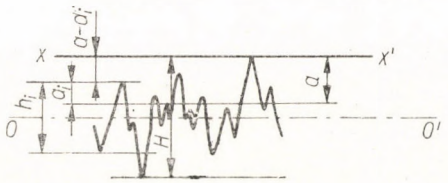


Fig. 4

From equation (4) and (6) the work done dW_P on all the asperities undergoing the same compressive deformation, a_i, is the following:

$$dW_P = C_P K a_i \left(a_i - \frac{a_k}{2} \right) [p - 2a(a - a_i)] \exp \{p(a - a_i) - q(a - a_i)^2\} da_i .$$

Thus, the total work W_P, when the compressive deformation of the asperities vary from a_k to a_p, i.e. till the onset of tearing, is the following:

$$W_P = \int_{a_k}^{a_p} C_P K a_i \left(a_i - \frac{a_k}{2} \right) \left[p - 2q(a - a_i) \right] \exp \{p(a - a_i) - (a - a_i)^2\} da_i . \qquad (7)$$

Partial integration and Gauss' integral together enable the integration of the above.

V. Work done on asperities in case of tearing

The work done in tearing an asperity is equivalent to the product of the contraction work i.e. the breaking work of the material A_c, and the volume of the asperity. A_c is constant for a given material.

In order to find the volume of an asperity let us consider any asperity of height, h_i, in a profile curve shown in Fig. 4. Let H be the maximum height of asperity, and a the approach to the surface, whose profile curve is given by

another surface. Then the latter will move through a distance of $(a - a_i)$, before it comes in contact with the asperity of height h_i. If OO' be the mean line of the profile curve, then it follows from what has just been said that

$$h_i = H - 2(a - a_i). \tag{8}$$

Assuming each asperity to be a cone with a circular base, and a base angle Θ the volume of the asperity, V_1, will be as follows:

$$V_1 = \frac{1}{3} \pi (h_i \cot \Theta)^2 h_i = \frac{1}{3} \pi \cot^2 \Theta h_i^3.$$

Substituting for h_i from equation (8) we obtain

$$V_1 = \frac{1}{3} \pi \cot^2 \Theta [H - 2(a - a_i)]^3.$$

Then the work W_{L_1} done on an asperity is the following:

$$W_{L_1} = \frac{1}{3} \pi \cot^2 \Theta [H - 2(a - a_i)]^3 A_c = C_L [H - 2(a - a_i)]^3 \tag{9}$$

where $$C_L = \frac{1}{3} \pi \cot^2 \Theta \cdot A_c.$$

Equation (4) gives the number of asperities corresponding to the compressive deformation a_i.

From equations (4) and (9) the work done dW_L on all the asperities undergoing the same compressive deformation, a_i, is the following:

$$dW_L = C_L K [H - 2(a - a_i)]^3 [p - 2a(a - a_i)] \exp \{p(a - a_i) - q(a - a_i)^2\} da_i.$$

Thus the total work W_L, when the compressive deformation of the asperity vary from a_p to a, where a is the approach between the surfaces, is the following:

$$W_L = \int_{a_p}^{a} C_L K [H - 2(a - a_i)]^3 [p - 2q(a - a_i)] \exp [p(a - a_i) - q(a - a_i)^2] da_i. \tag{10}$$

The integration of the above takes place in a manner similar to the previous case, i.e. the integration of W_P.

VI. Volume of wear particles

With the means of determining W_R, W_P, W_L at disposal, it would now be possible to find the total work of tearing in the course of sliding. This is equal to the product of the total work expended during sliding and the factor $W_L\,(W_R + W_P + W_L)$. Then a simple division of the value so obtained by the contraction work, A_c of the material, gives the volumes of the wear particles.

VII. Conclusion

A method for determining the amount of wear during sliding has been discussed. Suitable experimental data in support of the above do not exist. An attempt to verify experimentally, by using lubricated sliding pairs in which one of the sliding elements was radioactive, showed that the experimental value of wear was well below the calculated one. Further investigations under semi-fluid lubrication, followed by those under dry conditions of sliding would throw more light on the validity of the above method.

REFERENCES

1. Gillemot, L.: Contribution to the Analytical Determination of the "Breaking Work". *Materialprüfung* **3** (1961), No. 9.
2. Yoshimoto, C.—Tsukizoe, T.: On the Mechanism of Wear between Metal Surface. *Wear* **1** (1957/58), 472—490.
3. Kragelskii, I. V.: Friction and Wear. Butterworths, London 1965.
4. Marochkin, V. N.: Axially-Symmetrical Stresses, the Real Area of Contact between Rough Surfaces in Conditions of Full Plasticity; (Thesis) 1958.

Theoretische Bestimmung des Betrages des Abriebs zwischen gleitenden Oberflächen. Die Arbeit behandelt ein Modell des Abriebs, in welchem einige der Rauheiten in den Berührungsflächen während des Gleitens losgerissen werden, so daß lose Partikeln entstehen. In diesem Modell besteht der Abrieb nur aus diesen Partikeln. Das Losreißen der Partikeln beansprucht ein gewißes Maß von mechanischer Arbeit. Diese Arbeit kann mit der Brucharbeit des Materials in Zusammenhang gebracht werden, welche notwendig ist, um eine Raumeinheit des Materials loszureißen. Die Geometrie der Rauheiten und die auf dieselben wirkenden Kräfte führen zu einer Methode für die Berechnung der gesamten Losreißbarkeit für eine bestimmte Gleitlänge. Die gesamte Reißbarkeit und die Brucharbeit des Werkstoffs erlauben dann die Berechnung des Volumens der abgetragenen Teilchen für eine gegebene Gleitfläche.

Détermination théorique approchée de l'usure entre surfaces de frottement. L'étude examine un modèle de frottement dans lequel certaines aspérités de la surface de contact sont arrachées pendant le frottement et constituent des particules libres. Dans le modèle, l'usure est représentée uniquement par ces particules, dont l'arrachement exige une certaine quantité de travail mécanique. Ce travail peut être associé au travail de rupture du matériau, soit le travail nécessaire pour détacher une unité de volume de la matière. La géométrie des aspérités et les forces agissant sur elles conduisent à une méthode qui permet de déterminer le travail d'arrachement total pour une longueur de glissement donnée. Ce travail d'arrachement et le travail de rupture de la matière permettent ensuite l'évaluation du volume des particules enlevées par usure pour une distance de glissement donnée.

Теоретический приближенный метод определения степени износа при трении между поверхностями *(Пранаб Кумар Де)*. Данная работа рассматривает такую модель трения, при которой ряд шероховатостей, встречающихся на соприкасающихся поверхностях, в процессе трения отрывается и, таким образом, возникают отдельностящие частицы. В этой модели только эти частицы представляют износ. Отрыв частиц требует некоторой определенной механической работы, развитой на них. Эту работу можно связать с работой разрыва материала, которая соответствует работе, необходимой для отрыва объема в единицу материала. Геометрия шероховатостей и действующие на них силы привели к такому методу, при помощи которого можно определить работу разрыва, отнесенную для общей, некоторой определенной длины скольжения. Общая работа разрыва и работа разрыва материала после этого позволяют определить объем отрывающихся при износе частиц для некоторой данной длины скольжения.

Acta Technica Academiae Scientiarum Hungaricae; Tomus 61 (3—4), pp. 439—466 (1968)

TOOTH ROOT THICKNESS BETWEEN TROCHOIDAL ROOT PROFILES ON INVOLUTE-SHAPED GEARS

F. HORVÁTH

RESEARCH INSTITUTE FOR ELECTRONICS AND PRECISION ENGINEERING, BUDAPEST

[Manuscript received June 30, 1967]

The paper deals with the calculation of the root thickness of involute gearing made by generating, with a tool based on a rounded or on a sharp-edged rack. In the course of the calculation of the root thickness which can be determined by trochoidal geometry the author generalizes the results of his investigations. He gives the conditions which must be fulfilled so that the critical root thickness can be determined by trochoidal geometry only. The gearing generated by a sharp-edged tool is considered as a special case of the gearing generated by a rounded tool and the root thickness limited by trochoids is determined for an arbitrary location. If the module, the number of teeth and the addendum modification coefficient are known, the root thickness for an arbitrary place of a tooth root limited by trochoids can be calculated. In the case of a sharp-edged generating tool the final result can be presented in an easily calculated explicit form.

List of symbols

m	module;
x	addendum modification coefficient;
z	number of teeth;
r_a	base circle radius;
r_0	$= m \cdot z/2$ pitch circle radius;
c_0	$= mc_0'$ where c_0' denotes the tool-addendum clearance factor; in our example its value equals 0,25;
α_0	inclination angle of the cutting edge of the generating tool i.e. basic profile angle (in our example, $\alpha_0 = 20°$);
f_s	$= m(1 + c_0' - x)$ addendum height of the generating tool;
r_l	$= r_0 - f_s$ radius of the dedendum circle;
r_t	radius of the point at the locus of maximum root stress;
s_t	arc length measured between the root points at the radius r_t;
r_H	radius of the limit circle, i.e. distance between the centre and the point of intersection of the involute and the trochoidal profile curves;
ϱ	rounding radius of the rack profile.

I. Introduction

It has to be borne in mind that the total tooth profile is composed of the involute section for power transmission, and of the trochoidal root section. When the task is to find the tooth thickness between corresponding points of the two involute curves, it can be determined by means of involute-geometry. When we have to find the root thickness belonging to the locus of the maximum stress i.e. to determine the critical cross section, this can be found, as a rule, only on the basis of trochoidal geometry.

According to HOFER [3] the critical cross section is determined by the chord belonging to a point on the trochoid at which the tangent line forms an angle of 30° with the tooth's centre line. In his paper, HOFER presents large scale diagrams showing the respective calculated values; these also are gathered together in tables; both diagrams and tables represent the chord length as a function of the number of teeth and the addendum modification coefficient.

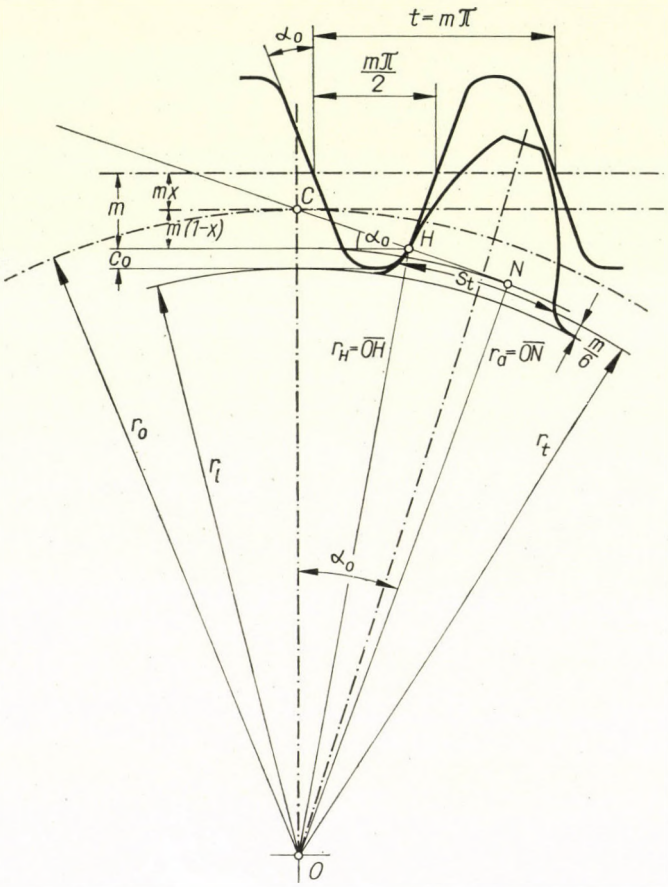


Fig. 1. Involute tooth profile generated by a round-cornered rack-tool, with a positive addendum modification ($z = 17$; $m = 10$ mm; $x = 0{,}4716$; $\alpha_0 = 20°$; $c_0' = 0{,}25$)

Due to some recent experimental tests, a more exact determination of the critical thickness is possible. According to the photo-elastic investigations made by JACOBSON [1], the maximum root stress is to be found at a point of the trochoid belonging to a radius which is by 1/6 module larger than the dedendum radius.

