# VORLESUNGEN ÜBER DIFFERENTIAL- UND INTEGRALRECHNUNG

VON


R. COURANT

O. PROFESSOR AN DER UNIVERSITÄT
GÖTTINGEN


ERSTER BAND

FUNKTIONEN EINER VERÄNDERLICHEN

ZWEITE, VERBESSERTE AUFLAGE

MIT 126 TEXTFIGUREN


BERLIN
VERLAG VON JULIUS SPRINGER
1930

ISBN-13: 978-3-642-98739-7 e-ISBN-13: 978-3-642-99554-5
DOI: 10.1007/978-3-642-99554-5



Softcover reprint of the hardcover 2nd edition 1927

MEINER FRAU
GEWIDMET

# Vorwort zur ersten Auflage.

Gewiß ist die mathematische Literatur nicht arm an guten Werken über Differential- und Integralrechnung; und doch wird der Anfänger nur schwer ein Buch finden, das ihm einen geraden Weg in das lebendige Wesen der Wissenschaft öffnet und ihm verständnisvolle Bewegungsfreiheit gegenüber den Anwendungen gibt. Der Anfänger will weder durch Weitschweifigkeit und Inhaltslosigkeit ermüdet werden, noch kann er jene Pedanterie ertragen, welche keinen Unterschied zwischen Wesentlichem und Unwesentlichem kennt und — der axiomatischen Systematik zuliebe — vor die eigentlichen Triebkräfte der Wissenschaft, vor ihren gegenständlichen Kern, einen undurchsichtigen Schleier zieht.

Sicherlich ist es leichter, Mängel zu sehen und zu fühlen, als sie abzustellen. Ich bin weit entfernt von der Vorstellung, dem Anfänger das ideale Lehrbuch darbringen zu können. Trotzdem glaube ich nicht, daß die Herausgabe meiner Vorlesungen überflüssig ist; sie weichen in der Anordnung und der Auswahl des Stoffes, in der Tendenz und vielleicht auch in der Darstellungsform erheblich von der landläufigen Literatur ab.

Am meisten wird auffallen, daß der Bruch mit der überlebten Tradition vollzogen ist, Differentialrechnung und Integralrechnung voneinander zu trennen. Diese sachlich wie didaktisch unbegründete Trennung, ein Produkt von historischen Zufälligkeiten, verhindert die Klarlegung des Kernpunktes: des Zusammenhanges zwischen bestimmtem Integral, unbestimmtem Integral und Differentialquotienten. Im mündlichen Vorlesungsbetrieb hat sich seit dem Vorgange von Felix Klein und anderen schon mehr und mehr die gemeinsame Behandlung durchgesetzt. Hier nun wird versucht, dieser auch einen Platz in unserer Literatur zu sichern. Der vorliegende erste Band behandelt Integral- und Differentialrechnung für Funktionen einer Veränderlichen; der zweite weniger umfangreiche Band wird den Funktionen mehrerer Veränderlicher gewidmet sein und einige weitere Ergänzungen enthalten.

Es ist mein Bestreben, dem Leser eine deutliche Einsicht in die enge Verbundenheit der Analysis mit den Anwendungen zu vermitteln und — bei aller Wahrung mathematischer Strenge und Präzision — der Anschauung als dem Urquell mathematischer Wahrheiten volle Gerechtigkeit widerfahren zu lassen. Gewiß, die Darstellung der Wissenschaft als geschlossenes System in sich ruhender Wahrheiten ohne eine Erinnerung an Herkunft und Ziel hat einen ästhetischen Reiz und bedeutet die Erfüllung eines tiefen philosophischen Erkenntnisdranges. Aber als ausschließliche grundsätzliche Einstellung oder als didaktisches Prinzip gegenüber Anfängern ist der Standpunkt der abstrakt logischen, in sich gekehrten Wissenschaft eine große Gefahr. Mathematische Analysis treiben und dabei den Anwendungen und der Anschauung den Rücken drehen, das heißt, die Wissenschaft rettungslos dem Schicksal der Vertrocknung und Verkümmerung preisgeben. Es scheint mir eine überaus wichtige Aufgabe, den Lernenden von Anfang an vor einem dünkelhaften allzu bequemen Purismus zu bewahren; nicht zuletzt diesem Zwecke soll mein Buch dienen.

Es wendet sich an jedermann, der sich auf der Grundlage normaler Schulkenntnisse ernstlich um die Wissenschaft und ihre Anwendungen bemühen will, sei er Student an einer Universität oder technischen Hochschule, sei er Lehrer oder Ingenieur. Es verspricht nicht, dem Leser das eigene Nachdenken zu ersparen, aber es führt ohne Zögern und ohne überflüssige Umwege direkt zu interessanten und fruchtbaren Gegenständen und versucht das Verständnis zu erleichtern, indem es nicht bloß Schritt für Schritt beweist, sondern auch die Zusammenhänge und Motive des Ganzen beleuchtet.

Für den jungen Leser, der sich naiv der Führung dieses Buches anvertrauen will, sei noch folgendes bemerkt: Ich habe es vermieden, den Zugang zu den konkreten Tatsachen der Differential- und Integralrechnung durch Grundlagenbetrachtungen zu verbarrikadieren, deren Notwendigkeit man doch erst hinterher ganz begreifen kann; statt dessen sind diese Dinge in Anhängen zu den einzelnen Kapiteln zusammengefaßt, und der Anfänger, dem es in erster Linie um die rasche Durchdringung des Stoffes oder um die Anwendungen zu tun ist, mag ruhig die Lektüre dieser Teile hinausschieben, bis das Bedürfnis dazu erwacht ist. Im übrigen enthalten die Anhänge stoffliche Ergänzungen zur Entlastung der fortlaufenden Darstellung in den einzelnen Kapiteln. Sie sind verhältnismäßig knapp gefaßt. Auch sonst wird der Leser bemerken, daß die anfangs breite Schreibweise gegen den Schluß des Bandes hin in eine knappere übergeht. Er sollte sich aber durch vereinzelte Schwierigkeiten, die er vielleicht in den letzten beiden Kapiteln findet, nicht abschrecken lassen; solche Lücken des Verständnisses pflegen sich später von selbst zu schließen, wenn sie nicht allzusehr gehäuft auftreten.

Ich kann diesen Anlaß nicht vorübergehen lassen, ohne in Dankbarkeit den Namen meines großen Vorgängers im Lehramte Felix Klein zu nennen; was ich hier versuche, liegt ganz in der Richtung seiner Bestrebungen. Auch meinem Freunde Otto Toeplitz in Kiel, der wie kaum ein anderer die hier vorliegenden didaktischen Probleme in ihrer Tiefe durchdacht hat, verdanke ich bewußt und unbewußt empfangene Anregungen.

Die Niederschrift und Drucklegung dieses Buches an Hand einer Vorlesungsausarbeitung in so kurzer Zeit wäre mir inmitten anderer Arbeiten unmöglich gewesen, wenn ich nicht das Glück hätte, in einer Schar hilfsbereiter junger Kollegen zu wirken. Ihnen allen, die kritisch und tätig mir beigestanden haben, gilt mein herzlicher und freundschaftlicher Dank.

Göttingen, im Juni 1927.

**R. Courant.**

## Vorwort zur zweiten Auflage.

Die vorliegende zweite Auflage dieses Bandes unterscheidet sich von der ersten lediglich durch Beseitigung von Druckfehlern und Irrtümern und durch Änderungen bei Einzelheiten der Darstellung.

