MATH. SCAND. 25 (1969), 49–58

ALGORITHMEN FÜR EINFACHE KURVEN AUF FLÄCHEN II

HEINER ZIESCHANG

7. Einleitung.

Zur Entscheidung, ob eine gegebene Homotopieklasse von Wegen auf einer Fläche eine einfache Kurve enthält oder nicht, war in dem ersten Teil [18] ein Algorithmus angegeben, und zwar ein rein gruppentheoretischer. Das Problem wird auf eine Frage in freien Gruppen zurückgeführt, welche mit der Whiteheadschen Methode algorithmisch entscheidbar ist. Allerdings sind dabei Prozesse zugelassen, die geometrisch keinen Sinn haben: es werden Wechsel zu Erzeugenden gestattet, die nicht in Verbindung mit Zerschneidungen der Fläche stehen. In diesem Teil wollen wir zeigen, daß zur Durchführung des Whiteheadschen Verfahrens nur solche Prozesse benötigt werden, die Zerschneidungswechseln entsprechen. — Obgleich nun ein sehr viel geringerer Aufwand zur Durchführung des Algorithmus als vorher vonnöten ist, scheint mir für eine praktische Durchführung wohl die sich in [18] anbietende geometrische Methode empfehlenswert. Der Zusatz hat somit überwiegend theoretisches Interesse, und erst hier wird [18] befriedigend abgeschlossen.

Die Arbeit folgt in den Bezeichnungen dem ersten Teil. In ihm waren die Standardtypen für einfache Kurven auf nicht-orientierbaren Flächen unvollständig angegeben und somit ist Satz 4 nicht korrekt. Im nächsten Abschnitt geben wir alle Typen an. Satz 4 und sein Beweis lassen sich ohne Schwierigkeiten übertragen. Für die Korrektur danke ich Herrn Elmar Vogt.

8. Typen einfach-geschlossener Kurven auf Flächen.

Ist die Fläche orientierbar, so sind je zwei nicht-zerlegende Kurven durch einen Homöomorphismus ineinander abbildbar, zwei zerlegende genau dann, wenn sie in gleicher Weise zerlegen. Abgelesen an einem kanonischen Kurvensystem ergeben sich als Repräsentanten der Typen die Worte T_1 bzw.

Eingegangen am 8. Juli 1968.

$$S_1 \ldots S_q \prod_{i=1}^{k} [T_i, U_i], \quad 0 \leqq q \leqq m, \ 0 \leqq k \leqq g .$$

Ist die Fläche nicht-orientierbar, so können nicht-zerlegende Kurven ein- oder zweiseitig sein und die Fläche in eine orientierbare oder eine nicht-orientierbare Fläche aufschneiden. Erhalten wir nach Aufschneiden an einer einseitigen Kurve eine nicht-orientierbare Fläche, so hat sie V_1 als Repräsentanten $(g > 1)$; erhalten wir eine orientierbare Fläche, so repräsentiert $V_1 \cdot \ldots \cdot V_g$ (g ungerade). Ist die Kurve zweiseitig und schneidet sie in eine nicht-orientierbare Fläche auf, so repräsentiert $V_1 V_2$ $(g > 2)$; schneidet sie in eine orientierbare Fläche auf, so ist g gerade und $V_1 \cdot \ldots \cdot V_g$ Repräsentant. Zerlegende Kurven sind wieder durch die Teile, in die sie zerlegen, bestimmt. Hier kann es sein, daß beide Teile nicht-orientierbar sind oder auch nur einer. Im ersten Falle bekommt man eine äquivalente Kurve zu einem Wort

$$S_1 \ldots S_q V_1^2 \ldots V_k^2 \quad \text{mit } 0 < k < g, \ 0 \leqq 1 \leqq m .$$

Im zweiten Fall führe man günstiger andere Erzeugende ein, derart daß die definierende Relation zu

$$S_1 \ldots S_m \prod_{i=1}^{\frac{1}{2}(g-1)} [T_i, U_i] V_g^2$$

für ungerades g bzw.

$$S_1 \ldots S_m \prod_{i=1}^{\frac{1}{2}(g-2)} [T_i, U_i] T V T^{-1} V$$

für gerades g wird. Die zerlegende Kurve ist dann, wenn man sie von der orientierbaren Seite her betrachtet, äquivalent zu einer Kurve vom Wort

$$S_1 \ldots S_q \prod_{i=1}^{k} [T_i, U_i] .$$

Will man von der anderen Seite schauen, so hat man neben Randerzeugenden und Kommutatoren das V_g^2 oder das Glied $T V T^{-1} V$ hinzuzunehmen.

