[11] C. Pomerance, *On the distribution of round numbers*, Abstracts A.M.S. 3 (1982), p. 414.
[12] S. Ramanujan, *Collected Papers*, Chelsea Publishing Company, 1962.
[13] L. G. Sathe, *On a problem of Hardy on the distribution of integers having a given number of prime factors*, J. Indian Math. Soc. 17 (1953), p. 63–141; 18 (1954); p. 27–81.
[14] A. Selberg, *Note on a paper by L. G. Sathe*, ibid. 18 (1954), p. 83–87.

DÉPARTEMENT DE MATHÉMATIQUES
UNIVERSITÉ DE LIMOGES
123 Avenue Albert Thomas
F-87060 Limoges Cédex
France

Reçu le 19. 11. 1982
et dans la forme modifiée le 18. 7. 1983 (1327)

ACTA ARITHMETICA
XLIV (1984)

Zur diophantischen Approximation von zwei reellen Zahlen

von

PETER THURNHEER (Zürich)

I. Einleitung, Resultat. Nach einem Satz von Dirichlet existieren zu gegebenen reellen Zahlen α, β unendlich viele, vom Ursprung verschiedene Gitterpunkte $\mathfrak{x} = (a, b)$, für die gilt

$$\|\alpha a + \beta b\| \leqslant \langle \mathfrak{x} \rangle^{-2},$$

wobei $\|\cdot\|$ den Abstand von der nächsten ganzen Zahl bezeichnet und $\langle \mathfrak{x} \rangle := \max(|a|, |b|)$ ist. Im folgenden wird ein analoges Resultat bewiesen, wobei aber die Wahl der Gitterpunkte $\mathfrak{x}$ eingeschränkt wird auf gewisse Teilgebiete von $\boldsymbol{R}^2$.

Sei $\boldsymbol{R}^2 := \{\mathfrak{y} \mid \mathfrak{y} = (\xi_1, \xi_2)\}$. Für nicht negative ϱ, τ setzt man

$$\Phi_0(\varrho, \tau) := \{\mathfrak{y} \mid |\xi_2| < |\xi_1|^{\varrho}\} \cup \{\mathfrak{y} \mid |\xi_1| < |\xi_2|^{\tau}\}.$$

Es bezeichne $\Phi(\varrho, \tau)$ das Bild von $\Phi_0(\varrho, \tau)$ unter einer beliebig vorgegebenen, regulären linearen Transformation. Im weitern seien $c_1, c_2, \ldots$ positive Schranken, die höchstens von α, β und $\Phi(\varrho, \tau)$ abhängen.

SATZ. *Sei $\varrho > 1$, $\tau \geqslant 0$ und $1 < t \leqslant r \leqslant 2$, wobei gelte*

$$(1)\qquad (1-\tau)\{\varrho(t^2 r - tr - t - r - 1) + t^2\} + (1-\varrho)(t^2 - 1) \leqslant 0.$$

Zu gegebenen reellen Zahlen α, β existieren dann unendlich viele Gitterpunkte $\mathfrak{x} = (a, b)$ mit entweder

$$(2)\qquad \mathfrak{x} \in \Phi(\varrho, 0) \quad \textit{und} \quad \|\alpha a + \beta b\| \leqslant c_1 \langle \mathfrak{x} \rangle^{-r},$$

oder

$$(3)\qquad \mathfrak{x} \in \Phi(1, \tau) \quad \textit{und} \quad \|\alpha a + \beta b\| \leqslant c_2 \langle \mathfrak{x} \rangle^{-t}.$$

Die im folgenden Korollar angegebenen Spezialfälle erhält man aus dem Satz, wenn man wählt $\varrho = 7/4$, $\tau = 0$, $r = t = 2$, beziehungsweise $1 < \varrho \leqslant 7/4$, $\tau = (7-4\varrho)/(4-\varrho)$, $r = t = 2$, beziehungsweise $1 < \varrho \leqslant 7/4$, $\tau = 0$, $r = t = s(\varrho)$, wobei $s(\varrho)$ die grösste reelle Nullstelle des Polynoms

$$f_\varrho(x) := \varrho x^3 - 2(\varrho-1)x^2 - 2\varrho x - 1$$

bezeichnet. Es ist $s(7/4) = 2$ und $s(1) = (\sqrt{5}+1)/2$.