Auch diesmal habe ich wieder treuen Helfern bei der Drucklegung meinen herzlichen Dank auszusprechen. Vor allen Dingen Herrn Dr. Werner Weber und Herrn Theodor Zech.

Göttingen, im Februar 1930.

**R. Courant.**

# Inhaltsverzeichnis.

Zweites Kapitel.

## Grundbegriffe der Integral- und Differentialrechnung.



### Anhang zum zweiten Kapitel.



Drittes Kapitel.

## Differential- und Integralrechnung der elementaren Funktionen.

## Anhang zum dritten Kapitel.

Seite

Viertes Kapitel.

## Weiterer Ausbau der Integralrechnung.



Fünftes Kapitel.

## Anwendungen.

**Anhang zum fünften Kapitel.**



Sechstes Kapitel.

## Die Taylorsche Formel und die Annäherung von Funktionen durch ganze rationale.



**Anhang zum sechsten Kapitel.**



Siebentes Kapitel.

## Exkurs über numerische Methoden.

# Vorbemerkungen.

Der Anfänger, welcher zum ersten Male mit der sogenannten höheren Mathematik in Berührung kommt, wird sich dem Gefühl einer gewissen Diskontinuität zwischen der Schulmathematik und der Mathematik, wie sie an den Hochschulen getrieben wird, nicht entziehen können. Dieses Gefühl hat seine innere Begründung nicht nur in den historisch gegebenen Verhältnissen, welche den Unterricht an den Hochschulen von dem an den Schulen so sehr verschieden gestaltet haben. Auch im Wesen der höheren oder besser neueren Mathematik, die sich in den letzten drei Jahrhunderten herangebildet hat, liegt ein Unterschied gegen die Elementarmathematik, welche bis vor kurzem den Schulunterricht völlig beherrschte und deren Inhalt vielfach fast unmittelbar der klassischen Mathematik der Griechen entnommen ist.

Der elementaren Mathematik eigentümlich ist zunächst ihre enge Beziehung zur Geometrie. Auch dort, wo die Entwicklung aus dem geometrischen Gebiet in das arithmetische hinüberwächst, bleibt doch die Geometrie fast immer das Fundament. Als zweiten charakteristischen Zug der alten Mathematik müssen wir wohl ihre Tendenz ansehen, das Augenmerk auf die einzelnen mathematischen Objekte gerichtet zu halten. Dinge, die wir heute als spezielle Fälle einer allgemeinen Erscheinung unterordnen würden, stehen dort häufig unvermittelt ohne sichtbare gegenseitige Beziehung nebeneinander. Die enge Beziehung zur geometrischen Anschauung und das Haften an den individuellen Einzelheiten verleiht der alten Mathematik einen eigentümlichen Reiz. Aber es war doch ein entscheidender Fortschritt, daß sich mit dem Beginn der Neuzeit in der Mathematik ganz andersartige Tendenzen durchsetzten und zu einer großen neuen Entwicklung Anlaß gaben, nachdem viele Jahrhunderte hindurch während des Mittelalters trotz mancher Fortschritte im einzelnen doch ein gewisser Stillstand geherrscht hatte.

Die Grundtendenz aller neueren Mathematik ist die Ersetzung von Einzelbetrachtungen durch immer allgemeinere systematische Methoden, welche vielleicht nicht immer den individuellen Zügen des einzelnen Falles in vollem Maße gerecht werden, aber doch durch die Kraft ihrer Allgemeinheit eine Fülle neuer Resultate versprechen.

Auf der anderen Seite gewinnt die Zahl, die analytische Behandlung der Dinge, immer mehr ein selbständiges Recht, um schließlich gänzlich zur Herrschaft über die Geometrie zu gelangen. Ihren deutlichsten Ausdruck finden diese neuen Tendenzen nach mannigfachen vorangehenden Ansätzen durch die Ausbildung der analytischen Geometrie, um deren Entwicklung *Fermat* und *Descartes* die größten Verdienste haben, und der Differential- und Integralrechnung, als deren Väter man gewöhnlich *Leibniz* und *Newton* betrachtet.

Die neuere Mathematik hat in den 300 Jahren ihres Bestehens eine so großartige, sowohl für die reine Wissenschaft als auch für die mannigfachsten technischen und naturwissenschaftlichen Anwendungen bedeutungsvolle Entwicklung genommen, daß ihre Grundbegriffe, vor allem der Funktionsbegriff, allmählich sich weiteste Verbreitung erzwangen und schließlich auch in den Schulbetrieb eindringen mußten.

Ich will in diesen Vorlesungen, ohne auf Vorkenntnisse der höheren Mathematik aus der Schule zurückzugreifen, die wichtigsten Tatsachen aus der Differential- und Integralrechnung so weit entwickeln, daß die Zuhörer am Schluß einerseits zum Studium der höheren mathematischen Disziplinen und zur Vertiefung der Grundlagen, andererseits zur Handhabung der Differential- und Integralrechnung in den verschiedenen Anwendungsgebieten gerüstet sind.

Dabei möchte ich auf eine Gefahr, die aus der eingangs erwähnten Diskontinuität entspringt, besonders hinweisen. Der elementare Standpunkt der Schulmathematik verleitet dazu, an den Einzelheiten haften zu bleiben und den Blick für die allgemeinen Zusammenhänge und systematischen Methoden zu verlieren. Der „höhere Standpunkt" der allgemeinen Methoden birgt aber die umgekehrte Gefahr in sich, daß der Zusammenhang mit dem konkreten Einzelnen verloren geht und daß man den einfachsten individuellen Schwierigkeiten ratlos gegenübersteht, weil man in der Welt allgemeiner Begriffe verlernt hat, das Konkrete zu sehen und zu fassen. Der Leser muß mit eigenen Kräften dafür sorgen, durch dieses Dilemma hindurchzukommen. Nur das immer wiederholte selbständige Durchdenken der Einzelheiten und das vollständige Erfassen der allgemeinen Gedanken im speziellen Beispiel kann zu diesem Ziele führen, und hierin liegt die Hauptaufgabe für jeden, der sich um das Studium der Wissenschaft bemüht.

Erstes Kapitel.

# Vorbereitungen.

Die Differential- und Integralrechnung und mit ihr die ganze höhere Analysis beruht, abgesehen von dem Zahlbegriff, vor allen Dingen auf zwei Begriffsbildungen, nämlich dem Begriff der *Funktion* und dem Begriff des *Grenzwertes*, die zwar schon im klassischen Altertum gelegentlich erkennbar werden, aber doch ihre charakteristische Prägung und Bedeutung erst in der modernen Mathematik erhalten. In diesem einleitenden Kapitel wollen wir versuchen, diese Begriffe auf möglichst einfache und anschauliche Art zu erfassen.

