Über die analytischen Eigenschaften von Gruppen linearer Transformationen und ihrer Darstellungen.

Von

J. v. Neumann in Berlin.

Einleitung.

§ 1.

Der Gegenstand dieser Arbeit ist, die Rolle der Stetigkeits-, Differentiierbarkeits- und Analytizitätsannahmen klarzustellen, die in der Theorie der (kontinuierlichen) Gruppen linearer Transformationen und ihrer Darstellungen eine entscheidende Rolle spielen. Es wird sich im Laufe unserer Ausführungen zeigen, daß die Verhältnisse, die bei den Gruppen linearer Transformationen bestehen, in den meisten Beziehungen recht günstige sind.

Es wird uns nämlich gelingen, eine jede solche Gruppe (ohne Einschränkungen) auf diejenige Form zu bringen, die in der Theorie der kontinuierlichen Gruppen allein berücksichtigt wird; eine Gruppe linearer Transformationen hat stets einen Normalteiler, der durch eine endliche Zahl von Parametern analytisch und im kleinen ein-eindeutig darstellbar ist, und dessen Faktorgruppe diskret ist. (Diese Begriffe müssen und werden wir natürlich noch näher präzisieren.)

Ferner können wir jeder linearen Gruppe eine „Infinitesimalgruppe" auf einwandfreiem Wege zuordnen, und die Art und Weise, wie sich Gruppe und Infinitesimalgruppe gegenseitig bedingen, erweist sich als höchst einfach.

Demgegenüber zeigt sich, daß nicht jeder „Infinitesimalgruppe" im üblichen Sinne des Wortes eine Gruppe zuordnen läßt; dieser letztere Umstand soll im Anhang behandelt werden.

Für die Darstellungen gewinnen wir das folgende Resultat: alle Darstellungen einer linearen Gruppe, deren Schwankung in der Umgebung der

Einheit < 1 ist (bei geeigneter Definition der „Entfernung" im Raume der linearen Transformationen), sind durchweg stetig, beliebig oft differentiierbar, ja sie können sogar in der Umgebung eines jeden Elementes der Gruppe durch Potenzreihen dargestellt werden. (In diesen Potenzreihen treten aber — bei komplexen Gruppen — die Real- und Imaginärteile der Transformationskoeffizienten gesondert auf, so daß hieraus keine Analytizität im gewöhnlichen Sinne folgt. Dieselbe ist ja im allgemeinen auch nicht vorhanden[1]).)

Ferner existiert zu jeder solchen Darstellung die sogenannte „Infinitesimaldarstellung" mit den gewohnten Eigenschaften.

Die Prämisse: Schwankung in der Umgebung der Einheit < 1 läßt sich aber nicht mehr wesentlich verschärfen: denn es gibt, wie wir zeigen werden, unstetige Darstellungen, bei denen diese Schwankung $= 2$ ist. (Diese Darstellungen sind übrigens überall unstetig.) Es kann sich also höchstens noch um Konstantenverbesserungen handeln.

§ 2.

Der innere Grund dieser, an den „pathologischen" Möglichkeiten der reellen Funktionentheorie gemessen, recht günstigen Resultate ist jedenfalls die folgenschwere Gruppeneigenschaft: eine Gruppe enthält mit zwei Transformationen stets auch ihr Produkt, und für die Darstellungen ist es analog die Funktionalgleichung, der sie zu genügen haben: wenn wir eine Darstellung D als Funktion der linearen Transformationen A auffassen, so gilt ja:

$$D(A \cdot B) = D(A) \cdot D(B).$$

Daß solche Funktionalgleichungen insbesondere die in § 1 skizzierte Wirkung haben, außer vollkommen regulären Lösungen nur überall unstetige zuzulassen, wollen wir uns an einem möglichst einfachen Beispiel klarmachen.

Wir betrachten die Funktionalgleichung

$$f(xy) = f(x) f(y)$$

(x, y positive Zahlen). Sie definiert ja alle 1-dimensionalen Darstellungen der 1-dimensionalen Transformationsgruppe, die sämtliche Dilatationen der Geraden um eine positive Zahl umfaßt.

Durch die Transformation

$$\varphi(\xi) = \ln(f(e^{\xi}))$$

---

[1]) Z. B. hat die Gruppe aller (komplexen) ebenen Lineartransformationen mit nichtverschwindender Determinante eine nichtanalytische Darstellung durch den Absolutwert der Determinante.

geht sie in die „lineare Funktionalgleichung“

$$\varphi(\xi+\eta)=\varphi(\xi)+\varphi(\eta)$$

über. Und von dieser hat bekanntlich Hamel gezeigt[2]), daß sie außer ganz regulären Lösungen ($\varphi(\xi)=a\xi$) nur überall unstetige zuläßt, und daß beide Fälle vorkommen.

Übrigens spielt die Exponentialfunktion und der Logarithmus in unseren Entwicklungen eine ebenso fundamentale Rolle, wie in der oben skizzierten Diskussion der Funktionalgleichung

$$f(xy)=f(x)f(y).$$

Wir werden diese Funktionen, in ihrer auf Matrizen erweiterten Form, anzuwenden haben; insbesondere das Verhältnis Gruppe — Infinitesimalgruppe, sowie Darstellung — Infinitesimaldarstellung ist völlig von ihnen beherrscht.

§ 3.

Durch unsere Resultate werden viele Einwände behoben, die vom strengen Standpunkte der reellen Funktionentheorie gegen die auf der Betrachtung der Infinitesimalgruppe fußenden Untersuchungen über Gruppen linearer Transformationen und ihrer Darstellungen erhoben werden könnten. Allerdings gebietet der Umstand, daß nicht jede „Infinitesimalgruppe“ aus einer Gruppe hergeleitet werden kann (vgl. Anhang), eine gewisse Vorsicht bei derartigen Betrachtungen.

In erster Linie sind damit die Resultate von Weyl über die Darstellungen halb-einfacher Gruppen[3]) von ihren weitgehenden Differentiierbarkeitsannahmen befreit. Der diesbezügliche Teil unserer Resultate ist übrigens bereits in einer anderen Arbeit dargelegt worden[4]).

I. Die Funktionen exp und ln.

§ 1.

Die linearen Transformationen in m Dimensionen ($m=1, 2, \ldots,$ diese Zahl sei zunächst fest) sind charakterisiert durch Angabe der Transformationsmatrix, d. h. einer m-dimensionalen quadratischen Matrix (mit eventuell komplexen Elementen). Diese ihrerseits wird durch $2m^2$ reelle Zahlen festgelegt: die Real- und Imaginärteile ihrer m^2 Elemente. Wir können diese als kartesische Koordinaten eines Punktes im $2m^2$-dimen-

[2]) Hamel, Math. Annalen **60**, S. 459.

[3]) Weyl, Math. Zeitschr. **23** (1925), S. 271; **24**, S. 328, 377.

[4]) Neumann, Sitzungsber. d. Preuß. Akad., Sitzung vom 17. März 1927.

sionalen (reellen) euklidischen Raume $\mathfrak{R}_m$ betrachten: dadurch sind die linearen Transformationen bzw. Matrizen ein-eindeutig auf $\mathfrak{R}_m$ bezogen. Wir werden beide Terminologien im folgenden benützen und die Transformationen sowohl als Matrizen als auch als Punkte des $\mathfrak{R}_m$ bezeichnen.

Wenn A, B zwei Matrizen sind, so ist der Sinn von

$$A \pm B, \quad AB, \quad \alpha A \qquad (\alpha \text{ eine komplexe Zahl})$$

der übliche, O und E sind die Null- bzw. Einheitsmatrix. Die Determinante von A bezeichnen wir mit $\det A$.

§ 2.

Unter $|A|$ verstehen wir die (gewöhnliche euklidische) Entfernung des A von O, also wenn A die Elemente $a_{\mu\nu}$ $(\mu, \nu = 1, 2, \ldots, m)$ hat,

$$|A| = \sqrt{\sum_{\mu,\nu=1}^{m} |a_{\mu\nu}|^2}.\ ^{5)}$$

$|A|$ heiße der Absolutwert von A.

Die Relationen

$$|O| = 0, \quad |E| = \sqrt{m}, \quad |\alpha A| = |\alpha|\,|A| \qquad (\alpha \text{ eine komplexe Zahl})$$

sind evident.

$$|A \pm B| \leqq |A| + |B|, \quad |A \pm B| \geqq |A| - |B|$$

ist nichts anderes, als der „Dreiecksatz" im euklidischen Raume $\mathfrak{R}_m$. Wir wollen noch zeigen, daß

$$|AB| \leqq |A|\,|B|$$

ist[6]). Dies ergibt sich so (die Elemente von A, B sind $a_{\mu\nu}$ bzw. $b_{\mu\nu}$):

$$\begin{aligned}
|AB|^2 &= \sum_{\mu,\nu=1}^{m} \Big| \sum_{\pi=1}^{m} a_{\mu\varrho} b_{\varrho\nu} \Big|^2 \leqq \sum_{\mu,\nu=1}^{m} \Big(\sum_{\pi=1}^{m} |a_{\mu\varrho}|\,|b_{\varrho\nu}| \Big)^2 \\
&= \sum_{\mu,\nu,\pi=1}^{m} |a_{\mu\varrho}|^2 |b_{\varrho\nu}|^2 + 2 \sum_{\mu,\nu,\varrho,\sigma=1}^{m} |a_{\mu\varrho}|\,|b_{\sigma\nu}|\,|a_{\mu\sigma}|\,|b_{\sigma\nu}| \\
&= \sum_{\mu,\nu,\pi=1}^{m} |a_{\mu\varrho}|^2 |b_{\varrho\nu}|^2 + \sum_{\substack{\mu,\nu,\varrho,\sigma=1 \\ \varrho<\sigma}}^{m} \left(|a_{\mu\varrho}|^2 |b_{\sigma\nu}|^2 + |a_{\mu\sigma}|^2 |b_{\varrho\nu}|^2 \right) \\
&= \sum_{\mu,\nu,\varrho,\sigma=1}^{m} |a_{\mu\varrho}|^2 |b_{\sigma\nu}|^2 = \sum_{\mu,\varrho=1}^{m} |a_{\mu\varrho}|^2 \cdot \sum_{\nu,\sigma=1}^{m} |b_{\nu\sigma}|^2 = |A|^2 |B|^2,
\end{aligned}$$

also $|AB| \leqq |A|\,|B|$, wie behauptet wurde.

(Man beachte übrigens, daß nicht $|E| = 1$ ist!)

---

[5]) Dieser Begriff wurde zuerst von Frobenius benutzt.

[6]) Der erste Beweis dieser Relation findet sich wohl bei Wedderburn.

Durch die Einführung des Absolutwertes haben wir der Menge aller linearer Transformationen (in m Dimensionen) die Metrik und Topologie des euklidischen Raumes auferlegt; es ist nun ohne weiteres klar, was unter abgeschlossenen, offenen, beschränkten, zusammenhängenden Mengen, sowie stetigen, differentiierbaren, analytischen Funktionen zu verstehen ist[7]. (In der Theorie der kontinuierlichen Gruppen spielt der Begriff der „abgeschlossenen Gruppe“ eine große Rolle, er wäre in unserer Terminologie wohl besser mit „beschränkt abgeschlossen“ oder „kompakt“ wiederzugeben.)

§ 3.

Das Konvergieren unendlicher Folgen oder Reihen (Summen) von Matrizen gegen eine bestimmte Matrix im Sinne unseres Absolutwertes (d. h. das gegen 0-Streben des Absolutwertes des Unterschiedes) ist offenbar dasselbe, wie das elementweise Konvergieren derselben. Also dürfen wir alle bekannten Konvergenzkriterien auf die Matrizen anwenden.

Folglich sind die Potenzreihen

$$\sum_{\nu=0}^{\infty} \frac{1}{\nu!} A^{\nu} \quad \text{und} \quad \sum_{\nu=1}^{\infty} \frac{(-1)^{\nu-1}}{\nu} (A-E)^{\nu}$$

für alle A bzw. für alle A mit $|A-E|<1$ konvergiert. (Die letztere Grenze ist genau; denn für die Matrix

$$H = \begin{vmatrix} 1, & 0, & \ldots, & 0 \\ 0, & 0, & \ldots. & 0 \\ \cdot & \cdot & \ddots & \vdots \\ 0, & 0, & \ldots, & 0 \end{vmatrix}$$

ist offenbar $|H|=1$ und $H^2=H$, so daß die zweite Reihe für $A=E-H$ divergiert[8]). Wir bezeichnen ihre Summen mit $\exp A$ bzw. $\ln A$. (Die Bezeichnung $\exp A$ benützen wir an Stelle der schwerfälligeren e^A.)

Die Konvergenz unserer Potenzreihen ist offenbar für $|A| \leqq a$ $(a>0)$ bzw. $|A-E| \leqq a$ $(0<a<1)$ gleichmäßig, und die einzelnen Summanden sind überall stetig, also sind beide Funktionen in ihrem ganzen Definitionsbereiche überall stetig. ($\exp A$ ist überall definiert, $\ln A$ nur für

---

[7]) Als zusammenhängend bezeichnen wir insbesondere eine Menge, wenn bei jeder Zerlegung derselben in zwei nicht leere Teile, mindestens einer dieser Teile einen Häufungspunkt des anderen enthält (Hausdorff). Die Komponente eines Punktes einer Menge ist die größte zusammenhängende Teilmenge derselben, die diesen Punkt enthält.

[8]) $E-H$ ist eine Singularität von $\ln A$, denn für $A=E-\vartheta H$, $0<\vartheta<1$, ist $\ln A = \ln(1-\vartheta)\cdot H$, so daß es für $\vartheta \to 1$ unendlich wird.

$|A-E|<1.$[9]) Wir wollen hieran gleich zwei genauere Abschätzungen der Größe und der Schwankung der Funktionen exp und ln anschließen.

Erstens ist

$$|\exp A - E| \leqq e^{|A|}-1, \qquad |\ln A| \leqq \ln \frac{1}{1-|A-E|} \qquad ((A-E)<1).$$

Beide Abschätzungen folgen unmittelbar aus der Potenzreihendarstellung.

Zweitens ist für $|A|\leqq a$, $|B|\leqq a$ $(a>0)$ bzw. $|A-E|\leqq a$, $|B-E|\leqq a$ $(0<a<1)$

$$|\exp A - \exp B| \leqq e^{a}|A-B|, \qquad |\ln A - \ln B| = \frac{1}{1-a}|A-B|.$$

Um dies zu beweisen, müssen wir zuerst die Richtigkeit der folgenden Hilfsformel einsehen: aus $|A|\leqq a$, $|B|\leqq a$ $(a>0)$ folgt

$$|A^p - B^p| \leqq p\, a^{p-1}|A-B| \qquad (p=0,1,2,\ldots).$$

Für $p=0,1$ ist sie evident. Und von p ist sie folgendermaßen auf $p+1$ zu übertragen (ohne Benützung der Vertauschbarkeit!):

$$\begin{aligned} |A^{p+1}-B^{p+1}| &= |(A^p-B^p)A + B^p(A-B)| \\ &\leqq |A^p-B^p|\,|A| + |B|^p\,|A-B| \\ &\leqq p\cdot a^{p-1}|A-B|\cdot a + a^p\cdot|A-B| = (p+1)a^p|A-B|. \end{aligned}$$

Die behaupteten Formeln folgen aber, bei Berücksichtigung dieser Hilfsformel, sofort aus der Potenzreihendarstellung von exp und ln.

§ 4.

