schließlich aus Bromsilber, das sich nur in sehr verschiedener Verteilung befindet. Wie beim Jodsilber ist diese Umwandlung des belichteten Bromsilbers von der Silberkeimbildung als solcher unabhängig. Man kann sogar die Silberkeime durch Bromwasser völlig zerstören und erhält doch die Entwicklung des „Reifungsbildes“ durch Ammoniak. Auch bei der „Entwicklung“ des Jodsilbers mit Joddampf wird ja jedes „chemische“ Bild, d. h. jede Spur von Silber, durch das Jod zerstört.

Von besonderem Interesse sind auch die Wirkungen der *Röntgenstrahlen* und verwandter Energiearten auf die Bromsilberplatte. Bestrahlt man bestimmte Sorten von Trockenplatten mit Röntgenstrahlen und legt sie darauf ins Tageslicht, so erscheint nach einiger Zeit das Röntgenbild ohne Anwendung irgendwelcher chemischer Agentien in rötlicher Farbe auf bläulichem Grunde. (*Lüppo-Cramer,* Die Röntgenographie [Halle 1909], farbiges Titelbild, daselbst auch die weitere Literatur.) Übergießt man ein solches durch Licht „entwickeltes“ Röntgenbild mit Chromsäuremischung oder dergl., so bleichen die von den X-Strahlen veränderten Bildstellen aus, während die nur belichteten fast unverändert bleiben. Diese Reaktionen führten zunächst auf die Theorie einer durch die Röntgenstrahlen erfolgten Zersplitterung des Bromsilbers, die sich bei weiteren Untersuchungen auch gut bewährte. Für die Eigenart der Röntgenstrahlenwirkung ist auch der folgende zuerst von *M. P. Villard* angegebene Versuch recht überzeugend. Man bestrahlt eine gewöhnliche hochempfindliche Trockenplatte in normaler Weise mit Röntgenstrahlen und legt dann die ganze Platte offen etwa eine Minute lang ins Tageslicht. Beim darauffolgenden Behandeln mit einem gewöhnlichen Entwickler entsteht ein deutliches positives Röntgenbild.

Würde nun die offensichtliche Eigenart des Röntgenstrahlenbildes darauf beruhen, daß die Zersplitterung des Bromsilbers hier sehr viel intensiver erfolgt als durch Licht, so sollte man annehmen, daß auch die ausschließlich auf der Zersplitterung des Bromsilbers beruhende „Entwicklung“ mit Ammoniak beim Röntgenstrahlenbilde leichter und durchgreifender erfolge, vor allem kürzere Belichtungen erfordere als beim gewöhnlichen Lichtbilde. Dies ist nun aber keineswegs der Fall. Auch erheblich über die Normalexposition hinaus mit Röntgenstrahlen behandelte Platten ergeben mit Ammoniak kein deutliches Reifungsbild, sondern es ist, genau wie beim Lichtbilde, eine außerordentlich starke Überexposition erforderlich, um ein direktes Reifungsbild mit Ammoniak zu erhalten. Legt man aber die mit X-Strahlen vorbestrahlte Platte darauf unter einer Sensitometerskala etwa eine halbe Minute lang ins Tageslicht und setzt sie dann in der beschriebenen Weise den Ammoniakdämpfen aus, so entsteht ein deutliches Reifungsbild des Röntgenstrahleneindruckes.

Die im Vergleich zu der starken Nachbelichtung geringe Menge an Röntgenstrahlenenergie hat also doch eine erhebliche Steigerung der Löslichkeit des Bromsilbers zur indirekten Folge. Zu dem Verständnis dieser eigenartigen Wirkung der Röntgenstrahlen kann man auf folgende Weise gelangen: Auch der X-Strahl wirkt nicht grundsätzlich anders als das Licht auf das Bromsilber: die primäre Wirkung ist immer Bromabspaltung und Silberkeimbildung. Aber die erwähnten und andere Tatsachen, besonders das Verhalten des Röntgenbildes gegen die Ammoniakentwicklung, lassen sich erklären, wenn man die bei der Eigenart der äußerst kurzwelligen und durchdringenden Röntgenstrahlen wohl erlaubte Hypothese aufstellt, daß der Röntgenstrahl im Bromsilber relativ zahlreichere Zersetzungszentren bildet als der Lichtstrahl, daß also das Keimsilber im Bromsilber feiner verteilt ist. Die Gesamtmenge an Silber kann in beiden Fällen gleich sein, aber die topographische Verteilung im Korn eine verschiedene. Wird nun aber die mit Röntgenstrahlen behandelte Platte dem gewöhnlichen Lichte ausgesetzt, so wirken die zahlreichen ursprünglichen Keime wie in anderen Fällen beschleunigend auf die weitere photochemische Zersetzung des Bromsilbers an diesen Punkten. Es wird daher an den zuerst von den X-Strahlen veränderten Bildstellen eine raschere Bromabspaltung und daher auch eine durchgreifendere Zersplitterung des Kornes erfolgen als an den nur vom Lichte getroffenen Teilen der Schicht.

Die vorgetragene Hypothese macht die Annahme einer primären Zerstäubung durch X-Strahlen ebenso entbehrlich wie bei der Lichtwirkung. In beiden Fällen führt erst die Halogenexplosion sekundär zur Änderung des sogen. Dispersitätsgrades des Bromsilbers. Anstatt der Hypothese von einer besonders starken direkten „mechanischen“ Wirkung der Röntgenstrahlen auf das Bromsilber wird aber die Annahme nötig, daß die X-Strahlen das erste Keimsilber in höher disperser Form abscheiden als der Lichtstrahl. Für die Berechtigung dieser Hypothese sprechen aber alle Tatsachen, die wohl früher zur Annahme einer direkten Zerstäubung führten, vor allem die starke optische Sensibilisierung durch die Röntgenstrahlen, die auf kaum etwas anderem beruhen kann als darauf, daß das bei der Röntgenbestrahlung gebildete Silber einen höheren Dispersitätsgrad besitzt, d. h. feiner verteilt ist als das bei der gewöhnlichen Belichtung entstehende.

Die physikalischen Voraussetzungen der Wahrscheinlichkeitsrechnung.

Von Hans Reichenbach, Berlin-Lichterfelde.

I. *Von der Anwendung der Wahrscheinlichkeitsgesetze auf die Dinge der Wirklichkeit.*

Man kann beobachten, daß die Wahrscheinlichkeitsgesetze zwei verschiedene Gruppen von Forschern beschäftigen. Einmal die Mathematiker. Sie entwickeln aus den einfachen Grundgesetzen der Theorie komplizierte Rechenformeln, stellen Beziehungen auf, die verschlungene Probleme zu lösen gestatten, führen auch neue Begriffe ein, wie Dispersion, mittleres Fehlerquadrat, zusammengesetzte Wahrscheinlichkeit usw. Ihnen treten die Statistiker aller Wissenschaften, der Physik, der Psychologie, der Soziologie usw. gegenüber; sie übernehmen gern den verzweigten Apparat der Mathematiker, nicht aber um ihn zu vervollständigen, sondern um ihn auf praktische Gegenstände anzuwenden, um Methoden aus ihm zu gewinnen, mit denen sich bestimmte empirische Sachverhalte darstellen lassen. Diese Zweiteilung der Arbeitsweisen entspricht einem tiefgehenden sachlichen Unterschied; es ist derselbe, der die reine mathematische Forschung von allen ihren Anwendungen trennt. Kein Zweifel, die Wahrscheinlichkeitsgesetze stellen ein geschlossenes mathematisches System dar, wie die Sätze der Infinitesimalrechnung oder wie die Sätze der Geometrie, und die strenge Sicherheit dieser Gebiete muß den Wahrscheinlichkeitssätzen ebenso zuerkannt werden, soweit sie geschlossene Relationen, Begriffsketten aus Verflechtungen der Elementarbegriffe, darstellen. Es sei an das Bernoullische Theorem erinnert, das die Häufigkeiten und die Dispersion einfacher Wiederholungsreihen berechnet, und das, mathematisch genommen, nichts anderes ist als eine Auszählung von Kombinationen. Niemand hat je an der Richtigkeit dieser Kombinationslehre gezweifelt. Um so mehr aber haben sich Zweifel erhoben, wenn es sich um die Anwendung der Wahrscheinlichkeitsgesetze auf Dinge der Wirklichkeit handelte; und die Statistiker der einzelnen Wissenschaften konnten sich niemals auf die mathematische Strenge der theoretischen Wahrscheinlichkeit berufen, weil es problematisch blieb, ob die wirklichen Dinge sich den berechneten Relationen unterordneten. Die Lage ist hier ähnlich wie bei der Geometrie: daß die geometrischen Sätze in sich richtig sind, wird von niemandem bezweifelt; aber ob sie die *wirklichen Dinge* beschreiben, ob der Raum, in dem wir die physikalischen Dinge messen, dreidimensional und euklidisch ist, darüber läßt sich mathematisch nichts aussagen, und erst die Methoden der Physik und der Philosophie können darüber die Entscheidung treffen.