9. Zerschneidungen und Zweiteilungen.

F sei eine kompakte Fläche mit Randkurven $\varrho_1, \ldots, \varrho_m$. Ferner sei $\hat{p}$ ein innerer Punkt von F, $\hat{p}_1, \ldots, \hat{p}_m$ auf $\varrho_1, \ldots, \varrho_m$ bzw. Als *Zerschnei-*

dung von F bezeichnen wir ein einfaches Kurvensystem $Z^* = \{\zeta_1^*, \ldots, \zeta_n^*\}$ auf F mit den folgenden Eigenschaften:

(9.1) Die Kurven haben den gemeinsamen Anfangspunkt $\hat{p}$.

(9.2) Für die Randkurve ϱ_j gibt es ein und nur ein ζ_i^*, welches $\hat{p}_j$ mit $\hat{p}$ verbindet. Alle übrigen ζ_i^* sind geschlossen.

(9.3) Schneidet man die Fläche an dem System Z^* auf, so erhält man eine Scheibe.

Spezielle Zerschneidungen sind die kanonischen (vergl. Abschnitt 4).

Eine Kurve ζ^* gehe von $\hat{p}$ aus und führe entweder zu $\hat{p}$ zurück oder zu einem $\hat{p}_j$, ohne dabei in inneren Punkten die Zerschneidung zu treffen. Ferner möge ζ^* die Fläche nicht zerlegen. Dann kann man eine Zerschneidungslinie aus Z^* entfernen, durch ζ^* ersetzen und erhält eine neue Zerschneidung Z'^*. Führt ζ^* zum Rand ϱ_i, so ist die zu ϱ_i gehörige Strecke aus Z^* wegzulassen, für ein geschlossenes ζ^* kann man irgendeine Kurve ζ_i^* wählen, die "durch ζ^* von ihrem Partner im Rande der Scheibe getrennt wird". Wir sagen, daß Z'^* aus Z^* durch *Zweiteilung* hervorgeht. — Nach Aufschneiden der Fläche an Z^* in eine Scheibe ergibt ζ^* eine einfache Kurve, die von einem aus $\hat{p}$ entstandenen Randpunkt ausgeht und zu einem anderen zu $\hat{p}$ bzw. $\hat{p}_1, \ldots, \hat{p}_m$ gehörigen läuft.

Dual zu einer Zerschneidung Z^* gehört ein "geometrisch" ausgezeichnetes Kurvensystem Z, dem wiederum ein Erzeugendensystem der Fundamentalgruppe $\mathfrak{F}$ entspricht; Zweiteilungen liefern dann gewisse Erzeugendenwechsel. Für uns ist es vorteilhaft, diese Situation erst nach dem Ausboren eines "nicht störenden" kleinen Loches um $\hat{p}$ zu beschreiben (vergl. Skizze in [18, S. 30]). Der Aufpunkt p von Z liege auf dem Rand des Loches, und die Kurven $\zeta_i \in Z$ haben außer ihrem Anfangs- und Endpunkt p keinen weiteren Punkt mit dem Loch und den Randkurven der Fläche gemein. Dem Z entspricht dann ein Erzeugendensystem $E = \{E_1, \ldots, E_n\}$ der Fundamentalgruppe $\hat{\mathfrak{G}}$ der zusätzlich gelochten Fläche, wobei E_j das Durchsetzen von ζ_i^* oder das "Entlang-ζ_i-Laufen" zählt. Mit Π sei das Element aus $\hat{\mathfrak{G}}$ zu dem Randweg des neuen Loches bezeichnet. Π_E sei das reduzierte Wort von Π in den Erzeugenden E, und $l^\circ_E(X)$ gebe die Länge der Konjugationsklasse des Elementes $X \in \hat{\mathfrak{G}}$ bezüglich des Erzeugendensystemes E an.