KOROLLAR. *Sei $1 < \varrho \leqslant 7/4$. Zu gegebenen reellen Zahlen α, β existieren dann unendlich viele Gitterpunkte $x = (a, b)$ mit*

$$x \in \Phi(7/4, 0) \quad \textit{und} \quad \|\alpha a + \beta b\| \leqslant c_3 \langle x\rangle^{-2},$$
$$x \in \Phi(\varrho, (7-4\varrho)/(4-\varrho)) \quad \textit{und} \quad \|\alpha a + \beta b\| \leqslant c_4 \langle x\rangle^{-2},$$
$$x \in \Phi(\varrho, 0) \quad \textit{und} \quad \|\alpha a + \beta b\| \leqslant c_5 \langle x\rangle^{-s(\varrho)}.$$

Der Satz ist zum Beispiel auch gültig für $\varrho = 3/2$, $\tau = 0$, $r = 2$, $t = (9+\sqrt{305})/14 = 1{,}890\ldots$ Da man hat $2 > s(3/2)$ und $1{,}890\ldots > s(\varrho)$ für $1 \leqslant \varrho < 1{,}47\ldots$, zeigt schon dieser Spezialfall: *Für ein festes Paar α, β ist der im Korollar gegebene Exponent $s(\varrho)$ nicht bestmöglich für alle ϱ, $1 < \varrho < 7/4$.*

Bemerkungen. (i) Die dritte Aussage des Korollars ist auch richtig für $\varrho = 1$, falls $1, \alpha, \beta$ linear unabhängig sind über den rationalen Zahlen $\boldsymbol{Q}$. Dies wurde – für ein beliebiges Winkelfeld mit Scheitel im Ursprung anstelle von $\Phi(\varrho, 0)$ und mit beliebig positivem c_5 – gezeigt von W. M. Schmidt [2]. Als Grundlage für den nachfolgenden Beweis diente die dabei von ihm eingeführte Methode. Der obige Satz ist eine Verschärfung eines in [3] erhaltenen Resultates.

(ii) Professor K. Chandrasekharan möchte ich danken, dass er mir diese Arbeit ermöglichte, Professor W. M. Schmidt für die Anregung und verschiedene hilfreiche Bemerkungen dazu.

II. Beweis des Satzes. Im weiteren bezeichnen Grossbuchstaben in deutscher Schrift Elemente aus dem auf die Koordinaten ξ_1, ξ_2, ξ_3 bezogenen $\boldsymbol{R}^3$, entsprechende Kleinbuchstaben stehen für ihre Projektionen parallel ξ_3 in die (ξ_1, ξ_2)-Ebene. Für $\mathfrak{Y} = (\xi_1, \xi_2, \xi_3)$ sei

$$L(\mathfrak{Y}) := \alpha\xi_1 + \beta\xi_2 + \xi_3.$$

In der (ξ_1, ξ_2)-Ebene sei (u, v) dasjenige Koordinatensystem, in dem sich $\Phi(\varrho, \tau)$ darstellen lässt in der Form

$$\Phi(\varrho, \tau) = \{\mathfrak{y} \mid |v(\mathfrak{y})| < |u(\mathfrak{y})|^\varrho\} \cup \{\mathfrak{y} \mid |u(\mathfrak{y})| < |v(\mathfrak{y})|^\tau\}.$$

Unter einem Gitterpunkt versteht man auch im weiteren ein Element, dessen Koordinaten im ξ-System ganze Zahlen sind. Die im folgenden in den $\ll$-Abschätzungen auftretenden Schranken sind stets von der Art der $c_1, c_2, \ldots$ Es bezeichne $[x]$ den Ganzteil der reellen Zahl x, $\boldsymbol{N}$ die Menge der natürlichen Zahlen und $\mathfrak{O}$ den Ursprung des $\boldsymbol{R}^3$.