Die Geometrie hat den Vorzug, daß sie ein entwickeltes Axiomensystem besitzt, und wir können heute die Frage nach der Geltung ihrer Sätze ersetzen durch die Frage nach der Geltung ihrer Axiome. Wenn die Axiome von dem wirklichen Raum befolgt werden, so muß dasselbe für alle geometrischen Sätze gelten; für die Untersuchung ist es aber viel einfacher, allein die Geltung der Axiome zu problematisieren. Nun gibt es allerdings für die Wahrscheinlichkeitsrechnung auch Axiomensysteme. Aber diese sind vollständig nur als Grundlage der rein arithmetischen Beziehungen der Wahrscheinlichkeitsrechnung. In ihrer Anwendung jedoch wollen die Wahrscheinlichkeitsgesetze wirkliche *Vorgänge* beschreiben, über *zeitliche Abläufe Bestimmtes aussagen,* und es muß deshalb noch eine durch das Auftreten des Zeitbegriffs gekennzeichnete Axiomgruppe geben, die die Anwendung der Wahrscheinlichkeitsgesetze auf wirkliche Vorgänge behandelt, und die wir deshalb als *Axiome der Anwendbarkeit* bezeichnen wollen. Im Gegensatz zu den mathematischen Axiomen lassen sie sich auch als *physikalische Axiome* bezeichnen, wenn man Physik im allgemeinsten Sinne als Wissenschaft von raum-zeitlichen Vorgängen auffaßt.

Die Aufstellung dieser Axiomgruppe ist von mir in einer Arbeit[1]) durchgeführt worden, auf die ich für eine ausführliche Begründung der hier dargestellten Gedanken verweisen muß. An dieser Stelle sollen die Resultate der Untersuchung mitgeteilt und der Weg ihrer Ableitung in seinen wesentlichen Zügen gezeigt werden. Da es sich bei den Wahrscheinlichkeitsgesetzen stets um Näherungsgesetze handelt, derart, daß mit wachsender Zahl der Wiederholungen eine engere Annäherung an die geforderte Verteilung stattfindet, so müssen auch die Axiome der Anwendbarkeit als Gesetze über ein Näherungsverhalten formuliert werden. Wir werden finden, daß sich diese Axiome auf ein einziges reduzieren. Es kann allerdings vorläufig nicht behauptet werden, daß nicht in gewissen Hypothesen der Physik, z. B. *Boltzmanns* Ergodenhypothese als Grundlage der Molekularstatistik, noch andere Voraussetzungen enthalten sind, und in diesem Sinne darf die Untersuchung noch nicht als abgeschlossen gelten. Auch beschränkt sich die Untersuchung auf physikalische Probleme und kann deshalb ein Urteil über die Anwendung der Wahrscheinlichkeitsrechnung in der Psychologie und in der Soziologie nicht geben. Aber es wird sich zeigen lassen, daß das aufgedeckte Axiom eine über den Rahmen der engeren Wahrscheinlichkeitsrechnung hinausgehende, philosophische Bedeutung besitzt, und daß seine Geltung im engsten Zusammenhang mit dem physikalischen Erkenntnisbegriff steht. Diese Frage wird in einem der nächsten Hefte dieser Zeitschrift behandelt werden, während sich die gegenwärtige Darstellung auf die Aufstellung des Axioms beschränken wird.

II. Das Axiom der Anwendbarkeit der Wahrscheinlichkeitssätze: die Hypothese der Wahrscheinlichkeitsfunktion.

Ein einfaches Beispiel der physikalischen Realisierung von Wahrscheinlichkeitsgesetzen bildet das Würfelspiel. Dort werden die sechs möglichen Lagen des Würfels als „mögliche Fälle“ bezeichnet, und jede Lage gilt als „gleich wahr-

[1]) *Reichenbach,* Der Begriff der Wahrscheinlichkeit für die mathematische Darstellung der Wirklichkeit. Ztschrft. f. Philos. u. philos. Kritik Bd. *161,* 1917.

scheinlich". Diese Klassifikation der möglichen Fälle in eine Anzahl gleich wahrscheinlicher Fälle ist charakteristisch für jede Wahrscheinlichkeitsberechnung. Unter gleicher Wahrscheinlichkeit versteht man, daß bei einer Wiederholung des Vorgangs die gleich wahrscheinlichen Fälle gleich oft realisiert werden, also z. B. jede Würfelseite gleich oft darankommt. Das Problem ist dabei: woher nimmt man das Recht, von bestimmten Fällen, z. B. dem Auftreten der Würfelseiten, zu sagen, daß sie gleich wahrscheinlich sind?

Man hat versucht, die Gleich-Wahrscheinlichkeit so zu definieren, daß sie eine Aussage über die Wiederholung des Vorgangs nicht enthält, und glaubte so das „Problem der großen Zahlen" von dem Wahrscheinlichkeitsproblem trennen zu können. Man definiert dann Gleich-Wahrscheinlichkeit als eine bestimmte physikalische Struktur, z. B. die räumliche Symmetrie des Würfels, und verzichtet auf die Behauptung, daß bei der Wiederholung jede Seite annähernd gleich oft darankommt. Durch derartige Definitionen kommt man natürlich um das Problem nicht herum. Die so definierte Gleich-Wahrscheinlichkeit bildet dann allerdings kein Problem mehr, man kann ihr Vorhandensein physikalisch konstatieren, die gleiche Größe der Würfelseiten und die Mittellage des Schwerpunkts kann man ausmessen. Aber das Merkwürdige bleibt, daß diesen geometrisch-physikalischen Verhältnissen gerade die gleiche Häufigkeit in der Wiederholung des Vorgangs entspricht; dies ist das Problem, um das es sich grundsätzlich handelt, und dessen Geltung in jeder Anwendung der Statistik vorausgesetzt wird. Durch das Stattfinden einer solchen Häufigkeitsbeziehung läßt sich auch erst jenes Erwartungsgefühl rechtfertigen, mit dem wir den Wahrscheinlichkeitsbegriff gewöhnlich verknüpfen, und das uns z. B. das Eintreffen zweier gleicher Würfe hintereinander als unwahrscheinlich empfinden läßt. Wenn es ausgemacht ist, daß eine derartige Kombination seltener vorkommt als andere Kombinationen, so ist das Spannungsgefühl, mit dem wir das Wahrscheinliche erwarten, das Unwahrscheinliche dagegen zurücksetzen, psychologisch gerechtfertigt; auf geometrisch-physikalische Verhältnisse aber läßt es sich nicht basieren. Noch weniger aber läßt sich dieses Spannungsgefühl zur Definition der Wahrscheinlichkeit verwenden, etwa indem man als gleichwahrscheinlich solche Verhältnisse definiert, die die gleiche „freie Erwartungsbildung" in uns hervorrufen. Derartige Definitionen, die die Wirkung auf den Zuschauer zum Ausgangspunkt nehmen, verführen dazu, in der Wahrscheinlichkeitssetzung eine lediglich subjektive Vermutung zu sehen, deren Inhalt von dem jeweiligen Stand unserer *subjektiven Kenntnisse* abhängt. Die Auffassung übersieht, daß es tatsächlich *objektive Sachverhalte* gibt, die durch Wahrscheinlichkeitsgesetze erschöpfend beschrieben werden, z. B. die Gesetzlichkeit des Würfelns. Es muß Aufgabe der Psychologie bleiben, zu erklären, wie aus der Kenntnis solcher objektiven Sachverhalte bestimmte Erwartungsgefühle entstehen; mit der Gesetzmäßigkeit der Wahrscheinlichkeit hat dies nichts zu tun.

Wir definieren deshalb als gleichwahrscheinlich solche Fälle, die bei der Wiederholung angenähert gleich oft realisiert werden, und wir müssen untersuchen, welche physikalischen Voraussetzungen für das Stattfinden eines solchen Phänomens gemacht werden müssen.