Es soll nun gezeigt werden, daß eine jede der Funktionen exp und ln die andere umkehrt, d. h. daß die Relationen

$$\exp \ln A = A, \qquad \ln \exp A = A$$

gelten. Natürlich muß dabei ln sinnvoll sein, also muß im ersten Falle $|A-E|<1$ sein, und im zweiten Falle $|\exp A - E|<1$. Dies letztere ist für $|A|<\ln 2$ sicher der Fall:

---

[9]) Natürlich wäre es möglich $\ln A$ analytisch fortzusetzen, indessen ist die allgemeine Theorie dieser Funktion ziemlich verwickelt. (Sie ist z. B. kontinuumfach vieldeutig). Erwähnenswert ist die folgende Reihenentwicklung, die über $|A-E|<1$ hinaus konvergiert:

$$\ln A = \sum_{\nu=1}^{\infty} \frac{2}{2\nu-1}\left(\frac{A-E}{A+E}\right)^{2\nu-1}$$

Dabei ist unter $\frac{A-E}{A+E}$ der gemeinsame Wert von $(A-E)\cdot(A+E)^{-1}$ und $(A+E)^{-1}\cdot(A-E)$ zu verstehen. Die Reihe konvergiert für

$$\left|\frac{A-E}{A+E}\right|<1.$$

$$|\exp A - E| \leqq e^{|A|} - 1 < e^{\ln 2} - 1 = 1.$$

Wir werden also die obigen Relationen unter der Voraussetzung $|A - E| < 1$ bzw. $|A| < \ln 2$ beweisen.

Das Verschwinden der Ausdrücke

$$\exp \ln A - A, \qquad \ln \exp A - A$$

soll bewiesen werden. Da die Potenzreihen dieser Funktionen konvergieren (nach den über A gemachten Annahmen), sind diese Ausdrücke die Limites von

$$\sum_{\mu=0}^{r} \frac{1}{\mu!} \left(\sum_{\nu=1}^{s} \frac{(-1)^{\nu-1}}{\nu} (A - E)^{\nu} \right)^{\mu} - A$$

bzw.

$$\sum_{\mu=1}^{r} \frac{(-1)^{\mu-1}}{\mu} \left(\sum_{\nu=1}^{s} \frac{1}{\nu!} A^{\nu} \right)^{\mu} - A,$$

wenn man zuerst s und dann r gegen ∞ streben läßt. Es genügt also jedenfalls zu zeigen, daß diese letzteren Ausdrücke für hinreichend große s, r beliebig klein werden: d. h. daß zu jedem $\varepsilon > 0$ ein $t = t(\varepsilon)$ existiert, so daß aus $s \geqq t$, $r \geqq t$ folgt, daß diese Ausdrücke absolut $< \varepsilon$ sind.

Wir können mit diesen Ausdrücken, da die in ihnen allein vorkommenden Potenzen von A und $A - E$ untereinander vertauschbar sind, ebenso rechnen, wie mit Polynomen gewöhnlicher Zahlen. Es sind offenbar Polynome rs-ten Grades in $A - E$ bzw. A.

Wenn $r \geqq t$, $s \geqq t$ ist, so sind die Koeffizienten der Potenzen bis zur t-ten (einschließlich) dieselben, wie in der Entwicklung von

$$e^{\ln(1+x)} - 1 - x \quad \text{bzw.} \quad \ln e^{x} - x$$

(x als Zahl gedacht!), d. h. 0. Und alle Koeffizienten sind absolut $\leqq$ als die entsprechenden Koeffizienten der Potenzreihen von

$$\sum_{\mu=0}^{\infty} \frac{1}{\mu!} \left(\sum_{\nu=1}^{\infty} \frac{1}{\nu} x^{\nu} \right)^{\mu} + 1 + x = e^{\ln \frac{1}{1-x}} + 1 + x = 1 + x + \frac{1}{1-x}$$

bzw.

$$\sum_{\mu=1}^{\infty} \frac{1}{\mu} \left(\sum_{\nu=1}^{\infty} \frac{1}{\nu!} x^{x} \right)^{\mu} + x = \ln \frac{1}{1-(e^{x}-1)} + x = x + \ln \frac{1}{2-e^{x}}.$$

Da diese Funktionen für $|x| < 1$ bzw. $|x| < \ln 2$ regulär sind, sind ihre Potenzreihen für $|x| < 1$ bzw. $|x| < \ln 2$ konvergent; d. h. für $0 < a < 1$ bzw. $0 < a < \ln 2$ ist der Koeffizient von x^{p} in einer jeden von ihnen

$$\leqq \frac{c_a}{a^{p}} \quad (c_a \text{ hängt nur von } a \text{ ab}).$$

(Die Koeffizienten der ersten sind übrigens fast alle $= 1$.) Um so mehr gilt dies für die Koeffizienten der uns interessierenden Ausdrücke.

Für $|A - E| = a' < a < 1$ bzw. $|A| = a' < a < \ln 2$ ist also für einen jeden der uns interessierenden Ausdrücke der Absolutwert

$$\leqq \sum_{p=t+1}^{rs} \frac{c_a}{a^p} \cdot a'^p = c_a \sum_{p=t+1}^{rs} \left(\frac{a'}{a}\right)^p \leqq \frac{a\, c_a}{a - a'} \left(\frac{a'}{a}\right)^{t+1},$$

sobald $r, s \geqq t$ ist. Für hinreichend große t ist dies beliebig klein und damit ist unsere Behauptung bewiesen; sie gilt offenbar stets, wenn ein solches a existiert, d. h. für alle A mit $|A - E| < 1$ bzw. $|A| < \ln 2$.

§ 5.

Die Relation $\exp(A + B) = \exp A \cdot \exp B$ kann keinesfalls für alle A, B gültig sein; denn dann wäre ja allgemein

$$\exp A \cdot \exp B = \exp B \cdot \exp A,$$

also insbesondere für $|A - E| < 1$, $|B - E| < 1$

$$\exp \ln A \cdot \exp \ln B = \exp \ln B \cdot \exp \ln A,$$
$$AB = BA,$$

was offenbar falsch ist. Wir werden aber zeigen, daß sie für alle vertauschbaren A, B gilt.

Der Ausdruck $\exp(A + B) - \exp A \cdot \exp B$ ist Limes der Ausdrücke

$$\sum_{\mu=0}^{r} \frac{1}{\mu!}(A + B)^\mu - \sum_{\mu=0}^{r} \frac{1}{\mu!} A^\mu \cdot \sum_{\mu=0}^{r} \frac{1}{\mu!} B^\mu$$

für $r \to \infty$. Es genügt also zu zeigen, daß diese gegen 0 streben. Da wir mit ihnen ebenso rechnen dürfen, wie mit Polynomen gewöhnlicher Zahlen (weil A, B vertauschbar sind), ergibt sich nach leichten Zwischenrechnungen:

$$\sum_{\mu=0}^{r} \frac{1}{\mu!}(A+B)^\mu - \sum_{\mu=0}^{r} \frac{1}{\mu!} A^\mu \cdot \sum_{\mu=0}^{r} \frac{1}{\mu!} B^\mu = \sum_{\substack{\mu,\nu=0 \\ \mu+\nu>r}}^{r} \frac{1}{\mu!\,\nu!} A^\mu B^\nu,$$

und folglich ist es dem Absolutwerte nach

$$\leqq \sum_{\substack{\mu,\nu=0 \\ \mu+\nu>r}}^{r} \frac{1}{\mu!\,\nu!} |A|^\mu |B|^\nu \leqq \sum_{\mu=0}^{\infty} \sum_{\nu=\frac{1}{2}r}^{\infty} \frac{1}{\mu!\,\nu!} |A|^\mu |B|^\nu + \sum_{\mu=\frac{1}{2}r}^{\infty} \sum_{\nu=0}^{\infty} \frac{1}{\mu!\,\nu!} |A|^\mu |B|^\nu$$

$$= \sum_{\mu=0}^{\infty} \frac{1}{\mu!} |A|^\mu \sum_{\nu=\frac{1}{2}r}^{\infty} |B|^\nu + \sum_{\mu=\frac{1}{2}r}^{\infty} \frac{1}{\mu!} |A|^\mu \sum_{\nu=0}^{\infty} \frac{1}{\nu!} |B|^\nu$$

$$= e^{|A|} \cdot \sum_{\nu=\frac{1}{2}r}^{\infty} \frac{1}{\nu!} |B|^\nu + e^{|B|} \cdot \sum_{\nu=\frac{1}{2}r}^{\infty} \frac{1}{\nu!} |A|^\nu.$$

Diese Schranke strebt aber für $r \to \infty$ offenbar gegen 0, und damit ist unsere Behauptung bewiesen.

Wir zeigen weiter: Für vertauschbare A, B ist

$$\ln(AB) = \ln A + \ln B,$$

wenn alle drei ln sinnvoll sind, d. h. wenn $|A - E| < 1$, $|B - E| < 1$, $|AB - E| < 1$ ist.

Zunächst werde festgestellt, daß es genügt, die Behauptung für alle vertauschbaren A, B mit $|A - E| < \delta$, $|B - E| < \delta$ (δ irgendeine positive Konstante) zu beweisen.

Es sei nämlich $|A - E| < 1$, $|B - E| < 1$, $|AB - E| < 1$. Dann gilt für alle ϑ mit $0 \leqq \vartheta \leqq 1$:

$$|\{E + \vartheta(A - E)\} - E| = |\vartheta(A - E)| \leqq |A - E| < 1,$$
$$|\{E + \vartheta(B - E)\} - E| = |\vartheta(B - E)| \leqq |B - E| < 1,$$

und weiter:

$$\begin{aligned} &|\{E + \vartheta(A - E)\}\{E + \vartheta(B - E)\} - E| \\ &\quad = |\vartheta(A - E) + \vartheta(B - E) + \vartheta^2(A - E)(B - E| \\ &\quad = |\vartheta(AB - E) - (\vartheta - \vartheta^2)(A - E)(B - E)| \\ &\quad \leqq \vartheta|AB - E| + (\vartheta - \vartheta^2)|A - E|\,|B - E|; \end{aligned}$$

dieser Ausdruck ist aber für $\vartheta \neq 0$

$$< \vartheta \cdot 1 + (\vartheta - \vartheta^2)\cdot 1 \cdot 1 = 2\vartheta - \vartheta^2 = 1 - (1 - \vartheta)^2 \leqq 1,$$

und für $\vartheta = 0$

$$= 0 < 1,$$

also jedenfalls < 1. Folglich ist

$$\begin{aligned} F(\vartheta) = {} & \ln\{E + \vartheta(A - E)\}\{E + \vartheta(B - E)\} \\ & - \ln\{E + \vartheta(A - E)\} - \ln\{E + \vartheta(B - E)\} \end{aligned}$$

für alle $0 \leqq \vartheta \leqq 1$ sinnvoll. Wir müssen zeigen, daß

$$\ln(AB) - \ln A - \ln B = F(1) = 0$$

ist. Nun konvergieren in $F(\vartheta)$ alle ln-Potenzreihen, also sind die Real- und Imaginärteile der Elemente von $F(\vartheta)$ für $0 \leqq \vartheta \leqq 1$ konvergente Potenzreihen von ϑ, d. h. analytische Funktionen desselben.

Und für $0 \leqq \vartheta \leqq \delta$ ist

$$|\{E + \vartheta(A - E)\} - E| = |\vartheta(A - E)| \leqq \delta(A - E) < \delta,$$
$$|\{E + \vartheta(B - E)\} - E| = |\vartheta(B - E)| \leqq \delta(B - E) < \delta,$$

und $E + \vartheta(A - E)$, $E + \vartheta(B - E)$ sind vertauschbar, weil A, B es sind. Nach Annahme gilt also $F(\vartheta) = 0$. Folglich verschwindet $F(\vartheta)$ identisch, und es ist $F(1) = 0$, d. h. die ursprüngliche Behauptung ist bewiesen.

Nun soll die als hinreichend erkannte Relation für $\delta = 1 - \sqrt{\frac{1}{2}}$ bewiesen werden.

$\ln A$, $\ln B$ sind vertauschbar, denn sie sind Limites von Polynomen in A bzw. B, die wegen der Vertauschbarkeit von A und B vertauschbar sein müssen. Folglich ist

$$\exp(\ln A + \ln B) = \exp \ln A \cdot \exp \ln B = AB\,.$$

Ferner ist

$$|A - E|,\ |B - E| < 1 - \sqrt{\frac{1}{2}},$$

$$|\ln A|,\ |\ln B| < \ln \frac{1}{1 - \left(1 - \sqrt{\frac{1}{2}}\right)} = \ln \sqrt{2} = \frac{1}{2} \ln 2\,,$$

$$|\ln A + \ln B| < \ln 2\,.$$

Also gilt

$$\ln \exp(\ln A + \ln B) = \ln A + \ln B\,,$$

d. h.

$$\ln(AB) = \ln A + \ln B\,.$$

§ 6.

Wir gehen nun daran, aus den zwei in § 5 bewiesenen Formeln die naheliegendsten Konsequenzen zu ziehen.

Da αA, βA stets vertauschbar sind (α, β komplexe Zahlen), ist

$$\exp((\alpha + \beta) A) = \exp \alpha A \cdot \exp \beta A\,.$$

Insbesondere gilt allgemein:

$$\exp A \cdot \exp(-A) = \exp O = E\,,$$
$$\det \exp A \cdot \det \exp(t A) = \det E = 1\,,$$

also

$$\det \exp A \neq 0\,, \qquad \exp(-A) = (\exp A)^{-1}\,.$$

(Die Ungleichheit $\det \exp A \neq 0$ wird in Evidenz gesetzt durch die Relation

$$\det \exp A = e^{\operatorname{Spur} A}\,.$$

Ferner ist die Gleichung $\exp A = B$ für die und nur die B lösbar, für welche $\det B \neq 0$ ist. Auf den Beweis dieser Tatsachen gehen wir hier nicht ein.)

Aus diesen Gleichungen folgt für alle A und alle $p = 0, \pm 1, \pm 2, \ldots$ sofort:

$$\exp(pA) = (\exp A)^p\,.$$

Eine Umkehrung dieses Satzes ist der folgende: Wenn $p = 0, \pm 1, \pm 2, \ldots$ ist, und

$$|A^r - E| < 1$$

ist, für alle r zwischen 1 und p (d. h. $1 \leqq r \leqq p$ bzw. $1 \geqq r \geqq p$), so gilt

$$\ln(A^p) = p \ln A.$$

Dies folgt ohne weiteres aus der zweiten Formel in § 5, wenn man berücksichtigt, daß alle Potenzen von A untereinander vertauschbar sind.

Wir stellen noch einige leicht beweisbare Formeln auf. Es ist

$$\exp(S^{-1}AS) = S^{-1}\exp(A)\cdot S, \quad \ln(S^{-1}AS) = S^{-1}\ln(A)S$$

(das erstere für $\det S \neq 0$, das letztere für $\det S \neq 0$, $|A - E| < 1$, $|S^{-1}AS - E| < 1$). Denn die entsprechenden Relationen gelten für die Partialsummen der Potenzreihen (die ja Polynome sind), und folglich auch für die Limites derselben.

Wenn A eine Diagonalmatrix mit den Diagonalelementen $\alpha_1, \alpha_2, \ldots, \alpha_m$ ist, so sind $\exp A$ und $\ln A$ auch Diagonalmatrizen, und zwar mit den Diagonalelementen

$$e^{\alpha_1}, e^{\alpha_2}, \ldots, e^{\alpha_m} \quad \text{bzw.} \quad \ln\alpha_1, \ln\alpha_2, \ldots, \ln\alpha_m.$$

(Für ln muß natürlich $|A - E| < 1$ sein; von den $\ln\alpha_\mu$ ist jeweils der Ast zu nehmen, der bei 1 gleich 0 ist.) Diese Behauptung ist trivial.

§ 7.

Schließlich beweisen wir noch einige auf das Verhalten von exp und ln in der Umgebung von O bzw. E bezüglichen Abschätzungen.