Man definiert V, W, X, Y und Z als Funktionen von ϱ, τ, r und t:

$$V := \frac{t}{t-1}, \quad W := \frac{1-\varrho\tau}{\varrho-1}, \quad X := \frac{V}{r(1+W)+W}, \quad Y := \frac{X+V-1}{t}, \quad Z := \frac{\varrho(1-Y)}{\varrho-1}.$$

Ist $\varrho\tau \leqslant 1$, so gilt

$$Z - V - X \leqslant -rZ, \tag{4}$$

und

$$Y > 1/\varrho. \tag{5}$$

Die Ungleichung (4) ist äquivalent zu der nach Voraussetzung erfüllten Ungleichung (1). Man zeigt dies etwa wie folgt:

Unter Verwendung der Definitionen von Z und Y sieht man, dass (4) äquivalent ist zur Ungleichung

$$\varrho(1+r)(t+1) - (X+V)\{\varrho(1+r) + t(\varrho-1)\} \leqslant 0,$$

respektive, indem man noch die Definitionen von X und V beachtet, zur Ungleichung

$$\varrho(1+r)(t^2-1)\{r(1+W)+W\} - t(1+W)(1+r)\{\varrho(1+r)+t(\varrho-1)\} \leqslant 0.$$

Nun klammert man auf der linken Seite den Faktor $1+r$ aus, fasst die Vielfachen von W zusammen, setzt die Definition von W ein und sieht, dass es genügt, die Ungleichung

$$(1-\varrho\tau)\big(\varrho(t^2-1)(r+1) - t\{\varrho(1+r)+t(\varrho-1)\}\big) +$$
$$+ (\varrho-1)\big(r\varrho(t^2-1) - t\{\varrho(1+r)+t(\varrho-1)\}\big) \leqslant 0$$

zu beweisen. Multipliziert man die grossen mit den kleinen runden Klammern aus, so hebt sich der Term $t\{\varrho(1+r)+t(\varrho-1)\}$ weg. Dies erlaubt, einen Faktor ϱ auszuklammern, wobei der andere Faktor gleich der linken Seite von (1) ist, was die Aequivalenz von (4) mit (1) beweist.

Nach der Definition von Y ist $X+V = Yt+1$. Setzt man zudem die Definition von Z in (4) ein, ergibt sich

$$Y \geqslant \frac{r\varrho+1}{t(\varrho-1)+\varrho(1+r)}. \tag{6}$$

Zur Verifikation von (5) genügt es also zu zeigen, dass die rechte Seite in (6) grösser als $1/\varrho$ ist, was äquivalent ist zum Beweis der Ungleichung

$$r\varrho^2 - r\varrho - t(\varrho-1) > 0.$$

Diese ist aber wegen $\varrho > 1$ und $1 < t \leqslant r$ erfüllt, denn die linke Seite ist nicht kleiner als $r(\varrho^2 - 2\varrho + 1) = r(\varrho-1)^2$.

Der Satz wird nun indirekt bewiesen, das heisst, man geht aus von der

GEGENANNAHME.

(7) Sowohl (2) als auch (3) ist höchstens für endlich viele Gitterpunkte $x = (a, b)$ erfüllt.

Daraus folgt:

(8) Die Zahlen $1, \alpha, \beta$ sind linear unabhängig über $\boldsymbol{Q}$.

Wären $1, \alpha, \beta$ linear abhängig über Q, das heisst bestände eine Beziehung $\alpha p+\beta q+m=0$, p, q, m ganz, nicht alle 0, so wäre $\|\alpha pn+\beta qn\|=0$ für alle ganzen n. Da $\varrho>1$ ist, enthält $\Phi(\varrho, 0)$ aber unendlich viele Gitterpunkte der Form $\mathfrak{x}=(pn, qn)$. Dies widerspricht der Gegenannahme. Damit ist (8) gezeigt. Genau wie in [2], Lemma 1, überlegt man sich mit Hilfe des ersten Satzes von Minkowski aus der Geometrie der Zahlen, dass die Gegenannahme folgende Aussage impliziert:

(9) Zu genügend grossem N existiert ein Gitterpunkt $\mathfrak{G} \neq \mathfrak{O}$, mit $|\mathfrak{G}| \ll N$ und $|L(\mathfrak{G})| \leqslant N^{-V}$.

Sei nun $\mathfrak{G}_m$, $m=1, 2, \ldots$ eine Folge von Gitterpunkten mit $v(\mathfrak{g}_m) \geqslant 0$, $|\mathfrak{G}_m| < |\mathfrak{G}_{m+1}|$ und so, dass für jedes $m=1, 2, \ldots$ gilt: Es gibt keinen Gitterpunkt $\mathfrak{G} \neq \mathfrak{O}$ mit $|\mathfrak{G}|<|\mathfrak{G}_m|$ und $|L(\mathfrak{G})| \leqslant |L(\mathfrak{G}_m)|$.