Wir wählen als Beispiel eines Wahrscheinlichkeitsmechanismus das Roulette. *Poincaré*[1]) hat dieses Spiel behandelt und auf eine sehr einfache Voraussetzung zurückgeführt. Der rotierende Zeiger wird seinen Kreis vielmal durchlaufen, bis er stehen bleibt; wir können seinen Weg durch einen Winkel messen, wenn wir nach der ersten Umdrehung über 360° hinauszählen und so fortlaufend bis zu jener Gradzahl Ω zählen, bei der der Zeiger stehen bleibt. Wäre Ω um ein Stückchen $\Delta\Omega$ größer, das gerade der Größe eines farbigen Sektors entspricht, so würde der Zeiger auf der anderen Farbe stehen geblieben sein; denken wir uns diesen Weg nochmals um $\Delta\Omega$ vergrößert, so hält der Zeiger wieder über der ersten Farbe, usw. Lassen wir den Zeiger mehrere Male spielen, so wird der Weg Ω jedesmal anders sein, weil die Kraft, mit der der Zeiger fortgeschnellt wird,

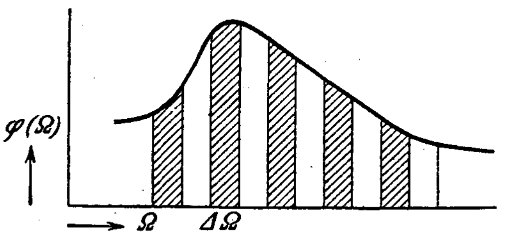

Fig. 1.
Zur Zurückführung der Gleich-Wahrscheinlichkeit auf die Stetigkeit einer Kurve.

niemals genau die gleiche ist; aber ob rot oder schwarz getroffen wird, hängt nur davon ab, in welchem Intervall $\Delta\Omega$ der Zeiger zur Ruhe kommt. Wir denken uns nun während einer größeren Reihe von Versuchen für jedes einzelne Intervall $\Delta\Omega$ gezählt, wie oft der Zeiger gerade in diesem Intervall anhält, und diese Zahl h durch die Gesamtzahl N der Versuche dividiert, so daß wir die relative Häufigkeit h/N des Intervalls erhalten. Dies werde graphisch aufgetragen, wie es Fig. 1 darstellt. Darin ist als Abszisse Ω aufgetragen, und die Intervalle $\Delta\Omega$ sind abgeteilt; die relative Anzahl der Treffer h/N ist für jedes Intervall durch den schmalen rechteckähnlichen Flächenstreifen dargestellt, der sich darüber erhebt. Man sieht aus der unregelmäßigen Form der Kurve, daß der Zeiger keinesfalls jeden Wert Ω gleich oft erreicht, im Gegenteil bevorzugt er ein gewisses Gebiet in der Mitte und nimmt selten

[1]) *Poincaré*, Calcul des Probabilités, Paris 1912, Gauthier-Villars, p. 149.

kleine oder sehr große Werte Ω an. Nun ist die Anzahl der Fälle, daß der Zeiger auf einem schwarzen Sektor anhält, durch die Summe der schraffierten Flächenstreifen gegeben, und die Anzahl der Treffer „rot“ durch die Summe der nichtschraffierten Streifen. Jeder Sektor des Roulettespiels entspricht dabei bereits einer Summe von Flächenstreifen der Figur, und die Farbe „rot“ oder „schwarz“ entspricht einer noch größeren Summe. Es sind aber zwei nebeneinanderliegende Streifen nahezu gleich groß, und bei der Summenbildung wird jedem kleinen schraffierten Flächenstreifen ein kleiner nichtschraffierter, jedem großen schraffierten Streifen ein großer nicht-schraffierter entsprechen, so daß die insgesamt schraffierte Fläche nahezu gleich wird der nicht-schraffierten. Das wird um so genauer erfüllt sein, je kleiner die Teilung $\Delta\Omega$ ist, je größer also die Anzahl der Intervalle ist, und es läßt sich leicht zeigen, daß in der Grenze für unendlich kleine $\Delta\Omega$ die beiden Flächen genau gleich werden. Voraussetzung ist dabei nur, daß die Kurve der Fig. 1 *stetig* verläuft, also keine anomalen Sprünge macht; ihre Form kann ganz beliebig sein, sie kann auf- und absteigen und beliebig gekrümmt sein. Allerdings muß die Summe der Rechtecke immer endlich bleiben, und die Kurve muß deshalb an beiden Enden asymptotisch zur Abszissenachse verlaufen, d. h. *sehr große* und *sehr kleine* Werte von Ω müssen äußerst selten vorkommen[1]). Sind diese Voraussetzungen erfüllt, so folgt, daß ebenso oft die rote wie die schwarze Farbe getroffen wird; damit ist die Gleich-Wahrscheinlichkeit der beiden Farben zurückgeführt auf die Existenz einer solchen Kurve.

Was bedeutet nun diese Kurve? Wir müssen uns klarmachen, daß ihre Existenz keineswegs bewiesen war. Wir hatten nur gesagt, daß wir die Trefferzahlen zählen und nach dem genannten Verfahren eintragen wollten. Genau genommen, erhalten wir dabei überhaupt keine Kurve. Wir können zunächst nur über jedem $\Delta\Omega$ ein Rechteck zeichnen, dessen Flächeninhalt gleich der relativen Trefferzahl $\frac{h}{N}$ dieses Intervalls ist, und die Oberkanten der Rechtecke werden dann einen Treppenweg bilden. Vergrößern wir die Anzahl N der Versuche, so wird der Treppenweg ausgeglichener werden; wir können dann die Intervalle $\Delta\Omega$ kleiner denken, und dadurch wird der Treppenweg einer Kurve ähnlicher werden. Allzu klein dürfen wir die $\Delta\Omega$ bei diesem graphischen Verfahren zunächst nicht wählen. Ist z. B. die Anzahl der $\Delta\Omega$ größer als die Anzahl N der Versuche, so kann unmöglich in jedem $\Delta\Omega$ ein Treffer liegen, an vielen Stellen also würde das Rechteck $= 0$ zu zeichnen sein, und der Treppenweg würde höchst unregelmäßig werden. Erst wenn die Anzahl der Versuche größer genommen wird, wird wieder ein regelmäßiger Treppenweg entstehen, der nun einer stetigen Kurve noch ähnlicher ist als der frühere. Zu einer stetigen Kurve selbst aber können wir mit einer endlichen Anzahl N von Versuchen — und nur solche endlichen Anzahlen stehen uns zur Verfügung — niemals kommen. Wenn wir trotzdem für die Umdrehungen des Zeigers die Existenz einer stetigen Häufigkeitskurve annehmen, so bedeutet dies die *Hypothese, daß bei dem geschilderten graphischen Verfahren mit wachsendem N eine Annäherung an eine solche stetige Kurve mit asymptotischen Enden zustandekommt.* Derartige Kurven $\varphi(\Omega)$ nennt man Wahrscheinlichkeitsfunktionen, weil die Wahrscheinlichkeit W, daß Ω in dem Intervall von Ω_1 bis Ω_2 liegt, also die relative Häufigkeit $\frac{h}{N}$ für beliebige Intervalle, durch den Ausdruck

$$\left(\frac{h}{N}\right)_{\lim N=\infty} = W = \int_{\Omega_1}^{\Omega_2} \varphi(\Omega)\, d\,\Omega$$

gegeben ist. (Aus der Summation der Rechtecke entsteht für $\lim N = \infty$ das Integral.) Und wir können sagen, daß sich die Gleichwahrscheinlichkeit der beiden Farben im Roulettespiel zurückführen läßt auf die *Hypothese,* daß für den Umdrehungswinkel des Zeigers *eine Wahrscheinlichkeitsfunktion existiert.*

Diese Erkenntnis bedeutet einen wesentlichen Fortschritt für das Wahrscheinlichkeitsproblem. Vorher standen wir vor der Frage, die Gleichwahrscheinlichkeit der roten und schwarzen Sektoren zu erklären; dabei erschien diese Gleichwahrscheinlichkeit als eine mysteriöse Eigenschaft der Farbenstreifen, und es schien gar kein Grund vorhanden, warum man über die Streifen eine derartig weitgehende Aussage machen sollte. Ja, man hat sogar versucht, aus dem Vorhandensein *keines* Grundes ein philosophisches Prinzip zu machen, indem man sagte, es sei kein Grund vorhanden, einen Streifen zu bevorzugen, und darum müßten die Streifen gleiche Wahrscheinlichkeit besitzen. Dieses „Prinzip des mangelnden Grundes“ übersieht, daß man ebenso keinen Grund hat, die Streifen gleich wahrscheinlich zu nennen, und daß man also auch das Gegenteil folgern könnte; Schlüsse auf *keinen* Grund zu basieren, ist eben immer sehr mißlich. Die Hypothese der Wahrscheinlichkeitsfunktion enthebt uns mit einem Schlage dieser Schwierigkeit. Denn sie nimmt die verlangte Eigenschaft von den Streifen weg und überträgt sie auf den rotierenden Zeiger; über die Natur des Rotationsvorgangs sagt sie etwas aus, und die Streifen übernehmen dabei nur die Aufgabe, diese eigentümliche Natur des Rotationsvorgangs *sichtbar* zu machen, in besonderer Zuspitzung zu veranschaulichen. Jetzt haben wir allerdings Grund genug, die Streifen gleich *wahrscheinlich* zu nennen, des-