In der Umgebung von O bzw. E gilt

$$\exp A = E + A + O(|A^2|), \qquad \ln A = (A - E) + O(|A - E|^2).$$

Es sei nämlich etwa $|A| < 1$ bzw. $|A - E| < \frac{1}{2}$, dann gilt:

$$|\exp A - A - E| = \Big|\sum_{\nu=2}^{\infty} \frac{1}{\nu!} A^\nu\Big| \leqq \sum_{\nu=2}^{\infty} \frac{1}{\nu!}|A|^\nu \leqq |A|^2 \sum_{\nu=2}^{\infty} \frac{1}{\nu!} \leqq |A|^2,$$

$$|\ln A - (A - E)| = \Big|\sum_{\nu=2}^{\infty} \frac{(-1)^{\nu-1}}{\nu}(A - E)^\nu\Big| \leqq \sum_{\nu=2}^{\infty} \frac{1}{\nu}|A - E|^2$$

$$= |A - E|^2 \sum_{\nu=2}^{\infty} \frac{1}{\nu}\,\frac{1}{2^{\nu-2}} \leqq |A|^2,$$

womit die Behauptung bewiesen ist.

In der Umgebung von O gilt

$$\ln(\exp A \cdot \exp B) = A + B + O(|A|\,|B|).$$

(Für vertauschbare A, B ist das Zusatzglied O!) Wir nehmen

$$|A| < \frac{1}{2}\ln\frac{3}{2}, \quad |B| < \frac{1}{2}\ln\frac{3}{2}.$$

an. Dann gilt:

$$|\exp A - E| < e^{\frac{1}{2}\ln\frac{3}{2}} - 1 = \sqrt{\frac{3}{2}} - 1, \quad |\exp B - E| < e^{\frac{1}{2}\ln\frac{3}{2}} - 1 = \sqrt{\frac{3}{2}} - 1,$$

$$\begin{aligned} |\exp A \cdot \exp B - E| &= |(\exp A - E) + (\exp B - E) + (\exp A - E)(\exp B - E)| \\ &\leqq |\exp A - E| + |\exp B - E| + |\exp A - E|\,|\exp B - E| \\ &< 2\left(\sqrt{\frac{3}{2}} - 1\right) + \left(\sqrt{\frac{3}{2}} - 1\right)^2 = \frac{1}{2} \end{aligned}$$

und:

$$|A + B| < \ln\frac{3}{2}, \quad |\exp(A + B) - E| < e^{\ln\frac{3}{2}} - 1 = \frac{1}{2}.$$

Folglich ist

$$\begin{aligned} \ln(\exp A \cdot \exp B) - A - B &= \ln(\exp A \cdot \exp B) - \ln\exp(A + B) \\ &= O(|\exp A \cdot \exp B - \exp(A + B)|), \end{aligned}$$

es gilt also, den Ausdruck $\exp A \cdot \exp B - \exp(A + B)$ abzuschätzen. Er ist der Limes von

$$\sum_{\mu=0}^{r} \frac{1}{\mu!} A^\mu \cdot \sum_{\mu=0}^{r} \frac{1}{\mu!} B^\mu - \sum_{\mu=0}^{r} \frac{1}{\mu!}(A + B)^\mu$$

für $r \to \infty$. Wenn man ausmultipliziert (diesmal ohne Benützung der, nicht vorausgesetzten, Vertauschbarkeit), so heben sich erstens offenbar alle Glieder A^μ und B^ν fort. Wir wollen nun die übrigen Glieder (die sowohl A als B enthalten) abschätzen.

Wie groß ist die Anzahl der Glieder, in denen A ϱ-mal und B σ-mal als Faktor auftritt? $(\varrho + \sigma \leqq r.)$

Im ausmultiplizierten

$$\sum_{\mu=0}^{r} \frac{1}{\mu!} A^\mu \sum_{\mu=0}^{r} \frac{1}{\mu!} B^\mu$$

kommt offenbar ein einziges solches Glied vor, und zwar mit dem Koeffizienten $\frac{1}{\varrho!\,\sigma!}$. Im ausmultiplizierten

$$\sum_{\mu=0}^{r} \frac{1}{\mu!}(A + B)^\mu$$

müssen alle solchen Glieder aus $\frac{1}{(\varrho + \sigma)!}(A + B)^{\varrho + \sigma}$ stammen, und hier kommen $\frac{(\varrho + \sigma)!}{\varrho!\,\sigma!}$ solche Glieder vor, ein jedes mit dem Koeffizienten $\frac{1}{(\varrho + \sigma)!}$. Der Absolutwert der Summe aller derartigen Glieder ist folglich

$$\leqq \frac{1}{\varrho!\,\sigma!}|A|^\varrho |B|^\sigma + \frac{(\varrho + \sigma)!}{\varrho!\,\sigma!} \frac{1}{(\varrho + \sigma)!}|A|^\varrho |B|^\sigma = \frac{2}{\varrho!\,\sigma!}|A|^\varrho |B|^\sigma.$$

Der Absolutwert des ganzen Ausdruckes ist also

$$\leqq \sum_{\varrho, \sigma = 1}^{\infty} \frac{2}{\varrho! \, \sigma!} |A|^{\varrho} |B|^{\sigma} = 2(e^{|A|} - 1)(e^{|B|} - 1).$$

Dies gilt für alle r, und somit auch für den Limes bei $r \to \infty$, wir haben also:

$$|\exp A \cdot \exp B - \exp(A + B)| \leqq 2(e^{|A|} - 1)(e^{|B|} - 1) = O(|A||B|).$$

Und hieraus folgt sofort

$$|\ln(\exp A \cdot \exp B) - A - B| = O(|A||B|),$$

womit unsere Behauptung bewiesen ist.

II. Gruppen und ihre Infinitesimalgruppen.

§ 1.

Als Gruppe bezeichnen wir, wie allgemein üblich, eine Menge in $\mathfrak{R}_m$ (d. h. eine Menge von Matrizen) $\mathfrak{G}$, die die folgenden Eigenschaften hat:

a) E gehört zu $\mathfrak{G}$.

b) Wenn A zu $\mathfrak{G}$ gehört, so ist $\det A \neq 0$, und A^{-1} gehört auch zu $\mathfrak{G}$.

c) Wenn A, B zu $\mathfrak{G}$ gehören, so gehört auch AB zu $\mathfrak{G}$.

Als Hülle von $\mathfrak{G}$ bezeichnen wir die Menge $\overline{\mathfrak{G}}$ aller Elemente von $\mathfrak{G}$, und aller Häufungspunkte von $\mathfrak{G}$ mit nichtverschwindender Determinante. $\overline{\mathfrak{G}}$ ist offenbar auch eine Gruppe, und seine eigene Hülle. (Man beachte, daß $\overline{\mathfrak{G}}$ keineswegs „abgeschlossen" sein muß, in dem Sinne wie dieses Wort in der Theorie der kontinuierlichen Gruppen benutzt wird: es muß nicht beschränkt und abgeschlossen sein.)

Als Infinitesimalgruppe $\mathfrak{J}$ von $\mathfrak{G}$ bezeichnen wir die Menge aller Matrizen U, zu denen eine Folge von zu $\mathfrak{G}$ gehörenden Matrizen A_1, $A_2, \ldots$, sowie eine Folge gegen 0 strebender positiver Zahlen $\varepsilon_1, \varepsilon_2, \ldots$ existiert, so daß

$$\frac{1}{\varepsilon_p}(A_p - E) \to U$$

für $p \to \infty$.

Man sieht leicht ein: $\mathfrak{G}$ und $\overline{\mathfrak{G}}$ haben dieselbe Infinitesimalgruppe. Denn $\mathfrak{G}$ ist in $\overline{\mathfrak{G}}$ enthalten, also ist es auch seine Infinitesimalgruppe in der von $\overline{\mathfrak{G}}$. Es bleibt noch übrig zu zeigen, daß jedes U, welches zur Infinitesimalgruppe von $\overline{\mathfrak{G}}$ gehört, auch zu der von $\mathfrak{G}$ gehören muß

$A_1, A_2, \ldots$ sollen also zu $\overline{\mathfrak{G}}$ gehören, $\varepsilon_1, \varepsilon_2, \ldots$ sei eine Folge gegen 0 strebender positiver Zahlen, und es sei

$$\frac{1}{\varepsilon_p}(A_p - E) \to U$$

für $p \to \infty$. Wir wählen für jedes p ein A'_p aus $\mathfrak{G}$ so aus, daß

$$|A'_p - A_p| < \varepsilon_p^2$$

ist; das ist möglich, weil A_p zu $\overline{\mathfrak{G}}$ gehört. Dann gilt für die Folge $A'_1, A'_2, \ldots$ aus $\mathfrak{G}$:

$$\left|\frac{1}{\varepsilon_p}(A_p - E) - \frac{1}{\varepsilon_p}(A'_p - E)\right| = \left|\frac{1}{\varepsilon_p}(A_p - A'_p)\right| < \varepsilon_p,$$

$$\frac{1}{\varepsilon_p}(A_p - E) - \frac{1}{\varepsilon_p}(A'_p - E) \to 0$$

für $p \to \infty$. Also strebt auch $\frac{1}{\varepsilon_p}(A'_p - E)$ gegen U, d. h. U gehört zur Infinitesimalgruppe von $\mathfrak{G}$.

§ 2.

Ehe wir die Infinitesimalgruppe weiter untersuchen, müssen wir zwei Sätze über Folgen $\frac{1}{\varepsilon_p}(A_p - E)$ beweisen.

Wir behaupten erstens: Wenn eine Folge von Matrizen $A_1, A_2, \ldots$ und eine Folge gegen 0 strebender positiver Zahlen $\varepsilon_1, \varepsilon_2, \ldots$ gegeben ist, so zieht die Konvergenz einer jeden der beiden Reihen

$$\frac{1}{\varepsilon_p}(A_p - E) \quad \text{und} \quad \frac{1}{\varepsilon_p}\ln A_p$$

die der anderen nach sich, und sie haben denselben Limes.

Wenn nämlich $\frac{1}{\varepsilon_p}(A_p - E)$ konvergiert, so ist $\left|\frac{1}{\varepsilon_p}(A_p - E)\right|$ beschränkt, und $|A_p - E|$ strebt gegen 0, folglich ist $\ln A_p$ für fast alle p sinnvoll, und es ist:

$$\frac{1}{\varepsilon_p}\ln A_p = \frac{1}{\varepsilon_p}(A_p - E) + \frac{1}{\varepsilon_p}O(|A_p - E|^2) = \frac{1}{\varepsilon_p}(A_p - E) + \varepsilon_p O\left(\left|\frac{1}{\varepsilon_p}(A_p - E)\right|^2\right)$$

$$= \frac{1}{\varepsilon_p}(A_p - E) + \varepsilon_p O(1) = \frac{1}{\varepsilon_p}(A_p - E) + O(\varepsilon_p).$$

Der erste Summand hat einen Limes, der zweite strebt gegen 0, also hat die Summe denselben Limes.

Konvergiert umgekehrt $\frac{1}{\varepsilon_p}\ln A_p$, so ist wieder $\left|\frac{1}{\varepsilon_p}\ln A_p\right|$ beschränkt, und es ist:

$$\frac{1}{\varepsilon_p}(A_p - E) = \frac{1}{\varepsilon_p}(\exp \ln A_p - E) = \frac{1}{\varepsilon_p}\left(\ln A_p + \frac{1}{\varepsilon_p}O(|\ln A_p|)^2\right)$$

$$= \frac{1}{\varepsilon_p}\ln A_p + \varepsilon_p O\left(\left|\frac{1}{\varepsilon_p}\ln A_p\right|^2\right) = \frac{1}{\varepsilon_p}\ln A_p + \varepsilon_p O(1)$$

$$= \frac{1}{\varepsilon_p}\ln A_p + O(\varepsilon_p).$$

Wieder hat der erste Summand einen Limes, der zweite strebt gegen 0, so daß die Summe denselben Limes hat.

Wenn jedem $\varepsilon > 0$ (etwa für $\varepsilon \leqq \varepsilon_0$, $\varepsilon_0 > 0$) eine Matrix A_ε zugeordnet ist, so zieht auch die Konvergenz einer jeder der Reihen

$$\frac{1}{\varepsilon}(A_\varepsilon - E) \quad \text{und} \quad \frac{1}{\varepsilon}\ln A_\varepsilon$$

(für $\varepsilon \to 0$) die der anderen nach sich, und ihr Limes ist gemeinsam. Um dies einzusehen, brauchen wir nur zu berücksichtigen, daß hier die Konvergenz für $\varepsilon \to 0$ dies bedeutet: Konvergenz für jede gegen 0 strebende Folge $\varepsilon_1, \varepsilon_2, \ldots$.

Unsere zweite Behauptung bezieht sich auf die Gruppe $\mathfrak{G}$ und lautet so: Wenn eine Folge $A_1, A_2, \ldots$ von Matrizen aus $\mathfrak{G}$ existiert, und eine Folge gegen 0 strebender positiver Zahlen $\varepsilon_1, \varepsilon_2, \ldots$, so daß

$$\frac{1}{\varepsilon_p}(A_p - E) \to U,$$

so können wir auch jedem $0 < \varepsilon \leqq 1$ eine Matrix A_ε aus $\mathfrak{G}$ zuordnen, so daß

$$\frac{1}{\varepsilon}(A_\varepsilon - E) \to U$$

für $\varepsilon \to 0$.

Es ist bequemer, die völlig gleichwertigen Folgen

$$\frac{1}{\varepsilon_p}\ln A_p \quad \text{und} \quad \frac{1}{\varepsilon}\ln A_\varepsilon$$

zu betrachten.

Wir suchen zu jedem $0 < \varepsilon \leqq 1$ ein $p(\varepsilon) = 1, 2, \ldots$ auf, so daß

$$\varepsilon_{p(\varepsilon)} < \varepsilon^2$$

ist, und wählen $q(\varepsilon) = 1, 2, \ldots$ so, daß

$$(q(\varepsilon) - 1)\varepsilon_{p(\varepsilon)} < \varepsilon^2 \leqq q(\varepsilon)\varepsilon_{p(\varepsilon)}$$

ist. Für $\varepsilon \to 0$ konvergiert $\varepsilon_{p(\varepsilon)}$ (da es $< \varepsilon^2$ ist) gegen 0, also $p(\varepsilon)$ gegen ∞. Und da

$$q(\varepsilon) \geqq \frac{\varepsilon}{\varepsilon_{p(\varepsilon)}} > \frac{1}{\varepsilon}$$

ist, strebt auch $q(\varepsilon)$ gegen ∞.

Wir definieren weiter:

$$B_\varepsilon = (A_{p(\varepsilon)})^{p(\varepsilon)}.$$

Da $A_{p(\varepsilon)}$ zu $\mathfrak{G}$ gehört, gehört auch B_ε dazu.

Da $\frac{1}{\varepsilon_p}\ln A_p$ konvergiert, ist es beschränkt, es gilt also:

$$\frac{1}{\varepsilon_p}\ln A_p = O(1), \qquad \ln A_p = O(\varepsilon_p),$$

$$q(\varepsilon)\ln A_{p(\varepsilon)} = O(q(\varepsilon)\,\varepsilon_{p(\varepsilon)}) = O(\varepsilon).$$

Für alle hinreichend kleinen ε ist also

$$|q(\varepsilon)\ln A_{p(\varepsilon)}| < \ln 2,$$

und folglich für alle $r = 1, 2, \ldots, q(\varepsilon)$

$$|r \ln A_{p(\varepsilon)}| < \ln 2,$$

$$|\exp(r \ln A_{p(\varepsilon)}) - E| < e^{\ln 2} - 1 = 1.$$

Nun ist aber

$$\exp(r \ln A_{p(\varepsilon)}) = (\exp \ln A_{p(\varepsilon)})^r = (A_{p(\varepsilon)})^r,$$

also haben wir für alle $r = 1, 2, \ldots, q(\varepsilon)$

$$|(A_{p(\varepsilon)})^r - E| < 1.$$

Wegen dieser Ungleichheit ist

$$\ln B_\varepsilon = \ln (A_{p(\varepsilon)})^{q(\varepsilon)} = q(\varepsilon) \ln A_{p(\varepsilon)},$$

$$\frac{1}{\varepsilon} \ln B_\varepsilon = \frac{q(\varepsilon)}{\varepsilon} \ln A_{p(\varepsilon)} = \frac{q(\varepsilon)\, \varepsilon_{p(\varepsilon)}}{\varepsilon} \cdot \frac{1}{\varepsilon_{p(\varepsilon)}} \ln A_{p(\varepsilon)}.$$

Für $\varepsilon \to 0$ konvergieren, wie wir wissen, $p(\varepsilon)$ und $q(\varepsilon)$ gegen ∞; der zweite Faktor konvergiert also gegen U. Und der erste ist

$$\geqq 1, \quad \leqq \frac{q(\varepsilon)}{q(\varepsilon) - 1},$$

also konvergiert er gegen 1. Damit ist

$$\frac{1}{\varepsilon} \ln B_\varepsilon \to U,$$

d. h. unsere Behauptung, bewiesen.