Eine Zweiteilung ergibt dual einen Übergang von Z zu einem neuen Kurvensystem Z' und damit einen Wechsel der Erzeugenden von E zu einem $E' = \{E_1', \ldots, E_n'\}$. Er hat die folgende Form:

(9.4) Die weggelassene Kurve habe den Index i. Es gibt dann eine Permutation

$$\begin{pmatrix} 1 \dots n \\ \alpha_1 \dots \alpha_n \end{pmatrix}$$

mit $\alpha_1 = i$ sowie ein $\varepsilon = \pm 1$ und Zahlen $1 \leqq n_1 \leqq n_2 \leqq n_3 \leqq n$, so daß

$$\begin{array}{llll} \text{(a)} & E'_{\alpha_j} = & E_{\alpha_j}, & 1 \leqq j \leqq n_1, \\ \text{(b)} & E'_{\alpha_j} = & E_{\alpha_j} E_i^{\varepsilon}, & n_1 < j \leqq n_2, \\ \text{(c)} & E'_{\alpha_j} = & E_i^{-\varepsilon} E_{\alpha_j}, & n_2 < j \leqq n_3, \\ \text{(d)} & E'_{\alpha_j} = & E_i^{-\varepsilon} E_{\alpha_j} E_i^{\varepsilon}, & n_3 < j \leqq n. \end{array}$$

(9.5) Eine Erzeugende, die von einer Randkurve herrührt, bleibt entweder erhalten oder wird konjugiert.

(9.6) $$l^{\circ}_{E'}(\Pi) = l^{\circ}_{E}(\Pi).$$

Durch (9.4) sind alle Erzeugendenwechsel beschrieben, die den Whiteheadschen Automorphismen entsprechen. Daß der Erzeugendenwechsel geometrischer Herkunft ist, drücken (9.5) und (9.6) aus; umgekehrt entstehen alle solche Erzeugendenwechsel durch Zweiteilen. — Ein Erzeugendenwechsel heiße *zugelassen* oder *geometrisch verfolgbar* bzw. *definierbar*, wenn er entweder nur eine Permutation der Menge der Erzeugenden und ihrer Inversen bewirkt oder die Eigenschaften (9.4-6) hat. Alle Erzeugendensysteme, die wir aus einem kanonischen Erzeuggendensystem durch zugelassene Erzeugendenwechsel bekommen, mögen *zugelassen* oder auch *geometrisch definierbar* genannt werden.

Es sei eine Potenz E_j^{η}, $\eta = \pm 1$, ausgezeichnet. Bei den zugelassenen Wechseln nehmen wir die Auszeichnung mit: nach einem Wechsel der Form (9.4-6) sei $E_j'^{\eta}$ ausgezeichnet, nach einer Permutation das Bild von E_j^{η}. Die Auszeichnung beginnen wir (etwa) bei einem kanonischen Erzeugendensystem durch die Forderung (5.3). Offenbar erhält dasselbe algebraische Erzeugendensystem verschiedene Auszeichnungen. Jetzt verstehen wir unter E immer das Erzeugendensystem mit Auszeichnung.

Wir sagen, daß eine Homotopieklasse $w \in \mathfrak{F}$ durch ein Erzeugendensystem E von $\hat{\mathfrak{G}}$ und das Wort W in den Erzeugenden aus E *günstig beschrieben* wird, wenn folgendes erfüllt ist:

(9.7) Bei dem Standardhomomorphismus $\hat{\mathfrak{G}} \to \mathfrak{F}$ wird W auf ein zu w konjugiertes Element abgebildet.

(9.8) W ist zyklisch reduziert.

(9.9) W enthält kein Teilwort des zyklisch-reduzierten Wortes $\Pi_E{}^\varepsilon$, $\varepsilon = \pm 1$, welches mehr als die Hälfte von $\Pi_E{}^\varepsilon$ ausmacht. (Dabei sind W und $\Pi_E{}^\varepsilon$ als zyklische Worte gesehen.)

(9.10) Enthält W eine Hälfte von $\Pi_E{}^\varepsilon$, so enthält diese die ε-te Potenz des ausgezeichneten Elementes aus E.

Hierbei entstehen die Forderungen (9.8-10) aus (5.1-3) und (9.7) aus dem davorstehenden Text, wenn man aus der Automorphismensprache in die der Erzeugendenwechsel übersetzt. Wir wollen nun zeigen, daß es sich allein unter Verwendung geometrisch zugelassener Erzeugendenwechsel rechnerisch entscheiden läßt, ob eine Homotopieklasse einfache Kurven enthält oder nicht. Genauer

SATZ 5. *Enthält die Homotopieklasse $w \in \mathfrak{F}$ eine einfach-geschlossene Kurve, wird ihre Konjugationsklasse relativ zu einem geometrisch definierbaren Erzeugendensystem E durch ein Wort W günstig beschrieben und ist dabei die zyklische Länge noch nicht minimal unter den Längen bezüglich geometrisch definierbarer Erzeugendensysteme, so gibt es einen geometrisch verfolgbaren Erzeugendenwechsel, bei dem die Länge verringert wird.*

ANMERKUNG. Für zerlegende Kurven gibt es zwei "minimale" Längen in Abhängigkeit von der Seite, von der man schaut, und wir wollen beide zulassen. Dabei zeichnen wir für die Betrachtung willkürlich eine Seite aus.