[1]) Der Beweis läßt sich bereits führen, wenn die Kurve integrierbar ist und das Integral von $-\infty$ bis $+\infty$ einen endlichen Wert hat. Die Forderung der Stetigkeit geht also etwas zu weit; aber sie drückt am deutlichsten die verlangte Eigenschaft aus und soll deshalb im folgenden immer benutzt werden. Genauer müßte man von der Stetigkeit des Integrals sprechen.

halb nämlich, weil sie gleich *groß* sind; wären sie verschieden, so wären auch die Intervalle $\Delta\Omega$ nicht gleich, und die Häufigkeit der Farben nicht durch 1 : 1, sondern durch ein anderes Verhältnis gegeben. Die Streifen bewirken nur eine eigentümliche Zerlegung und Zuordnung der verschiedenen Drehungswinkel des Zeigers; der Zahlwert der dabei auftretenden Mengenverhältnisse ist durch die Größe der Streifen bestimmt — und das Prinzip des mangelnden Grundes ist mit dieser Verschiebung des Problems verschwunden.

Damit ist allerdings das Problem noch nicht gelöst. Wir werden die neue Hypothese erst zu rechtfertigen haben. Sie unterscheidet sich durchaus von anderen physikalischen Hypothesen; sie ist auch nicht mit den üblichen physikalischen Methoden zu kontrollieren, weil sie nicht durch *Messungen* bestätigt werden kann. Sie besagt eine Gesetzmäßigkeit der Natur, die im *Zählen* von Größen zum Ausdruck kommt, und auch hier geht sie viel weiter als alle Erfahrung, weil sie einen Grenzwert für unendlich viele Beobachtungen aufstellt. Aber wir können schon jetzt als ihren Vorzug erwähnen, daß sie die Form einer Stetigkeitsvoraussetzung hat und keine quantitativen Verhältnisse vorschreibt. Wir brauchen nicht mehr anzunehmen, daß endlichen Flächenstücken gleiche Wahrscheinlichkeit zukommt; unsere Hypothese lautet *nicht,* daß alle Werte für den Drehungswinkel Ω gleich wahrscheinlich sind, sondern nur, daß unendlich benachbarte Werte gleichwahrscheinlich sind. Darin, daß eine Annahme über den bestimmten Wert der Wahrscheinlichkeitsfunktion $\varphi(\Omega)$ nicht gemacht zu werden braucht, sondern nur ihre *Stetigkeit* vorausgesetzt werden muß, um die *gleich*wahrscheinlichen Fälle zu erklären, die die Grundlage der Wahrscheinlichkeitsberechnung bilden, liegt die Überlegenheit dieser Hypothese; das wird für die philosophische Seite des Problems wichtig werden. Zunächst soll jedoch gezeigt werden, daß dieselbe Voraussetzung auch für andere Probleme hinreichend ist.

III. Die Stetigkeit der Wahrscheinlichkeitsfunktion und die physikalischen Grundlagen einiger Glücksspiele.

Von jeher haben die Glücksspiele als Idealfall der Wahrscheinlichkeitsrechnung gegolten; es ist üblich geworden, an ihnen als Beispielen die Gesetze der Wahrscheinlichkeit zu erläutern, und nirgends scheinen die Voraussetzungen jener eigentümlichen Kombinationslehre, wie sie die Wahrscheinlichkeitsrechnung darstellt, die einzelnen gleich wahrscheinlichen und die Möglichkeiten erschöpfenden Fälle, ihre beliebige Kombinationsfähigkeit und die Regelmäßigkeit ihres Eintreffens, so klar und deutlich gegeben wie in diesen anerkannten Tummelplätzen des Zufalls. Sie erscheinen geradezu als eine Symbolisierung jener Rechenregeln, die das Gebäude der Wahrscheinlichkeitsrechnung ausmachen, während ihre physikalische Natur höchst uninteressant und unwichtig bleibt. Das ändert sich erst, wenn man, wie wir es für das Roulettespiel getan haben, nach den physikalischen Voraussetzungen sucht, die diese doch immer nur empirischen Vorgänge zu Musterbildern mathematischer Operationen machen; dabei gelingt es dann, jenes Axiom der Anwendbarkeit der Wahrscheinlichkeitssätze aufzudecken, das wir von vornherein als unser Ziel aufstellten und das wir jetzt als Hypothese der Wahrscheinlichkeitsfunktion formulieren konnten.

Es ist bei der großen Ähnlichkeit aller Glücksspiele leicht zu zeigen, daß diese Hypothese auch bei den anderen Spielen zur Anwendung kommt. Nehmen wir z. B. das Spiel mit der geworfenen Münze. Dabei gibt es zwei mögliche und gleichwahrscheinliche Fälle, je nachdem, ob *Kopf oder Wappen* oben liegt. Aber wieder liegt die wesentliche Hypothese nicht in der Münze, sondern in der Natur des *Bewegungsvorgangs.* Die Zeit, die von dem Abwerfen der Münze bis zu ihrem Niederfall verstreicht, und die für jeden Wurf verschieden ist, ergibt diesmal die Größe Ω, die wir als Abszisse der Figur auftragen. Die Einteilung in Intervalle erfolgt durch die Rotation der Münze; je nachdem, ob die Fallzeit etwas länger oder kürzer ist, kommt die Münze in dem durch Kopf oder Wappen charakterisierten Intervall zu Boden. Nun erfolgt allerdings die Rotation der Münze nicht mit gleichförmiger Geschwindigkeit, und dadurch werden die Intervalle ungleich groß. Aber wir dürfen annehmen, daß die Geschwindigkeit sich *stetig* ändert, und so werden benachbarte Intervalle nahezu gleich groß; die Figur sieht dann etwas anders aus als Figur 1, weil die Teilung $\Delta\Omega$ nach rechts immer größer wird, aber wegen der angenäherten Gleichheit benachbarter $\Delta\Omega$ läßt sich der Schluß auf Gleichheit der schraffierten und der nicht-schraffierten Fläche ebenso durchführen. Die beiden Seiten der Münze übernehmen also, den Sektoren des Roulettespiels entsprechend, nur *die Klassifizierung der Fallzeiten,* ihre Einteilung in zwei Scharen, und als charakteristische Hypothese bleibt die Existenz einer Wahrscheinlichkeitsfunktion für die Fallzeit. (Daß wir auch die Stetigkeit der Rotationsgeschwindigkeit voraussetzen müssen, ist keine für die Wahrscheinlichkeitsrechnung charakteristische Hypothese. Derartige Annahmen macht die Physik stets über ihre Größen, sie bedeuten, daß sich physikalische Größen nicht sprungweise ändern können, sondern alle Zwischenwerte durchlaufen. Wir dürfen also von dieser Voraussetzung Gebrauch machen, ohne damit ein neues Element in das Problem hineinzutragen. Es ist vielmehr zu erwarten, daß die allgemeinen Voraussetzungen, die die Physik jederzeit macht, auch für das Gebiet der Wahrscheinlichkeitsgesetze angewandt werden; nicht dies zu bestätigen ist unsere Aufgabe, sondern die *speziellen* Voraussetzungen zu finden, die für die Geltung der Wahrscheinlichkeitsgesetze notwendig sind.)