Let us consider Figs 1 and 2. The involute tooth profile shown in Fig. 1

is generated by a rack-shaped tool having rounded corners; that one represented in Fig. 2, originates from a sharp-cornered tool.

Expressed in a formula, the locus of maximum root stress, according to JACOBSON, is induced at a point of the trochoid which lies on the circle of the radius r_t:

$$r_t = r_l + \frac{1}{6} m.$$

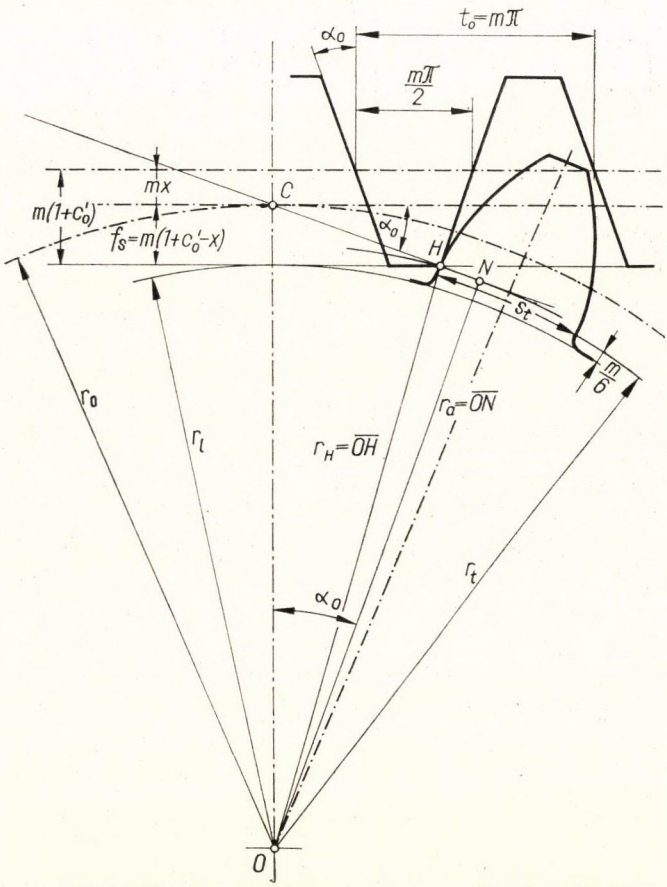


Fig. 2. Involute tooth profile generated by a sharp-cornered rack-tool with a positive addendum modification ($z = 17$; $m = 10$ mm; $x = 0{,}4716$; $\alpha_0 = 20°$; $c_0' = 0{,}25$)

According to Figs 1 and 2, the equation

$$r_t = r_l + \frac{1}{6} = r_0 - f_s + \frac{1}{6} = \frac{z}{2} - (1 + c_0' - x) + \frac{1}{6} \qquad (1)$$

refers to a profile developed by a tool with the unit module ($m = 1$ mm).

Again, for the general case of Fig 1, i.e. with a round-cornered tool, we obtain:

$$r_{H\,(\text{round-cornered})} = \sqrt{(\overline{CN} - \overline{CH})^2 + r_a^2} = \sqrt{\left(\frac{z}{2}\sin\alpha_0 - \frac{1-x}{\sin\alpha_0}\right)^2 + \left(\frac{z}{2}\cos\alpha_0\right)^2} \tag{2}$$

and for the special case of Fig. 2, i.e. with a sharp-cornered tool, we have:

$$r_{H\,(\text{sharp-cornered})} = \sqrt{(\overline{CN} - \overline{CH})^2 + r_a^2} = \sqrt{\left(\frac{z}{2}\sin\alpha_0 - \frac{1+c_0'-x}{\sin\alpha_0}\right)^2 + \left(\frac{z}{2}\cos\alpha_0\right)^2} \tag{3}$$

and tooth equations (2) and (3) refer to the unit module ($m = 1$ mm).

As follows from equations (1), (2), (3) and the diagrams 1 and 2, it is obvious that the critical root cross section belonging to r_t can be calculated *but* by means of the geometry of trochoids, when

$$r_t < r_H . \tag{4}$$

Now, the involute section of the tooth profile is undercut by the trochoidal root curve when the following condition is fulfilled:

$$x \gtreqless 0$$

and

$$r_0 - m(1-x) < r_a \cos\alpha_0 \tag{5}$$

namely with reference to Fig. 1; when

$$x \gtreqless 0$$

and

$$r_0 - m(1 + c_0' - x) < r_a \cos\alpha_0 \tag{6}$$

with reference to Fig. 2.

In both cases, point H represents the intersection point of the involute and the trochoid. Of course, the radius of the intersection point (limit point) H is always

$$r_H > r_a \tag{7}$$

i.e. larger than the base circle radius.

And, as inferred from (5) and (6), the relationship

$$r_l < r_a \cos\alpha_0 \tag{8}$$

holds true.

Again, when taking into consideration the relationship expressed by (7), (8) and (1), the validity of the inequality

$$r_t < r_H$$

can be proved even when the inequalities (5) and (6) hold true. But, in such a case the value of r_H cannot be calculated as established by equations (2) and (3). Thus, the determination of the root cross section can be effected ***only*** by applying the geometry of trochoids.

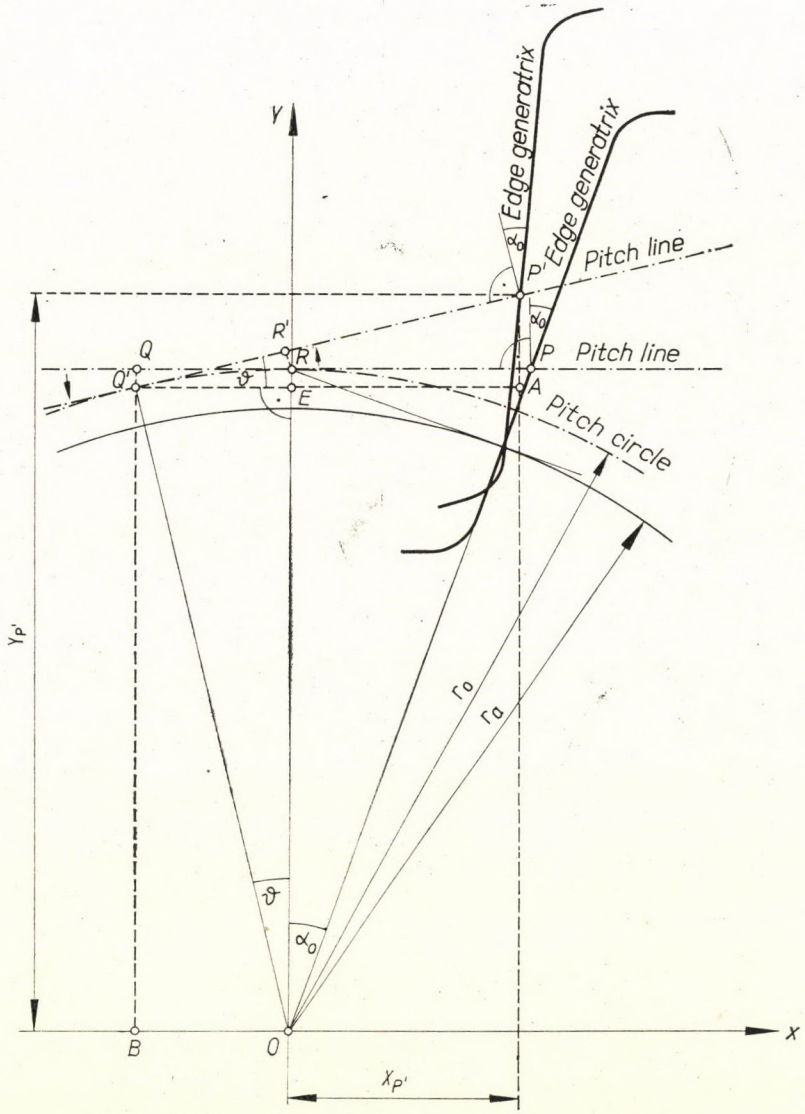


Fig. 3. Rolling of the round-cornered generating rack-tool

II. The formation of the trochoidal curve by the generating movement

As is shown in Figs 3 and 4, during the generating movement of the rack-shaped tool, the section $\overline{QR}$ of the pitch line is rolling along the peripheral line of the circle of the radius r_0 (in positive sense, according to the figure) and the tangential point R of the pitch line takes a new locus at R', and point Q becomes

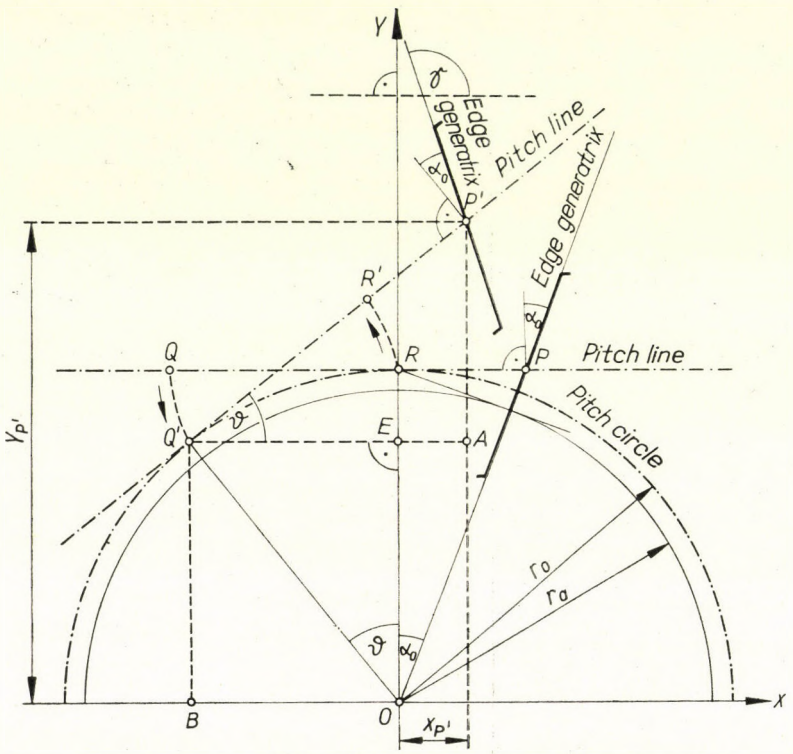


Fig. 4. Rolling of the sharp-cornered generating rack-tool

a tangential locus at Q'. Further, the intersection point P of the pitch line and the generatrix line directed at the angle α_0 reaches its new position at point P'. The coordinates $(X_{P'}; Y_{P'})$ can be determined, according to the diagram, as follows

$$X_{P'} = \overline{Q'A} - \overline{OB}\,; \tag{9}$$

$$Y_{P'} = \overline{BQ'} + \overline{AP'}\,. \tag{10}$$

In view of a true rolling movement, we have:

$$\overline{QR} = \overline{Q'R'} = \widehat{Q'R} = r_0\,\vartheta \tag{11}$$

where ϑ denotes the angle of turning.