Setzt man $u_m:=u(\mathfrak{g}_m)$, $v_m:=v(\mathfrak{g}_m)$, $L_m:=|L(\mathfrak{G}_m)|$, $D_m:=|u_m v_{m+1}-u_{m+1} v_m|$, $N_m:=|\mathfrak{G}_m|$ und bezeichnen μ_{m1}, μ_{m2} die sukzessiven Minima des von $\mathfrak{g}_m$ und $\mathfrak{g}_{m+1}$ aufgespannten Gitters Λ_m bezüglich der Eichfunktion $f(\mathfrak{y}):=\max(|u(\mathfrak{y})|, |v(\mathfrak{y})|)$, so gilt das folgende

Lemma. *Für genügend grosses m ist*

$$L_m \ll N_{m+1}^{-V}, \tag{10}$$

$$N_m^{\tau} \ll |u_m| \ll N_m^{1/\varrho}, \tag{11}$$

$$N_m \ll v_m \ll N_m, \tag{12}$$

und für unendlich viele Indizes j gilt

$$D_j \gg N_j^V, \quad \mu_{j2} \ll D_j N_j^{-1}. \tag{13}$$

Die Abschätzung (10) folgt aus (9) und der Definition der Folge $\mathfrak{G}_m$. Offensichtlich ist $v_m \ll N_m$.

Man nimmt nun an, es sei $|u_m| \gg N_m^{1/\varrho}$ für unendlich viele m. Mit genügend grossem $c_6 \in N$ gilt dann für den Gitterpunkt $c_6 \mathfrak{G}_m$

$$c_6 \mathfrak{g}_m \in \Phi(\varrho, 0) \quad \text{und} \quad |L(c_6 \mathfrak{G}_m)| \ll N_{m+1}^{-V} \ll N_m^{-r} \ll \langle c_6 \mathfrak{g}_m \rangle^{-r}. \tag{14}$$

Dabei wurde (10) benützt, und dass wegen $t \leqslant 2$ gilt $V \geqslant 2 \geqslant r$. Für genügend grosses m widerspricht (14) der Gegenannahme (7). Damit ist die zweite Ungleichung in (11) bewiesen.

Da $1/\varrho<1$ und offenbar $|\mathfrak{g}_m| \gg N_m$ ist, impliziert $|u_m| \ll N_m^{1/\varrho}$ die Abschätzung $v_m \gg N_m$, womit (12) vollständig gezeigt ist.

Mit Hilfe von (10) und $V \geqslant 2$ überlegt man sich, dass auf Grund der Gegenannahme (7) $\mathfrak{g}_m \notin \Phi(1, \tau)$ ist für genügend grosses m, woraus zusammen mit (12) folgt $N_m^{\tau} \ll |u_m|$. Es bleibt also nur noch (13) zu verifizieren. (Sind ϱ und τ so gewählt, dass $\tau>1/\varrho$ ist, so führt die Gegenannahme gemäss (11) auf einen Widerspruch, was den Satz beweist. Also darf man $\varrho\tau \leqslant 1$ annehmen, so dass (4) gilt).

Aus (10) bis (12), $V \geqslant 2$ und $1/\varrho<1$ folgt $|v_n L(\mathfrak{G}_{n+1})-v_{n+1} L(\mathfrak{G}_n)| \to 0$, $L_{n+1} D_{n-1} \to 0$, und $L_n|u_{n-1} v_{n+1}-u_{n+1} v_{n-1}| \to 0$, für $n \to \infty$. Mit der ersten dieser Aussagen, sowie (8), zeigt man wie in [1], Lemma 3, dass $\mathfrak{G}_{j-1}$, $\mathfrak{G}_j$, $\mathfrak{G}_{j+1}$ linear unabhängig sind für unendlich viele j. Unter Verwendung der andern beiden Aussagen folgert man daraus wie in [2], Lemma 2, dass $D_j \gg N_j^V$ ist. Die zweite Abschätzung unter (13) ergibt sich aus der ersten und dem Satz von Minkowski über die sukzessiven Minima, nach welchen gilt

$$N_j^2 \ll N_j^V \ll D_j \ll \mu_{j1} \mu_{j2} \ll D_j. \tag{15}$$

Da $|\mathfrak{g}_j| \gg N_j$ ist, folgt daraus, dass μ_{j1} von der Grössenordnung N_j sein muss, und die letzte Ungleichung in (15) impliziert $\mu_{j2} \ll N_j^{-1} D_j$. Damit ist das Lemma bewiesen.