Für das *Würfelspiel* gilt die gleiche Betrachtung. Die Größe, für deren Wiederholung eine

Wahrscheinlichkeitsfunktion angesetzt wird, ist wieder die Fallzeit, und die Rotation des Würfels teilt stetig wachsende Intervalle ab, unter denen benachbarte nahezu gleich sind. Der Unterschied ist allein der, daß entsprechend den 6 Seiten eine Klassifizierung in 6 Scharen von Intervallen eintritt; wir müssen also in der Figur 6 aufeinanderfolgende Intervalle verschieden schraffieren und dann mit der Schraffierung von neuem beginnen; aber man kann für die 6 Flächen, die von den gleichschraffierten Intervallen eingenommen werden, ganz entsprechend beweisen, daß sie in der Grenze für unendlich kleine Intervalle einander gleich werden. Daß die Genauigkeit mit kleineren Intervallen wächst, ist eine praktisch anerkannte Erscheinung; denn man nimmt allgemein an, daß die Verteilung um so regelmäßiger ausfällt, je rascher der Würfel rotiert. Auch für dieses Glücksspiel läßt sich also die Hypothese der Wahrscheinlichkeitsfunktion als hinreichende Voraussetzung betrachten.

Wesentlich bleibt dabei immer, daß die Hypothese nur Stetigkeit der Wahrscheinlichkeitsfunktion verlangt, über ihre spezielle Form jedoch nichts voraussetzt.

Nun gibt es Fälle, für die man „aus dem Gefühl heraus" eine Wahrscheinlichkeitsfunktion der Form

$$f(x) = \text{const.}$$

ansetzt; es ist wichtig, daß man diese Form auch aus der bloßen Stetigkeit einer Wahrscheinlichkeitsfunktion, aber für eine andere Größe, ableiten kann. Um dies an einem Beispiel zu verdeutlichen, greifen wir wieder auf das Roulettespiel zurück. Wir hatten dort den Umdrehungswinkel Ω über 360° hinaus gemessen und die Existenz einer stetigen Wahrscheinlichkeitsfunktion $\varphi(\Omega)$ angenommen (Fig. 1). Nun entspricht jeder rote oder schwarze Sektor bereits einer Summe von Intervallen $\Delta\Omega$, und wenn man nicht mehr 2 Scharen abteilt, sondern, ähnlich wie beim Würfelspiel, so viel Scharen nimmt, wie Sektoren da sind, so ergibt sich für *jeden einzelnen* Sektor bereits die gleiche Summe schraffierter Flächenstreifen und damit die gleiche Wahrscheinlichkeit. Wir brauchen nicht mehr bis zu der größeren *Summe* der roten und der schwarzen Sektoren zu gehen. Würde man die Sektoren ungleich groß machen, so würde dem größeren Sektor eine in diesem Verhältnis größere schraffierte Fläche und damit eine entsprechend größere Wahrscheinlichkeit zukommen. D. h. die Wahrscheinlichkeit ist proportional dem Winkel des Sektors, und wenn man den Winkel jetzt *nicht* über 360° hinaus zählt und mit ϑ bezeichnet, so bedeutet dies, daß

$$\int \varphi(\vartheta)\, d\,\vartheta = k \cdot \vartheta$$

ist, wo k eine Konstante darstellt. Daraus folgt

$$\varphi(\vartheta) = k = \text{const.}$$

Wir erhalten also für die Größe ϑ eine Wahrscheinlichkeitsfunktion von der speziellen Form $\varphi(\vartheta) = \text{const.}$, wenn wir für die Größe Ω eine Wahrscheinlichkeitsfunktion $\varphi(\Omega)$ von beliebiger Form annehmen. So ist unter Umständen die Stetigkeitshypothese hinreichende Voraussetzung für das Auftreten einer ganz speziellen Wahrscheinlichkeitsform.

IV. Die Ausdehnung der Hypothese der Wahrscheinlichkeitsfunktion auf die Kombination mehrerer — voneinander unabhängiger — Ereignisse.

Die Wahrscheinlichkeitsrechnung begnügt sich nicht damit, die Gleichwahrscheinlichkeit der einzelnen Fälle zu statuieren. Ihr ganzer Aufbau entsteht vielmehr erst dadurch, daß sie die Kombination solcher Fälle vornimmt und die Wahrscheinlichkeit beliebiger Kombinationen berechnet, wenn die Wahrscheinlichkeit der Einzelfälle gegeben ist. Dabei benutzt sie das „Gesetz der zusammengesetzten Wahrscheinlichkeit", welches besagt, daß die Wahrscheinlichkeit einer Kombination gleich dem Produkt der Einzelwahrscheinlichkeiten ist, wenn die betrachteten Fälle voneinander unabhängig sind (Multiplikationstheorem). Es entsteht die Frage, welche physikalische Hypothese wir für das Zutreffen dieses Rechenverfahrens machen müssen. Ehe wir dazu schreiten, müssen wir jedoch genauer erklären, was unabhängige Vorgänge sind; denn dieser Begriff ist wesentlich für die Geltung des Gesetzes.

Wir beschreiben einen physikalischen Vorgang dadurch, daß wir die ihn charakterisierenden Bestimmungsstücke zueinander in Beziehung setzen; wir beobachten etwa an einem fallenden Stein eine bestimmte Geschwindigkeit und vergleichen sie mit der Zeit, während der der Stein bereits gefallen ist, und die dabei gefundene Relation $v = g \,.\, t$ stellt eine Beschreibung des Fallvorganges dar. Wir sagen, daß die Geschwindigkeit eine Funktion der Fallzeit ist, oder einfacher, daß sie von der Fallzeit abhängig ist; und wir können das Erkenntnisverfahren der Physik geradezu als ein Suchen nach abhängigen Größen und der Art ihrer Abhängigkeit bezeichnen. Nun ist die Zahl der Abhängigkeitsrelationen sehr groß, so groß, daß es sogar ganz unmöglich ist, sie jemals zu erschöpfen; aber unter ihnen zeichnen sich einzelne dadurch aus, daß sie das Geschehen *vorherrschend charakterisieren,* und man gelangt zu Erkenntnissen bereits dadurch, daß man sich auf diese Relationen beschränkt. So besteht in unserem Beispiel des fallenden Körpers auch eine Relation zwischen der Geschwindigkeit und der Luftdichte, weil diese die Reibung beeinflußt; aber man kann feststellen, daß großen Änderungen der Luftdichte nur kleine Änderungen der Fallgeschwindigkeit entsprechen, und darum darf man diese Relation neben der ersten vernachlässigen. Bei einer genaueren Theorie des Falles wird man diese Einflüsse allerdings hinzuziehen; aber sie bedeuten doch nur eine Abhängigkeit von geringerem Grade und bleiben dadurch von der ersten Abhängigkeit unterschieden. Der Abhängigkeitsgrad kann jedoch noch weiter sinken. So ist

z. B. die Fallgeschwindigkeit auch abhängig von der Stellung des Mondes zur Erde, weil diese das Gravitationsfeld beeinflußt; aber hier entsprechen großen Änderungen des Mondortes bereits so überaus kleine Änderungen der Fallgeschwindigkeit, daß man praktisch von einer Abhängigkeit nicht mehr spricht. Prinzipiell muß allerdings festgehalten werden, daß es unabhängige physikalische Größen nicht gibt. Denn es gibt keine abgeschlossenen Systeme, jedes System steht durch seine Oberfläche mit anderen Systemen in Beziehung, diese wieder mit anderen usf., so daß alle Systeme schließlich in Beziehung zueinander stehen. Aber es lassen sich sehr niedrige, sogar beliebig niedrige Abhängigkeitsgrade aufzeigen, und es hat sich in der Physik eingebürgert, derartige gering verbundene Größen als unabhängig zu bezeichnen. Wir halten als Definition der Unabhängigkeit, richtiger des geringen Abhängigkeitsgrades, fest, daß großen Änderungen der einen Größe verschwindend kleine Änderungen der anderen Größe entsprechen.

Entsprechend können wir auch die Unabhängigkeit von *Vorgängen* definieren. Denken wir uns etwa 2 Körper nebeneinander zu Boden fallen. Jeden der beiden Fallvorgänge können wir dadurch charakterisieren, daß wir die Koordinaten des Schwerpunkts des Körpers als Funktion der Zeit darstellen, etwa indem wir angeben, zu dieser bestimmten Zeit besitzt der Körper diese bestimmte Höhe usw. In diesen Beziehungsgleichungen werden außer der Zeit noch andere Größen auftreten, die bestimmte Werte haben und die Funktion beeinflussen, z. B. die Anfangshöhe, der Luftwiderstand usw. Die Koordinaten selbst sind ebenfalls physikalische Größen, sie werden durch die Gleichung als unabhängig von den anderen Größen dargestellt. Was aber in diesen Gleichungen *nicht* vorkommt, ist eine Abhängigkeit zwischen den Koordinaten des einen Körpers und denen des anderen; der eine Körper mag eine beliebige Lage haben, die Koordinaten des anderen werden dadurch nicht beeinflußt. Das ist nun, streng genommen, wieder nicht richtig. So wird der eine Körper einen Luftstrom hervorrufen, der gleichzeitig seitlich saugt und den anderen Körper von seiner Bahn ablenken wird, so daß die Koordinaten dieses Körpers eine andere zeitliche Änderung zeigen, je nachdem, ob der erste Körper in der Nähe vorbeieilt oder nicht. Aber die Abhängigkeit dieser Größen wird von so geringem Grade sein, daß sie als Unabhängigkeit betrachtet werden kann. Wir werden danach zwei Vorgänge als unabhängig bezeichnen, wenn die den einen Vorgang charakterisierenden Bestimmungsstücke, z. B. die Koordinaten, unabhängig sind von den Bestimmungsstücken des anderen, d. h. nur verschwindende Änderungen erfahren, wenn die Bestimmungsstücke des anderen ganz verschiedene Werte annehmen.