From the initial position, the distance $\overline{RP}$ is expressed by

$$\overline{RP} = \overline{R'P'} = r_0 \tan \alpha_0 . \tag{12}$$

Now, because of $\overline{Q'P'} = \overline{Q'R'} + \overline{R'P'}$, and by substituting from (11) and (12) we obtain

$$\overline{Q'P'} = r_0 (\tan \alpha_0 + \vartheta) . \tag{13}$$

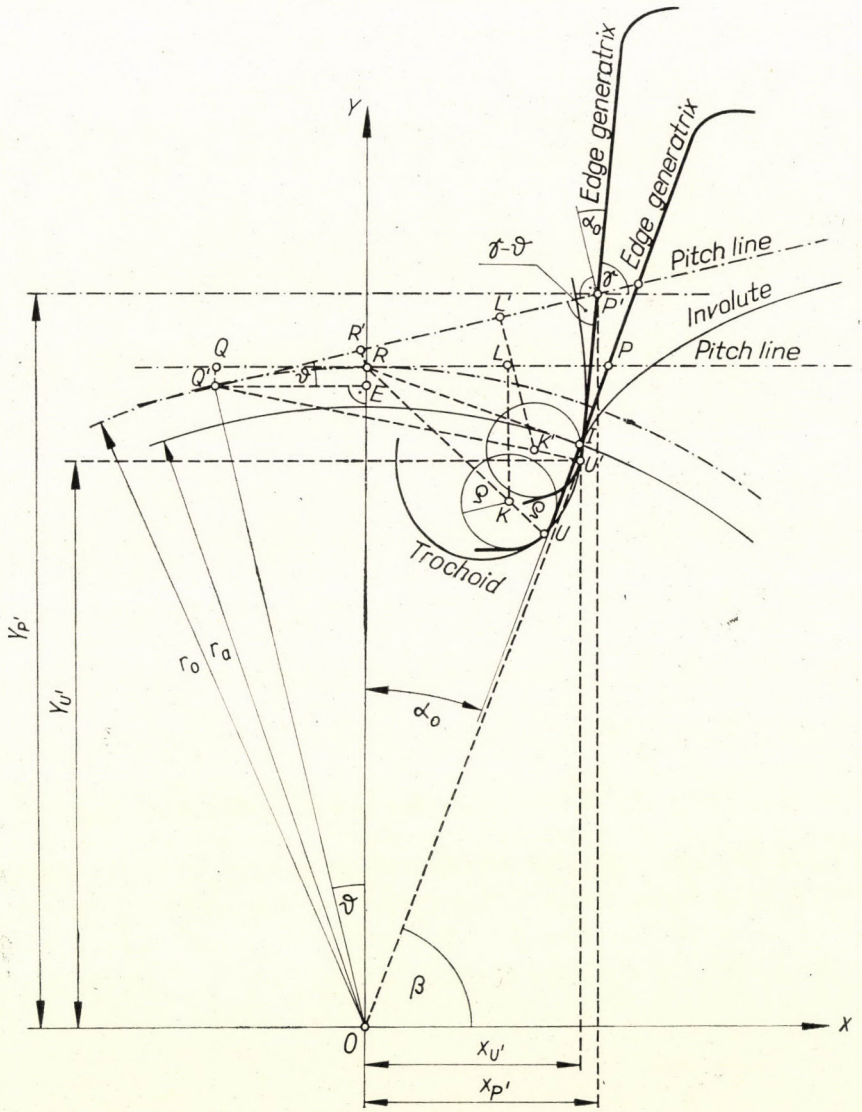


Fig. 5. Construction of the trochoidal curve, when generated by a round-cornered rack-tool

By taking equation (13) into account and by establishing the additive members in the coordinates of point P', according to equations (9) and (10), namely

$$\overline{Q'A} = \overline{Q'P'} \cos \vartheta = r_0 (\tan \alpha_0 + \vartheta) \cos \vartheta ;$$

furthermore

$$\overline{OB} = r_0 \sin \vartheta ,$$

$$\overline{BQ'} = r_0 \cos \vartheta ,$$

$$\overline{AP'} = \overline{Q'P'} \sin \vartheta = r_0 (\tan \alpha_0 + \vartheta) \sin \vartheta$$

the coordinates (9) and (10) will be expressed in the new form

$$X_{P'} = r_0 [(\tan \alpha_0 + \vartheta) \cos \vartheta - \sin \vartheta]; \tag{14}$$

$$Y_{P'} = r_0 [(\tan \alpha_0 + \vartheta) \sin \vartheta + \cos \vartheta]. \tag{15}$$

Let us now consider the Figs 5 and 6. Onto the straight line $\overline{OP}$ that is inclined at the angle α_0, a perpendicular straight line should be drawn from point R; thus, point I is the footpoint of this perpendicular, and $\overline{OI} = r_a$.

This point I on the base circle is the limit of undercutting, when in the initial position, this point is the tangential point of the straight generatrix as tangent line to the generating tool's rounding circle of the radius ϱ; or, when $\varrho = 0$ (with a sharp cornered tool) this point is the end point of the straight generatrix line.

It is known that by rolling the pitch line (in the diagram in negative sense) along the pitch circle having the radius r_0, the end point I of the generatrix moves along an involute curve developed from the starting point I of the base circle (Disteli's theorem, see [4]).

Let us consider the movement of the generating point U, that lies at the generating tool edge i.e. at the circle-arc having the radius ϱ, so we have

$$\overline{OU} < \overline{OI} = r_a \tag{16}$$

and bearing in mind that as a rule of geometric construction the points R, K, U are always collinear and so are the points Q', K', U' [5], it is obvious that the undercutting curve described by point U is located in that part of the plane below the involute curve i.e. below the straight line $\overline{OI}$.

Now, when one can find a generating point U, for which

$$\overline{OU} = \overline{OI} = r_a \tag{17}$$

is valid, then the same point represents the limit of undercutting.

It is clear that the character of the curve described by the moving point I is the same as that of the curve described by point U which fulfils equation (16), but — as already mentioned — no intersection occurs.

When equations (16) and (17), respectively, are fulfilled, the path of the moving point is represented by a trochoidal curve.

From what was described above, the trochoidal curve comes into being when there is some point U' on the generated line, for which the inequality

$$\frac{Y_{U'}}{X_{U'}} < \cot \alpha_0 \tag{18}$$

holds true.

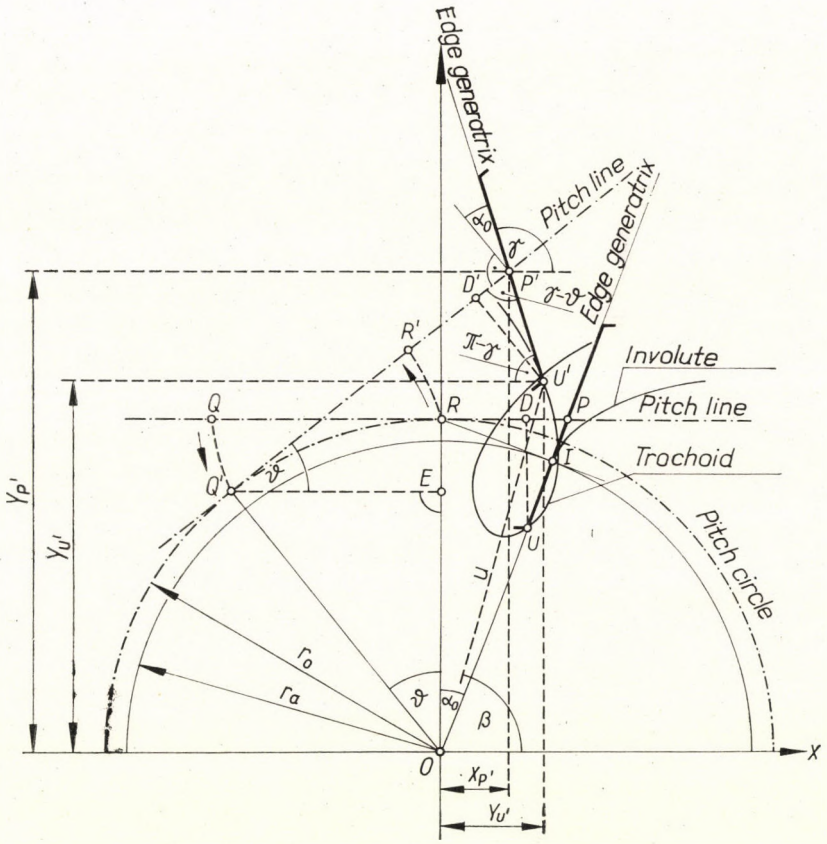


Fig. 6. Construction of the trochoidal curve, when generated by a sharp-cornered rack-tool

In the following, let us continue analyzing the trochoidal curve as delineated in Figs 5 and 6.

As can be seen in Fig. 5, the circle of the radius ϱ can be considered, during the movement, as a circle that rolls along the trochoid around the centre K', for which the tangential point is denoted by U'.

In order to find the value of radius ϱ, a detail of diagram in Fig. 5 should be re-considered according to Fig. 7. With the known value of the tool addendum clearance c_0, we have:

$$\varrho = c_0 + \varrho \sin \alpha_0$$

from which

$$\varrho = \frac{c_0}{1 - \sin \alpha_0} = \frac{mc_0'}{1 - \sin \alpha_0} . \tag{19}$$

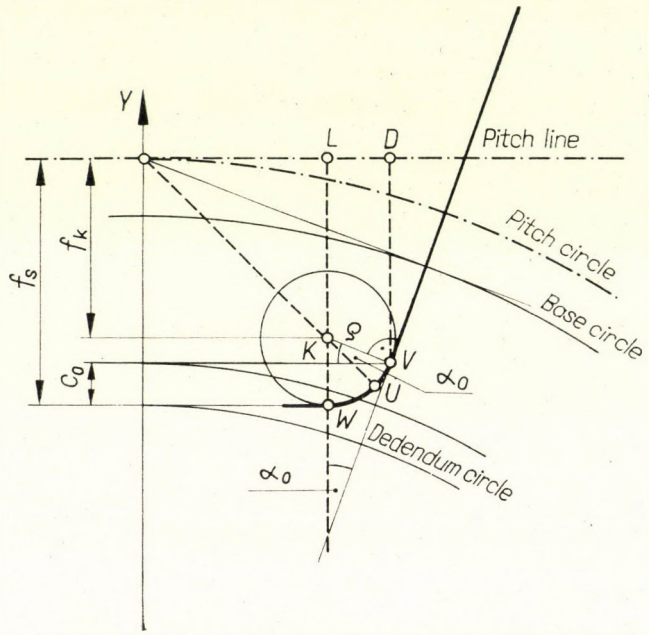


Fig. 7. A particular detail taken from Fig. 5, for the determination of the rounding radius

When the radius ϱ is reduced to the unit radius $[r_{0(1)} = 1]$ of the pitch circle we obtain, by making use of equation (19)

$$\varrho_{(1)} = \frac{\varrho}{r_0} = \frac{2c_0'}{z(1 - \sin \alpha_0)} . \tag{20}$$

As a general rule, the tool addendum height $f_s = \overline{WL}$ (when the tool is sharp-cornered, then $f_s = \overline{UD}$, see Fig. 6) can be determined when the addendum coefficient is known. Again, in the case of no modification

$$f_s = m + c_0 = m + mc_0' = m(1 + c_0'). \tag{21}$$

When $x < 0$, the tool addendum height is larger; again, when $x > 0$, it is shorter than its original value; expressed by a formula:

$$f_s = m(1 + c_0' - x). \tag{22}$$

Referring to a pitch circle with the unit radius:

$$f_{s(1)} = \frac{f_s}{r_0} = \frac{2}{z}(1 + c_0' - x). \tag{23}$$

According to Fig. 7, the distance between the centre point K of the circle with radius ϱ and the pitch line of the tool is

$$f_k = \overline{KL} = f_s - \varrho \tag{24}$$

and, reduced to the unit radius of the pitch circle:

$$f_{k(1)} = \frac{f_k}{r_0} = f_{s(1)} - \varrho_{(1)}. \tag{25}$$

One may clearly see, from (24) and (25) that with a sharp-cornered tool i.e. when $\varrho = 0$, we have:

$$f_{k(\varrho=0)} = f_s. \tag{26}$$

In order to respresent the trochoid, the system of polar coordinates seems to be suitable; by defining the angle β belonging to any given radius $u = \overline{OU'}$ [as for the angle β, the formula is to be found under (41)].