§ 3.

Aus dem zweiten Resultat von § 2 folgt: Wenn U ein Element von $\mathfrak{J}$ ist, und eine Folge $\eta_1, \eta_2, \ldots$ gegen 0 strebender positiver Zahlen willkürlich vorgegeben ist, so gibt es eine Folge von Matrizen $A_1, A_2, \ldots$ aus $\mathfrak{G}$, so daß für diese $\eta_1, \eta_2, \ldots$

$$\frac{1}{\eta_p}(A_p - E) \to U$$

ist. Wir wählen nämlich zuerst die B_ε aus $\mathfrak{G}$ so, daß

$$\frac{1}{\varepsilon}(B_\varepsilon - E) \to U$$

(für $\varepsilon \to 0$); und setzen dann $A_p = B_{\eta_p}$.

Auf Grund dieses Resultates können die wichtigsten Eigenschaften von $\mathfrak{J}$ unschwer entwickelt werden.

Wir behaupten: Wenn die Matrizen U, V zu $\mathfrak{J}$ gehören, so gehören auch die Matrizen αU (α eine beliebige *reelle* Zahl), $U + V$, $UV - VU$

zu $\mathfrak{J}$. (Daß $\mathfrak{J}$ überhaupt Elemente hat, ist klar: O gehört zu ihm, da man $A_p = E$ und etwa $\varepsilon_p = \frac{1}{p}$ setzen kann.)

Um dies zu beweisen, wählen wir irgendeine Folge $\varepsilon_1, \varepsilon_2, \ldots$ gegen 0 strebender positiver Zahlen aus und bestimmen zu ihnen solche Matrizen $A_1, A_2, \ldots$ und $B_1, B_2, \ldots$ aus $\mathfrak{G}$, daß

$$\frac{1}{\varepsilon_p}(A_p - E) \to U, \quad \frac{1}{\varepsilon_p}(B_p - E) \to V$$

ist. (Auf den Wert der ε_p kommt es uns nicht an, wichtig ist aber, daß wir für U, V dieselben ε_p haben!)

Dann ist erstens für $\alpha \neq 0$ für die ebenfalls gegen 0 strebende Folge $\frac{1}{\alpha}\varepsilon_1, \frac{1}{\alpha}\varepsilon_2, \ldots$

$$\frac{1}{\frac{1}{\alpha}\varepsilon_p}(A_p - E) \to \alpha U,$$

so daß αU zu $\mathfrak{J}$ gehört. Für $\alpha = 0$ gehört $\alpha U = O$ auch dazu.

Zweitens gehören die Matrizen $A_p B_p$ auch zu $\mathfrak{G}$, und es ist

$$\begin{aligned}\frac{1}{\varepsilon_p}(A_p B_p - E) &= \frac{1}{\varepsilon_p}(A_p - E) + \frac{1}{\varepsilon_p}(B_p - E) + \frac{1}{\varepsilon_p}(A_p - E)(B_p - E)\\ &= \frac{1}{\varepsilon_p}(A_p - E) + \frac{1}{\varepsilon_p}(B_p - E) + \varepsilon_p \cdot \frac{1}{\varepsilon_p}(A_p - E) \cdot \frac{1}{\varepsilon_p}(B_p - E).\end{aligned}$$

Dieser Ausdruck konvergiert offenbar gegen $U + V + O \cdot U \cdot V = U + V$, so daß auch $U + V$ zu $\mathfrak{J}$ gehört.

Drittens bilden wir die Matrizen $A_p B_p A_p^{-1} B_p^{-1}$, die auch zu $\mathfrak{G}$ gehören, wir wollen

$$\frac{1}{\varepsilon_p^2}(A_p B_p A_p^{-1} B_p^{-1} - E) \to UV - VU$$

beweisen, damit ist dann gezeigt, daß auch $UV - VU$ zu $\mathfrak{G}$ gehört.

$A_p - E$, $B_p - E$ streben gegen 0, also A_p, B_p gegen E, und somit auch $A_p^{-1} B_p^{-1}$. Wegen

$$\frac{1}{\varepsilon_p^2}(A_p B_p A_p^{-1} B^{-1} - E) = \frac{1}{\varepsilon^2}(A_p B_p - B_p A_p) \cdot A_p^{-1} B_p^{-1}$$

genügt es also

$$\frac{1}{\varepsilon_p^2}(A_p B_p - B_p A_p) \to UV - VU$$

zu beweisen. Dies ist aber leicht verifizierbar:

$$\begin{aligned}\frac{1}{\varepsilon_p^2}(A_p B_p - B_p A_p) &= \frac{1}{\varepsilon_p^2}((A_p - E)(B_p - E) - (B_p - E)(A_p - E))\\ &= \frac{1}{\varepsilon_p}(A_p - E) \cdot \frac{1}{\varepsilon_p}(B_p - E) - \frac{1}{\varepsilon_p}(B_p - E) \cdot \frac{1}{\varepsilon_p}(A_p - E),\end{aligned}$$

woraus die Behauptung ohne weiteres folgt.

§ 4.

Wir wählen unter den Elementen von $\mathfrak{J}$ eine maximale Anzahl linear unabhängiger aus. Die lineare Unabhängigkeit verstehen wir so, daß nur reelle Koeffizienten zugelassen werden; dann ist die höchstmögliche Anzahl $2m^2$. Diese Elemente seien etwa $\overline{U}_1, \overline{U}_2, \ldots, \overline{U}_k$.

Jedes Element U von $\mathfrak{J}$ muß von diesen linear abhängen, also

$$U = \alpha_1 \overline{U}_1 + \alpha_2 \overline{U}_2 + \ldots + \alpha_k \overline{U}_k \quad (\alpha_1, \alpha_2, \ldots, \alpha_k \text{ reelle Zahlen})$$

sein, wegen der Unabhängigkeit der $\overline{U}_1, \overline{U}_2, \ldots, \overline{U}_k$ sind die $\alpha_1, \alpha_2, \ldots, \alpha_k$ eindeutig bestimmt. Nach den Resultaten von § 3 gehört aber umgekehrt auch jedes

$$U = \alpha_1 \overline{U}_1 + \alpha_2 \overline{U}_2 + \ldots + \alpha_k \overline{U}_k \quad (\alpha_1, \alpha_2, \ldots, \alpha_k \text{ reelle Zahlen})$$

zu $\mathfrak{J}$. Also ist $\mathfrak{J}$ die durch die linear unabhängigen $\overline{U}_1, \overline{U}_2, \ldots, \overline{U}_k$ aufgespannte lineare Mannigfaltigkeit.

Dabei ist $0 \leqq k \leqq 2m^2$, $k = 0$ bedeutet, daß $\mathfrak{J}$ aus der O allein besteht.

Unter den linearen Mannigfaltigkeiten ist aber $\mathfrak{J}$ noch dadurch ausgezeichnet, daß es mit zwei Elementen U, V stets auch ihren sogenannten Kommutator $UV - VU$ enthält. In der Theorie der kontinuierlichen Gruppen werden die Infinitesimalgruppen in der Regel durch diese Eigenschaft independent definiert. Es wird aber oft als selbstverständlich angenommen, daß eine solche „Infinitesimalgruppe" stets auch eine im eigentlichen Sinne ist: d. h. daß sie aus einer Gruppe $\mathfrak{G}$ gewonnen werden kann. Diese Ausnahme ist indessen falsch, wie wir im Anhang an Hand sehr einfacher Beispiele zeigen werden.

Uns sollen hier nur die aus gewöhnlichen Gruppen $\mathfrak{G}$ hergeleiteten Infinitesimalgruppen $\mathfrak{J}$ beschäftigen.

III. Zusammenhangseigenschaften und Parameterzahl.

§ 1.

Da die Gruppe $\overline{\mathfrak{G}}$ eine Punktmenge im euklidischen Raume $\mathfrak{R}_m$ ist, können wir ihre Zusammenhangseigenschaften untersuchen[10]). Wir bezeichnen die Komponente von A (A ist irgendein Element von $\overline{\mathfrak{G}}$) mit $\overline{\mathfrak{G}}(A)$.

Die Abbildungen

$$X' = AX \quad \text{und} \quad X' = XA$$

sind stetige Abbildungen von $\overline{\mathfrak{G}}$ auf sich selbst, folglich bilden beide jede Komponente auf eine Komponente ab. Die Komponente von C geht also in die von AC bzw. CA über.

---

[10]) Vgl. Fußnote [7]), S. 7.

Wenn A, B zu $\overline{\mathfrak{G}}(E)$ gehören, so bildet

$$X' = AX$$

demnach $\overline{\mathfrak{G}}(E)$ auch $\overline{\mathfrak{G}}(A)$ ab, d. h. auf $\overline{\mathfrak{G}}(E)$. Da B zu $\overline{\mathfrak{G}}(E)$ gehört, gehört AB zum Bilde, also wieder zu $\overline{\mathfrak{G}}(E)$. Also ist $\overline{\mathfrak{G}}(E)$ eine Gruppe.

Wenn S irgendein Element von $\overline{\mathfrak{G}}$ ist, so können wir die Abbildung

$$X'' = S^{-1}XS$$

aus

$$X' = XS, \qquad X'' = S^{-1}X$$

zusammensetzen, sie bildet also $\overline{\mathfrak{G}}(E)$ auf $\overline{\mathfrak{G}}(S^{-1} \cdot E \cdot S) = \overline{\mathfrak{G}}(E)$ ab, folglich gehört für jedes A von $\overline{\mathfrak{G}}(E)$ auch $S^{-1}AS$ zu $\overline{\mathfrak{G}}(E)$.

Also ist $\overline{\mathfrak{G}}(E)$ ein Normalteiler von $\overline{\mathfrak{G}}$, und $\overline{\mathfrak{G}}(A)$ ist die zu A gehörige Nebengruppe von $\overline{\mathfrak{G}}(E)$: die Komponenten von $\overline{\mathfrak{G}}$ bilden seine Faktorgruppe.

Als Komponente von $\overline{\mathfrak{G}}$ ist $\overline{\mathfrak{G}}(E)$ in $\overline{\mathfrak{G}}$ relativ abgeschlossen, es enthält also alle seine Häufungspunkte mit nichtverschwindender Determinante.

§ 2.

U gehöre zu $\mathfrak{J}$. Wir wählen eine Folge $A_1, A_2, \ldots$ aus $\mathfrak{G}$ so aus, daß

$$p(A_p - E) \to U \quad \left(\text{d. h. } \varepsilon_p = \frac{1}{p}\right),$$

also

$$p \ln A_p \to U.$$

Es ist

$$\exp(p \ln A_p) = (\exp \ln A_p)^p = (A_p)^p,$$

und $\exp X$ ist an der Stelle U stetig, also gilt

$$(A_p)^p \to \exp U.$$

Da alle $(A_p)^p$ zu $\mathfrak{G}$ gehören, ist $\exp U$ ein Häufungspunkt von $\mathfrak{G}$. Dabei ist seine Determinante $\neq 0$, also gehört es zu $\overline{\mathfrak{G}}$.

Hieraus folgt sofort: wenn $U_1, U_2, \ldots, U_j$ zu $\mathfrak{J}$ gehören, so gehört auch

$$\exp(U_1) \cdot \exp(U_2) \cdot \ldots \cdot \exp(U_j)$$

zu $\overline{\mathfrak{G}}$. Wir behaupten aber: es gehört sogar zu $\overline{\mathfrak{G}}(E)$.

Denn für jedes $0 \leqq \vartheta \leqq 1$ gehören auch die $\vartheta U_1, \vartheta U_2, \ldots, \vartheta U_j$ zu $\mathfrak{J}$, also gehört

$$F(\vartheta) = \exp(\vartheta U_1) \cdot \exp(\vartheta U_2) \cdot \ldots \cdot \exp(\vartheta U_j)$$

zu $\overline{\mathfrak{G}}$. Wir haben also eine in $\overline{\mathfrak{G}}$ liegende analytische Kurve, die $F(0) = E$ mit $F(1)$, d. h. der untersuchten Matrix, verbindet. Diese muß folglich zu $\overline{\mathfrak{G}}(E)$ gehören.

Im folgenden wird sich unter anderem zeigen, daß $\overline{\mathfrak{G}}(E)$ keine weiteren Elemente hat.

§ 3.

Wir nehmen an, es wäre eine Folge von Matrizen $A_1, A_2, \ldots$ aus $\overline{\mathfrak{G}}$ gegeben, derart, daß alle $\ln A_p$ sinnvoll sind und keines zu $\mathfrak{J}$ gehört, und die A_p konvergieren gegen E. Es soll gezeigt werden, daß dies unmöglich ist.

Da O zu $\mathfrak{J}$ gehört, ist durchweg $\ln A_p \neq O$; wir setzen

$$|\ln A_p| = \delta_p > 0.$$

Da $\ln A_p$ nicht zu $\mathfrak{J}$ gehört und $\mathfrak{J}$ als lineare Mannigfaltigkeit eine abgeschlossene Menge ist, so hat es eine positive minimale Entfernung ε_p von $\ln A_p$, die in einem Punkte U_p von $\mathfrak{J}$ angenommen wird:

$$|\ln A_p - U_p| = \varepsilon_p > 0.$$

Da O zu $\mathfrak{J}$ gehört, ist insbesondere $\delta_p \geqq \varepsilon_p$. A_p konvergiert gegen E, also $\ln A_p$ gegen 0, folglich gilt

$$\delta_p \to 0, \qquad \varepsilon_p \to 0.$$

Also konvergiert auch U_p gegen 0.

Insbesondere sind A_p, $\ln A_p$, U_p beschränkt, wir dürfen also schließen:

$$|\exp \ln A_p - \exp U_p| = O(|\ln A_p - U_p|) = O(\varepsilon_p),$$
$$|A_p - \exp U_p| = O(\varepsilon_p),$$

und weiter:

$$\begin{aligned} |A_p \exp(-U_p) - E| &= |(A - \exp U_p) \exp(-U_p)| \\ &\leqq |A - \exp U_p|\,|\exp(-U_p)| = O(\varepsilon_p)\, O(1) = O(\varepsilon_p). \end{aligned}$$

$A_p \exp(-U_p)$ strebt demnach gegen E, was die folgende Abschätzung motiviert:

$$|\ln(A_p \exp(-U_p))| = O(|A_p \exp(-U_p) - E|) = O(\varepsilon_p).$$

Man beachte, daß für fast alle p (nämlich sobald U_p nahe genug zu O ist)

$$\ln(A_p \exp(-U_p)) \neq O$$

sein muß, denn das Gegenteil hätte

$$A_p \exp(-U_p) = E, \qquad A_p = \exp U_p,$$
$$\ln A_p = \ln \exp U_p = U_p$$

zur Folge. Wir setzen

$$|\ln(A_p \exp(-U_p))| = \eta_p > 0.$$

Wir haben gezeigt, daß $\eta_p = O(\varepsilon_p)$ ist.