Im weiteren bezeichne j einen Index, für den alle Aussagen des Lemmas gelten.

Fall 1: $N_j \ll N_{j+1}^X$. Mit $\delta_j:=c_7[N_j^W]$, $c_7 \in N$, setzt man $\mathfrak{X}_j:=\delta_j \mathfrak{G}_j$, so dass $\mathfrak{X}_j$ ein Gitterpunkt ist. Dann gilt nach (10) bis (12)

$$|L(\mathfrak{X}_j)| \ll N_j^W N_{j+1}^{-V} \ll N_j^{W-V/X} \ll N_j^{-r(W+1)} \ll \langle \mathfrak{x}_j \rangle^{-r}. \tag{16}$$

Zudem zeigen (11) und (12), dass für genügend grosses c_7 der Punkt $\mathfrak{x}_j \in \Phi(\varrho, 0)$ ist. Zusammen mit (16) stellt dies für grosses j einen Widerspruch zu (7) dar. Dies beweist den Satz im Fall 1.

Fall 2.1: $N_j \gg N_{j+1}^X$, $\mu_{j2} \ll N_{j+1}^Y$. Unter diesen Bedingungen kann wie in [2] vorgegangen werden. Jede Kreisscheibe in der (ξ_1, ξ_2)-Ebene vom Radius $2\mu_{j2}$ enthält einen Punkt des Gitters Λ_j. Also existiert ein Gitterpunkt $\mathfrak{x}_j \in \Lambda_j$, für den gilt

$$\mathfrak{x}_j \in \Phi(1, \tau), \quad |u(\mathfrak{x}_j)| \ll \mu_{j2}, \quad |v(\mathfrak{x}_j)| \ll \mu_{j2}. \tag{17}$$

Mit ganzen Zahlen p_j, p_{j+1} lässt sich $\mathfrak{x}_j$ darstellen in der Form $\mathfrak{x}_j=p_j \mathfrak{g}_j+p_{j+1} \mathfrak{g}_{j+1}$, wobei man hat

$$|p_j| \ll N_{j+1} N_j^{-1}, \quad |p_{j+1}| \ll 1. \tag{18}$$

Es ist nämlich $D_j|p_j|=|u(\mathfrak{x}_j) v_{j+1}-u_{j+1} v(\mathfrak{x}_j)|$. Mit (11) bis (13), sowie (17) folgt daraus die Abschätzung für $|p_j|$. Eine analoge Rechnung ergibt diejenige für $|p_{j+1}|$. Für den Gitterpunkt $\mathfrak{X}_j:=p_j \mathfrak{G}_j+p_{j+1} \mathfrak{G}_{j+1}$, dessen Projektion der Punkt $\mathfrak{x}_j$ ist, erhält man aus (18), wenn man (10) und (17) beachtet:

$$|L(\mathfrak{X}_j)| \ll N_{j+1}^{1-V} N_j^{-1} \ll N_{j+1}^{1-V-X} \ll N_{j+1}^{-tY} \ll \langle \mathfrak{x}_j \rangle^{-t}.$$

Da gilt $\mathfrak{x}_j \in \Phi(1, \tau)$, ist dies für grosse j wiederum nicht vereinbar mit (7).

Fall 2.2: $N_j \gg N_{j+1}^X$, $\mu_{j2} \gg N_{j+1}^Y$. Aus (13), der Definition von D_j, (11), sowie (12) folgt

$$N_j N_{j+1}^Y \ll D_j \ll |u_j| N_{j+1}+N_j N_{j+1}^{1/\varrho}.$$

Also ist wegen (5)

$$N_j N_{j+1}^{Y-1} \ll |u_j|. \tag{19}$$

Man wählt $\delta_j := c_8[1+N_{j+1}^Z N_j^{-1}]$, $c_8 \in N$, und setzt $\mathfrak{X}_j := \delta_j \mathfrak{G}_j$. Dann zeigen (19) und (12), dass für genügend grosses c_8 gilt