## § 1. Der Zahlbegriff.

Was für eine Art von Dingen eigentlich die Zahlen sind, das ist eine Frage, die mehr den Philosophen als den Mathematiker angeht und mit der sich auch die Philosophen viel beschäftigt haben. Die Mathematik muß jedoch darauf bedacht sein, sich von dem Einfluß widerstreitender philosophischer Meinungen frei zu halten; sie braucht glücklicherweise keine erkenntnistheoretischen Vorstudien über das tiefere Wesen des Zahlbegriffes. So wollen wir die Zahlen, und zwar zunächst die ganzen positiven oder natürlichen Zahlen $1, 2, 3, \ldots$ als etwas Gegebenes hinnehmen; ebenfalls wollen wir die Regeln, nach denen man mit diesen Zahlen rechnen kann, als etwas Gegebenes betrachten und uns nur noch einmal kurz in Erinnerung zurückrufen, in welcher Weise man den Begriff der ganzen positiven Zahlen oder der natürlichen Zahlen notgedrungen hat erweitern müssen.

### 1. Das System der reellen Zahlen.

Im Bereich der natürlichen Zahlen sind die Grundoperationen der Addition und Multiplikation immer unbeschränkt ausführbar, d. h. Summe und Produkt zweier natürlicher Zahlen ist immer wieder eine natürliche Zahl. Aber die Umkehrung dieser Operationen, d. h. Subtraktion und Division, ist im Bereiche der natürlichen Zahlen nicht mehr ausnahmslos durchführbar, und diese Tatsache drängte die mathematische Erfindungskraft schon frühzeitig zur Einführung der Zahl 0, der negativen Zahlen und der positiven und negativen Brüche. Die Gesamtheit aller dieser Zahlen pflegt man als die *rationalen Zahlen* zu bezeichnen, weil sie alle aus der Einheit durch Anwendung der „*rationalen Rechenoperationen*“ Addition, Multiplikation, Subtraktion und Division entstehen.

Man pflegt die Zahlen anschaulich durch Punkte einer geraden Linie, der „*Zahlengeraden*“, zu repräsentieren, indem man auf dieser geraden Linie einen beliebigen Punkt als den Nullpunkt festsetzt, einen anderen als den Punkt 1; die Strecke zwischen diesen beiden Punkten dient dann als Maßstab, um jeder positiven oder negativen rationalen Zahl eine bestimmte Stelle auf der Zahlengeraden zuzuordnen, wobei üblicherweise die positiven Zahlen auf der rechten, die negativen Zahlen auf der linken Seite des Nullpunktes eingetragen werden (vgl. Fig. 1). Versteht man in der üblichen Weise unter dem absoluten Betrag $|a|$ einer Zahl $a$ den Wert $a$ selbst, wenn $a \geqq 0$ [1]), dagegen den Wert $-a$, wenn $a < 0$ ist, so bedeutet $|a|$ einfach die Entfernung des betreffenden Punktes auf der Zahlengeraden vom Anfangspunkt.

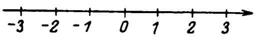



Fig. 1. Zahlengerade.

Die geometrische Deutung der rationalen Zahlen durch Punkte auf der Zahlengeraden weist uns auf eine wichtige Eigenschaft hin, die man durch die Wendung bezeichnet: „*Die Menge der rationalen Zahlen ist überall dicht.*" Dies besagt, daß es zwischen zwei beliebig nahe aneinanderliegenden rationalen Zahlen noch weitere rationale Zahlen gibt oder daß in jedem noch so kleinen Intervall der Zahlengeraden noch rationale Zahlen liegen; geometrisch gesprochen, daß jede noch so kleine Strecke auf der Zahlengeraden rationale Punkte enthält. Diese Dichtigkeit der rationalen Zahlen wird sofort klar, wenn wir von der Bemerkung ausgehen, daß die Zahlen $\frac{1}{2}, \frac{1}{2^2}, \frac{1}{2^3}, \ldots, \frac{1}{2^n}, \ldots$ immer kleiner werden und sich mit wachsendem $n$ der Null immer mehr nähern. Teilen wir nun, vom Nullpunkt beginnend, die Zahlengerade in gleiche Teile der Länge $\frac{1}{2^n}$ ein, so stellen die Endpunkte dieser Intervalle $\frac{1}{2^n}, \frac{2}{2^n}, \frac{3}{2^n}, \ldots$ rationale Zahlen der Form $\frac{m}{2^n}$ dar; über die Zahl $n$ dürfen wir dabei noch beliebig verfügen. Haben wir irgend ein fest vorgegebenes Intervall der Zahlengeraden, mag es auch noch so klein sein, so brauchen wir nur $n$ hinreichend groß zu wählen — nämlich so groß, daß $\frac{1}{2^n}$ kleiner als die Intervallänge wird —, die obige Einteilung also entsprechend fein zu machen, um sicher zu sein, daß Punkte der Einteilung in das Intervall hineinfallen.

Trotz dieser Eigenschaft der Dichtigkeit aber reichen die rationalen Zahlen nicht aus, um alle Punkte auf der Zahlengeraden darzustellen. Schon die Griechen haben erkannt, daß es Strecken gibt, welche nach Wahl einer Strecke der Maßzahl 1 nicht durch eine rationale Zahl repräsentiert werden können, sog. mit der Einheit inkommensurable Strecken. So z. B. ist die Hypotenuse eines rechtwinklig-gleichschenk-

[1]) Durch das Zeichen $\geqq$ soll angedeutet werden, daß *entweder* das Zeichen $>$ *oder* das Zeichen $=$ gelten soll. Entsprechendes gilt für das hernach einzuführende Zeichen $\pm$ (bzw. $\mp$).

ligen Dreieckes, dessen Katheten die Länge 1 haben, nicht kommensurabel mit der Strecke 1, d. h. ihre Länge ist nicht durch eine rationale Zahl gegeben. Das Quadrat dieser Länge $l$ muß nämlich nach dem pythagoreischen Lehrsatz gleich 2 sein; es müßte also, wenn $l$ rational, also von der Form $\frac{p}{q}$ mit ganzzahligen, selbstverständlich von 0 verschiedenen $p$ und $q$ wäre, $p^2 = 2q^2$ sein. Dabei können wir annehmen, daß $p$ und $q$ keinen gemeinsamen Teiler haben, weil wir einen solchen von vornherein hätten wegheben können. Da $p^2$ infolge der obigen Gleichung eine gerade Zahl ist, so muß auch $p$ eine gerade Zahl sein, etwa $p = 2p'$; setzen wir diesen Ausdruck ein, so erhalten wir $4p'^2 = 2q^2$ oder $q^2 = 2p'^2$; es müßte also auch $q^2$ und daher ebenso $q$ selbst eine gerade Zahl sein, entgegen unserer Voraussetzung, daß $p$ und $q$ keinen gemeinsamen Teiler, also auch nicht beide den Teiler 2 besitzen. Unsere Annahme, daß die Hypotenuse durch einen Bruch $\frac{p}{q}$ darstellbar ist, hat sich daher als widerspruchsvoll, mithin als falsch erwiesen.

Die eben durchgeführte — für das Wesen eines „indirekten Beweises" charakteristische — Schlußweise lehrt uns, daß dem Zeichen $\sqrt{2}$ keine rationale Zahl entsprechen kann.

Wir sehen aus dieser Betrachtung, daß wir genötigt sind, außer den rationalen Zahlen noch andere, „*irrationale*" Zahlen einzuführen, wenn wir jedem Punkt der geraden Linie eine Zahl zuordnen wollen. Und diese Forderung, daß den Punkten einer geraden Linie in umkehrbar eindeutiger Weise — nach Wahl einer Einheitsstrecke — Zahlen zugeordnet sein sollen, wollen wir durchaus festhalten. Das System dieser in umkehrbar eindeutiger Weise den Strecken zugeordneten rationalen und irrationalen Zahlen nennt man das System der *reellen* Zahlen.