A_p gehört zu $\overline{\mathfrak{G}}$, U_p und $-U_p$ gehören zu $\mathfrak{J}$, also $\exp(-U_p)$ zu $\overline{\mathfrak{G}}$, folglich gehört $A_p \exp(-U_p)$ zu $\overline{\mathfrak{G}}$. Alle Elemente der Folge

$$\frac{1}{\eta_p} \ln(A_p \exp(-U_p))$$

haben den Absolutwert 1, sie muß also mindestens einen Häufungspunkt W haben. Wir wählen eine Teilfolge $\nu(1), \nu(2), \ldots$ so aus, daß

$$\frac{1}{\eta_{\nu(p)}} \ln (A_{\nu(p)} \exp(-U_{\nu(p)})) \to W.$$

W gehört also zu $\mathfrak{J}$.

Aus der Definition von W und der $\nu(p)$ folgt

$$\frac{1}{\eta_{\nu(p)}} \ln (A_{\nu(p)} \exp(-U_{\nu(p)})) = W + o(1),$$
$$\ln (A_{\nu(p)} \exp(-U_{\nu(p)})) - \eta_{\nu(p)} W = o(\eta_{\nu(p)}).$$

Da beide Glieder der linken Seite gegen 0 streben, also beschränkt sind, so folgt daraus:

$$\begin{aligned} |A_{\nu(p)} \exp(-U_{\nu(p)}) - \exp(\eta_{\nu(p)} W)| &= O(|\ln(A_{\nu(p)} \exp(-U_{\nu(p)})) - \eta_{\nu(p)} W|) \\ &= o(\eta_{\nu(p)}), \end{aligned}$$

$$\begin{aligned} A_{\nu(p)} - \exp(\eta_{\nu(p)} W) \exp(U_{\nu(p)})| &= |(A \exp(-U_{\nu(p)}) - \exp(\eta_{\nu(p)} W)) \exp U_{\nu(p)}| \\ &\leqq |A \exp(-U_{\nu(p)}) - \exp(\eta_{\nu(p)} W)| \, |\exp U_{\nu(p)}| \\ &= o(\eta_{\nu(p)})\, O(1) = o(\eta_{\nu(p)}). \end{aligned}$$

Und hieraus folgt weiter, weil beide Glieder der linken Seite gegen E streben:

$$\begin{aligned} |\ln A_{\nu(p)} - \ln(\exp(\eta_{\nu(p)} W) \exp(U_{\nu(p)}))| &= O(|A_{\nu(p)} - \exp(\eta_{\nu(p)} W) \exp(U_{\nu(p)})|) \\ &= o(\eta_{\nu(p)}). \end{aligned}$$

Wir berücksichtigen nun, daß andererseits

$$\begin{aligned} |\ln(\exp(\eta_{\nu(p)} W) \exp(U_{\nu(p)})) - \eta_{\nu(p)} W - U_{\nu(p)}| &= O(|\eta_{\nu(p)} W| \, |U_{\nu(p)}|) \\ &= O(\eta_{\nu(p)})\, o(1) = o(\eta_{\nu(p)}) \end{aligned}$$

ist. Dies ergibt, kombiniert mit dem vorigen Resultat:

$$|\ln A_{\nu(p)} - (\eta_{\nu(p)} W + U_{\nu(p)})| = o(\eta_{\nu(p)}) = o(\varepsilon_{\nu(p)}).$$

Der Ausdruck $\eta_{\nu(p)} W + U_{\nu(p)}$ gehört aber zu $\mathfrak{J}$, weil W und $U_{\nu(p)}$ dazu gehören, es muß also nach Definition von $\varepsilon_{\nu(p)}$

$$|\ln A_{\nu(p)} - (\eta_{\nu(p)} W + U_{\nu(p)})| \geqq \varepsilon_{\nu(p)}$$

sein. Und dies ist unmöglich; denn derselbe Ausdruck kann nicht sowohl $o(\varepsilon_{\nu(p)})$ als auch für alle p gleichzeitig $\geqq \varepsilon_{\nu(p)}$ sein.

Damit ist die Unverträglichkeit unserer Annahmen bewiesen.

§ 4.

Das Resultat von § 3 kann auch so formuliert werden: Es gibt ein festes (aber von $\mathfrak{G}$ abhängiges!) $\Delta > 0$, so daß für alle A von $\mathfrak{G}$ mit $|A - E| < \Delta$ der $\ln A$ zu $\mathfrak{J}$ gehört. Dies hat aber sehr weitgehende Folgen.

Erstens ist $A = \exp(\ln A)$, und $\ln A$ gehört zu $\mathfrak{J}$: nach § 2 gehört also A zu $\overline{\mathfrak{G}}(E)$. D. h.: es gibt eine Umgebung von E, in der $\overline{\mathfrak{G}}(E)$ mit $\overline{\mathfrak{G}}$ übereinstimmt. Nach § 1 entsteht $\overline{\mathfrak{G}}(A)$ aus $\overline{\mathfrak{G}}(E)$ durch die Abbildung

$$X' = AX,$$

also gibt es eine Umgebung von A (die außer von $\overline{\mathfrak{G}}$ auch von A abhängt), in der $\overline{\mathfrak{G}}$ mit $\overline{\mathfrak{G}}(A)$ übereinstimmt.

Wir können dies auch so formulieren: Jede Komponente von $\overline{\mathfrak{G}}$ ist in $\overline{\mathfrak{G}}$ relativ offen, oder auch: jedes A von $\mathfrak{G}$ hat eine positive minimale Entfernung von der Komplementären seiner Komponente $\overline{\mathfrak{G}}(A)$ in $\overline{\mathfrak{G}}$.

Zweitens kann die am Schlusse von § 2 ausgesprochene Behauptung bewiesen werden: alle Elemente von $\overline{\mathfrak{G}}(E)$ haben die Form

$$\exp(U_1) \cdot \exp(U_2) \cdot \ldots \cdot \exp(U_j), \qquad (U_1, U_2, \ldots, U_j \text{ aus } \mathfrak{J}).$$

(Daß diese Matrizen alle zu $\overline{\mathfrak{G}}(E)$ gehören, wissen wir schon.)

Es sei nämlich $\overline{\mathfrak{G}}'$ die Menge aller dieser Matrizen und $\overline{\mathfrak{G}}''$ ihre Komplementäre in $\overline{\mathfrak{G}}(E)$. Wäre nicht $\overline{\mathfrak{G}}'$ gleich $\overline{\mathfrak{G}}(E)$, so wäre keine der beiden Mengen $\overline{\mathfrak{G}}'$, $\overline{\mathfrak{G}}''$ leer; da $\overline{\mathfrak{G}}(E)$ zusammenhängend ist, müßte eine von ihnen einen Häufungspunkt der anderen enthalten.

Dieser Punkt sei etwa A_0, A_0 gehört allenfalls zu $\overline{\mathfrak{G}}$, und jede Umgebung von A_0 enthält Punkte aus $\overline{\mathfrak{G}}'$ und aus $\overline{\mathfrak{G}}''$. Es gibt offenbar eine Umgebung $\mathfrak{U}$ von A_0, so daß für alle B, C von $\mathfrak{U}$

$$|B^{-1}C - E| < \Delta$$

ist (wegen $\det A_0 \neq 0$).

Wir wählen aus $\mathfrak{U}$ ein zu $\overline{\mathfrak{G}}'$ gehöriges B und ein zu $\overline{\mathfrak{G}}''$ gehöriges C aus, dann gehören B, C und $B^{-1}C$ zu $\overline{\mathfrak{G}}$, und es ist

$$|B^{-1}C - E| = \Delta.$$

Folglich gehört

$$\ln(B^{-1}C) = V$$

zu $\mathfrak{J}$. Da B zu $\overline{\mathfrak{G}}'$ gehört, ist

$$B = \exp(U_1) \cdot \exp(U_2) \cdot \ldots \cdot \exp(U_j) \qquad (U_1, U_2, \ldots, U_j \text{ aus } \mathfrak{J}),$$

und hieraus folgt

$$C = B \cdot B^{-1}C = \exp(U_1) \cdot \exp(U_2) \cdot \ldots \cdot \exp(U_j) \cdot \exp(V).$$

Also gehört auch C zu $\overline{\mathfrak{G}}'$, entgegen der Annahme.

Damit ist unsere Behauptung bewiesen.

Da $\overline{\mathfrak{G}}(A)$ aus $\overline{\mathfrak{G}}(E)$ durch die Abbildung

$$X' = AX$$

entsteht, so ist $\overline{\mathfrak{G}}(A)$ die Menge aller

$$A \cdot \exp(U_1) \cdot \exp(U_2) \cdot \ldots \cdot \exp(U_j) \qquad (U_1, U_2, \ldots, U_j \text{ aus } \mathfrak{J}).$$

§ 5.

Die soeben gewonnene Darstellung von $\overline{\mathfrak{G}}(A)$ erlaubt uns, den folgenden Satz zu beweisen: A und B gehören dann und nur dann zur gleichen Komponente von $\overline{\mathfrak{G}}$, wenn sie in $\overline{\mathfrak{G}}$ durch eine analytische Kurve verbunden werden können.

Daß die Bedingung hinreichend ist, ist klar; es bleibt zu zeigen, daß sie notwendig ist.

Wenn A, B zur selben Komponente gehören, so gehört B zu $\overline{\mathfrak{G}}(A)$, also ist

$$B = A\cdot\exp(U_1)\cdot\exp(U_2)\cdot\ldots\cdot\exp(U_j) \qquad (U_1, U_2, \ldots, U_j \text{ aus } \mathfrak{J}).$$

Wir setzen für $0 \leqq \vartheta \leqq 1$

$$F(\vartheta) = A\cdot\exp(\vartheta U_1)\cdot\exp(\vartheta U_2)\cdot\ldots\cdot\exp(\vartheta U_j).$$

$F(\vartheta)$ ist eine analytische Kurve; da $\vartheta U_1, \vartheta U_2, \ldots, \vartheta U_j$ zu $\mathfrak{J}$ gehören, gehören alle $F(\vartheta)$ zu $\overline{\mathfrak{G}}(A)$, also zu $\overline{\mathfrak{G}}$, und sie verbindet in der Tat $F(0) = A$ mit $F(1) = B$.

Trotz ihres durchsichtigen Charakters ist aber diese Darstellung von $\overline{\mathfrak{G}}(A)$ nicht geeignet, über die „Parameterzahl" von $\overline{\mathfrak{G}}$ Aufschlüsse zu geben; denn die $U_1, U_2, \ldots, U_j$ (und j selbst) sind ja keineswegs eindeutig bestimmt.

Diese ermitteln wir wie folgt:

Zu jedem A von $\overline{\mathfrak{G}}$ gibt es eine Umgebung $\mathfrak{U}$, welche das Bild von $|X - E| < \bar{\delta}$ bei der Abbildung

$$X' = AX$$

ist. Ein B aus $\mathfrak{U}$ gehört dann und nur dann zu $\overline{\mathfrak{G}}$, wenn $A^{-1}B$ zu $\overline{\mathfrak{G}}$ gehört, und dies ist wegen

$$|A^{-1}B - E| < \bar{\delta}$$

dann und nur dann der Fall, wenn $\ln(A^{-1}B)$ zu $\mathfrak{J}$ gehört. Wir können also auch sagen: Es gibt eine Umgebung $\mathfrak{U}$ von A, derart, daß der in $\mathfrak{U}$ gelegene Teil von $\overline{\mathfrak{G}}$ (oder, was offenbar dasselbe ist, von $\overline{\mathfrak{G}}(A)$) durch die Abbildung

$$Y = \ln(A^{-1}X), \qquad X = A\exp Y$$

ein-eindeutig auf den in einer gewissen Umgebung $\mathfrak{B}$ von O gelegenen Teil von $\mathfrak{J}$ abgebildet wird.

Das allgemeine Element von $\mathfrak{J}$ ist aber durch Angabe von k reellen Parametern charakterisiert (vgl. II § 4), es ist ja gleich

$$\alpha_1\overline{U}_1 + \alpha_2\overline{U}_2 + \ldots + \alpha_k\overline{U}_k \qquad (\alpha_1, \alpha_2, \ldots, \alpha_k \text{ beliebige reelle Zahlen}),$$

wo $\overline{U}_1, \overline{U}_2, \ldots, \overline{U}_k$ linear unabhängige Elemente von $\mathfrak{J}$ sind, in der größtmöglichen Anzahl k. Also ist auch $\overline{\mathfrak{G}}$ im Kleinen durch k Parameter analytisch beschreibbar: in einer Umgebung $\mathfrak{U}$ von A wird es durch die

$$A \cdot \exp(\alpha_1 \overline{U}_1 + \alpha_2 \overline{U}_2 + \ldots + \alpha_k \overline{U}_k)$$

ein-eindeutig dargestellt ($\alpha_1, \alpha_2, \ldots, \alpha_k$ reell, in einer gewissen Umgebung von $0, 0, \ldots, 0$).

Der sonderbare Umstand, daß $\overline{\mathfrak{G}}(A)$ im Kleinen durch

$$A \cdot \exp(U) \qquad (U \text{ in } \mathfrak{J})$$

erschöpft wird, im Großen dies aber zunächst nur für

$$A \cdot \exp(U_1) \cdot \exp(U_2) \cdot \ldots \cdot \exp(U_j) \qquad (U_1, U_2, \ldots, U_j \text{ in } \mathfrak{J})$$

bewiesen ist, legt die Vermutung nahe, ob man nicht auch im Großen mit $j = 1$ durchkommt.

Daß dies nicht so ist, soll im Anhang gezeigt werden. Es liegt jedoch durchaus im Bereiche der Möglichkeit, daß $j = 2$ stets hinreicht, um $\overline{\mathfrak{G}}(A)$ zu erschöpfen; wir können dies weder beweisen noch widerlegen.

§ 6.

Wir fassen unsere Resultate zusammen:

Satz I. *$\mathfrak{J}$ ist eine lineare Mannigfaltigkeit, d. h. es gibt k linear unabhängige Matrizen $\overline{U}_1, \overline{U}_2, \ldots, \overline{U}_k$ (k ist eine der Zahlen $0, 1, \ldots, 2m^2$), so daß $\mathfrak{J}$ die Menge aller*

$$\alpha_1 \overline{U}_1 + \alpha_2 \overline{U}_2 + \ldots + \alpha_k \overline{U}_k \qquad (\alpha_1, \alpha_2, \ldots, \alpha_k \text{ reelle Zahlen})$$

ist. Wenn $\mathfrak{J}$ die Matrizen U, V enthält, so enthält es auch ihren Kommutator $UV - VU$.

Die Komponente von E in $\overline{\mathfrak{G}}$, $\overline{\mathfrak{G}}(E)$ ist ein Normalteiler von $\overline{\mathfrak{G}}$, die Komponente von A in $\overline{\mathfrak{G}}$, $\overline{\mathfrak{G}}(A)$, ist die zu A gehörige Nebengruppe von $\overline{\mathfrak{G}}(E)$. Die Komponenten von $\overline{\mathfrak{G}}$ bilden die Faktorgruppe von $\overline{\mathfrak{G}}(E)$ in $\overline{\mathfrak{G}}$.

Jedes $\overline{\mathfrak{G}}(A)$ ist relativ offen in $\overline{\mathfrak{G}}$, folglich ist die Faktorgruppe von $\overline{\mathfrak{G}}(E)$ in $\overline{\mathfrak{G}}$ diskret: zu keinem A mit nichtverschwindender Determinante gibt es unendlich viele Komponenten von $\overline{\mathfrak{G}}$ in beliebiger Nähe.

Zwei Punkte A, B von $\overline{\mathfrak{G}}$ gehören dann und nur dann zur selben Komponente von $\overline{\mathfrak{G}}$, wenn sie durch eine in $\overline{\mathfrak{G}}$ liegende analytische Kurve verbunden werden können.