Nach diesen grundsätzlichen Definitionen schreiten wir zu dem Problem der zusammengesetzten Wahrscheinlichkeit. Denken wir uns etwa zwei Geldstücke nebeneinander zu Boden geworfen; wir fragen nach der Wahrscheinlichkeit, daß bei beiden das Wappen nach oben kommt. Nach unserer Definition von Wahrscheinlichkeit bedeutet dies die Frage, wie oft in einer Reihe von Wiederholungen diese Kombination im Verhältnis zu den anderen möglichen Kombinationen vorkommt. Wir wissen, sofern unsere Hypothese der Wahrscheinlichkeitsfunktion gilt, daß für jede Münze Kopf und Wappen gleich häufig auftreten; können wir nun daraus ableiten, daß die Kombination Wappen-Wappen nach der bekannten Formel gerade in ¼ der Fälle eintrifft? Daß wir dies *nicht* können, erhellt aus folgender Überlegung. Wir denken uns die erste Münze wiederholt geworfen und die Resultate so ausfallend, daß sie dem Gesetz einer Wahrscheinlichkeitsfunktion für diese Münze entsprechen. Würde nun die zweite Münze, gleichfalls geworfen, stets denselben Wurf ergeben wie die erste, so würden ihre Resultate ebensogut einer Wahrscheinlichkeitsfunktion entsprechen; aber für die Kombinationen ergäbe sich das befremdliche Gesetz, daß nur entweder Kopf-Kopf oder Wappen-Wappen auftritt. Wir hätten also eine Verteilung, bei der jeder einzelne Vorgang durch eine Wahrscheinlichkeitsfunktion geregelt ist, die aber nicht dem Multiplikationstheorem der Wahrscheinlichkeiten entspricht. Damit ist bewiesen, daß die Hypothese der Wahrscheinlichkeitsfunktion für das Multiplikationstheorem nicht die hinreichende Voraussetzung ist.

Seien x und y die Werte der Fallzeit für die erste und zweite Münze. Wir machen nun folgende Hypothese: Es existiert eine Wahrscheinlichkeitsfunktion

$$\varphi(x,\, y),$$

welche jeder Kombination von Werten x, y eine Wahrscheinlichkeit genau so zuordnet, wir wir es für die einfache Wahrscheinlichkeitsfunktion definiert haben, d. h. die Wahrscheinlichkeit, daß x zwischen den Grenzen a und b und gleichzeitig y zwischen den Grenzen c und d liegt, sei gegeben durch das Doppelintegral

$$W = \int_a^b \int_c^d \varphi(x,\, y)\; dx\, dy$$

Mit dieser Voraussetzung können wir das Multiplikationstheorem der Wahrscheinlichkeiten ableiten. Wir denken uns die Werte x auf der X-Achse, die y auf der Y-Achse eines Koordinatenkreuzes aufgetragen (Fig. 2), jede Achse in Intervalle[1]) geteilt und diesen entsprechend ein Netz von Rechtecken gezogen. Über jedem Rechteck denken wir uns die zugehörige Wahrscheinlichkeit, daß also x gerade in diesem Intervall Δx und gleichzeitig y gerade in dem zuge-

[1]) Der Einfachheit halber zeichnen wir die Intervalle gleich groß, d. h. wir denken uns die Rotationsgeschwindigkeit der Münze konstant. Dasselbe Resultat läßt sich natürlich auch für die allgemeinere Voraussetzung der Stetigkeit dieser Geschwindigkeit ableiten.

hörigen Δy liegt, durch eine prismatische Säule dargestellt. Die Kuppen dieser Säulen werden in einer gekrümmten Fläche liegen; wir nehmen noch die Voraussetzung hinzu, daß diese Fläche zwar von beliebiger Krümmung, aber stetig ist, und daß sie sich für entfernte Gebiete der x-y-Ebene asymptotisch nähert. Die Streifen der Intervalle in der x-y-Ebene sind wieder abwechselnd schraffiert. Wir betrachten nun die stark umrahmte Schnittfigur von vier Streifen; darin entspricht das weiße Rechteck der Kombination Wappen-Wappen, das doppelt schraffierte Rechteck der Kombination Kopf-Kopf, und die beiden einfach schraffierten Rechtecke den beiden gemischten Kombinationen. Da $\varphi(x, y)$ durch eine stetige Fläche dargestellt ist, so werden die prismatischen Säulen über diesen vier Rechtecken bei genügender Kleinheit der Intervalle nahezu gleich, und wenn auch die Säulen an verschiedenen Stellen der Ebene ganz verschieden hoch sind, so werden doch die 4 Summen, die bei der Addition der Säulen über gleich schraffierten

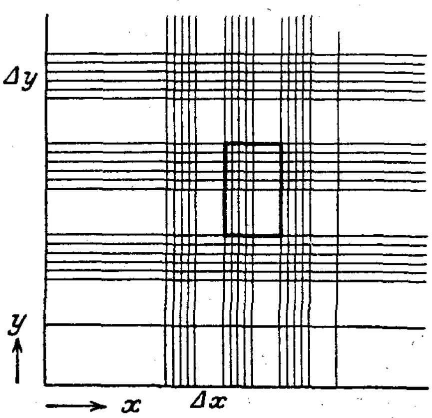

Fig. 2.
Zur Wahrscheinlichkeit des gleichzeitigen Eintretens zweier voneinander unabhängiger Ereignisse.

Rechtecken entstehen, nahezu gleich; in der Grenze für unendlich kleine Intervalle werden sie genau gleich groß. Endlich sind diese Summen, weil wegen des asymptotischen Verlaufs der Fläche der Raum zwischen ihr und der x-y-Ebene endlich ist. Die Wahrscheinlichkeiten der Kombinationen entsprechen diesen Summen; für die Kombination Wappen-Wappen ergibt sich $\frac{1}{4}$, für die gemischten Kombinationen $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$, und für Kopf-Kopf entsteht ebenfalls $\frac{1}{4}$.

Wir können jedoch aus unseren Voraussetzungen noch ein weiteres Resultat ableiten. Wir müssen verlangen, daß neben der Existenz von $\varphi(x, y)$ unsere alte Hypothese gilt, daß also *außerdem* für jede Größe, x sowohl wie y, eine Wahrscheinlichkeitsfunktion $f(x)$ bzw. $f(y)$ existiert, welche die Verteilung dieser Größen *gleichzeitig* bestimmt. Wir können dann für $\varphi(x, y)$ die spezielle Form $\varphi(x, y) = f(x) \cdot f(y)$ ableiten, die das Multiplikationstheorem der Wahrscheinlichkeitsfunktionen vorschreibt, wenn wir noch annehmen, daß die beiden Größen voneinander *unabhängig* sind[1]).