Die obige Einteilung der Zahlengeraden in Intervalle der Länge $\frac{1}{2^n}$ zeigt uns: Eine Irrationalzahl $\alpha$ läßt sich durch rationale Zahlen beliebig genau annähern, d. h. wir können immer eine rationale Zahl finden, deren Abstand von $\alpha$ kleiner ist als eine beliebig kleine von uns gewünschte Schranke, z. B. kleiner als $\frac{1}{100}$ oder $\frac{1}{1000}$. Ebenso kann man natürlich jede irrationale Zahl in ein beliebig kleines Intervall einschließen, dessen Grenzen durch rationale Zahlen gegeben werden.

Wir setzen als eine fundamentale Tatsache voraus, daß man mit den reellen Zahlen nach den üblichen formalen Rechengesetzen rechnen darf. Dabei soll nicht unerwähnt bleiben, daß die Einführung der irrationalen Zahlen und die Rechtfertigung der Gültigkeit der Rechengesetze für sie Gegenstand einer genaueren Untersuchung sein kann und sein muß. Ebenso jedoch wie die Entwicklung der modernen

Mathematik sich bis in die zweite Hälfte des 19. Jahrhunderts hinein ohne eine solche genaue Betrachtung ungestört vollzogen hat, ist es ganz naturgemäß, sich den Weg zu der eigentlichen höheren Analysis nicht erst durch eine axiomatische Untersuchung des Zahlbegriffes hindurch zu bahnen oder — zu versperren. Ich möchte dem Leser aber den Rat geben, nach Erlangung einer gewissen Reife der Ausbildung eine der vielen guten Darstellungen dieser Fragen zu studieren [1]).

Was die Rechengesetze selbst anbetrifft, so will ich nur an die einfachsten Regeln für das *Rechnen mit Ungleichungen* erinnern, das in der ganzen Analysis eine viel größere Rolle spielt als in der Elementarmathematik. Es ist stets

$$|a \pm b| \leqq |a| + |b|, \quad |a \pm b| \geqq |a| - |b|.$$

Ferner darf man Ungleichungen addieren und mit positiven Faktoren multiplizieren; bei der Multiplikation mit einer negativen Zahl kehrt sich der Sinn der Ungleichung um. Weiter folgt aus $a > b > 0$ und $0 < a' < b'$ die neue Ungleichung

$$\frac{a}{a'} > \frac{b}{b'}.$$

## 2. Die Zahlensysteme.

Man pflegt für numerische Rechnungen die Zahlen im Dezimalsystem folgendermaßen zu schreiben:

$$a = \ldots a_2\, a_1\, a_0,\, a_{-1}\, a_{-2} \ldots .$$

Es ist hier $a_0$ die Anzahl der Einer, $a_1$ die Anzahl der Zehner, $a_2$ die Anzahl der Hunderter, $a_{-1}$ die Anzahl der Zehntel usw. Dies bedeutet nichts weiter als eine Darstellung der Zahl $a$ in folgender Form:

$$a = a_0 + a_1 \cdot 10 + a_2 \cdot 10^2 + \cdots + a_{-1} \cdot 10^{-1} + a_{-2} \cdot 10^{-2} + \cdots,$$

wobei

$$a_0,\ a_1,\ a_2,\ \ldots,\ a_{-1},\ a_{-2},\ \ldots$$

Ziffern der Folge $0, 1, 2, \ldots, 9$ sind.

Dem Leser wird aus der Schule noch das einfache Resultat geläufig sein, daß eine rationale Zahl durch einen rein periodischen oder gemischt periodischen Dezimalbruch dargestellt wird (speziell durch einen abbrechenden), während ein Dezimalbruch ohne Periode stets eine irrationale Zahl darstellt.

Die dezimale Schreibweise trägt in gewisser Hinsicht den Charakter des Zufälligen, z. B. hinsichtlich der Grundzahl 10. Für rein theoretische Zwecke empfiehlt sich häufig die Darstellung der Zahlen im sog. *Dual-*

---

[1]) Z. B. *K. Knopp*: Theorie und Anwendung der unendlichen Reihen, 2. Aufl., Berlin 1924, insbes. Kap. 1.

*system,* d. h. in einem System mit der Grundzahl 2. In einem solchen System würde eine ganze positive Zahl $a$ in der Form

$$a = \alpha_0 + \alpha_1 \cdot 2 + \alpha_2 \cdot 2^2 + \alpha_3 \cdot 2^3 + \cdots$$

dargestellt werden, wobei die Ziffern $\alpha_0, \alpha_1, \alpha_2, \ldots$ jeweils nur 0 oder 1 sein können. Die Zahl 9 würde z. B. im Zweiersystem folgendermaßen zu schreiben sein: $1001 = 1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0$. Man sieht, daß im Zweiersystem die Zahlen verhältnismäßig lange Ausdrücke haben. Gleichviel aber in welchem Zahlensystem wir die Zahlen schreiben, stets lassen sich in ihm die sämtlichen rationalen und irrationalen Zahlen ausdrücken [1]).

## § 2. Der Funktionsbegriff.

Der Begriff der *Funktion* verdankt seine klare Erfassung und Hervorhebung erst der neueren Zeit. Wir wollen ihn zunächst an einer Reihe von Beispielen erläutern.

### 1. Beispiele.

a) Wenn ein ideales Gas durch einen Stempel in einem Gefäß eingeschlossen ist, so besteht zwischen dem Druck $p$ und dem Volumen $v$ (wenn die Temperatur stets denselben Wert behält) die Beziehung

$$v \cdot p = C,$$

wo $C$ eine Konstante bedeutet. Dieses sog. *Boylesche Gesetz* sagt nichts über die Größen $v$ und $p$ selbst aus, sondern bedeutet: Wenn $p$ einen bestimmten, innerhalb gewisser Grenzen beliebig wählbaren Wert hat, so kann ich daraus $v$ bestimmen und umgekehrt:

$$v = \frac{C}{p}, \qquad p = \frac{C}{v}.$$

Man sagt dann: $v$ ist eine Funktion von $p$, resp.: $p$ ist eine Funktion von $v$.

b) Erhitzen wir einen Metallstab, der bei der Temperatur von Null Grad die Länge $l_0$ besitzt, auf die Temperatur $\vartheta$ Grad, so wird unter den einfachsten Annahmen seine Länge $l$ durch das Gesetz

$$l = l_0 (1 + \beta\vartheta)$$

gegeben, wo der „Ausdehnungskoeffizient" $\beta$ eine Konstante ist. Wir sagen wieder: $l$ ist eine Funktion von $\vartheta$.

[1]) Übrigens hat man im Dezimalsystem an sich zwei Möglichkeiten, die $10^m$-fachen ($m$ ganz) der ganzen Zahlen darzustellen. So ist z. B. $1 = 0{,}999\ldots$; wir wählen in diesem Falle stets die erste Art der Darstellung. Analoges gilt in den anderen Zahlensystemen.

## 2. Begriffliche Formulierung.

Um den mathematischen Funktionsbegriff allgemein zu formulieren, fassen wir ein bestimmtes Intervall unserer Zahlenskala ins Auge, etwa das Intervall zwischen den Zahlen $a$ und $b$, und betrachten die Gesamtheit der Zahlen $x$, welche diesem Intervall angehören, d. h. der Beziehung

$$a \leqq x \leqq b$$

genügen.