$\overline{\mathfrak{G}}(A)$ ist die Menge aller

$$A \cdot \exp(U_1) \cdot \exp(U_2) \cdot \ldots \cdot \exp(U_j) \qquad (U_1, U_2, \ldots, U_j \text{ aus } \mathfrak{J}).$$

Aber jedes A hat eine Umgebung $\mathfrak{U}$, derart, daß hier $\overline{\mathfrak{G}}$ (oder, was hier dasselbe ist, $\overline{\mathfrak{G}}(A)$) bereits durch die

$$A \cdot \exp(U) \qquad (U \textit{ aus } \mathfrak{J})$$

erschöpft wird, ja die Abbildung

$$Y = \ln(A^{-1}X), \qquad X = A\exp(Y)$$

bildet $\mathfrak{U}$ ein-eindeutig auf eine gewisse Umgehung $\mathfrak{V}$ von 0 in $\mathfrak{J}$ ab. $\overline{\mathfrak{G}}$ kann also überall im Kleinen ein-eindeutig und analytisch auf k reelle Parameter bezogen werden; es besteht in $\mathfrak{U}$ aus den

$$A \cdot \exp(\alpha_1 \overline{U}_1 + \alpha_2 \overline{U}_2 + \ldots + \alpha_k \overline{U}_k)$$

$$(\alpha_1, \alpha_2, \ldots, \alpha_k \textit{ reell, in einer Umgebung von } 0, 0, \ldots, 0).$$

$\overline{\mathfrak{G}}(E)$ ist also für $k > 0$ eine kontinuierliche Gruppe im üblichen Sinne des Wortes, mit allen Analytizitätseigenschaften, die man von einer solchen erwartet. $\overline{\mathfrak{G}}$ selbst ist nur dann als eigentlich kontinuierliche Gruppe anzusprechen, wenn es zusammenhängend ist, d. h. wenn sich die Faktorgruppe von $\overline{\mathfrak{G}}(E)$ auf die Einheit reduziert.

Für $k = 0$ besteht $\mathfrak{J}$ aus O allein, also $\overline{\mathfrak{G}}(A)$ aus A allein, d. h. $\overline{\mathfrak{G}}$ ist mit der Faktorgruppe von $\overline{\mathfrak{G}}(E)$ (das ja E allein enthält) identisch. Folglich ist es diskret: es hat keinen einzigen Häufungspunkt mit nichtverschwindender Determinante[11]. (Insbesondere ist es endlich oder abzählbar, weil es keinen seiner Häufungspunkte enthält.)

In diesem Falle ist also $\overline{\mathfrak{G}}$ eine diskrete Gruppe im üblichen Sinne des Wortes.

IV. Darstellungen und Stetigkeit.

§ 1.

Mit dem Satze I haben unsere auf die Gruppe $\mathfrak{G}$ allein bezüglichen Betrachtungen einen Abschluß erreicht, wir wenden uns nunmehr den Darstellungen von $\mathfrak{G}$ zu.

n sei eine Zahl $= 1, 2, \ldots$, wir werden neben den Matrizen aus $\mathfrak{R}_m$ nunmehr auch solche aus $\mathfrak{R}_n$ betrachten. (Um Verwechslungen zu vermeiden, schreiben wir in $\mathfrak{R}_n$ für O, E, ln, exp stets O_*, E_*, $\ln_*$, $\overline{\exp_*}$.)

Eine Darstellung von $\mathfrak{G}$ ist eine in $\mathfrak{G}$ definierte Funktion D, die Werte aus $\mathfrak{R}_n$ annimmt und der folgenden Funktionalgleichung genügt:

$$D(A \cdot B) = D(A) \cdot D(B) \qquad (A, B \text{ aus } \mathfrak{G}).$$

---

[11]) Häufungspunkte mit verschwindender Determinante können sehr wohl vorhanden sein, z. B. $\mathfrak{G}$ sei die Menge aller 2^ν ($\nu = 0, \pm 1, \pm 2, \ldots$), diese Zahlen können ja als eindimensionale Matrizen gelten.

Wir verlangen außerdem noch

$$D(E) = E_*,$$

dies ist übrigens, wie aus einfachen Sätzen über Matrizen folgt, keine wesentliche Beschränkung[12]).

Eine Darstellung D von $\mathfrak{G}$ muß keineswegs stetig sein, Beispiele hierfür werden wir im Anhang geben. Es soll aber im folgenden gezeigt werden, daß Darstellungen, deren Schwankung in der Umgebung von E kleiner als 1 ist (d. h. für die ein $\bar{\delta} > 0$ und ein $\bar{\eta} > 0$ existiert, so daß für alle A aus $\mathfrak{G}$ mit $|A - E| < \bar{\delta}$

$$|D(A) - E_*| < 1 - \bar{\eta}$$

ist), nicht nur überall auf $\mathfrak{G}$ stetig und auf $\bar{\mathfrak{G}}$ stetig erweiterbar sind, sondern noch viel weitergehende Analytizitätseigenschaften besitzen müssen.

Diese Prämisse (Schwankung bei E kleiner als 1) kann übrigens nicht mehr wesentlich verschärft werden: denn es gibt unstetige Darstellungen, deren Schwankung in der Umgebung von E gleich 2 ist (vgl. Anhang).

Die Prämisse ist sicher erfüllt, wenn $D(A)$ für $A = E$ stetig ist (Schwankung $= 0$), und dies ist wegen

$$D(A) = D(A A_0) D(A_0^{-1})$$

der Fall, wenn $D(A)$ für irgendein $A = A_0$ in $\mathfrak{G}$ stetig ist. Darstellungen, welche sie verletzen, müssen also überall in $\mathfrak{G}$ unstetig sein.

§ 2.

Wir nehmen also die Existenz der in § 1 erwähnten $\bar{\delta} > 0$, $\bar{\eta} > 0$ an. p sei irgendeine Zahl $= 1, 2, \ldots$. Wenn A zu $\mathfrak{G}$ gehört, und

$$|A - E| < \varepsilon$$

ist ($\varepsilon > 0$, wir werden später genau über ε verfügen), so gilt:

$$|\ln A| < \ln \frac{1}{1-\varepsilon}$$

und für alle $r = 1, 2, \ldots, p$:

$$|r \ln A| = r |\ln A| < p \ln \frac{1}{1-\varepsilon},$$

$$|\exp(r \ln A) - E| < e^{p \ln \frac{1}{1-\varepsilon}} - 1 = \frac{1}{(1-\varepsilon)^p} - 1,$$

---

[12]) Da jedenfalls $D(E)^2 = D(E)$ ist, zeigt man leicht, daß $D(E)$ auf die Diagonalform transformiert werden kann und daß alle Diagonalelemente $= 1$ oder 0 sind. Also ist die ganze Darstellung einer solchen äquivalent, bei der $D(E)$ gleich einer mit Nullen geänderten Einheitsmatrix ist. Wegen

$$D(A) = D(E) D(A) D(E)$$

kommt dieselbe Änderung bei allen $D(A)$ vor; wenn wir sie fortlassen, bleibt eine Darstellung mit $D(E) = E_*$ übrig.

d. h.

$$|A^r - E| < \frac{1}{(1-\varepsilon)^p} - 1.$$

Wenn also

$$\frac{1}{(1-\varepsilon)^p} - 1 \leqq \bar{\delta}, \qquad \varepsilon \leqq 1 - \frac{1}{\sqrt[p]{1+\bar{\delta}}}$$

ist, so gilt

$$|A^r - E| < \bar{\delta}, \qquad |D(A^r) - E_*| < 1 - \bar{\eta}, \qquad |(D(A))^r - E_*| < 1 - \bar{\eta}.$$

Hieraus folgt erstens

$$\ln_*(D(A))^p = p \ln_* D(A),$$

und zweitens

$$|\ln_*(D(A))^p| < \ln\frac{1}{\bar{\eta}}.$$

Folglich ist

$$|\ln_* D(A)| < \frac{1}{p}\ln\frac{1}{\bar{\eta}}, \qquad |D(A) - E_*| < e^{\frac{1}{p}\ln\frac{1}{\bar{\eta}}} - 1 = \sqrt[p]{\frac{1}{\bar{\eta}}} - 1.$$

Nun gibt es sicher eine Konstante $c_1 > 0$ (c_1 und ebenso c_2, c_3 sind von p und A unabhängig), so daß stets

$$\frac{c_1}{p} \leqq 1 - \frac{1}{\sqrt[p]{1+\bar{\delta}}}$$

ist, wir können also $\varepsilon = \frac{c_1}{p}$ setzen.

Wenn A zu $\mathfrak{G}$ gehört und

$$|A - E| < c_1$$

ist, so sei zunächst $A \neq E$. Dann gibt es ein $p = 1, 2, \ldots$ mit

$$\frac{c_1}{2p} \leqq |A - E| < \frac{c_1}{p},$$

und folglich ist

$$|D(A) - E_*| < \sqrt[p]{\frac{1}{\bar{\eta}}} - 1 \leqq \frac{c_2}{p}$$

(für ein geeignet gewähltes c_2), also

$$\leqq c_2 \cdot \frac{2}{c_1} \cdot |A - E| = c_3 |A - E|,$$

Aus $|A - E| < c_1$ folgt also $|D(A) - E_*| < c_3 |A - E|$, wenn $A \neq E$ ist. Dies ist natürlich auch für $A = E$ gültig, nur tritt an Stelle von $<$ das Gleichheitszeichen.

D. h. $D(A)$ genügt in der Umgebung von E der Lipschitzschen Bedingung, es ist also dort stetig. Dies werden wir nun auf ganz $\bar{\mathfrak{G}}$ ausdehnen.

§ 3.

A_0 sei ein Punkt von $\overline{\mathfrak{G}}$. Wir behaupten: es gibt eine Umgebung $\mathfrak{U}$ von A_0, in der (d. h. in deren mit $\mathfrak{G}$ gemeinsamen Teile) $D(A)$ beschränkt ist.

Es gibt ein $\varepsilon > 0$, so daß aus $|B - A_0| < \varepsilon$, $|C - A_0| < \varepsilon$

$$|B^{-1}C - E| < \bar{\delta}$$

folgt (wegen $\det A_0 \neq 0$). Wir definieren $\mathfrak{U}$ durch $|X - A_0| < \varepsilon$ und wählen ein zu $\mathfrak{G}$ gehöriges B_0 in $\mathfrak{U}$ fest aus (dies ist möglich, weil A_0 zu $\overline{\mathfrak{G}}$ gehört).

Dann gilt für jedes zu $\mathfrak{G}$ gehörige C in $\mathfrak{U}$:

$$|B_0^{-1}C - E| < \bar{\delta}, \quad |D(B_0^{-1}C) - E_*| < 1,$$

$$\begin{aligned} |D(C) - D(B_0)| &= |D(B_0)(D(B_0^{-1}C) - E_*)| \\ &\leqq |D(B_0)|\,|D(B_0^{-1}C) - E_*| \leqq |D(B_0)|, \end{aligned}$$

$$|D(C)| \leqq 2|D(B_0)|,$$

also ist $D(A)$ in $\mathfrak{U}$ in der Tat beschränkt,

Nun sei $\overline{\mathfrak{H}}$ irgendeine beschränkt-abgeschlossene Teilmenge von $\overline{\mathfrak{G}}$. Wir behaupten: $D(A)$ ist in $\overline{\mathfrak{H}}$ (d. h. im gemeinsamen Teile von $\overline{\mathfrak{H}}$ und $\mathfrak{G}$) beschränkt.

Jeder Punkt A_0 von $\overline{\mathfrak{H}}$ ist ja in einer Umgebung $\mathfrak{U}$ enthalten, in welcher $D(A)$ beschränkt ist; $\overline{\mathfrak{H}}$ wird also durch diese Umgebungen überdeckt. Da $\overline{\mathfrak{H}}$ beschränkt-abgeschlossen ist, ist der Borelsche Überdeckungssatz anwendbar: $\overline{\mathfrak{H}}$ wird bereits durch endlich viele $\mathfrak{U}$ überdeckt. Folglich ist $D(A)$ in ganz $\overline{\mathfrak{H}}$ beschränkt.

Hieraus kann aber weiter gefolgert werden, daß $D(A)$ in $\overline{\mathfrak{H}}$ gleichmäßig der Lipschitzschen Bedingung genügt.

Wir wählen nämlich c_4 so, daß in $\overline{\mathfrak{H}}$ stets

$$|D(A)| \leqq c_4$$

ist (c_4 und ebenso $c_5, c_6, \ldots$ hängen nicht von A, B ab). Ferner ist in $\overline{\mathfrak{H}}$ (weil es Teil von $\overline{\mathfrak{G}}$ ist) stets

$$\det A \neq 0.$$

Da $\overline{\mathfrak{H}}$ beschränkt abgeschlossen ist, gibt es ein $c_5 > 0$, so daß in $\overline{\mathfrak{H}}$ durchweg

$$|\det A| \geqq c_5$$

ist. Und wegen der Beschränktheit von $\overline{\mathfrak{H}}$ gibt es ein c_6, so daß in $\mathfrak{H}$ stets

$$|A| \leqq c_6$$

ist.

Aus den beiden letzten Ungleichheiten folgt, daß es ein c_7 gibt, so daß für alle A, B von $\overline{\mathfrak{H}}$

$$|AB^{-1} - E| \leqq c_7 |A - B|$$

ist.

Nun sollen A, B zu $\overline{\mathfrak{H}}$ (und gleichzeitig zu $\mathfrak{G}$) gehören. Wenn

$$|A - B| < \frac{c_1}{c_7}$$

ist, so ist

$$|AB^{-1} - E| \leqq c_7 |A - B| < c_1,$$

$$|D(AB^{-1}) - E_*| \leqq c_3 |AB^{-1} - E| \leqq c_3 c_7 |A - B|,$$

$$|D(A) - D(B)| = |(D(AB^{-1}) - E_*)D(B)| \leqq |D(AB^{-1}) - E_*| \, |D(B)| \leqq c_3 c_4 c_7 |A - B|.$$

Ist hingegen

$$|A - B| \geqq \frac{c_1}{c_7},$$

so gilt

$$|D(A) - D(B)| \leqq 2 c_4 \leqq \frac{2 c_4 c_7}{c_1} |A - B|,$$

Mit $c_8 = \operatorname{Max}\left(c_3 c_4 c_7, \frac{2 c_4 c_7}{c_1}\right)$ gilt also für alle A, B von $\overline{\mathfrak{H}}$ (und $\mathfrak{G}$)

$$|D(A) - D(B)| \leqq c_8 |A - B|.$$

§ 4.

Nach dem Schlußresultate von § 3 ist $D(A)$ nicht nur in ganz $\mathfrak{G}$ stetig, sondern auch in allen Punkten von $\overline{\mathfrak{G}}$ (es braucht dort nicht definiert zu sein, es ist aber in beliebig nahen Punkten definiert, und hat die Schwankung 0). Folglich kann $D(A)$ auf eine ganz eindeutige Weise auf $\overline{\mathfrak{G}}$ erweitert werden, unter Wahrung der Stetigkeit.

Weil die Erweiterung stetig ist, so wird die Gleichung

$$D(AB) = D(A) D(B)$$

für alle A, B von $\overline{\mathfrak{G}}$ fortbestehen: das erweiterte D ist also eine Darstellung von $\overline{\mathfrak{G}}$. Auch die zunächst nur im gemeinsamen Teile von $\overline{\mathfrak{H}}$ und $\mathfrak{G}$ gültige Relation

$$|D(A) - D(B)| \leqq c_8 |A - B|$$

muß, wegen der Stetigkeit, in ganz $\overline{\mathfrak{H}}$ bestehen. (Eigentlich braucht nicht jeder Punkt von $\overline{\mathfrak{H}}$ Häufungspunkt des Durchschnittes mit $\mathfrak{G}$ zu sein. Wir müssen darum eigentlich zuerst eine umfassendere beschränkt-abgeschlossene Teilmenge $\overline{\mathfrak{H}}'$ von $\overline{\mathfrak{G}}$ aufstellen — so daß $\overline{\mathfrak{H}}$ mit $\mathfrak{G} - \overline{\mathfrak{H}}'$ keinen gemeinsamen Häufungspunkt hat — und dann c_8 entsprechend wählen. Dann ist der Übergang auf $\overline{\mathfrak{H}}$ ohne weiteres zu vollziehen.)