Zum Beweise denken wir uns zunächst $x = x_0 =$ konst. festgehalten, so daß die Funktion die Form $\varphi(x_0, y)$ annimmt. Dabei entsteht eine Verteilung von Kombinationen x_0, y, bei der x überall dasselbe ist und allein y variiert, so daß dies allein für die y eine Verteilung bedeutet. Die Anzahl der Werte y für jedes beliebige Intervall ist aber gleich der Anzahl der Werte (x_0, y). Sind nun die Vorgänge unabhängig, so werden (nach Definition) die Werte y nicht beeinflußt durch die Werte x; es muß also dieselbe Verteilung y entstehen, gleichgültig, ob x variiert oder $x = x_0$ konstant bleibt, oder der Vorgang x überhaupt nicht stattfindet. Darum muß die durch $\varphi(x_0, y)$ dargestellte Verteilung der y in jedem Punkt der durch $f(y)$ gegebenen Verteilung entsprechen, ihr also proportional sein; der dabei auftretende Proportionalitätsfaktor k kann noch von x_0, nicht aber von y abhängen. Also gilt, als charakteristisch für unabhängige Vorgänge,

$$\varphi(x_0 \cdot y) = k(x_0)\, f(y)$$

Da dies für jedes beliebige $x = x_0$ gilt (das ist wieder die Forderung der Unabhängigkeit), muß dies eine Identität sein, und wir schreiben

$$\varphi(x \cdot y) \equiv k(x)\, f(y).$$

Dieselbe Überlegung gilt, wenn $y = y_0$ konstant bleibt und x variiert, so daß ebenfalls gilt:

$$\varphi(x \cdot y) \equiv f(x)\, k(y).$$

Daraus folgt:

$$k(x) = f(x)\,; \qquad k(y) = f(y)\,;$$

also:

$$\varphi(x \cdot y) = f(x)\, f(y).$$

Wesentlich für diesen Beweis ist die Forderung der Unabhängigkeit der beiden Vorgänge. Denken wir uns z. B. einen Körper auf rauhem Boden beliebig hin- und hergestoßen, so daß seine Lage durch eine Wahrscheinlichkeitsfunktion bestimmt ist, und mit ihm durch eine federnde Kupplung verbunden einen zweiten Körper, der auf dem rauhen Boden sprungweise gleitet. Dann wird auch für den zweiten Körper eine Wahrscheinlichkeitsfunktion existieren, und es wird sogar für die Kombination ihrer Lagen eine Wahrscheinlichkeitsfunktion $\varphi(x, y)$ existieren. Aber diese wird nicht die spezielle Form $f(x) \cdot f(y)$ haben, sondern bestimmte Kombinationen, für die der räumliche Abstand von x bis y der mittleren Länge der Kupplung entspricht, werden bevorzugt sein. In diesem Fall ist eben die Bedingung der Unabhängigkeit nicht erfüllt. Es ist nur natürlich, daß diese als Forderung in unsere Ableitung eingeht, denn sie wird auch von der geltenden Wahrscheinlichkeitsrechnung für das Multiplikationstheorem vorausgesetzt.

Das Multiplikationstheorem zwingt uns also, die Hypothese der Wahrscheinlichkeitsfunktion zu erweitern und auf die Kombination mehrerer Argumente auszudehnen. Es leuchtet ein, daß wir nicht bei der Anzahl zwei stehen bleiben dürfen, da auch die Kombination mehrerer Ereignisse geregelt werden muß. Wieder aber

[1]) Da diese Ableitung in der genannten Arbeit von mir nicht scharf genug formuliert worden ist, sei sie hier ausführlich gegeben.

braucht nur die Stetigkeit der Funktion, nicht irgendein spezieller Wert, gefordert zu werden. Die spezielle Produktform läßt sich für den besonderen Fall der Unabhängigkeit aus der allgemeinen Form ableiten. Indem wir die Erweiterung in den ursprünglichen Begriff aufnehmen, dürfen wir jetzt sagen: *die Existenz der Wahrscheinlichkeitsfunktion für eine oder mehrere Veränderliche stellt die hinreichende Voraussetzung dar für die Anwendbarkeit der Wahrscheinlichkeitsgesetze.*

Es ist zu beachten, daß wir aus der Stetigkeit von $\varphi(x, y)$ bereits das Multiplikationstheorem für Wahrscheinlichkeiten ableiten konnten, die sich nach dem Schema der geworfenen Münze darstellen lassen, daß wir dazu also die spezielle Produktform von $\varphi(x, y)$, d. i. das Multiplikationstheorem der Wahrscheinlichkeits*funktionen,* nicht brauchten. Es scheint deshalb, als ob für das erste Multiplikationstheorem die Unabhängigkeit der Vorgänge nicht vorausgesetzt zu werden brauchte, die für das zweite verlangt wird. Das ist jedoch ein Irrtum. Mit der Stetigkeit von $\varphi(x, y)$ wird bereits vorausgesetzt, daß eine Abhängigkeit zwischen den Intervallen der einen Größe und denen der anderen nicht existiert, d. h. daß nicht ein Intervall „Wappen“ der einen Reihe wieder ein Intervall „Wappen“ der anderen Reihe zur Folge hat. Wäre dies der Fall, so würde bei Verkleinerung der Intervalle eine Annäherung der prismatischen Säulen niemals stattfinden. Durch die Stetigkeit der Funktion ist dies ausgeschlossen. Dies ist die Unabhängigkeit, die allein für die Multiplikation einfacher Wahrscheinlichkeiten zu gelten hat. Im übrigen dürften die Größen allerdings abhängig sein, und aus dem geschilderten Mechanismus der federnd gekuppelten Körper ließe sich bei geeigneter Teilung in Intervalle ebenfalls das Schema für die einfache Multiplikation ableiten. Erst das Multiplikationstheorem der Funktionen verlangt Unabhängigkeit der ganzen Vorgänge.

V. Die Wahrscheinlichkeitsfunktion in der Theorie der Messungsfehler (Gaußsche Fehlerfunktion).

In der Theorie der Beobachtungsfehler spielen die Wahrscheinlichkeitsfunktionen eine bevorzugte Rolle. Diese Theorie hat die Aufgabe, aus zahlreichen voneinander verschiedenen Meßresultaten denjenigen Wert zu berechnen, der der gesuchten Größe am besten entspricht; dazu muß sie eine Annahme über die Verteilung der Messungsfehler machen, und obgleich sie über diese Fehler nur sehr allgemeine Vermutungen aufstellen kann, muß sie auf ganz spezielle Formen des Verteilungsgesetzes schließen, wenn sie überhaupt zu Resultaten kommen will. Dabei muß beachtet werden, daß der Fehler der einzelnen Messung auf dem Zusammenwirken sehr vieler Fehlerquellen (Elementarfehler) beruht. Die bedeutendste Lösung des Problems stellt die Gaußsche Fehlerfunktion dar, in welcher die Wahrscheinlichkeit eines Fehlers in exponentieller Form von seiner Größe abhängig gemacht wird[1]).

Es gibt verschiedene Voraussetzungen, unter denen man diese spezielle Form ableiten kann. Für unsere Betrachtungen wichtig ist die Tatsache, daß man das Gaußsche Gesetz auf folgende drei Bedingungen bauen kann:

1. Die Häufigkeit jedes Elementarfehlers ist durch eine Wahrscheinlichkeitsfunktion von beliebiger Form bestimmt.
2. Diese Funktionen setzen sich nach dem Multiplikationstheorem zusammen.
3. Es müssen sehr viele, voneinander unabhängige Fehler gleicher Größenordnung zusammenwirken.

Man erkennt, daß die erste Bedingung mit unserer ersten Hypothese identisch ist, und die zweite nach der gegebenen Ableitung auf die erweiterte Hypothese für die Wahrscheinlichkeitsfunktion von Kombinationen zurückgeführt werden kann. Die dritte Bedingung stellt im Gegensatz dazu keine prinzipielle Voraussetzung, sondern eine Annahme dar, die nur unter gewissen Umständen erfüllt ist; ob diese Umstände gegeben sind, läßt sich empirisch konstatieren, und nur für diese Fälle gilt dann das Gaußsche Gesetz. Dieses Gesetz ist eben eine Spezialform, die keineswegs ein allgemeines Prinzip darstellt und nicht in den Vordergrund der Betrachtung geschoben werden darf. Seine praktische Bedeutung hat das Gesetz daher, daß die modernen Meßinstrumente in ihrem komplizierten Bau zahlreiche Fehlerquellen zusammenführen und so die dritte Bedingung erfüllen; daher die Vielheit seiner Verwendung.

Ähnlich liegt es mit der sogenannten „Hypothese des arithmetischen Mittels“, nach welcher der Mittelwert der Messungen mit größter Wahrscheinlichkeit der gesuchten Größe entspricht. Dieses Gesetz ist immer dann erfüllt, wenn das Gaußsche Fehlergesetz gilt; aus diesem folgt es durch eine sehr einfache mathematische Operation. Aber damit ist auch gesagt, daß es an dieselben Bedingungen geknüpft ist wie das Gaußsche Exponentialgesetz; ist die dritte Bedingung nicht erfüllt, so ist auch das Verfahren des arithmetischen Mittels nicht anwendbar. Danach ist es nicht zweckmäßig, von einer Hypothese des arithmetischen Mittels zu sprechen. Hypothetisch nennt man besser nur die weiter zurückliegenden Voraussetzungen.