In the general case of a tool with *rounded corners*, no simple formula is to be found to express the coordinates of the generating point U' as functions of the turning angle ϑ. This will be clearer, when first we establish the calculating method of the same coordinates expressed by ϑ in the case of $\varrho = 0$. According to Fig. 6, we have:

$$X_{U'} = X_{P'} + \overline{U'P'} \cos(\pi - \gamma); \tag{27}$$

$$Y_{U'} = Y_{P'} + \overline{U'P'} \sin(\pi - \gamma) \tag{28}$$

where γ denotes the angle of the straight generatrix line in its turned position with the horizontal, i.e.:

$$\gamma = \frac{\pi}{2} + \vartheta - \alpha_0. \tag{29}$$

The distance $\overline{U'P'}$ in (27) and (28) equals:

$$\overline{U'P'} = \overline{UP} = \frac{\overline{UD}}{\cos \alpha_0} = \frac{f_s}{\cos \alpha_0} \tag{30}$$

and with a substitution according to (22) we obtain

$$\overline{U'P'} = \frac{m\,(1 + c_0' - x)}{\cos \alpha_0}. \tag{31}$$

As a last step, let the expressions (14), (15), (29) and (31) be substituted into (27) and (28), and let us write the new expression of the coordinates of any point of the trochoid as a function of angle ϑ (only in the special case of a sharp cornered tool):

$$X_{U'} = r_0 \left[(\tan \alpha_0 + \vartheta) \cos \vartheta - \sin \vartheta\right] + \frac{m\,(1 + c_0' - x)}{\cos \alpha_0} \sin(\vartheta - \alpha_0); \tag{32}$$

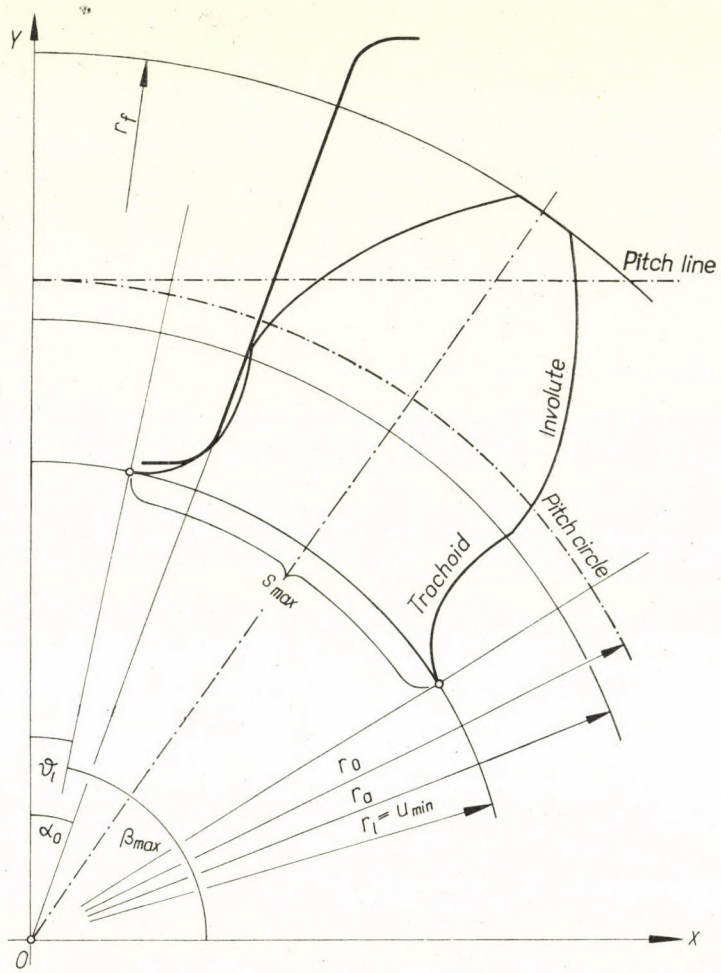


Fig. 8. Involute and trochoidal curves forming the tooth profile generated by a round-cornered rack-tool

$$Y_{U'} = r_0 \left[(\tan \alpha_0 + \vartheta) \sin \vartheta + \cos \vartheta \right] - \frac{m\,(1 + c_0' - x)}{\cos \alpha_0} \cos (\vartheta - \alpha_0).$$

This rather simple calculating method cannot be applied to the trochoid developed by a tool with rounded corners, because the generating point does

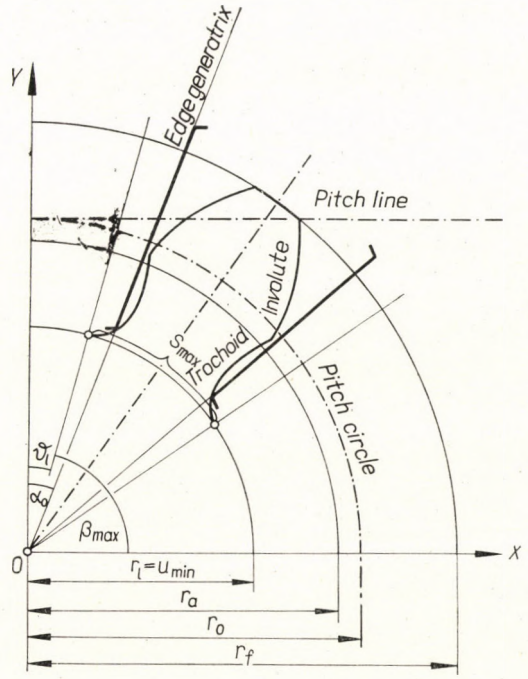


Fig. 9. Involute- and trochoidal curves forming the tooth profile generated by a sharp-cornered rack-tool

not lie on the generatrix line, but it moves along the circular arc of the rounding radius ϱ. Obviously, the simple system of polar coordinates is the suitable means for this general case.

III. Calculation of the root-thickness

The problem can be considered as solved, when the length of the arc s on the circle of a given radius u is known: from this, the chord length i.e. the root thickness can simply be calculated.

Let us consider Figs 8 and 9. Obviously, when $u = r_l$, the arc s between the trochoids assumes the maximum value on the same circle. The dedendum circle with the radius $r_l = u_{\min}$ is tangential to the trochoid.

As can be seen from Fig. 10, when the generating point moves along the rounding arc of the tool, its radial distance from the pitch circle will assume

the maximum value, only then when on the pitch line the new locus L of the circle centre point K coincides with the tangential point L' of the pitch line and the pitch circle.

Now, let us calculate the rolling angle ϑ_l by which s_{max} on the dedendum circle is determined:

$$\overline{RL} = \overline{RP} - \overline{LP} = \overline{RL'} = r_0 \vartheta_l, \tag{33}$$

in which

$$\overline{RP} = r_0 \tan \alpha_0,$$
$$\overline{LP} = \overline{LD} + \overline{DP},$$
$$\overline{LD} = \varrho \cos \alpha_0,$$
$$\overline{DP} = (f_s - c_0) \tan \alpha_0 = m(1 - x) \tan \alpha_0.$$

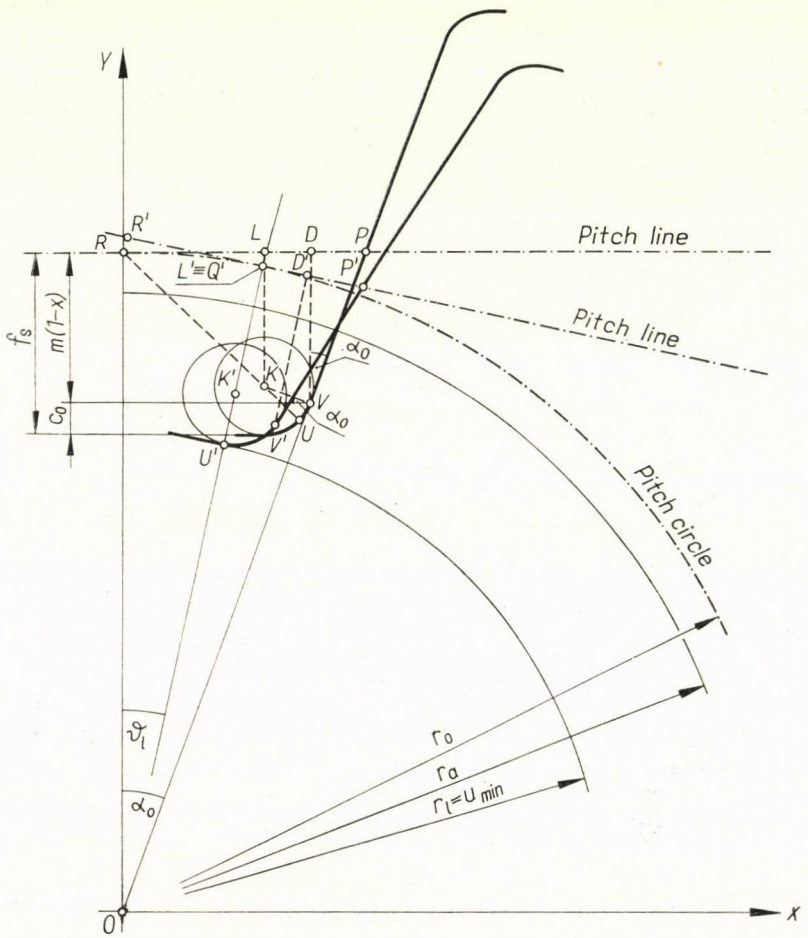


Fig. 10. The maximum root thickness generated by a round-cornered rack-tool

With the corresponding substitutions into (33), we obtain:

$$\overline{RL} = \widehat{RL'} = r_0 \vartheta_l = r_0 \tan \alpha_0 - [\varrho \cos \alpha_0 + m(1 - x) \tan \alpha_0]; \quad (34)$$

from which

$$\vartheta_l = \tan \alpha_0 - \frac{1}{r_0} [\varrho \cos \alpha_0 + m (1 - x) \tan \alpha_0]. \quad (35)$$

By reduction to the unit pitch circle radius, and making use of (20), we have:

$$\vartheta_l = \tan \alpha_0 - \frac{2}{z} \left[\frac{c_0'}{1 - \sin \alpha_0} \cos \alpha_0 + (1 - x) \tan \alpha_0 \right]. \quad (36)$$

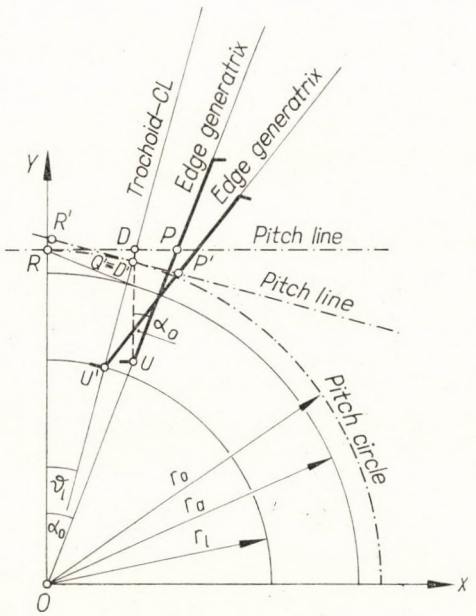


Fig. 11. The maximum root thickness generated by a sharp-cornered rack-tool

This course of calculation is applied for finding the angle ϑ_l when the trochoid is generated by a sharp-cornered tool (Fig. 11). When considering the tool in the initial position, the trochoid-generating point U of the cutting edge reaches the position at the maximum distance from the pitch circle, only then when by rolling the pitch line along the pitch circle, the point D of the tool addendum height on the pitch line at the same time becomes the tangential point of pitch circle and pitch line. In this position, point U reaches the dedendum circle.