$$\mathfrak{x}_j \in \Phi(\varrho, 0). \tag{20}$$

Wäre $N_{j+1}^Z < N_j$ für unendlich viele j, so $\delta_j = c_8$ und $c_8 \mathfrak{g}_j \in \Phi(\varrho, 0)$, sowie nach dem Lemma $|L(c_8 \mathfrak{G}_j)| \ll N_{j+1}^{-Y} \ll N_j^{-2} \ll \langle c_8 \mathfrak{g}_j\rangle^{-r}$ unendlich oft. Dies widerspricht (7), so dass man

$$N_{j+1}^Z \geqslant N_j$$

annehmen darf für alle genügend grossen j. Dann ist $\delta_j \ll N_{j+1}^Z N_j^{-1}$ und mit (10) bis (13), sowie (4)

$$\langle \mathfrak{x}_j\rangle \ll N_{j+1}^Z,$$

$$|L(\mathfrak{X}_j)| \ll N_{j+1}^{Z-Y} N_j^{-1} \ll N_{j+1}^{Z-Y-X} \ll N_{j+1}^{-rZ} \ll \langle \mathfrak{x}_j\rangle^{-r}.$$

Für grosse j widerspricht dies, zusammen mit (20), erneut (7). Damit ist der Satz bewiesen.

Literaturangaben

[1] H. Davenport and W. M. Schmidt, *Approximation to real numbers by quadratic irrationals*, Acta Arith. 13 (1967), S. 169–176.

[2] W. M. Schmidt, *Two questions in diophantine approximation*, Monatsh. Math. 82 (1976), S. 237–245.

[3] P. Thurnheer, *Eine Verschärfung des Satzes von Dirichlet über diophantische Approximation*, Comment. Math. Helv. 57 (1982), S. 60–78.

FORSCHUNGSINSTITUT FÜR MATHEMATIK
ETH ZÜRICH, Hg. G.66.2
Rämistr. 101
8092 Zürich

Eingegangen am 4. 2. 1983
und in revidierter Form am 31. 5. 1983 (1340)

ACTA ARITHMETICA
XLIV (1984)

Dedekind type sums and Hecke operators

by

CHIAKI NAGASAKA (Fukuoka, Japan)

1. Introduction. For $\alpha, \gamma \in \boldsymbol{Z}$ with $\gamma > 0$, the Dedekind sum $s(\alpha, \gamma)$ is defined by

$$s(\alpha, \gamma) = \sum_{\mu \pmod{\gamma}} \left(\left(\frac{\mu}{\gamma}\right)\right)\left(\left(\frac{\alpha\mu}{\gamma}\right)\right),$$

where $((x)) = x-[x]-1/2$ is $x \notin \boldsymbol{Z}$ and $((x)) = 0$ if $x \in \boldsymbol{Z}$. We regard $s(\alpha, \gamma)$ as a function on $\boldsymbol{Z} \times \boldsymbol{N}$ and call it the *homogeneous Dedekind sum*. In fact, it is well known [10] that

$$s(\alpha+\gamma, \gamma) = s(\alpha, \gamma), \tag{1}$$

$$s(n\alpha, n\gamma) = s(\alpha, \gamma), \qquad n \in \boldsymbol{N}. \tag{2}$$

For this function Dedekind [3] discovered the following relation: for a prime number p

$$s(p\alpha, \gamma) + \sum_{b=0}^{p-1} s(\alpha+b\gamma, p\gamma) = (p+1)s(\alpha, \gamma). \tag{3}$$

Recently Knopp [5] obtained an extension of (3): for every $n \in \boldsymbol{N}$

$$\sum_{\substack{ad=n\\ d>0}} \sum_{b=0}^{d-1} s(a\alpha+b\gamma, d\gamma) = \sigma(n)s(\alpha, \gamma), \tag{4}$$

where $\sigma(n)$ means the sum of all positive divisors of n. However, before Knopp, Subrahmanyam [11] had given another extension of (3): for every $n \in N$

$$\sum_{b=0}^{n-1} s(\alpha+b\gamma, n\gamma) = \sum_{d|n} \mu(d)\sigma\left(\frac{n}{d}\right)s(d\alpha, \gamma), \tag{5}$$

where $\mu(n)$ means Möbius' function. We shall see later that (4) is equivalent to (5). Goldberg [4] has already derived (4) from (5).

Following Dedekind we define

$$D\left(\frac{\alpha}{\gamma}\right) = s(\alpha, \gamma).$$