Um zu einer Schreibung der gegebenen Homotopieklasse mit minimaler Länge zu gelangen, sind also nur geometrisch zugelassene Prozesse vonnöten. Hat man die minimale Länge erst einmal erreicht, so folgt aus Satz 4, daß weiterhin benötigte Wechsel die Längen weder des Wortes noch der definierenden Relation zu erhöhen brauchen, somit also auch geometrischer Natur sind. Damit haben wir einen algebraischen Algorithmus, welcher nur geometrisch verfolgbare Schritte zu verwenden braucht.

10. Beweis von Satz 5.

Sei $\varkappa$ eine einfach-geschlossene Kurve aus w im Innern von F. Die Zerschneidung Z^* definiere das Erzeugendensystem E. Durch isotope Deformationen können wir $\varkappa$ in eine günstige Lage bezüglich Z^* bringen (Hilfssatz 1). Schneiden wir nun die Fläche an Z^* auf, so erhalten wir eine Scheibe, in deren Rand die Kurven aus Z^* je doppelt auftreten, die Randkurven der Fläche je einmal. Aus $\varkappa$ entsteht ein einfaches System von Bögen, die jeweils von einer Randstrecke zu einer *anderen* laufen, dabei aber nie zu $\varrho_1, \ldots, \varrho_m$ gelangen. Verbinden zwei Bögen dasselbe Paar von Randstrecken miteinander, so befindet sich zwischen ihnen ein Band; jeder Bogen, der in es eintritt, verläuft ganz in ihm und verbindet dieselben beiden Randstrecken. Eine maximale Schar solcher paralleler Bögen bezeichnen wir als *Streifen.*

Sei nun ein Streifen gegeben, der von der Randstrecke ξ^* der Scheibe zu ζ^* führt und y Bögen enthält. Seien x und z die Zahlen der von ξ^* bzw. ζ^* ausgehenden Bögen vor dem Streifen, x' und z' die nach dem Streifen. "Vor" und "nach" bezieht sich auf irgendeinen Umlaufsinn des Scheibenrandes. Wir setzen voraus, daß ξ^* und ζ^* von verschiedenen Kurven der Zerschneidung Z^* herrühren. Je nach der Verteilung der zu ξ^* und ζ^* gehörigen Strecken ξ^*_+ und ζ^*_+ im Rande der Scheibe ergeben sich im Prinzip die beiden Situationen, die durch die ausgezogenen Linien in den Abbildungen A und B beschrieben werden.

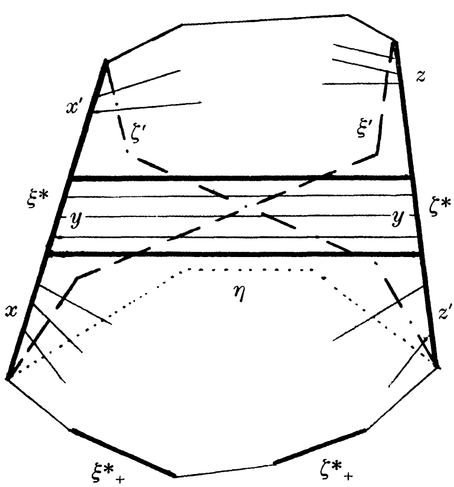

Abbildung A

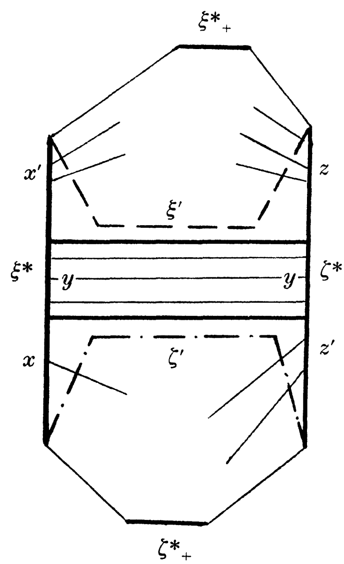

Abbildung B

Durch Zweiteilungen lassen sich ξ^* durch ξ' oder η und ζ^* durch ζ' oder η ersetzen. Nehmen wir einmal an, daß wir bei keiner dieser Zweiteilungen die Schnittpunktzahl verringern können. Dann bekommen wir Ungleichungen:

$$\begin{array}{ll} \text{A} & \text{B} \\ x+z' \geqq x+y+x', & x+z' \geqq z+y+z', \\ x+z' \geqq z+y+z', & z+x' \geqq x+y+x', \\ x+y+z \geqq x+y+x', & \\ x'+y+z' \geqq z+y+z', & \end{array}$$

oder

$$\begin{array}{ll} z' \geqq x'+y, & x > z, \\ x \geqq z+y, & z > x, \\ x' = z, & \end{array}$$

Tritt also Fall B ein, so läßt sich die Schnittzahl durch eine Zweiteilung verringern. Im Falle A ist die Lage komplizierter. Nehmen wir an, daß wir den Streifen der Breite y so haben wählen können, daß $x'+y>x$ und $x+y>x'$ erfüllt ist. Dann müßte auch $z+y>x\geqq z+y$ sein. Gehen wir also von einer Strecke ξ^* aus und finden einen Streifen, der die "Mitte" einschließt, so läßt sich die Zahl der Schnitte wieder durch eine Zweiteilung verringern.

Eine Verkleinerung der Schnittzahl läßt sich also höchstens dann nicht erreichen, wenn kein Streifen die Mitte irgendeiner Strecke echt enthält. (Insbesondere müßte die einfache Kurve zerlegen.) Dann beginnen wir mit einem ξ^*, welches maximal viele Schnittpunkte hat, und wählen den Streifen so, daß $x=x'+y$ ist. (Derer gibt es zwei.) Tritt dann die Situation A ein, so ist wegen

$$z+y+z' \leqq x+y+x' \quad \text{und} \quad z' \geqq x'+y$$

auch

$$z' = x = z+y = x'+y .$$

Der Streifen verbindet also zwei Strecken, die beide die Maximalzahl an Schnittpunkten besitzen.

Wir haben damit: ist ein Verkürzen durch Zweiteilen nicht möglich, so treten entweder nie Situationen A oder B auf, oder es treten "genügend viele" Situationen A auf, bei denen mit einer der Strecken ξ^* oder ζ^* auch die andere maximale Schnittpunktzahl hat und für die dann

$$x'+y = x = z' = z+y$$

ist. Gibt es keine Stellen der Form A oder B, so kann es sich in $\varkappa$ nur um eine Kurve handeln, die nur eine Zerschneidungslinie trifft und diese einmal oder eventuell — im nicht-orientierbaren Fall — zweimal. Jedenfalls hat dann $\varkappa$ die minimale Länge. Im zweiten Fall erhalten wir, von einem Bogen "neben" der Mitte einer Strecke mit maximaler

Schnittpunktzahl ausgehend, eine einfach-geschlossene Kurve, die immer "neben" der Mitte läuft und eine Zerschneidungslinie, wenn überhaupt, so höchstens zweimal trifft. Es muß sich um $\varkappa$ handeln. Trifft $\varkappa$ eine Zerschneidungskurve einmal, so läßt sich die Schnittzahl durch eine Zweiteilung verringern. Aber auch im anderen Fall gibt es eine verkürzende Zweiteilung. Zum Beweis vernachlässige man zunächst die nicht geschnittenen Zerschneidungslinien. Danach kann kein Bogen ein Streckenpaar trennen, wenn es keine verkürzende Zweiteilung gibt. Deshalb haben alle Situationen A das strengere Aussehen C. $\varkappa$ umläuft

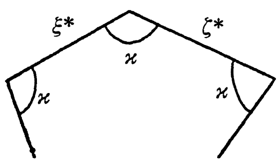

Abbildung C

somit den Stern des Basispunktes einmal. Nach dem Einfügen der unbeachteten Strecken läßt sich eine eventuell gefundene verkürzende Zweiteilung durch eine verkürzende induzieren. Da W günstig beschrieben war und keine Zweiteilungen verkürzen sollten, hat das Wort eine der minimalen Längen. (Hier spielt hinein, daß man je nach der Seite, von der aus man betrachtet, zwei "minimale" Längen bekommt.) Damit ist Satz 5 bewiesen.