According to Fig. 11, we have

$$\overline{RD} = \overline{RP} - \overline{DP} = \widehat{RD'} = r_0 \vartheta_l \quad (37)$$

with

$$\overline{RP} = r_0 \tan \alpha_0,$$
$$\overline{DP} = f_s \tan \alpha_0 = m(1 + c_0' - x) \tan \alpha_0$$

and by substitution into (37)

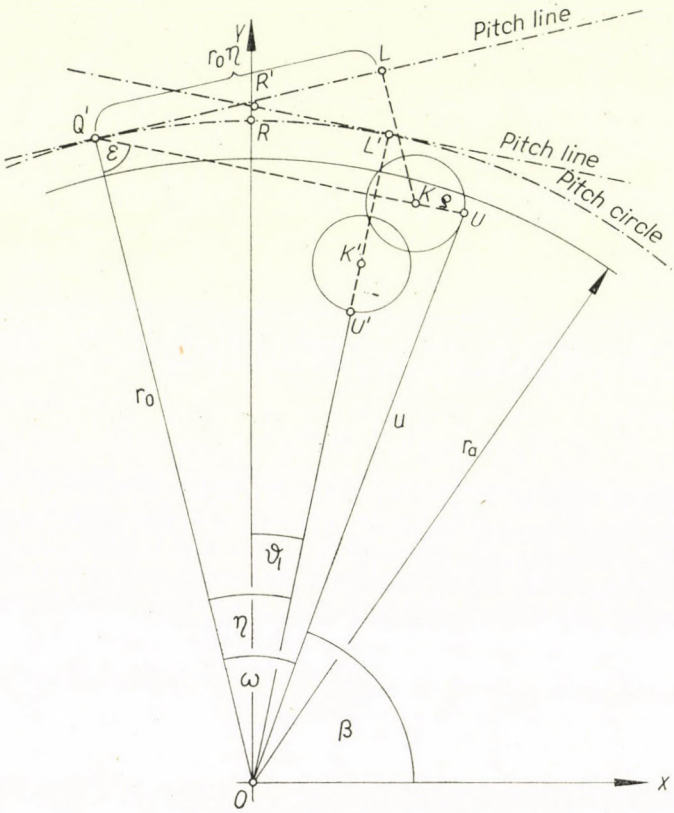


Fig. 12. Diagram for the calculation of the root thickness generated by a round-cornered rack tool

$$\overline{RD} = \widehat{RD'} = r_0 \vartheta_l = r_0 \tan \alpha_0 - m(1 + c_0' - x) \tan \alpha_0; \tag{38}$$

from where

$$\vartheta_l = \left[1 - \frac{m}{r_0}(1 + c_0' - x)\right] \tan \alpha_0 \tag{39}$$

and by reduction to the unit radius of the pitch circle:

$$\vartheta_l = [1 - f_{s(1)}] \tan \alpha_0 = \left[1 - \frac{2(1 + c_0' - x)}{z}\right] \tan \alpha_0. \tag{40}$$

In a simple way, we can see that the formulae which are valid for gears generated by a tool with rounded corners can easily be transformed for the case of a sharp-cornered tool by taking $\varrho = 0, f_{k(\varrho=0)} = f_s$, and putting $m(1 + c_0' - x)$ instead of $m(1 - x)$ [e.g. the formulae (35) and (39) might be compared].

Let us consider the point U of the trochoid, lying simultaneously on the circle with radius u — as is shown in Figs 12 and 13; and now the angle β formed by the straight $\overline{OU}$ and the abscissa-axis should be determined. As can

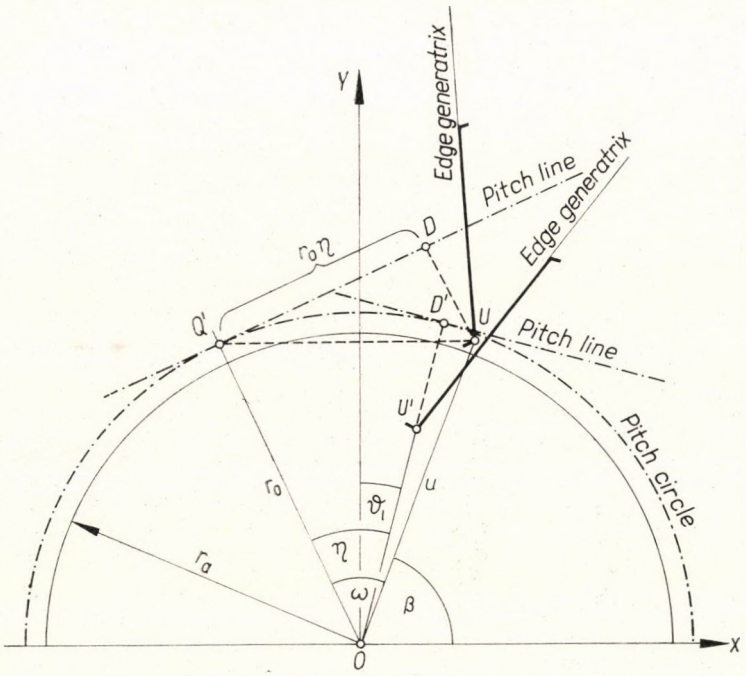


Fig. 13. Diagram for the calculation of the root thickness generated by a sharp-cornered rack-tool

be seen in the diagram :

$$\beta + \omega = \frac{\pi}{2} + (\eta - \vartheta_l),$$

from which

$$\beta = \frac{\pi}{2} - (\omega - \eta + \vartheta_l) \tag{41}$$

when ω denotes the angle between the straight lines $u = \overline{OU}$, and $r_0 = \overline{OQ'}$ η denotes the angle of roll belonging to the arc $\widehat{L'Q'}$, or $\widehat{D'Q'}$, respectively, and ϑ_l is the angle of roll already defined by formulae (35), (36), or (39), (40).

By observing in Fig. 12 the rolling movement of the pitch line of the generating tool, we obtain:

$$\overline{LQ'} = \widehat{L'Q'} = r_0 \eta; \tag{42}$$

and from Fig. 13:

$$\overline{DQ'} = \widehat{D'Q'} = r_0 \eta. \tag{43}$$

In order to determine the angles ω and η [as occurring in (41)], it is advisable to take out from Fig. 12 the configuration O, U, K, L, Q' (Fig. 14).

First, we can define the angle ε formed by the straight lines $\overline{LK} = f_k$ and $\overline{KQ'}$ namely as

$$\tan \varepsilon = \frac{r_0 \eta}{f_k}$$

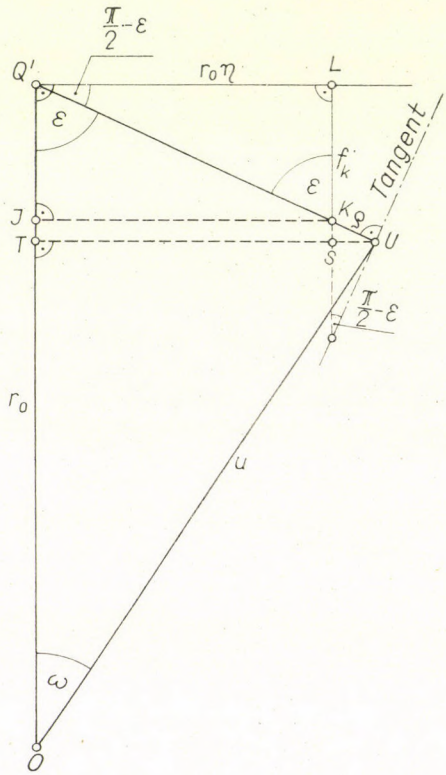


Fig. 14. Detail taken from Fig. 12

and by reduction to the unit radius of the pitch circle:

$$\eta = f_{k(1)} \tan \varepsilon. \tag{44}$$

Again, considering the rolling movement as seen in the diagram, we have

$$r_0 \eta = u \sin \omega - \varrho \sin \varepsilon. \tag{45}$$

From this, by taking the values reduced to the unit radius of the pitch circle into account, namely, $u_{(1)} = u/r_0$ and $\varrho_{(1)} = \varrho/r_0$ the following equation can be derived:

$$\omega = \arcsin \frac{\eta + \varrho_{(1)} \sin \varepsilon}{u_{(1)}} . \tag{46}$$

Finally, according to the diagram the radius u can be expressed

$$u^2 = \overline{OU}^2 = \overline{TU}^2 + \overline{OT}^2 = (r_0 \eta + \varrho \sin \varepsilon)^2 + (r_0 - f_k - \varrho \cos \varepsilon)^2. \tag{47}$$

Now, the next step is to determine the angle ε. When the subsequent operations and the reduction to the pitch circle unit radius are made, we obtain

$$u_{(1)}^2 = \eta^2 + 2\,\eta \varrho_{(1)} \sin \varepsilon + \varrho_{(1)}^2 + 1 - 2 f_{k(1)} + f_{k(1)}^2 - 2[1 - f_{k(1)}]\varrho_{(1)} \cos \varepsilon. \tag{48}$$

In the above (48) equation, the expression (44) is substituted instead of η:

$$\begin{aligned} u_{(1)}^2 = {} & f_{k(1)}^2 \tan^2 \varepsilon + 2 f_{k(1)} \varrho_{(1)} \tan \varepsilon \cdot \sin \varepsilon + \varrho_{(1)}^2 + \\ & + 1 - 2 f_{k(1)} + f_{k(1)}^2 - 2\,[1 - f_{k(1)}]\,\varrho_{(1)} \cos \varepsilon. \end{aligned} \tag{49}$$

In the following both sides of equation (49) are multiplied by $\cos^2 \varepsilon$; instead of $\sin^2 \varepsilon$, the function $1 - \cos^2 \varepsilon$ is substituted; the whole expression is arranged according to powers of $\cos \varepsilon$ and reduced to zero:

$$\begin{aligned} 2\,\varrho_{(1)} \cos^3 \varepsilon + [u_{(1)}^2 + 2 f_{k(1)} - \varrho_{(1)}^2 - 1] \cos^2 \varepsilon - \\ - 2 f_{k(1)}\,\varrho_{(1)} \cos \varepsilon - f_{k(1)}^2 = 0. \end{aligned} \tag{50}$$

Thus, angle ε can be calculated from the roots of a cubic equation referring to $\cos \varepsilon$.

The equation is divided by $2\,\varrho_{(1)}$ and the following denotations will be introduced:

$$\frac{u_{(1)}^2 + 2 f_{k(1)} - \varrho_{(1)}^2 - 1}{2\,\varrho_{(1)}} = a,$$

$$\cos \varepsilon = \xi, \qquad -f_{k(1)} = b, \qquad -\frac{f_{k(1)}^2}{2\,\varrho_{(1)}} = c$$

and thus, equation (50) assumes the form:

$$\xi^3 + a\,\xi^2 + b\,\xi + c = 0. \tag{51}$$

In the usual way, with the substitution

$$\chi = \xi + \frac{a}{3}$$

equation (51) can be transformed into

$$\chi^3 + p\chi + q = 0 \tag{52}$$

where

$$p = b - \frac{a^2}{3},$$

$$q = \frac{2a^3}{27} - \frac{ab}{3} + c.$$

For finding the roots of the cubic equation (52), we might apply the so-called quadratic resolvent (root-formula) or the Hornerian arrangement. As far as root relations are concerned, the following ranges are known:

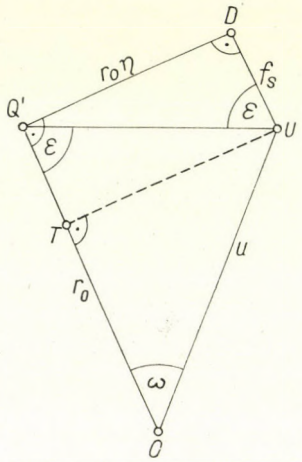


Fig. 15. Detail taken from Fig. 13

when

$$\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3 \begin{cases} > 0 & \text{one of the roots is a real, and the two others are conjugate-complex quantities;} \\ = 0 & \text{the roots are real quantities, and at least two of them are identic;} \\ < 0 & \text{all three roots are real, non-identic quantities.} \end{cases}$$

In our case, one real root is to be expected.

Once the value of ε is known, the angle η is found according to (44), the angle ω is found according to (46), and equation (41) serves to calculate the angle β.