Das erweiterte D ist also nicht nur überall in $\overline{\mathfrak{G}}$ stetig, sondern es genügt auch in jedem beschränkt-abgeschlossenen Teile $\overline{\mathfrak{H}}$ von $\overline{\mathfrak{G}}$ gleichmäßig der Lipschitzschen Bedingung.

Die Möglichkeit der Erweiterung auf $\overline{\mathfrak{G}}$ hängt wesentlich an der Stetigkeit; denn es gibt unstetige Darstellungen, die auf keine Art auf $\overline{\mathfrak{G}}$ erweitert werden können (vgl. Anhang). Im folgenden verstehen wir unter D stets die erweiterte Darstellung.

V. Die Infinitesimaldarstellung und der Entwicklungssatz.

§ 1.

Wir werden zeigen: Wenn $A_1, A_2, \ldots$ eine Folge von Matrizen aus $\overline{\mathfrak{G}}$ ist, und $\varepsilon_1, \varepsilon_2, \ldots$ eine Folge gegen 0 strebender Zahlen, so folgt aus der Konvergenz von

$$\frac{1}{\varepsilon_p}(A_p - E)$$

auch die von

$$\frac{1}{\varepsilon_p}(D(A_p) - E_*).$$

Natürlich dürfen an Stelle dieser Folgen die Folgen

$$\frac{1}{\varepsilon_p}\ln A_p \qquad \text{und} \qquad \frac{1}{\varepsilon_p}\ln_* D(A_p)$$

betrachtet werden (vgl. II § 2; die dortigen Resultate gelten natürlich ebenso in $\mathfrak{R}_n$). Der Limes der ersteren Folge sei etwa U. Wir nehmen eine Zahl $k = 1, 2, \ldots$ als gegeben an, sie soll erst später genau festgelegt werden. Ferner wählen wir (was offenbar möglich ist) eine Folge $q(1), q(2), \ldots$ positiver ganzer Zahlen so aus, daß

$$q(p)\,\varepsilon_p \to \frac{1}{k}$$

für $p \to \infty$. Dann gilt offenbar

$$q(p)\ln A_p \to \frac{1}{k}U,$$

und es genügt jedenfalls, die Konvergenz der Folge

$$q(p)\ln_* D(A_p)$$

zu beweisen.

Wir wählen nun k: Es soll

$$\left|\frac{1}{k}U\right| < \ln(1+\bar\delta), \qquad k > \frac{|U|}{\ln(1+\bar\delta)}$$

sein. Folglich gilt für alle hinreichend großen p

$$|q(p)\ln A_p| < \ln(1+\bar\delta)$$

und folglich für alle $r = 1, 2, \ldots, q(p)$

$$|r \ln A_p| \leqq |q(p) \ln A_p| < \ln(1+\bar\delta),$$
$$|\exp(r \ln A_p) - E| < e^{\ln(1+\bar\delta)} - 1 = \bar\delta,$$
$$|A_p^r - E| < \bar\delta.$$

Hieraus folgt weiter

$$|D(A_p^r) - E_*| < 1, \qquad |(D(A_p))^r - E_*| < 1,$$

also

$$\ln_* D(A_p^{q(p)}) = \ln_*(D(A_p))^{q(r)} = q(r) \ln_* D(A_p).$$

Es genügt also die Konvergenz der Folge $\ln_* D(A_p^{q(r)})$ zu beweisen.

Aus

$$q(p) \ln A_p \to \frac{1}{k} U$$

folgt

$$\exp(q(p) \ln A_p) \to \exp\left(\frac{1}{k} U\right),$$
$$A_p^{q(p)} \to \exp\left(\frac{1}{k} U\right).$$

Die A_p, also auch die $A_p^{q(p)}$ gehören zu $\bar{\mathfrak{G}}$, also ist $\exp\left(\frac{1}{k} U\right)$ ein Häufungspunkt von $\bar{\mathfrak{G}}$. Ferner ist die Determinante von $\exp\left(\frac{1}{k} U\right)$ keinesfalls 0, also gehört $\exp\left(\frac{1}{k} U\right)$ zur Hülle von $\bar{\mathfrak{G}}$, d. h. zu $\bar{\mathfrak{G}}$ selbst. Folglich ist $D(A)$ an der Stelle $\exp\left(\frac{1}{k} U\right)$ definiert und stetig, so daß

$$D(A_p^{q(p)}) \to D\left(\exp\left(\frac{1}{k} U\right)\right)$$

gilt.

Ferner ist

$$\left|\frac{1}{k} U\right| < \ln(1+\bar\delta), \qquad \left|\exp\left(\frac{1}{k} U\right) - E\right| < e^{\ln(1+\bar\delta)} - 1 = \bar\delta,$$
$$\left|D\left(\exp\left(\frac{1}{k} U\right)\right) - E_*\right| < 1,$$

also $\ln_* A$ an der Stelle $D\left(\exp\left(\frac{1}{k} U\right)\right)$ definiert und stetig. Und dies hat

$$\ln_* D(A_p^{q(p)}) \to \ln_* D\left(\exp\left(\frac{1}{k} U\right)\right),$$
$$q(p) \ln_* D(A_p) \to \ln_* D\left(\exp\left(\frac{1}{k} U\right)\right)$$

zur Folge, womit unsere Behauptung bewiesen ist. (Wir könnten schon jetzt die Limites berechnen, indessen bestimmen wir sie später auch höchst einfach.)

§ 2.

U sei irgendein Element von $\mathfrak{J}$. Wir betrachten alle Folgen $A_1, A_2, \ldots$ aus $\overline{\mathfrak{G}}$, zu denen eine Folge $\varepsilon_1, \varepsilon_2, \ldots$ gegen 0 strebender positiver Zahlen existiert, so daß

$$\frac{1}{\varepsilon_p}(A_p - E) \to U.$$

Für alle diese Folgen hat auch

$$\frac{1}{\varepsilon_p}(D(A_p) - E_*)$$

einen Limes; und zwar für alle denselben; denn wir können ja zwei derartige Folgen stets zu einer ebensolchen kombinieren. Dieser Limes hängt also nur von U ab, wir bezeichnen ihn mit $J(U)$. J ist also in $\mathfrak{J}$ definiert, es ist die zu D gehörende Infinitesimaldarstellung.

Wir wollen nun die Eigenschaften von J etwas näher untersuchen.

U, V seien zwei Elemente von $\mathfrak{J}$, α eine reelle Zahl. Wir wählen eine Folge $\varepsilon_1, \varepsilon_2, \ldots$ gegen 0 strebender Zahlen, dann gibt es zwei Folgen $A_1, A_2, \ldots$ und $B_1, B_2, \ldots$ von Matrizen aus $\overline{\mathfrak{G}}$, derart, daß

$$\frac{1}{\varepsilon_p}(A_p - E) \to U, \qquad \frac{1}{\varepsilon_p}(B_p - E) \to V.$$

Natürlich ist dann

$$\frac{1}{\varepsilon_p}(D(A_p) - E_*) \to J(U), \qquad \frac{1}{\varepsilon_p}(D(B_p) - E_*) \to J(V).$$

Wir haben in II § 3 bewiesen, daß dann

$$\frac{1}{\frac{1}{\alpha}\varepsilon_p}(A_p - E) \to \alpha U,$$

$$\frac{1}{\varepsilon_p}(A_p B_p - E) \to U + V,$$

$$\frac{1}{\varepsilon_p^2}(A_p B_p A_p^{-1} B_p^{-1}) \to UV - VU.$$

Wörtlich ebenso können wir schließen, daß

$$\frac{1}{\frac{1}{\alpha}\varepsilon_p}(D(A_p) - E_*) \to \alpha J(U),$$

$$\frac{1}{\varepsilon_p}(D(A_p B_p) - E_*) = \frac{1}{\varepsilon_p}(D(A_p) D(B_p) - E_*) \to J(U) + J(V),$$

$$\frac{1}{\varepsilon_p^2}(D(A_p B_p A_p^{-1} B_p^{-1}) - E_*)$$
$$= \frac{1}{\varepsilon_p^2}(D(A_p) D(B_p) D(A_p)^{-1} D(B_p)^{-1} - E_*)$$
$$\to J(U) J(V) - J(V) J(U).$$

D. h.: Es ist

$$J(\alpha U) = \alpha J(U) \qquad (\alpha \text{ reell}),$$
$$J(U+V) = J(U) + J(V),$$
$$J(UV - VU) = J(U)\,J(V) - J(V)\,J(U).$$

(Die erste Relation gilt allerdings zunächst nur für $\alpha \neq 0$. Aber wenn man etwa $\alpha = 1$ und $\alpha = -1$ setzt und die zweite heranzieht, so folgt sie auch für $\alpha = 0$.)

Wenn nun $\mathfrak{J}$ etwa durch $\overline{U}_1, \overline{U}_2, \ldots, \overline{U}_k$ aufgespannt wird, so ist jedes U von $\mathfrak{J}$ eindeutig als

$$U = \alpha_1 \overline{U}_1 + \alpha_2 \overline{U}_2 + \ldots + \alpha_k \overline{U}_k \qquad (\alpha_1, \alpha_2, \ldots, \alpha_k \text{ reell})$$

darstellbar, und es gilt dann

$$J(U) = \alpha_1 J(\overline{U}_1) + \alpha_2 J(\overline{U}_2) + \ldots + \alpha_k J(\overline{U}_k).$$

Da die $\alpha_1, \alpha_2, \ldots, \alpha_k$ durch U eindeutig bestimmt sind, sind sie lineare Ausdrücke der Real- und Imaginärteile von U, also gilt (weil die $J(\overline{U}_1)$, $J(\overline{U}_2), \ldots, J(\overline{U}_k)$ konstant sind) dasselbe von den Real- und Imaginärteilen von $J(U)$. D. h.: $J(U)$ ist linear in U. Insbesondere ist es also stetig und beliebig oft differenzierbar.

Mit der Linearität sind bloß die zwei ersten Eigenschaften von J berücksichtigt, die dritte Eigenschaft

$$J(UV - VU) = J(U)\,J(V) - J(V)\,J(U)$$

ist eine weitere, recht wirksame Einschränkung. In der Darstellungstheorie bildet sie vielfach den Ausgangspunkt zur Bestimmung aller J, und so aller D.

§ 3.

Wir werden auf Grund der Resultate von § 2 beweisen können, daß die Darstellung $D(A)$ in einer Umgebung eines jeden Punktes A_0 von $\overline{\mathfrak{G}}$ durch Potenzreihen (der Real- und Imaginärteile der Elemente von U) erzeugt wird. Indessen ist hierzu die Heranziehung des vollen Satzes I nötig: wir müssen im Besitz der Erzeugung von $\overline{\mathfrak{G}}$ durch

$$A_0 \cdot \exp(\alpha_1 \overline{U}_1 + \alpha_2 \overline{U}_2 + \ldots + \alpha_k \overline{U}_k) \qquad (\alpha_1, \alpha_2, \ldots, \alpha_k \text{ reell})$$

in einer Umgebung von A_0 (bzw. $0, 0, \ldots, 0$) sein.

Wir wollen aber zeigen, wie ohne Heranziehung dieser Tatsache die beliebig oftmalige Differenzierbarkeit von D aus den Resultaten von § 2 direkt erschlossen werden kann.

A sowie $A_1, A_2, \ldots$ seien Matrizen aus $\overline{\mathfrak{G}}$, $\varepsilon_1, \varepsilon_2, \ldots$ sei eine gegen 0 strebende Folge positiver Zahlen. Aus

$$\frac{1}{\varepsilon_p}(A_p - A) \to U$$

folgt dann:

$$\frac{1}{\varepsilon_p}(A_p A^{-1} - E) = \frac{1}{\varepsilon_p}(A_p - A)\cdot A^{-1} \to UA^{-1}$$

so, daß UA^{-1} zu $\mathfrak{J}$ gehört; und weiter:

$$\frac{1}{\varepsilon_p}(D(A_p A^{-1}) - E_*) \to J(UA^{-1}),$$

$$\frac{1}{\varepsilon_p}(D(A_p) - D(A)) = \frac{1}{\varepsilon_p}(D(A_p A^{-1}) - E_*)\cdot D(A) \to J(UA^{-1})D(A).$$

Dies bedeutet aber, daß D an der Stelle A und in der Richtung U differenzierbar ist und daß der Differentialquotient $J(UA^{-1})D(A)$ ist.

Wenn nun D als bereits ν-mal differenzierbar erkannt ist, so kann man so weiter schließen: $J(UA^{-1})$ ist linear in UA^{-1}, also beliebig oft differenzierbar. Also ist $J(UA^{-1})D(A)$ auch ν-mal differenzierbar; da es aber der Differentialquotient von $D(A)$ ist, ist dieses selbst $\nu+1$-mal differenzierbar.

Also ist D beliebig oft differenzierbar.

§ 4.

Wir führen nun unsere Betrachtungen systematisch zu Ende.

A gehöre zu $\overline{\mathfrak{G}}$, und es gehöre zu der Umgebung von E, in der hieraus folgt, daß auch $\ln A$ zu $\mathfrak{J}$ gehört (vgl. Satz I).

Dann gibt es eine Folge $A_1, A_2, \ldots$ aus $\overline{\mathfrak{G}}$ so, daß

$$p(A_p - E) \to \ln A \quad \left(\text{also } \varepsilon_p = \frac{1}{p}\right),$$

$$p \ln A_p \to \ln A.$$

Folglich ist

$$p(D(A_p) - E_*) \to J(\ln A),$$

$$p \ln_* D(A_p) \to J(\ln A).$$

Hieraus folgt:

$$\exp(p \ln A_p) \to \exp(\ln A), \qquad A_p^p \to A,$$

$$\exp_*(p \ln_* D(A_p)) \to \exp_*(J(\ln A)), \quad D(A_p)^p = D(A_p^p) \to \exp_*(J(\ln A)).$$

Da D an der Stelle A stetig ist, gilt also:

$$D(A) = \exp_*(J(\ln A)).$$

Nun sei A_0 irgendein Element von $\overline{\mathfrak{G}}$, und $\mathfrak{U}$ diejenige Umgebung von A_0, die bei der Abbildung

$$X' = A_0 X$$

aus der oben erwähnten Umgebung von E entsteht. Dann ist $\ln(A_0^{-1}A)$ sinnvoll und gehört zu $\mathfrak{J}$, und es gilt:

$$D(A) = D(A_0)\, D(A_0^{-1}A) = D(A_0) \exp_*\left(J(\ln(A_0^{-1}A))\right).$$

Jetzt ist ein leichtes die Existenz der Potenzreihenentwicklung nachzuweisen; denn die Funktion

$$D(X) = D(A_0) \cdot \exp_*\left(J(\ln(A_0^{-1}X))\right)$$

setzt sich aus den folgenden Funktionen zusammen:

$$X' = A_0^{-1}X, \qquad X'' = \ln X', \qquad X''' = J(X''), \qquad X^{\mathrm{IV}} = \exp_*(X'''),$$
$$X^{\mathrm{V}} = D(A_0) \cdot X^{\mathrm{IV}}$$

Die erste, dritte und fünfte dieser Funktionen sind linear. Die vierte hat eine Potenzreihe, die stets konvergiert. Und die zweite hat eine Potenzreihe, die für die zugelassenen X (aus $\mathfrak{U}$) konvergiert: ist doch für diese $\ln(A_0^{-1}X)$ sinnvoll, d. h.

$$|X' - E| = |A_0^{-1}X - E| < 1.$$

Damit ist bewiesen, daß $D(X)$ eine in $\mathfrak{U}$ konvergente Potenzreihe hat.

Man beachte, daß hieraus die Analytizität im gewöhnlichen Sinne nicht folgt (vgl. Fußnote [1]) S. 4), weil Real- und Imaginärteile getrennt behandelt werden müssen. Allerdings erfüllen die zwei ersten und die zwei letzten unserer fünf Funktionen die Cauchy-Riemannsche Differentialgleichung, aber für die dritte, J, braucht dies nicht der Fall zu sein.