11. Bemerkungen über die praktische Durchführung.

Schon das Ergebnis des ersten Teiles liefert einen endlichen algebraischen Algorithmus. Jedoch ist die Zahl der anzusetzenden Erzeugendenwechsel sehr groß und damit das Verfahren aufwendig, ganz zu schweigen davon, daß eventuell Schritte ohne topologischen Sinn getan würden. Durch unseren Satz 5 wird die Zahl der anzuwendenden Erzeugendenwechsel sehr eingeschränkt, es wird im Beweis sogar angegeben, welche Erzeugendenwechsel verkürzen müßten. Die nötigen Erzeugendenwechsel sind leicht zu beschreiben: Man dualisiere zunächst Π_E und beschreibe damit den Rand der Scheibe durch ein Wort. Danach zerlege man dieses Wort oder sein Inverses in zwei Teile P_1, P_2 und ersetze nun die erste in P_1 auftretende Strecke durch eine neue. Ihr Partner muß sich allerdings in P_2 befinden. Dann bekommt man einen Erzeugendenwechsel (9.4), bei dem E_{α_1} von der zu ersetzenden Strecke herrührt. Erzeugende werden nach (a) bzw. (d) abgeändert, wenn das zugehörige Paar ganz in P_1 bzw. P_2 (oder umgekehrt) liegt, und es werden Erzeugende nach (b) bzw. (c) ersetzt, wenn das Paar getrennt liegt. Diese Erzeugendenwechsel lassen sich leicht bis auf einen inneren mit ${E_{\alpha_1}}^{\pm 1}$ als Faktor genau beschreiben. Aus dem Beweis geht hervor, mit welchen Aufteilungen P_1, P_2 man zweckmäßig beginnt, um schnell zu verkürzen oder nachzuweisen, daß die Homotopieklasse einfache Kurven nicht enthalten kann.

Trotzdem scheint mir dieser Algorithmus überwiegend theoretische Bedeutung zu haben. Für eine praktische Nachprüfung scheint mir der folgende Weg günstiger: Zunächst repräsentiere man die Homotopieklasse durch ein Wort in den Erzeugenden mit den Eigenschaften (5.1-3) bzw. (9.8-10). Wie aus dem ersten Teil folgt, bringt es die Existenz überhaupt einer einfachen Kurve in der Homotopieklasse mit sich, daß gerade dieses Wort das Schnittverhalten einer einfachen Kurve der Klasse mit der Zerschneidung beschreiben kann. Dann schneide man an der Zerschneidung auf und versuche, die durch das Wort geforderten Streifen ohne gegenseitige Schnitte auf die Scheibe zu zeichnen. Gelingt es, so ersetze man jeden Streifen durch die nötige Zahl paralleler Bögen und identifiziere ihre Anfangs- und Endpunkte, wie es durch das Verkleben aufgezwungen wird. Jetzt bleiben noch zwei Fragen: Entsteht *eine* Kurve? Und stammt dann diese aus der Homotopieklasse? Läßt sich die Zeichnung durchführen und werden die Fragen bejaht, so enthält die Homotopieklasse offenbar eine einfache Kurve, andernfalls aber nicht, wie im ersten Teil [18] gezeigt wurde.

Einen anderen praktisch durchführbaren Weg gibt D. Chillingworth [17] an, der das Problem auf die Berechnung von Windungszahlen nach B. L. Reinhart zurückführt. Ferner siehe [16]. — Man kann probieren, und das erscheint mir reizvoll, die Frage auf das Geschlecht 2 zu reduzieren. Das dürfte Betrachtungen von Überlagerungen ergeben.

LITERATUR

7. R. C. Lyndon, *On Dehn's algorithm*, Math. Ann. 166 (1966), 208–228.

14. H. Zieschang, *Ein Satz von Nielsen, einige Anwendungen und Verallgemeinerungen*, Trudy IV, Allunionskonferenz über Topologie in Taschkent 1963, S. 184–201, FAN-Verlag der Usbekischen Republik, Taschkent, 1967.

15. H. Zieschang, *Über Automorphismen ebener diskontinuierlicher Gruppen*, Math. Ann. 166 (1966), 148–167.

16. G. Călugăreanu, *Sur les courbes fermées simples tracées sur une surface fermée orientable*, Mathematika (Cluj) 8 (31) (1966), 29–38.
17. D. Chillingworth, *Simple closed curves on surfaces*, Bull. London, Math. Soc. (to appear).
18. H. Zieschang, *Algorithmen für einfache Kurven auf Flächen*, Math. Scand. 17 (1965), 17–40.

RUHR-UNIVERSITÄT, BOCHUM, DEUTSCHLAND