For calculating the angles ω and η with a sharp-cornered rack-tool, it is practicable to analyse the figure *OUDQ'* (Fig. 15) as a detail of Fig. 13. By considering that the substitutions

$$\varrho = 0$$

and

$$f_{k(\varrho=0)} = f_s$$

can be used to find the corresponding relationships for a sharp-cornered tool from the formulae that has been set up for the tool with rounded corners, we obtain from (44)

$$\eta = f_{s(1)} \tan \varepsilon, \tag{53}$$

from (46)

$$\omega = \arcsin \frac{\eta}{u_{(1)}} \tag{54}$$

and from (48)

$$u_{(1)}^2 = \eta^2 + f_{s(1)}^2 - 2 f_{s(1)} + 1 \tag{55}$$

which, of course, can be derived from Fig. 13, too. Yet, when the expression according to (53) is substituted for η into (55) we have:

$$u_{(1)}^2 = f_{s(1)}^2 (1 + \tan^2 \varepsilon) - 2 f_{s(1)} + 1 \tag{56}$$

this latter can also be obtained from (49), by adequate substitutions.
Further, from (56) we have:

$$\tan \varepsilon = \frac{\sqrt{u_{(1)}^2 - f_{s(1)}^2 + 2 f_{s(1)} - 1}}{f_{s(1)}} = \frac{\sqrt{u_{(1)}^2 - [f_{s(1)} - 1]^2}}{f_{s(1)}}. \tag{57}$$

On the basis of (53), the angle of roll η is:

$$\eta = \sqrt{u_{(1)}^2 - [f_{s(1)} - 1]^2}. \tag{58}$$

By substituting this expression into (54) we obtain

$$\omega = \arcsin \frac{\sqrt{u_{(1)}^2 - [f_{s(1)} - 1]^2}}{u_{(1)}}. \tag{59}$$

Thus, the value of β can be found on the basis of known parameters for the given gear. In concrete terms, when the formulae (40), (58) and (59) are substituted into (41), we obtain the arcual value:

$$\operatorname{arc} \beta = \frac{\pi}{2} - \left\{ \arcsin \frac{\sqrt{u_{(1)}^2 - [f_{s(1)} - 1]^2}}{u_{(1)}} - \sqrt{u_{(1)}^2 - [f_{s(1)} - 1]^2} + \right.$$
$$\left. + [1 - f_{s(1)}] \tan \alpha_0 \right\} \tag{60}$$

or by adequate substitutions:

$$\operatorname{arc} \beta = \frac{\pi}{2} - \left\{ \arcsin \sqrt{1 - \left(\frac{mz}{2u}\right)^2 \left[\frac{2(1 + c_0' - x)}{z} - 1 \right]^2} - \right.$$
$$\left. - \sqrt{\left(\frac{2u}{mz}\right)^2 - \left[\frac{2(1 + c_0' - x)}{z} - 1 \right]^2} + \left[1 - \frac{2(1 + c_0' - x)}{z} \right] \tan \alpha_0 \right\}. \tag{61}$$

Let us consider Figs 16 and 17. As will be seen from what follows, we can determine the arc length s of the circle arc as having the radius u and being limited in a simple way from the known angle β; (which corresponds to the determination of the root thickness).

As is known, the arc length s_0 on the pitch circle is:

$$s_0 = m\left(\frac{\pi}{2} + 2\,x\tan\alpha_0\right), \tag{62}$$

and the corresponding angle:

$$2\,\text{arc}\,\psi_0 = \frac{s_0}{r_0} = \frac{2\left(\frac{\pi}{2} + 2\,x\tan\alpha_0\right)}{z}. \tag{63}$$

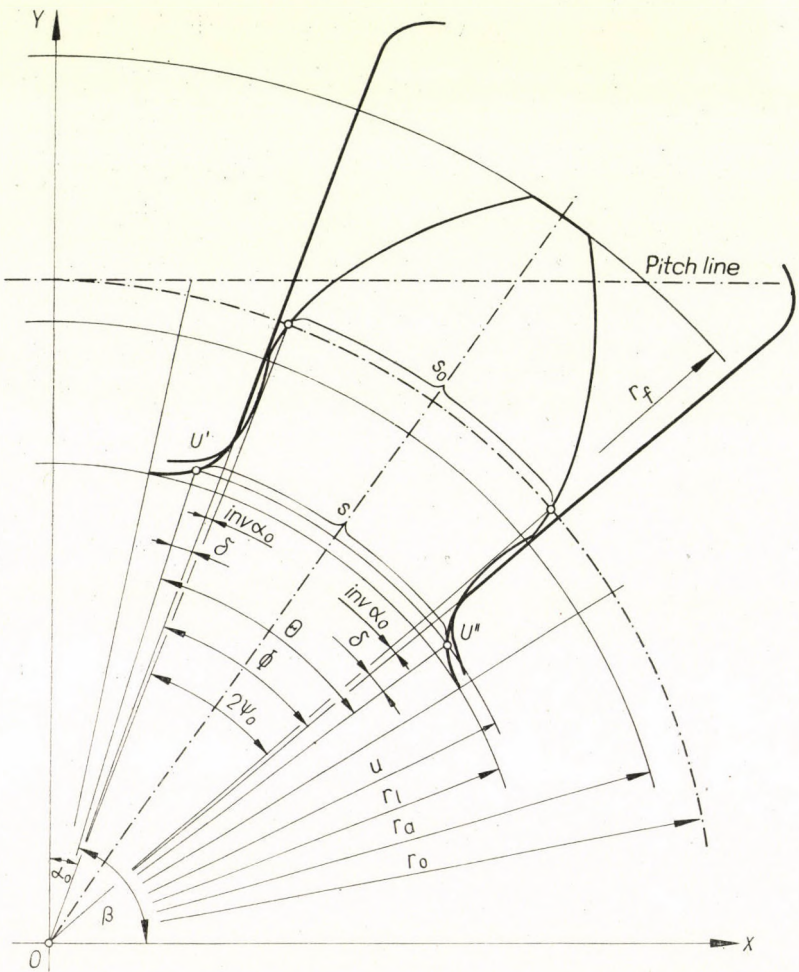


Fig. 16. Diagram for the calculation of the arc length on a circle of any optional radius when the limiting trochoids are generated by a round-cornered rack-tool

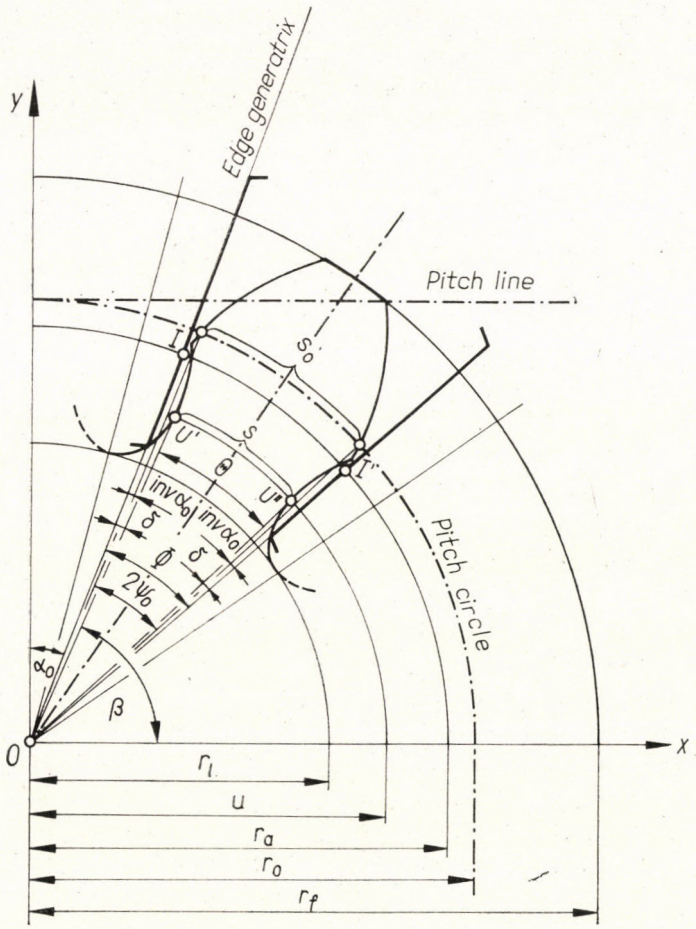


Fig. 17. Diagram for the calculation of the arc length on a circle of any optional radius when the limiting trochoids are generated by a sharp-cornered rack-tool

Now, the angle δ, that is formed by the straight line $\overline{OU'}$ belonging to the trochoidal point U', and by the straight $\overline{OI}$ belonging to the starting point I of the profile involute, developed from the base circle with radius r_a, is expressed according to Figs 16 and 17 as follows:

$$\text{arc}\,\delta = \frac{\pi}{2} - \text{arc}\,\alpha_0 - \text{arc}\,\beta \tag{64}$$

again, belonging to the arc $\widehat{II'}$ measured on the base circle of the radius r_a, the angle is:

$$\text{arc}\,\Phi = 2\ \text{arc}\,\psi_0 + 2\ \text{inv}\,\alpha_0 \tag{65}$$

and, belonging to the arc $s = \widehat{U'U''}$, measured between the trochoidal points

U' and U'' on the circle of the radius u, the angle Θ is:

$$\operatorname{arc} \Theta = \operatorname{arc} \Phi - 2 \operatorname{arc} \delta, \tag{66}$$

from which, the arc s lying on an optional circle of a given radius u, is:

$$s = u \operatorname{arc} \Theta . \tag{67}$$

By adequate substitutions, into (67) the expressions given in (62), . . . , (66), and by taking the relationship:

$$\operatorname{inv} \alpha_0 + \operatorname{arc} \alpha_0 = \tan \alpha_0$$

into account we obtain the generalized equation:

$$s = 2u \left(\frac{\frac{\pi}{2} + 2x \tan \alpha_0}{z} + \tan \alpha_0 - \frac{\pi}{2} + \operatorname{arc} \beta \right). \tag{68}$$

Obviously, in a certain case, at a given locus, the cross section of a tooth as measured between trochoids generated by a rack-tool with rounded corners is larger than it would be when generated by a sharp-cornered rack-tool. Consequently, for dimensioning, we can confine ourselves for calculating the root cross section of a toothing generated by a sharp-cornered rack-tool: by using this simplified formula, the factor of safety becomes greater. In manufacturing practice, it is usual to have the tooth profile generated by a round-cornered rack-tool, and to finish it by a sharp-cornered grinding wheel; so one has to take root cross section determined by trochoidal geometry into account.

Thus we obtain, when calculating the arc s at a given radius u, derived from the generalized formula (68) by substituting the expression of arc β according to (61), the new equation [2]:

$$s = 2u \left\{ \frac{\pi/2 + 2x \tan \alpha_0}{z} - \operatorname{arc} \sin \sqrt{1 - \left(\frac{mz}{2u}\right)^2 \left[\frac{2(1 + c_0' - x)}{z} - 1 \right]^2} + \right.$$
$$\left. + \sqrt{\left(\frac{2u}{mz}\right)^2 - \left[\frac{2(1 + c_0' - x)}{z} - 1 \right]^2} + \frac{2(1 + c_0' - x)}{z} \tan \alpha_0 \right\}. \tag{69}$$

In a given toothing system, α_0 and c_0' are constants, wherefore the arc s at an optimal radius u, limited by trochoids can be calculated from the given parameters of the gear viz.: number of teeth, module, and addendum modification coefficient.

Briefly, the course of calculation is the following:

1. For teeth generated by a round-cornered rack-tool:

a) from (50), the third root of $\cos^3 \varepsilon$ is determined;

b) from (44) and (46), the angles η and ω are calculated;

c) by using the obtained angle values, and the value of angle ϑ_l according to (36), and by substituting these into the formula (41) of the angle β, the arc length s can be calculated according to (68).

2. For teeth generated by a sharp-cornered rack-tool: the arc s can be calculated from the explicit equation (69). This calculation is adequate for the dimensioning of a tooth with the view of mechanical strength even when the root thickness is limited between trochoids generated by a round-cornered rack-tool, because this approximation is tantamount to a larger value of the safety factor.

IV. Numeric examples

We obtained, for the pinion of a given pair of gears, on the basis of $\alpha_0 = 20°$ and $c_0' = 0{,}25$, the following parameters:

$$z = 17;\ m = 10 \text{ mm};\ r_0 = m\frac{z}{2} = 85 \text{ mm};$$

$$x = 0{,}4716 \text{ mm};\ f_s = m(1 + c_0' - x) = 7{,}7838 \text{ mm};$$

$$r_l = r_0 - f_s = 77{,}2152 \text{ mm}.$$

We calculate the arc length s_t, measured between trochoids on a circle of the radius r_t, which pertains to the locus of the maximum root stress:

First, for a tooth profile generated by a round-cornered rack-tool;

second, for a tooth profile generated by a sharp-cornered rack-tool.