Wir fassen unsere Resultate über Darstellungen zusammen:

Satz II. *Wenn D eine Darstellung von $\mathfrak{G}$ ist, die in der Umgebung von E eine Schwankung < 1 hat (dies ist sicher der Fall, wenn es nicht überall in $\mathfrak{G}$ unstetig ist), so ist es überall in $\mathfrak{G}$ stetig und kann eindeutig auf $\overline{\mathfrak{G}}$ stetig erweitert werden.*

Die Erweiterung ist eine überall auf $\overline{\mathfrak{G}}$ stetige Darstellung von $\overline{\mathfrak{G}}$, genügt in jedem beschränkt-abgeschlossenen Teile von $\overline{\mathfrak{G}}$ der Lipschitzschen Bedingung und ist überall in $\overline{\mathfrak{G}}$ beliebig oft differenzierbar. Sie kann sogar in einer Umgebung eines jeden A_0 von $\overline{\mathfrak{G}}$ in konvergente Potenzreihen der Real- und Imaginärteile der Elemente des Arguments entwickelt werden.

Anhang.

§ 1.

Wir wollen hier die im Laufe der Arbeit angekündigten Gegenbeispiele anführen, die den Text störend unterbrochen hätten. Es sind die folgenden:

a) Eine unstetige Darstellung mit der Schwankung 2. ($\mathfrak{G} = \overline{\mathfrak{G}}$, für beliebiges m und n.) Vgl. Einl. § 1, IV § 1.

b) Eine (unstetige) Darstellung von $\mathfrak{G}$, die zu keiner Darstellung von $\overline{\mathfrak{G}}$ erweitert werden kann ($m = 3$, $n = 1$). Vgl. IV § 4.

c) Eine Gruppe $\mathfrak{G} = \overline{\mathfrak{G}} = \overline{\mathfrak{G}}(E)$, die durch die $\exp U$, U von J, nicht erschöpft wird (für beliebiges $m \geqq 2$). Vgl. III § 5.

d) Eine lineare Mannigfaltigkeit $\mathfrak{J}$, die mit U, V stets auch ihren Kommutator $UV - VU$ enthält und trotzdem nicht die Infinitesimalgruppe eines $\mathfrak{G}$ ist (für $m \geqq 2$). Vgl. Einl. § 1, II § 4.

Sie lauten so:

a) Wir bezeichnen die (p-dimensionale) Matrix

$$\begin{bmatrix} e^{2\pi\alpha i}, & 0, & 0, & \ldots, & 0 \\ 0 & , 1, & 0, & \ldots, & 0 \\ 0 & , 0, & 1, & \ldots, & 0 \\ \cdot & \cdot & \cdot & \ddots & \cdot \\ 0 & , 0, & 0, & \ldots, & 1 \end{bmatrix} \qquad (\alpha \text{ eine reelle Zahl})$$

mit $E_p(\alpha)$. $\mathfrak{G}$ sei die Menge aller $E_m(\alpha)$ (α reell); dann ist $\mathfrak{G}$ offenbar eine Gruppe, und es ist $\mathfrak{G} = \overline{\mathfrak{G}} = \overline{\mathfrak{G}}(E)$. Wir definieren $D(A)$ in $\mathfrak{G}$ folgendermaßen: Wenn $A = E_m(\alpha)$ ist, so sei $D(A) = E_n(\varphi(\alpha))$. Dies ist jedenfalls eine (eindeutige) Darstellung von $\mathfrak{G}$, wenn

$$\varphi(1) = 0, \qquad \varphi(\alpha + \beta) = \varphi(\alpha) + \varphi(\beta)$$

ist. Wir wählen nun eine unstetige Hamelsche Lösung der Funktionalgleichung

$$\psi(\alpha + \beta) = \psi(\alpha) + \psi(\beta)$$

aus (vgl. Fußnote [2]), S. 5) und setzen

$$\varphi(\alpha) = \psi(\alpha) - \alpha\psi(1).$$

Dann genügt φ unseren Anforderungen, D ist also eine Darstellung.

Ferner ist auch $\varphi(\alpha)$ unstetig und ebenfalls eine Lösung der obigen Funktionalgleichung; nach Hamel (a. a. O.) liegen also die Punkte $\alpha, \varphi(\alpha)$ in der Ebene überall dicht.

Insbesondere liegen solche Punkte beliebig nahe an $0, 0$ und $0, \frac{1}{2}$, die entsprechenden A und $D(A)$ werden also beliebig nahe an $E_m(0) = E$, $E_n(0) = E_*$ bzw. $E_m(0) = E$, $E_n(\frac{1}{2})$ liegen. Wegen

$$\left| E_n\left(\frac{1}{2}\right) - E_* \right| = |e^{\pi i} - 1| = 2$$

ist also die Schwankung von D bei E mindestens 2. Aber sie ist genau 2, da ja stets

$$|E_n(\alpha) - E_n(\beta)| = |e^{2\pi\alpha i} - e^{2\pi\beta i}| \leqq 2$$

ist.

Da die Schwankung $\neq 0$ ist, ist D natürlich unstetig.

b) Hausdorff hat gezeigt[13]): Wenn im dreidimensionalen Raume zwei durch den Nullpunkt gehende Achsen gegeben sind, die miteinander einen Winkel mit transzendentem Kosinus einschließen; und wenn S die 180°-Drehung um die eine und T die 120°-Drehung um die andere Achse ist, so kann keine Gleichung von einer der Formen

$$\left.\begin{aligned} S T^{u_1} S T^{u_2} \ldots S T^{u_\nu} S &= E, \\ S T^{u_1} S T^{u_2} \ldots S T^{u_\nu} &= E, \\ T^{u_1} S T^{u_2} \ldots S T^{u_\nu} S &= E, \\ T^{u_1} S T^{u_2} \ldots S T^{u_\nu} &= E, \end{aligned}\right. \qquad (u_1, u_2, \ldots, u_\nu = 1, 2)$$

bestehen.

Wir setzen

$$A = TST, \qquad B = STSTS.$$

Auch A, B sind Drehungen, und zwar um denselben Winkel, da ja

$$B = S^{-1} A S$$

ist. Und dieser Winkel ist mit π inkommensurabel, da sonst

$$A^\nu = E, \qquad (TST)^\nu = E, \qquad TST^2ST^2 \ldots T^2ST = E$$

wäre. Ferner sieht man leicht ein, daß überhaupt keine Gleichung der Form

$$A^{u_1} B^{v_1} A^{u_2} B^{v_2} \ldots A^{u_\nu} B^{v_\nu} = E$$

(alle u_μ, v_μ bis auf u_1 und v_ν sind $\neq 0$, und wenn $u_1 = v_\nu = 0$ ist, $\nu > 1$) bestehen kann.

Wir nennen die Menge aller $A^{u_1} B^{v_1} \ldots A^{u_\nu} B^{v_\nu}$ $\mathfrak{G}$, dies ist eine Gruppe. Da $\mathfrak{G}$ lauter Drehungen (um durch den Nullpunkt gehende Achsen) enthält, so gilt dasselbe von $\overline{\mathfrak{G}}$; wir wollen zeigen, daß $\overline{\mathfrak{G}}$ alle diese Drehungen enthält. Da $\mathfrak{G}$ Drehungen um mit π inkommensurable Winkel um die Achse von A und die von B enthält, muß $\overline{\mathfrak{G}}$ alle Drehungen um diese zwei Achsen enthalten. Da aber diese Achsen verschieden sind (sonst wäre wegen der Gleichheit der Drehwinkel $A = B$ oder $A = B^{-1}$), kann jede Drehung aus solchen Drehungen zusammengesetzt werden; hieraus folgt die Behauptung.

Nun kann jedes Element von $\mathfrak{G}$ auf eine einzige Art auf die Form

$$X = A^{u_1} B^{v_1} \ldots A^{u_\nu} B^{v_\nu} \quad (\text{alle } u_\mu, v_\mu \text{ bis auf } u_1 \text{ und } v_\nu \text{ sind } \neq 0)$$

gebracht werden. Wir können also

$$D(X) = 2^{u_1 + u_2 + \cdots + u_\nu}$$

---

[13]) Hausdorff, Grundzüge der Mengenlehre, Berlin-Leipzig 1914 (1. Auflage).

setzen ($D(X)$ ist eine Zahl, d. h. eine eindimensionale Matrix). Da offenbar

$$D(E)=1, \quad D(XY)=D(X)\,D(Y)$$

gilt, ist D eine Darstellung von $\mathfrak{G}$.

Es ist

$$D(A)=2, \quad D(B)=1, \quad \text{d. h. } D(A)\neq D(B)$$

Da A und B in $\overline{\mathfrak{G}}$ konjugiert sind (denn S gehört zu $\overline{\mathfrak{G}}$), muß jede Darstellung von $\overline{\mathfrak{G}}$ für A und B denselben Wert haben.

D kann also nicht auf $\overline{\mathfrak{G}}$ erweitert werden.

c) Für $m=1$ gibt es kein solches Beispiel, da dann alle Matrizen vertauschbar sind, also

$$\exp(U_1)\exp(U_2)\ldots\exp(U_j)=\exp(U_1+U_2+\ldots+U_j)$$

gilt. Wir geben ein Beispiel für $m=2$ an, welches folgendermaßen auf alle $m>2$ übertragen werden kann:

Man rändere alle Matrizen mit $m-2$ Zeilen und Spalten, und zwar versehe man die von $\mathfrak{J}$ mit lauter Nullen, die von $\mathfrak{G}$ mit 1 auf der Diagonale und im übrigen mit Nullen.

$\mathfrak{G}$ sei die Menge aller Matrizen mit der Determinante 1 ($m=2$). Man sieht leicht ein, daß $\mathfrak{G}=\overline{\mathfrak{G}}=\overline{\mathfrak{G}}(E)$ ist und daß $\mathfrak{J}$ aus allen Matrizen mit der Spur 0 besteht.

Die Matrix

$$A=\begin{bmatrix}-1 & 1\\ 0 & -1\end{bmatrix}$$

gehört zu $\mathfrak{G}$. Wir wollen zeigen, daß sie von allen $\exp U$, U von $\mathfrak{J}$, verschieden ist.

U hat bekanntlich eine der Formen

$$S^{-1}\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}S \quad \text{und} \quad S^{-1}\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}S.$$

Im ersten Falle ist

$$\exp(U)=S^{-1}\exp\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}S=S^{-1}\begin{bmatrix}e^{\lambda_1}, & 0\\ 0, & e^{\lambda_2}\end{bmatrix}S.$$

Dies ist jedenfalls $\neq A$, da A bekanntlich nicht auf die Diagonalform transformiert werden kann. Im zweiten Falle stellen wir zunächst fest, daß $\mathfrak{J}$ die Spur 2λ hat, also muß $\lambda=0$ sein. Dann ist aber

$$\exp(U)=S^{-1}\exp\begin{bmatrix}0, & 1\\ 0, & 0\end{bmatrix}S=S^{-1}\begin{bmatrix}1, & 1\\ 0, & 1\end{bmatrix}S.$$

Auch dies ist $\neq A$; denn es hat die Multiplikatoren $1, 1$, während die von A $-1, -1$ sind.

d) Wir geben das Beispiel für $m = 2$ an, auf $m > 2$ überträgt es sich durch dieselben Ränderungen wie in c).

τ sei irgendeine irrationale Zahl, wir setzen

$$\overline{U} = \begin{bmatrix} i, & 0 \\ 0, & \tau i \end{bmatrix}.$$

$\mathfrak{J}$ sei die Menge aller $\alpha \overline{U}$ (α reell). $\mathfrak{J}$ ist eine lineare Mannigfaltigkeit, alle Elemente von $\mathfrak{J}$ sind vertauschbar, so daß ihr Kommutator $= 0$ ist und zu $\mathfrak{J}$ gehört.

Wir werden zeigen, daß trotzdem $\mathfrak{J}$ nicht die Infinitesimalgruppe eines $\mathfrak{G}$ sein kann.

Denn wäre dies der Fall, so enthielte $\overline{\mathfrak{G}}$ alle

$$\exp(\alpha \overline{U}) = \exp \begin{bmatrix} \alpha i, & 0 \\ 0, & \tau \alpha i \end{bmatrix} = \begin{bmatrix} e^{\alpha i}, & 0 \\ 0, & e^{\tau \alpha i} \end{bmatrix}.$$

Da τ irrational ist, kann man nach dem bekannten Satze von Kronecker zu jedem Paare reeller Zahlen η, ϑ ein solches α finden, daß sich α von η und $\tau \alpha$ von ϑ mod(1) beliebig wenig unterscheidet. Die Matrizen

$$\begin{bmatrix} e^{\eta i}, & 0 \\ 0, & e^{\vartheta i} \end{bmatrix} \quad (\eta, \vartheta \text{ reell})$$

sind also Häufungspunkte von $\overline{\mathfrak{G}}$ mit nichtverschwindender Determinante, so daß sie zu $\overline{\mathfrak{G}}$ gehören müssen. Da $\mathfrak{J}$ auch Infinitesimalgruppe von $\overline{\mathfrak{G}}$ ist, müßte es die Matrix

$$\lim_{n \to 0} n \left(\begin{bmatrix} e^{\frac{1}{n} i}, & 0 \\ 0, & 1 \end{bmatrix} - E \right) = \begin{bmatrix} i, & 0 \\ 0, & 0 \end{bmatrix}$$

enthalten, was offenbar nicht der Fall ist.

§ 2.

Eine Infinitesimalgruppe $\mathfrak{J}$ im gewöhnlichen Sinne der Theorie der kontinuierlichen Gruppen, d. h. eine lineare Mannigfaltigkeit, die mit zwei Matrizen U, V stets auch ihren Kommutator $UV - VU$ enthält, wollen wir eine formale Infinitesimalgruppe nennen, und ein $\mathfrak{J}$, das Infinitesimalgruppe einer Gruppe $\mathfrak{G}$ ist, eine echte Infinitesimalgruppe.

Das Beispiel d) zeigte, daß nicht jede formale Infinitesimalgruppe eine echte sein muß. Immerhin können auch formalen Infinitesimalgruppen gewisse gruppenähnliche Gebilde zugeordnet werden; wie das geschehen kann, wollen wir hier in großen Zügen skizzieren.

$\mathfrak{U}$ sei irgendeine Umgebung von E. Wir nennen eine Teilmenge $\mathfrak{G}^*$ von $\mathfrak{U}$ ein Gruppenelement (in $\mathfrak{U}$), wenn es die folgenden Eigenschaften hat:

a') E gehört zu $\mathfrak{G}^*$.

b') Wenn A zu $\mathfrak{G}^*$ gehört, so ist $\det A \neq 0$, und A^{-1} gehört zu $\mathfrak{G}^*$, falls es zu $\mathfrak{U}$ gehört.

c') Wenn A, B zu $\mathfrak{G}^*$ gehören, so gehört auch AB zu $\mathfrak{G}^*$, falls es zu $\mathfrak{U}$ gehört.

Für solche Gruppenelemente kann man natürlich die Infinitesimalgruppe analog zu II § 1 definieren, und fast alle unsere Sätze lassen sich auch auf diesen Fall übertragen[14]).

Man kann nun zeigen, daß jede formale Infinitesimalgruppe $\mathfrak{J}$ gleichzeitig Infinitesimalgruppe eines geeignet gewählten Gruppenelementes ist; dabei kann $\mathfrak{U}$ stets als die Menge aller X mit

$$|X - E| < \frac{1}{2} \ln 2$$

gewählt werden. Dies Gruppenelement ist übrigens mit seiner Hülle identisch, und zwar ist es die Menge aller $\exp(U)$ (U von $\mathfrak{J}$), die in $\mathfrak{U}$ liegen.

---

[14]) Diese Begriffsbildung ist mit dem von O. Schreier untersuchten Begriff des Gruppenkeimes eng verwandt. Vgl. Hamb. Abh. 4 (1926), S. 15–32 und S. 233–244.

(Eingegangen am 2. Februar 1927.)