Wenn wir die Größe $x$ als beliebig in diesem Intervall wählbar betrachten, so nennen wir sie eine *(stetige) Veränderliche* oder *Variable* in diesem Intervall.

Ist nun durch irgendeine Vorschrift jedem Werte $x$ dieses Intervalles ein bestimmter Wert $y$ zugeordnet, so sagen wir: $y$ ist eine *Funktion von $x$*, und schreiben symbolisch:

$$y = f(x) \quad \text{oder} \quad y = F(x) \quad \text{oder} \quad y = g(x)$$

oder ähnlich. Man bezeichnet dann $x$ als die *unabhängige*, $y$ als die *abhängige* Größe oder Veränderliche; oder auch $x$ als das *Argument* der Funktion $y$.

Es sei schon hier bemerkt, daß es für gewisse Betrachtungen einen Unterschied macht, ob man zu dem Intervall von $a$ bis $b$ die Endpunkte hinzurechnet, wie wir es oben getan haben, oder nicht; im letzteren Falle hätten wir die Größe $x$ durch die Ungleichungen

$$a < x < b$$

zu beschränken. Wir sprechen, wenn sonst Mißverständnisse möglich sind, im ersten Fall von einem *abgeschlossenen*, im zweiten Fall von einem (beiderseits) *offenen* Intervall. Wird nur ein Endpunkt, nicht der andere, zum Intervalle mit gerechnet (z. B. $a < x \leqq b$), so sprechen wir von einem *halb* oder *einseitig offenen* Intervalle. Endlich kann man offene Intervalle auch nach einer oder nach beiden Seiten hin ins Unendliche ausgedehnt denken. Man sagt dann: die stetige Veränderliche $x$ läuft in einem *unendlichen* (offenen) Intervalle, und schreibt symbolisch: $a < x < \infty$ oder $-\infty < x < b$ oder $-\infty < x < \infty$.

In der allgemeinen Definition des Begriffes einer in einem Intervall gegebenen Funktion ist über die Art der Vorschrift, durch welche die abhängige Variable zu der unabhängigen Variablen konstruiert wird, nichts Näheres gesagt. In der Tat enthält der Funktionsbegriff bei dieser Allgemeinheit noch etwas Unbestimmtes und Vages. Er bedarf für alle Anwendungen einer weitgehenden Einschränkung in der Hinsicht, daß an das Abhängigkeitsgesetz präzisere Forderungen gestellt werden müssen.

### 3. Graphische Darstellung. Eindeutigkeit und Mehrdeutigkeit. Stetigkeit.

Zu einer solchen vernünftigen Einschränkung, durch welche der allgemeine Funktionsbegriff erst wirklich anwendbar wird, führt uns der Zusammenhang mit der Geometrie. Der Grundgedanke der analytischen Geometrie ist ja, eine geometrisch gegebene Kurve dadurch analytisch zu charakterisieren, daß man eine der beiden rechtwinkeligen Koordinaten, etwa $y$, als Funktion $y = f(x)$ der andern Koordinate $x$ betrachtet; z. B. wird eine Parabel durch die Funktion $y = x^2$, der Kreis mit dem Radius 1 um den Nullpunkt durch die Funktion $y = \sqrt{1 - x^2}$ dargestellt. Beim ersten Beispiel können wir die Funktion im unendlichen Intervall $-\infty < x < \infty$ definiert denken; beim zweiten ist es notwendig, sich auf das Intervall $-1 \leqq x \leqq 1$ zu beschränken, da außerhalb die Funktion (im Reellen) keinen Sinn hat.

Gehen wir umgekehrt nicht von einer geometrisch gegebenen Kurve aus, sondern betrachten eine Funktion $y = f(x)$ als das primär Gegebene, so können wir uns die funktionale Abhängigkeit der Größe $y$ von der Größe $x$ graphisch darstellen, indem wir dazu in der bekannten Weise ein rechtwinkliges *Koordinatensystem* (siehe Fig. 2) benutzen. Trägt man zu jeder Abszisse $x$ die zugehörige Ordinate $y = f(x)$ auf, so erhält man als geometrisches Bild der Funktion eine *Kurve*. Die Einschränkung, welche wir dem Funktionsbegriff auferlegen wollen, besteht nun darin, daß dieses geometrische Bild wirklich auch eine anschaulich erfaßbare Kurve darstellt. Dann werden wir z. B. folgende Funktion auszuschließen haben: Für jeden rationalen Wert $x$ sei $y = 1$; für jeden irrationalen Wert $x$ sei $y = 0$. Man erkennt, daß bei dieser Funktion der Funktionswert unendlich oft zwischen 0 und 1 hin- und herspringen muß, wenn man auch nur ein noch so kleines Intervall von $x$ durchläuft.

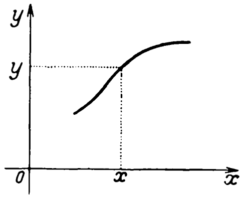
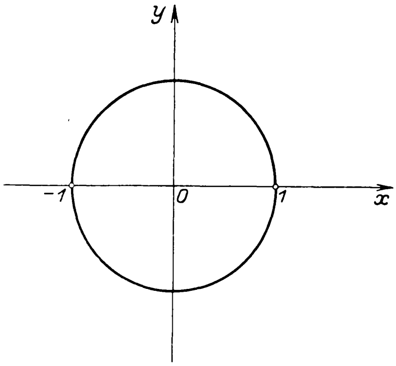


Fig. 2. Koordinatensystem.

An Hand der graphischen Veranschaulichung der Funktionen durch Kurven möchte ich auf einen Punkt hinweisen, welcher dem Anfänger oft gewisse Schwierigkeiten bereitet. Das Gesetz, welches der Variablen $x$ die Funktion $y$ zuordnet, soll, wenn nicht das Gegenteil ganz ausdrücklich gesagt ist oder aus dem Zusammenhang hervorgeht, diese Zuordnung in e i n d e u t i g e r Weise vermitteln; z. B. die Zuordnung $y = x^2$ oder $y = \sin x$. Will man diese Eindeutigkeitsforderung hervorheben, so spricht man von „eindeutig definierten" oder „*eindeutigen*" Funktionen. Es kann jedoch vorkommen, daß das Zuordnungsgesetz uns zunächst in einer Form entgegentritt, bei welcher eine Mehrdeutigkeit vorliegt; z. B. $y = \sqrt{x}$ oder $y = \sqrt{1 - x^2}$, wo ja wegen der Doppeldeutigkeit der Quadratwurzel die Wahl zwischen zwei Vorzeichen offen bleibt. Läßt man diese Doppeldeutigkeit in der Funktionsdefinition bestehen, so spricht man von einer „*mehrdeutigen*“ Funktion, in diesem Falle von einer zweideutigen.