According to our derived formulae:

$$u = r_t = r_l + \frac{1}{6}m = 78{,}88286 \text{ mm}$$

or

$$u_{(1)} = r_{t(1)} = \frac{r}{r_0} = 0{,}92803.$$

For the first alternative, the necessary parameters are:

$$\varrho = \frac{mc_0'}{1 - \sin \alpha_0} = 3{,}79951 \text{ mm}$$

or

$$\varrho_{(1)} = \frac{\varrho}{r_0} = 0{,}04470;$$

and

$$f_k = f_s - \varrho = 3{,}98429 \text{ mm}$$

or

$$f_{k(1)} = \frac{f_k}{r_0} = 0{,}04687.$$

1. Calculation for the case of a round-cornered generating rack-tool

a) According to (50)

$$2\,\varrho_{(1)}\cos^3\varepsilon + [u_{(1)}^2 + 2f_{k(1)} - \varrho_{(1)}^2 - 1]\cos^2\varepsilon - 2f_{k(1)}\,\varrho_{(1)}\cos\varepsilon - f_{k(1)}^2 = 0$$

we have numerically

$$0{,}08940\cos^3\varepsilon - 0{,}04700\cos^2\varepsilon - 0{,}00419\cos\varepsilon - 0{,}00220 = 0,$$

from which (e.g. according to Horner's scheme, by a gradual approximation) one real root is obtained:

$$\cos\varepsilon = 0{,}6547$$

viz.:

$$\varepsilon = 49^\circ\,6'\,10''.$$

b) According to Equ. (44)

$$\eta = f_{k(1)}\tan\varepsilon = 0{,}04687 \cdot 1{,}15455 = 0{,}05412.$$

c) According to Equ. (46)

$$\omega = \arcsin\frac{\eta + \varrho_{(1)}\sin\varepsilon}{u_{(1)}} = \arcsin\frac{0{,}05412 + 0{,}04470 \cdot 0{,}75589}{0{,}92803} = 0{,}09486.$$

d) According to Equ. (36)

$$\vartheta_l = \tan\alpha_0 - \frac{2}{z}\left[\frac{c_0'}{1-\sin\alpha_0}\cos\alpha_0 + (1-x)\tan\alpha_0\right] =$$

$$= 0{,}36397 - \frac{2}{17}\left[\frac{0{,}25}{1-0{,}34202}\,0{,}93969 + (1-0{,}4716)\cdot 0{,}36397\right] = 0{,}29934.$$

e) According to (41)

$$\beta = \frac{\pi}{2} - (\omega - \eta + \vartheta_l) = 1{,}57080 - (0{,}09486 - 0{,}05412 + 0{,}29934) = 1{,}23072.$$

f) According to (68) the arc length s_t is, on the circle arc of the radius $u = r_t$:

$$s_t = 2u\left(\frac{\pi/2 + 2x\tan\alpha_0}{z} + \tan\alpha_0 - \frac{\pi}{2} + \operatorname{arc}\beta\right) =$$

$$= 2 \cdot 78{,}88286\left(\frac{1{,}57080 + 2\cdot 0{,}4716\cdot 0{,}36397}{17} + 0{,}36397 - 1{,}57080 + 1{,}23072\right)$$

the final result of which equals

$$s_t = 21{,}533 \text{ mm}.$$

This is, as can be seen from Fig. 1, identic with the geometrically constructed length.

2. Calculation for the case of a sharp-cornered generating rack-tool

For the arc length s_t, as measured on the circle of the radius $u = r_t$, the explicit formula (69) goes as follows:

$$s_t = 2u\left\{\frac{\pi/2 + 2x\tan\alpha_0}{z} - \arcsin\sqrt{1 - \left(\frac{mz}{2u}\right)^2\left[\frac{2(1 + c_0' - x)}{z} - 1\right]^2} + \right.$$

$$\left. + \sqrt{\left(\frac{2u}{mz}\right)^2 - \left[\frac{2(1 + c_0' - x)}{z} - 1\right]^2} + \frac{2(1 + c_0' - x)}{z}\tan\alpha_0\right\} =$$

$$= 2 \cdot 78{,}88286 \left\{ \frac{1{,}57080 + 2 \cdot 0{,}4716 \cdot 0{,}36397}{17} - \right.$$

$$- \arcsin \sqrt{1 - \left(\frac{170}{2 \cdot 78{,}88286} \right)^2 \left[\frac{2(1 + 0{,}25 - 0{,}4716)}{17} - 1 \right]^2} +$$

$$+ \sqrt{\left(\frac{2 \cdot 78{,}88286}{170} \right)^2 - \left[\frac{2\,(1 + 0{,}25 - 0{,}4716)}{17} - 1 \right]^2} +$$

$$\left. + \frac{2\,(1 + 0{,}25 - 0{,}4716)}{17}\ 0{,}36397 \right\}$$

the final result of which equals

$$s_t = 20{,}412$$

that is, again, fairly identical with the geometric construction in Fig. 2.

So we can state, that

$$s_{t,\,\text{cornered}} < s_{t,\,\text{rounded}}\,.$$

REFERENCES

1. Jacobson, M. A.: Bending Stresses in Spur Gear Teeth. Proposed New Design Factors Based on a Photo-Elastic Investigation. *Proceedings of the Mechanical Engineers* **169** (1955). Jacobson, M. A.: Diskussionbeitrag zum Aufsatz von Kelley, B. W.—Pedersen, R. „Zahnfußfestigkeit bei neuzeitlichen Getriebekonstruktionen." Getriebe-Kupplungen, Antriebselemente. Vieweg, Braunschweig 1957.
2. Horváth, F.: Berechnung der Zahnwurzeldicke an der durch Trochoiden begrenzten Zahnwurzel. *Acta Techn. Hung.* **53** (1966).
3. Hofer, H.: *Automobil Technische Zeitschrift* (1947), 2. März.
4. Disteli, M.: Über einige Sätze der kinematischen Geometrie, welche der Verzahnungslehre zylindrischer und konischer Räder zugrunde liegen. *Zeitschrift für Mathematik und Physik* **56** (1908), 233.
5. Rothe, R.: Höhere Mathematik für Mathematiker, Physiker und Ingenieure (7. §.). B. G. Teubner Verlagsgesellschaft, Leipzig 1952.

Zahnwurzeldicke einer Evolventenverzahnung an der von Trochoiden begrenzten Zahnwurzel. In der Abhandlung wird die Berechnung der von Trochoiden begrenzten Zahnwurzelstärke einer mit Hilfe eines eckigen bzw. abgerundeten Zahnstangenwerkzeuges im Abwälzverfahren erzeugten Verzahnung besprochen. Die Ableitung der Berechnung nach Gesetzen der Trochoiden-Geometrie gibt dem Verfasser Gelegenheit zur Verallgemeinerung der Untersuchung. Es werden die Bedingungen bestimmt, auf deren Grund der kritische Zahnwurzelquerschnitt nur durch Anwendung der Trochoiden-Geometrie ermittelt werden kann. Das mittels eines eckigen Abwälzstangenwerkzeuges erzeugte Zahnprofil wird als Grenzfall des mittels einer abgerundeten Zahnstange erzeugten Profils betrachtet, und die von Trochoiden begrenzte Zahnstärke wird an beliebiger Stelle der Profilwurzel bestimmt. Zur Berechnung der Zahnwurzelstärke an beliebiger Stelle der von Trochoiden begrenzten Verzahnung genügt die Kenntnis des Moduls, der Zähnezahl und des Profilverschiebungsfaktors. Für den Fall eines eckigen Abwälzinstrumentes läßt sich das Ergebnis in expliziter Form ausdrücken, die sich leicht für Berechnung eignet.

L'épaisseur de base des dentures à développante, limitée par des trochoïdes. L'étude traite du calcul de l'épaisseur des pieds de dents limités par des trochoïdes d'une denture à développante taillée avec des crémaillères de profil arrondi et à angles vifs. Au cours du calcul de la surface de base déterminable par la géométrie des trochoïdes, les examens sont généralisés par l'étude. Elle définit les conditions nécessaires à ce que la surface de base critique ne puisse être déterminée que par la géométrie des trochoïdes. La denture générée par une crémaillère à angles vifs est considérée comme un cas spécial de celle générée par un outil à angles arrondis et la surface de base limitée par des trochoïdes est déterminée à des endroits librement choisis. L'épaisseur des pieds de dent limités par des trochoïdes peut être calculée à un endroit quelconque si le module, le nombre des dents et le coefficient de déport sont connus. En cas d'un outil générateur à angles vifs, le résultat final peut être facilement calculé sous une forme explicite.

Толщина корня зуба, ограниченного трохоидами, в случае эвольвентного зацепления *(Ф. Хорват)*. Предметом данной работы является расчет толщины корня зуба, ограниченного трохоидами, при изготовлении зубьев методом обкатки, используя инструмент с округленным и угловым профилем гребенки. По ходу расчета сечения корня зуба, определяемого с помощью трохоидной геометрии, в работе дается обобщение исследований. Указываются условия, при удовлетворении которых критическое сечение корня зуба может быть определено только при помощи трохоидной геометрии. Зубья, полученные при помощи инструмента с угловым профилем гребенки, автор считает специальным случаем зубьев, изготовленных при помощи инструмента с округленным профилем гребенки, и сечение корня зуба, ограниченное трохоидами, определяется им в произвольном месте. Толщина корня зуба, ограниченного трохоидами, в каком-нибудь произвольном месте, если известны модуль, число зубьев и коэффициент смещения профиля, может быть вычислена. В случае углового обрабатывающего инструмента конечный результат можно получить в легко вычисляемой эксплицитной форме.

INDEX

Verő, J. A.: The Yield Strength in Homogeneous Alloys of α-Iron — Einfluß der gelösten Legierungselemente auf die Fließgrenze des α-Eisens — L'influence des éléments d'alliage dissous sur la limite d'écoulement du fer α — *Й. Верэ:* Воздействие растворенных легирующих элементов на условный предел текучести α-железа 243

Lipka, I.: Geometrie der Abwälzfräser mit Grundschnecken evolventenförmiger Ausführung — The Geometry of Hobs with Involute Basic Worm — Géométrie des fraises vis-mère à vis de base à développante — *И. Липка:* Геометрия обкаточных червячных фрез с эвольвентным червяком 261

Szőke, B.: The Proper Conduction of the Flux of Mechanical Energy with Special Regard to the Self-Balanced Movable Clutch — Zweckmäßige Leitung des Energieflusses; die bewegliche, selbstausgleichende Wellenkupplung — Conduite efficace de flux d'énergie et accouplement auto-équilibrant mobile — *Б. Сёке:* Целесообразное проведение потока энергии и самоуравновешивающаяся подвижная муфта сцепления 277

Tóth, L.: Distribution of Stresses Acting on the Roll Surface in Strip Rolling — Verteilung der Druckspannung an der Walzenoberfläche beim Auswalzen von dünnem Stahlblech — Répartition des pressions à la surface d'un rouleau au laminage des feuillards — *Л. Тот:* Распределение напряжений, действующих на поверхность валков, при прокатке лент 351

Lipka, I.: Geometrie der Schneckenwälzfräser mit Grundschnecken schleifenevolventischer Ausführung — The Geometry of Generating Cutters with Convolute Basic Worm — Géométrie des fraises vis-mère à vis de base convolutée — *И. Липка:* Геометрия обкаточных фрез с конволютным червяком 371

Wawra, H.: Zur Frage der Konzentrationsabhängigkeit der Poisson-Zahl — The Problem of Poisson's Ratio as a Function of Concentration — Le problème de la constante de Poisson en fonction de la concentration — *Х. Вавра:* К вопросу зависимости от концентрации числа Пуассона 401

Horváth, Z.: The Determination of Equilibrium and Kinetics in the Processes Occurring in Oxidation-Smelting — Kinetische und Gleichgewichtsuntersuchung der Vorgänge beim oxydierenden Schmelzen — Examen de l'équilibre et de la cinétique des phénomènes accompagnant la fusion oxydante — *З. Хорват:* Равновесные и кинетические исследования процессов при окислительной плавке 413

Pranab Kumar De: A Theoretical Approach to the Determination of the Amount of Wear during Sliding between Surfaces — Theoretische Bestimmung des Betrages des Abriebs zwischen gleitenden Oberflächen — Détermination théorique approchée de l'usure entre surfaces de frottement — *Пранаб Кумар Де:* Теоретический приближенный метод определения степени износа при трении между поверхностями 429

Horváth, F.: Tooth Root Thickness between Trochoidal Root Profiles on Involute-Shaped Gears — Zahnwurzeldicke einer Evolventenverzahnung an der von Trochoiden begrenzten Zahnwurzel — L'épaisseur de base des dentures à développante, limitée par des trochoïdes — *Ф. Хорват:* Толщина корня зуба, ограниченного трохоидами, в случае эвольвентного зацепления 439

Third Conference on Dimensioning and Strength Calculation

October 27 to November 3, 1968
Organized by the Technical Section of the
Hungarian Academy of Sciences

Subject matter of the Conference is: problems connected with fatigue and brittle failure of structures in mechanical and civil engineering practice, further, theoretical questions referring to the same phenomena.