Daraus ergibt sich, daß auch die Fehlertheorie keine neuen Hypothesen für die Anwendbarkeit von Wahrscheinlichkeitsgesetzen enthält. Sie führt uns vielmehr auf diejenigen Voraussetzungen, die wir aus einfachen Beispielen von Wahrscheinlichkeitsverteilungen bereits entwickelten.

---

[1]) Sie lautet, wenn x die Fehlergröße bezeichnet:

$$\varphi(x) = \frac{h}{\sqrt{\pi}} \cdot e^{-h^2(x-a)^2}$$

a ist der „systematische Fehler“, h das „Präzisionsmaß“.

Wir wollen damit unsere Durchmusterung physikalischer Wahrscheinlichkeitsprobleme abschließen. Wir sind zu dem Resultat gelangt, daß alle diese Probleme eine eigentümliche Hypothese einschließen, die wir als Prinzip der Wahrscheinlichkeitsfunktion formulieren konnten und die das Gesetz der Wahrscheinlichkeit darstellt. Es wird unsere nächste Aufgabe sein, die Berechtigung dieser Hypothese zu prüfen; die Kritik der Hypothese wird die Antwort auf das philosophische Problem der Wahrscheinlichkeit darstellen.

Besprechungen.

Brockmann-Jerosch, H., Baumgrenze und Klimacharakter. Zürich, Rascher & Cie., 1919. Pflanzengeographische Kommission der Schweizerischen Naturforschenden Gesellschaft; Beiträge zur geobotanischen Landesaufnahme 6. Mit einer farbigen Tafel und 18 Textfiguren. Preis Frcs. 8,—.

Das Problem der Baumgrenze hat die Botanik seit langem beschäftigt, ohne daß es bisher gelungen wäre, es befriedigend zu lösen. *Brockmann-Jerosch* untersucht nun, ausgehend von den Verhältnissen in den Schweizer Alpen, die alpine Baumgrenze und setzt sie in Vergleich mit der polaren.

Die Baumgrenze verläuft in den nördlichen Alpenketten zwischen 1725 m ü. M. (Säntis) und 1880 m (Guttannen), erhebt sich in den zentralen Alpen auf 2000 m (Churwalden, Braggio, Leukerbad, Leysin) bis 2400 m (Zermatt, Simplon), um in den südlichen Alpenketten auf 1950 m (Pasturo, Sotto-Ceneri) abzusinken. Dieses auffällige starke Ansteigen der alpinen Baumgrenze in den Zentralalpen läßt sich nun weder durch Temperatur-, noch durch Niederschlags-, noch durch die Bodenverhältnisse erklären. Auch mit dem Verlauf der Schneegrenze hängt die Baumgrenze nicht zusammen, denn man kann z. B. am Mt. Rainier im Kaskadengebirge feststellen, daß die bei etwa 2000 m liegende Schneegrenze um etwa 250 m hinter der Baumgrenze zurückbleibt. Auch der Wind kommt als alleinwirkender Faktor nicht in Frage. Nur die Gesamtwirkung der klimatischen Faktoren, der *Klimacharakter*, kann eine ausreichende Erklärung geben. Massenerhebung, Klimacharakter, Wald- und Baumgrenze stehen in engstem Zusammenhange: Größere Massenerhebung, wie in den Zentralalpen, bedingt einen kontinentaleren Klimacharakter und dieser wieder eine Förderung der Ausdehnung des Baumwuchses. Die Wirkungsweise des Klimacharakters läßt sich dahin zusammenfassen, daß bei gleicher, verhältnismäßig niedriger Durchschnittstemperatur eine Funktion des Pflanzenkörpers, z. B. Assimilation, im kontinentalen Klima viel eher möglich ist als im ozeanischen. Der kontinentale Klimacharakter mit seinen größeren Temperaturschwankungen, reicherer Sonnenwirkung und günstigerer Verteilung der Niederschläge wirkt fördernd auf die Entwicklung des Baumwuchses, dagegen wirkt das ozeanische Klima mit seiner geringeren Belichtung infolge stärkerer Bewölkung und seinen gleichmäßig über das ganze Jahr verteilten Niederschlägen bei nicht ausreichend hohen Temperaturen ungünstig.

Das mehr ozeanische Klima der nördlichen und südlichen Alpenketten drückt die alpine Baumgrenze gegen die kontinentalen Zentralalpen herab.

Vergleicht man hiermit nun den Verlauf der polaren Baumgrenze, so zeigt sich ein ausgesprochener Parallelismus. Nicht der Meereswind wirkt baumfeindlich, sondern das ozeanische Küsten- oder Inselklima. Der ozeanische Klimacharakter drückt in den Gebirgen Englands (Penninen) die klimatische Baumgrenze auf etwa 600 m herab (gegenwärtig durch den Einfluß des Menschen liegt die Baumgrenze bei 4—500 m), macht die noch gar nicht sehr nördlich (bei 59 °, 60 °, 62 ° n. Br.) liegenden Orkney-, Shetlandsinseln und Faröer trotz ihres milden Klimas baumlos. Auch auf Island zeigt sich der gleiche Einfluß: der wärmere, mildere, stärker ozeanische Westen ist baumlos, der rauhere, mehr kontinentale Norden, Osten und das Innere zeigen stellenweise bis 550 m ü. M. Baumwuchs. Auch in Fennoskandien, auf Kola, im europäischen Rußland und in Sibirien läßt sich die entsprechende, die polare Baumgrenze stark äquatorwärts verschiebende Wirkung des ozeanischen Klimacharakters verfolgen. Der ewig gefrorene Boden ist auf den Verlauf der polaren Baumgrenze ganz *ohne* Einfluß. Wenn auch im europäischen Nordrußland die polare Baumgrenze mit der Südgrenze des ewig gefrorenen Bodens auf große Strecken zusammenfällt, so ist hier die Ursache des Zurückdrängens der Baumgrenze nach Süden das ozeanische Klima. Denn im arktischen Sibirien finden wir unter dem Einfluß des kontinentalen Klimacharakters ein außerordentlich weites Vorstoßen der polaren Baumgrenze nach Norden (bis 70 ° 10 ′ n. Br. in Nordsibirien bei Tolstoj-Noß, und sogar bis 72 ° 40 ′ n. Br. in Ostsibirien an der unteren Chatanga), während gleichzeitig unter denselben Längengraden die Südgrenze des ewig gefrorenen Bodens im oberen Amurgebiet bis 47 ° n. Br. nach Süden vorstößt. Üppige Waldgebiete, die ein Areal bedecken, das größer ist als das ganze europäische Rußland, stehen in Sibirien auf ewig gefrorenem Boden.

Verfolgt man den Verlauf der polaren Baumgrenze in Eurasien, so zeigt sich, daß die weitesten Vorstöße nach Norden zusammenfallen mit den großen Landkomplexen Sibiriens mit ihrem kontinentalen Klimacharakter und daß überall dort, wo sich ozeanischer Klimacharakter bemerkbar macht, die polare Baumgrenze weit nach Süden ausbiegt. So senkt sich die Baumgrenze unter dem Einfluß des Pazifischen Ozeans scharf nach Süden und biegt sogar bei 70 ° östl. L. jäh nach SW um und erreicht bei 60 ° n. Br. das Nordende von Kamtschatka. Bis 50 ° n. Br. senkt sich die Baumgrenze auf den Aleuten. Entsprechend liegen die Verhältnisse im arktischen Nordamerika, wo die polare Baumgrenze unter dem Einfluß des Seeklimas an der Nordspitze von Neufundland bei 51 ° 53 ′ n. Br. ihren südlichsten Punkt erreicht.

Wesentlich anders liegen dagegen die Verhältnisse bei der südpolaren Baumgrenze: alle Kontinente laufen nach Süden spitz aus oder bleiben schon in sehr niedrigen Breiten zurück (Afrika und Australien schon bei 40 ° s. Br.). Trotz seiner sturmgepeitschten Küsten ist der am weitesten polwärts reichende südamerikanische Kontinent bis an die Südostspitze, bis auf die letzte Insel, mit immergrünem Wald bestanden. Dagegen sind die Falklandsinseln baumlos. Die polare Baumgrenze verläuft also zwischen diesen Inseln und dem Festland. Die Inseln zwischen Südamerika und Afrika (Tristan da Cunha und Diego Alvarez) liegen zwar noch im baumhaften Gebiete, können aber von der polaren Baumgrenze nicht weit entfernt sein. Dagegen liegen die Inseln südlich von Afrika (Bouvet-, Prinz-Edward- und Marion-, Croset-Inseln) jenseits der