In vielen Fällen wird es jedoch angenehm sein, diese Mehrdeutigkeit zu beseitigen. Dies kann man ohne weiteres tun, indem man durch eine besondere Vorschrift eine der verschiedenen Möglichkeiten auszeichnet, in unserem Falle z. B. durch die Vorschrift $y \geqq 0$. Man spricht dann von einem „*eindeutigen Zweig*“ einer mehrdeutigen Funktion; der andere eindeutige Zweig wird in unserem Falle durch die Beziehung $y \leqq 0$ charakterisiert. Jeder solche Zweig der Funktion



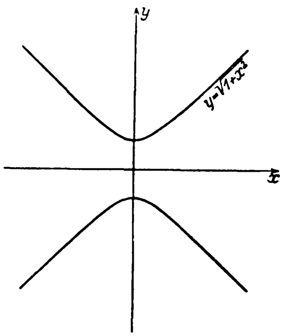




Fig. 3. Mehrdeutige Funktionen. Fig. 4.

$y = \sqrt{1 - x^2}$ ist dann in dem Intervalle $-1 \leqq x \leqq 1$ eine eindeutige Funktion von $x$. Wenn im folgenden nichts Besonderes gesagt ist, wollen wir unter „Funktion“ immer einen derartigen eindeutigen Zweig verstehen, insbesondere bei mehrdeutigen Wurzeln den nicht negativen.

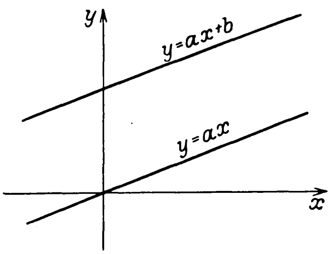



Fig. 5. Ganze lineare Funktionen.

Ist die Funktion in ihrem ganzen Verlauf eindeutig definiert, so wird die sie darstellende Kurve von jeder Parallelen zur $y$-Achse nur in einem Punkte geschnitten. Handelt es sich jedoch um eine mehrdeutige Funktion, wie z. B. $y = \pm \sqrt{1 - x^2}$, so werden solche Parallelen mehrere Schnittpunkte mit der Kurve haben können, und verschiedenen Schnittpunkten werden verschiedene eindeutige Funktionszweige entsprechen. Die Kurvenstücke, welche zu verschiedenen eindeutigen Funktionszweigen gehören, können dabei miteinander zusammenhängen und so den Gesamtverlauf einer geometrisch mit einem Schlage definierbaren Kurve bilden wie beim Kreise (vgl. Fig. 3) $y = \pm \sqrt{1 - x^2}$; die verschiedenen Zweige können aber auch völlig getrennt verlaufen, wie bei der Hyperbel $y = \pm \sqrt{1 + x^2}$ (vgl. Fig. 4).

Ich gebe im folgenden noch einige Beispiele für die graphische Darstellung von Funktionen.

α) $$y = ax.$$

$y$ ist $x$ proportional. Das graphische Bild (vgl. Fig. 5) ist eine Gerade durch den Nullpunkt des $xy$-Systemes.

β) $$y = ax + b.$$

$y$ ist eine „ganze lineare Funktion" von $x$. Das Bild ist eine Gerade durch den Punkt $x = 0$, $y = b$, die im Falle $a \neq 0$ auch noch durch den Punkt $x = -\frac{b}{a}$, $y = 0$ geht, im Falle $a = 0$ horizontal läuft.

γ) $$y = \frac{a}{x}.$$

$y$ ist $x$ umgekehrt proportional. Ist speziell $a = 1$, also

$$y = \frac{1}{x},$$

so ist z. B.

für $x = 1$ $y = 1$, für $x = \frac{1}{2}$ $y = 2$, für $x = 2$ $y = \frac{1}{2}$.

Das Bild (vgl. Fig. 6) ist eine zu den Winkelhalbierenden der Achsen symmetrisch liegende Kurve, eine gleichseitige Hyperbel.

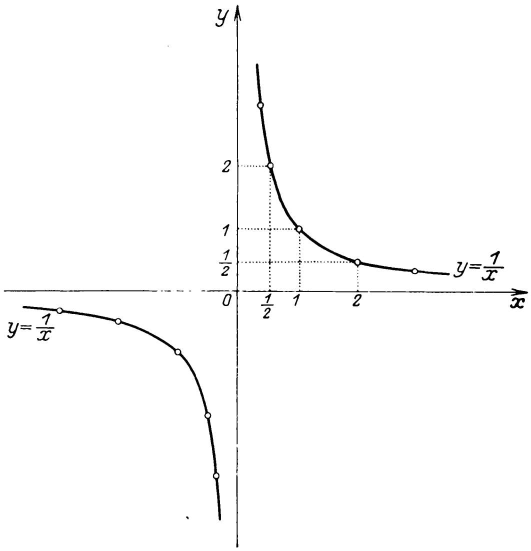


Fig. 6. Unendlichkeitsstelle.

Diese letztere Funktion ist für den Wert $x = 0$ offenbar nicht definiert; denn die Division durch 0 hat keinen Sinn. Die Stelle $x = 0$, die ein Ausnahmepunkt für die Funktion ist, in dessen Umgebung positiv und negativ beliebig große Funktionswerte vorkommen, stellt das einfachste Beispiel einer *Unendlichkeitsstelle* dar, worauf wir später (vgl. § 8, Nr. 2) noch ausführlich zurückkommen.

δ) $$y = x^2.$$

Diese Funktion wird bekanntlich durch eine Parabel dargestellt (vgl. Fig. 7).

Ebenso wird die Funktion $y = x^3$ anschaulich durch die sog. kubische Parabel beschrieben (vgl. Fig. 8).

Die eben betrachteten Funktionen und ihre graphischen Bilder zeigen eine Eigenschaft, welche von der größten Bedeutung für die Diskussion jeder Funktion ist, nämlich die Eigenschaft der *Stetigkeit*. Wir werden diesen Begriff später (vgl. § 8) noch genauer analysieren; sein anschauliches Wesen, mit dessen Hervorhebung ich mich an dieser Stelle begnügen kann, besagt, daß bei einer kleinen Änderung von $x$ der Wert von $y$ auch nur eine kleine Änderung erleidet und nicht plötzlich springt, d. h. daß die repräsentierende Kurve nirgends zerreißt.

Eine Funktion, die für alle Werte von $x$ in einem Intervall denselben Wert $y = a$ besitzt, nennt

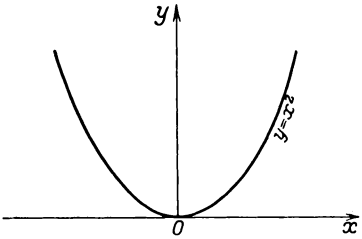



Fig. 7. Parabel.