Applications for attendance should be sent to the following address:

Organizing Committee of the Third Conference on Dimensioning and Strength Calculation, P. O. B. 91, Budapest 112, Hungary.

Any further informations may also be obtained from the same Committee.

antriebs-technik international



Unabhängige Zeitschrift für Maschinen, Getriebe und Antriebselemente

Die Umformung von Bewegung und Kräften, die Konstruktion und Koppelung der vielfältigen Kraft- und Verarbeitungsmaschinen werden von „ant" unter Zugrundelegung der neuesten technischen Erkenntnisse behandelt.

Wirtschaftliche Gesichtspunkte erfordern die Anwendung moderner Konstruktionsprinzipien und technischer Verfahren. Zeitgemäße Fertigungsmethoden und der Einsatz von Standardelementen und Aggregaten sind notwendig geworden.

Dieser in fast allen Zweigen der Technik aktuellen Thematik ist die Zeitschrift gewidmet, die in wenigen Jahren eine Spitzenstellung auf dem europäischen Kontinent erringen konnte.

1967 im 6. Jahrgang

Erscheinungsweise: monatlich

Jahresbezugspreis Ausland: DM 54,— einschl. Porto

Krausskopf-Verlag • D-65 Mainz • Lessingstraße 12-14

(Krausskopf-Haus)

Printed in Hungary

A kiadásért felel az Akadémiai Kiadó igazgatója — Műszaki szerkesztő: Farkas Sándor

A kézirat nyomdába érkezett: 1968. II. 5. — Terjedelem: 19,75 (A/5) ív, 102 ábra, 3 melléklet

68.65067 Akadémiai Nyomda, Budapest — Felelős vezető: Bernát György

Acta Techn. Hung. **61** (1968), 243—260

VERŐ, J.: *The Yield Strength in Homogeneous Alloys of Iron*

The equation derived by v. GÖLER, SACHS and HONEYCOMBE constitutes a simple relation between atomic dimensions and strengthening effect of substitutionally dissolved alloying elements. The constant standing in this equation was first determined making use of previous experimental results obtained with homogeneous alloys of α-iron. Then, assuming that interstitial carbon atoms displace the nearest iron atoms in the lattice and hereby cause a similar distortion of the slip planes, as if these nearest iron atoms would have been substituted by greater foreign atoms, it is shown that the yield strength is raised by one atomic per cent carbon dissolved in α-iron by 87—87,3 kp/mm², in excellent agreement with the most reliable earlier experimental results.

Acta Techn. Hung. **61** (1968), 261—276

LIPKA, I.: *The Geometry of Hobs with Involute Basic Worm*

The author deduced a criterion for deciding whether a gear profile produced by some generating tool is an ordinary involute of a circle or a curve differing from it. This involute criterion decides by the fulfilment of a partial differential equation the character of the generated tooth profile. In the present paper the author applies the involute criterion to a gear profile generated by a milling cutter with an involute basic worm. He proves that a tooth profile generated by such a cutter is not an ordinary (pointed) involute of a circle.

Acta Techn. Hung. **61** (1968), 277—350

SZŐKE, B.: *The Proper Conduction of the Flux of Mechanical Energy with Special Regard to the Self-Balanced Movable Clutch*

Up to now, it was not possible to find, in the technical literature such a special formula which has now been deduced by the author on plain theoretical basis inasmuch as the proposed equation serves as a conclusive explanation of the strange occurrence experienced with a series of mill rolls, when namely, the feed-back output disappears at a gear ratio approaching 1 : 1 or ∞ : 1. The author demonstrates a special grouping of the coupling in series of mill frames having rolls with various diameters in which the feed-back of the energy-flux does never charge the driving gear but its circulation occurs within the roller mill system: thus the whole train of mechanism (gears, clutches) has to bear a less load. Further, the self-balancing movable clutch is considered. By theoretic analysis it is shown that this type of clutch represents a newly modelled machine part that serves as an insulator of oscillations and not as a damper.

Acta Techn. Hung. **61** (1968), 351—370

TÓTH, L.: *Distribution of Stresses Acting on the Roll Surface in Strip Rolling*

The first part of this paper contains the description of the basic equations of rolling; the second part presents the solution of these basic equations. The approximate solution is based on knowledge of the enlargement of the pieces during rolling. The solution contains the flow diagram and the determination of the compression stresses acting on the surface of the roll. The correctness of the results is proved by comparing them with values determined by experiment.

Acta Techn. Hung. **61** (1968), 371—400

LIPKA, I.: *The Geometry of Generating Cutters with Convolute Basic Worm*

The author investigates gear profiles generated by a hob the basic worm of which is a convolute worm made on a lathe with a dipped tool. He shows that a tooth profile generated with such a hob approaches the ordinary involute of a circle the better, the smaller the helix angle of the cutter is. The author finally applies the analytical criterion for deciding the involute character of the generated tooth profile to the gear profile made with a convolute hob.

Acta Techn. Hung. **61** (1968), 401—412

WAWRA, H.: *The Problem of Poisson's Ratio as a Function of Concentration*

Poisson's ratio as a function of concentration according to indirect measurements on copper- and brass-cylinders is described and discussed in detail. We distinguish types μ I as concentration diagrams of deformed materials, types μ II os soft-annealed specimens, and mixed types, which are all distinguished by characteristic extreme values, chiefly as a result of the deviations of statistical distribution of the crystallites of the polycristallite aggregation.

Acta Techn. Hung. **61** (1968), 413—428

HORVÁTH, Z.: *The Determination of Equilibrium and Kinetics in the Processes Occurring in Oxidation Smelting*

The work describes the processes which take place during oxidation-smelting and then analyses the possibilities and the most suitable working methods of eliminating the impurity by the same process for different cases. After accounting for the equilibrium conditions, there is a treatment of the kinetics of the reactions taking place in individual equipment for different cases.

Acta Techn. Hung. **61** (1968), 429—438

PRANAB, KUMAR DE: *A Theoretical Approach to the Determination of the Amount of Wear During Sliding between Surfaces*

The present work takes into consideration a model of wear in which some of the asperities occurring in the surfaces of contact, tear away during sliding to give rise to loose particles. These particles alone constitute wear in this model. The tearing of particles would require a certain amount of mechanical work to be done on them. This work may be associated with the breaking work of the material which is the work necessary to break by tearing a unit volume of the material. The geometry of the asperities and the forces acting on them led to a method for finding the total work of tearing of a particular length of sliding. The total tearing work and the breaking work of the material then permits the evaluation of the volume of wear particles for a given distance of sliding.

Acta Techn. Hung. **61** (1968), 439—466

HORVÁTH, F.: *Tooth Root Thickness between Trochoidal Root Profiles on Involute-Shaped Gears*

The paper deals with the calculation of the root thickness of involute gearing made by generating with a tool based on a rounded or on a sharp-edged rack. In the course of the calculation of the root thickness which can be determined by trochoidic geometry the author generalizes the results of his investigations. He gives the conditions which must be fulfilled so that the critical root thickness can be determined by trochoid geometry only. The gearing generated by a sharp-edged tool is considered as a special case of the gearing generated by a rounded tool and the root thickness limited by trochoids is determined for an arbitrary location. If the module, the number of teeth and the profile displacement factor are known, the root thickness for an arbitrary place of a tooth root limited by trochoids can be calculated. In the case of a sharp-edged generating tool the final result can be presented in an easily calculated explicit form.

The Acta Technica publish papers on technical subjects in English, German, French and Russian.

The Acta Technica appear in parts of varying size, making up volumes.

Manuscripts should be addressed to

Acta Technica
Nádor u. 7.
Budapest V.
Hungary

Correspondence with the editors and publishers should be sent to the same address.

The rate of subscription to the Acta Technica is 165 forints a volume. Orders may be placed with "Kultúra" Foreign Trade Company for Books and Newspapers (Budapest I., Fő utca 32. Account No. 43-790-057-181) or with 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
Nádor u. 7.
Budapest V.
Hongrie

Toute correspondance doit être envoyée à cette même adresse.

Le prix de l'abonnement est de 165 forints par volume.

On peut s'abonner à l'Entreprise pour le Commerce Extérieur de Livres et Journaux «Kultúra» (Budapest I., Fő utca 32. Compte courant No. 43-790-057-181) ou à l'étranger chez tous les représentants ou dépositaires.

«*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 рукописи следует направлять по адресу:

Acta Technica
Nádor u. 7.
Budapest V.
Венгрия

По этому же адресу направлять всякую корреспонденцию для редакции и администрации.

Подписная цена «*Acta Technica*» — 165 форинтов за том. Заказы принимает предприятие по внешней торговле книг и газет «Kultúra» (Budapest I., Fő utca 32. Текущий счет № 43-790-057-181) или его заграничные представительства и уполномоченные.

88,— *Ft*

Reviews of the Hungarian Academy of Sciences are obtainable at the following addresses:

ALBANIA
Ndermarja Shtetnore e Botimeve
Tirana

AUSTRALIA
A. Keesing
Box 4886, GPO
Sydney

AUSTRIA
Globus Buchvertrieb
Salzgries 16
Wien I

BELGIUM
Office International de Librairie
30, Avenue Marnix
Bruxelles 5
Du Monde Entier
5, Place St. Jean
Bruxelles

BULGARIA
Raznoiznos
1, Tzar Assen
Sofia

CANADA
Pannonia Books
2, Spadina Road
Toronto 4, Ont.

CHINA
Waiwen Shudian
Peking
P. O. B. 88

CZECHOSLOVAKIA
Artia
Ve Směčkách 30
Praha 2
Poštovná novinová služba
Dovoz tisku
Vinohradská 46
Praha 2
Mad'arská Kultura
Václavské nám. 2
Praha I
Poštová novinová služba
Dovoz tlače
Leningradská 14
Bratislava

DENMARK
Ejnar Munksgaard
Nörregade 6
Copenhagen

FINLAND
Akateeminen Kirjakauppa
Keskuskatu 2
Helsinki

FRANCE
Office International de Documentation et Librairie
48, rue Gay Lussac
Paris 5

GERMAN DEMOCRATIC REPUBLIC
Deutscher Buch-Export und Import
Leninstraße 16
Leipzig 701
Zeitungsvertriebsamt
Clara Zetkin Straße 62
Berlin N. W.

GERMAN FEDERAL REPUBLIC
Kunst und Wissen
Erich Bieber
Postfach 46
7 Stuttgart S.

GREAT BRITAIN
Collet's Holdings Ltd.
Dennington Estate
London Rd.
Wellingborough, Northants.
Robert Maxwell and Co. Ltd.
Waynflete Bldg. The Plain
Oxford

HOLLAND
Swetz and Zeitlinger
Keizersgracht 471—487
Amsterdam C
Martinus Nijhof
Lange Voorhout 9
The Hague

INDIA
Current Technical Literature
Co. Private Ltd.
India House OPP
GPO Post Box 1374
Bombay I

ITALY
Santo Vanasia
Via M. Macchi 71
Milano
Libreria Commissionaria Sansoni
Via La Marmora 45
Firenze

JAPAN
Nauka Ltd.
92, Ikebukur O-Higashi 1-chome
Toshima-ku
Tokyo
Maruzen and Co. Ltd.
P. O. Box 605
Tokyo-Central
Far Eastern Booksellers
Kanda P. O Box 72
Tokyo

KOREA
Chulpanmul
Phenjan

NORWAY
Johan Grundt Tanum
Karl Johansgatan 43
Oslo

POLAND
Ruch
ul. Wronia 23
Warszawa

ROUMANIA
Cartimex
Str. Aristide Briand 14—18
Bucuresti

SOVIET UNION
Mezhdunarodnaya Kniga
Moscow G—200

SWEDEN
Almquist and Wiksell
Gamla Brogatan 26
Stockholm

USA
Stechert Hafner Inc.
31, East 10th Street
New York, N. Y. 10003
Walter J. Johnson
111, Fifth Avenue
New York, N. Y. 10003

VIETNAM
Xunhasaba
19, Tran Quoc Toan
Hanoi

YUGOSLAVIA
Forum
Vojvode Mišića broj 1
Novi Sad
Jugoslovenska Knjiga
Terazije 27
Beograd

18. VI. 1968. Index: 26.025