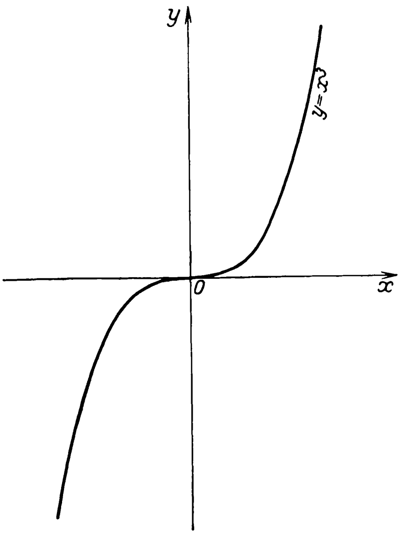



Fig. 8. Kubische Parabel.

man eine ***Konstante a***; sie wird durch eine horizontale gerade Linie repräsentiert. Eine Funktion $y = f(x)$, welche die Eigenschaft hat, daß in einem Intervall zu größeren Werten von $x$ immer größere Werte

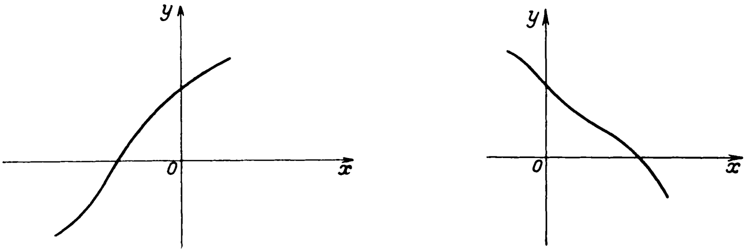



Fig. 9. Monotone Funktionen.

von $y$ gehören, heißt eine in diesem Intervall *monoton wachsende* Funktion; gehört aber zu größerem $x$ immer ein kleineres $y$, so heißt sie in diesem Intervall eine *monoton abnehmende* Funktion. Anschaulich werden solche monotone Funktionen durch Kurven dargestellt, welche in dem betreffenden Intervall immer ansteigen oder immer abfallen (vgl. Fig. 9).

Ist die durch $y = f(x)$ dargestellte Kurve symmetrisch zur $y$-Achse, d. h. gehört zu $x = -a$ derselbe Funktionswert wie zu $x = a$ oder ist

$$f(-x) = f(x),$$

so sprechen wir von einer *geraden* Funktion. Z. B. $y = x^2$ (vgl. Fig. 7). Ist dagegen die Kurve symmetrisch zum Nullpunkt, d. h. ist

$$f(-x) = -f(x),$$

so sprechen wir von einer *ungeraden* Funktion; z. B. $y = x$ oder $y = x^3$ (vgl. Fig. 8) oder $y = \frac{1}{x}$.

### 4. Umkehrfunktionen.

Wir konnten bereits bei unserm ersten Beispiel unter Nr. 1 bemerken, daß man eine formale Abhängigkeit zwischen zwei Größen verschieden auffassen kann, indem man ebenso berechtigt ist, die erste Größe als Funktion der zweiten, wie die zweite als Funktion der ersten

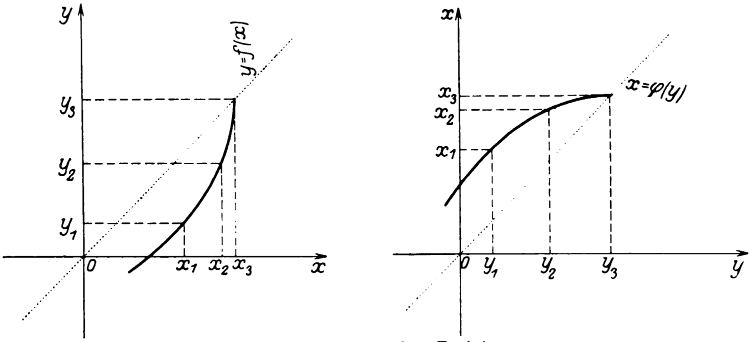



Fig. 10. Umkehrung einer Funktion.

anzusehen. Ist z. B. $y = ax + b$, wobei wir $a \neq 0$ annehmen, so wird $x$ als Funktion von $y$ durch die Gleichung $x = \frac{y-b}{a}$ dargestellt. Ebenso können wir den durch die Gleichung $y = x^2$ dargestellten funktionalen Zusammenhang zwischen $y$ und $x$ auch durch die Gleichung $x = \pm\sqrt{y}$ darstellen. Ganz ähnlich können wir versuchen, bei einer beliebigen Funktion $y = f(x)$ auch $x$ als Funktion von $y$ aufzufassen oder, wie wir sagen wollen, die Funktion $y = f(x)$ durch ihre „*Umkehrfunktion*“ $x = \varphi(y)$ zu ersetzen.

Geometrisch bedeutet dies folgendes: Man kann das graphische Bild einer Funktion $y = f(x)$ auch betrachten, indem man es um die Gerade umklappt, die den Winkelraum zwischen der positiven $x$-Achse und der positiven $y$-Achse halbiert[1] (vgl. Fig. 10). Man hat dann ohne weiteres graphisch $x$ als Funktion von $y$ dargestellt und somit die Umkehrfunktion $x = \varphi(y)$ repräsentiert.

Man erkennt aber schon aus dieser geometrischen Betrachtung sofort, daß die eindeutige Umkehrbarkeit einer Funktion $y = f(x)$ in einem

[1]) Statt diese Umklappung vorzunehmen, könnten wir zunächst das Koordinatensystem mit der Kurve $y = f(x)$ um einen rechten Winkel drehen und dann um die $x$-Achse klappen.

Intervall an gewisse Bedingungen gebunden ist. Sobald die zu der Funktion gehörige Kurve von einer Parallelen $y = c$ zur $x$-Achse in mehreren Punkten geschnitten wird, wird dem Werte $y = c$ mehr als ein Wert $x$ entsprechen, die Funktion in dem betrachteten Intervall also nicht eindeutig umkehrbar sein. Dieser Fall tritt nicht ein, wenn die Funktion $y = f(x)$ dort stetig und monoton ist. Dann lehrt uns Fig. 10, daß tatsächlich zu jedem Werte von $y$ in dem betrachteten Intervall $x_1 \leqq x \leqq x_3$ gerade ein Wert von $x$ gehört, und wir entnehmen der Figur die *Tatsache, daß sich eine in einem Intervalle monotone stetige Funktion in diesem Intervall stets durch eine eindeutig bestimmte stetige Funktion umkehren läßt* (Genaueres auf S. 52).

## § 3. Nähere Betrachtung der elementaren Funktionen.

### 1. Die rationalen Funktionen.

Wir gehen nun dazu über, die elementaren Funktionen, die dem Leser von der Schulmathematik her schon bekannt sind, noch einmal kurz zu durchmustern. Die einfachsten Klassen von Funktionen erhalten wir durch wiederholte Anwendung der elementaren Rechenoperationen: Addition, Multiplikation, Subtraktion. Wenden wir diese Rechenoperationen auf eine unabhängige Veränderliche $x$ und irgendwelche rationale oder irrationale Zahlen an, so erhalten wir als Resultat die *ganzen rationalen Funktionen oder Polynome*:

$$y = a_0 + a_1 x + \cdots + a_n x^n .$$

Die ganzen rationalen Funktionen sind wegen ihres einfachen Baues sozusagen die Grundfunktionen für die ganze Analysis.

Bilden wir noch die Quotienten derartiger Funktionen, d. h. Ausdrücke der Form

$$y = \frac{a_0 + a_1 x + \cdots + a_n x^n}{b_0 + b_1 x + \cdots + b_m x^m},$$

so gelangen wir zu den *allgemeinen* oder *gebrochenen rationalen Funktionen*, die überall definiert sind, wo der Nenner von Null verschieden ist.

Die einfachste ganze rationale Funktion ist die (ganze) *lineare Funktion*

$$y = ax + b .$$

Sie wird durch eine gerade Linie dargestellt. Jede *quadratische Funktion* von der Form

$$y = ax^2 + bx + c$$

wird durch eine Parabel dargestellt. Die Kurven, welche eine ganze rationale Funktion dritten Grades

$$y = ax^3 + bx^2 + cx + d$$

darstellen, nennt man gelegentlich Parabeln dritter Ordnung, usw.

Als anschauliche Beispiele fügen wir in Fig. 11 die graphischen Bilder der Funktion $y = x^n$ für die Exponenten $n = 1, 2, 3, 4$ bei.

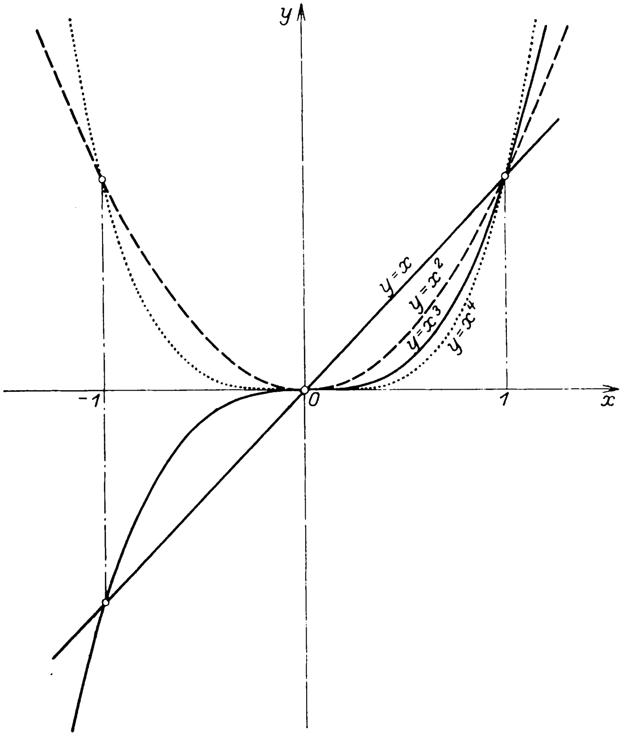



Fig. 11. Potenzfunktionen.

Man erkennt, daß für gerade Werte von $n$ die Funktion $y = x^n$ der Gleichung $f(-x) = f(x)$ genügt, also eine gerade Funktion ist, während für ungerades $n$ die Funktion die Bedingung $f(-x) = -f(x)$ befriedigt, d. h. ungerade ist.

Das einfachste Beispiel für eine rationale Funktion, die keine ganze rationale Funktion ist, ist die schon früher betrachtete Funktion $y = \frac{1}{x}$, deren Bild durch die gleichseitige Hyperbel gegeben wird. Ein anderes ist die Funktion $y = \frac{1}{x^2}$ (vgl. Fig. 12).

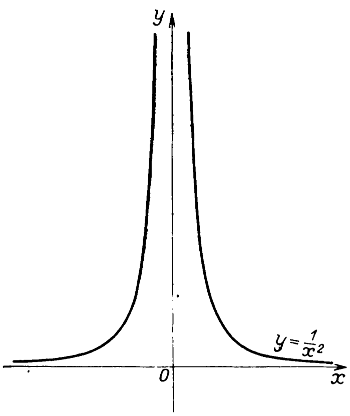



Fig. 12.

## 2. Algebraische Funktionen.

Aus dem Bereich der rationalen Funktionen führt uns sofort das Problem hinaus, Umkehrfunktionen von ihnen zu bilden; z. B. wird für die Potenz $y = x^n$ dieses Problem gelöst durch Einführung der $n$-ten Wurzel $\sqrt[n]{y} = x$ oder, wenn wir die Bezeichnungen von abhängiger und unabhängiger Variabler vertauschen, durch die Funktion

$$y = \sqrt[n]{x} = x^{\frac{1}{n}}.$$

Allgemeiner können wir

$$y = \sqrt[n]{R(x)}$$

betrachten, wo $R(x)$ eine rationale Funktion bedeutet. Zu weiteren Funktionen von ähnlichem Typus gelangt man, indem man auf eine oder mehrere solcher spezieller Funktionen rationale Rechenprozesse anwendet. So gewinnt man beispielsweise die Funktionen

$$y = \sqrt[m]{x} + \sqrt[m]{x^2 + 1}, \qquad y = x + \sqrt{x^2 + 1}.$$

Funktionen dieser Art sind spezielle „*algebraische Funktionen*".

### 3. Die trigonometrischen Funktionen.

Während die rationalen und die eben betrachteten algebraischen Funktionen unmittelbar durch die elementaren Rechenoperationen und ihre Umkehrungen definiert werden, ist die Quelle, aus welcher wir die Kenntnis der übrigen, der sog. *transzendenten*[1]) *Funktionen* schöpfen, zunächst die Geometrie. Wir wollen uns hier mit den *elementaren transzendenten Funktionen* beschäftigen, nämlich mit den trigonometrischen Funktionen, der Exponentialfunktion und dem Logarithmus.

Bei allen Betrachtungen der höheren Analysis, in welchen Winkel vorkommen, pflegt man diese Winkel nicht nach Grad, Minuten und Sekunden zu messen, sondern legt der Winkelmessung das *Bogenmaß* zugrunde. Man denkt sich den zu messenden Winkel mit seinem Scheitel in den Mittelpunkt eines Kreises vom Radius 1 gelegt und mißt die Größe des Winkels durch die Länge des Bogens, den der Winkel aus der Kreisperipherie ausschneidet. Ein Winkel von $180^0$ hat also das Bogenmaß $\pi$, ein Winkel von $90^0$ das Bogenmaß $\frac{\pi}{2}$, ein Winkel von $45^0$ das Bogenmaß $\frac{\pi}{4}$, ein Vollwinkel von $360^0$ das Bogenmaß $2\pi$. Umgekehrt berechnet sich ein Winkel mit dem Bogenmaß 1 in Grad zu

$$\frac{180^0}{\pi} \text{ oder angenähert } 57^0\,17'\,45''.$$

[1]) Das Wort „transzendent" bedeutet keineswegs etwas besonders Geheimnisvolles oder Schwieriges, sondern deutet lediglich die Tatsache an, daß die Definition dieser Funktionen mit Hilfe der elementaren Rechenoperationen nicht mehr möglich ist: „quod algebrae vires transcendit".

Grundsätzlich werden wir von nun ab immer, wenn wir von einem Winkel $x$ sprechen, einen Winkel meinen, dessen Bogenmaß $x$ ist.

Nach diesen Vorbemerkungen erinnere ich noch kurz an die Bedeutung der trigonometrischen Funktionen $\sin x$, $\cos x$, $\operatorname{tg} x$, $\operatorname{ctg} x$,

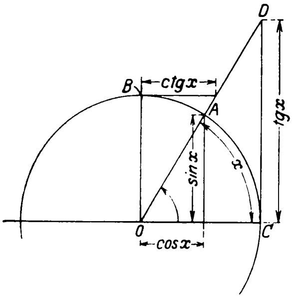



Fig. 13.
Die trigonometrischen Funktionen.

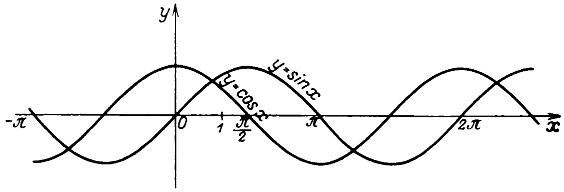



Fig. 14.

indem ich auf Fig. 13 verweise, in welcher der Winkel $x$ von dem Schenkel $OC$ (Länge 1) aus abgetragen ist und als positiver Drehsinn der durch die Pfeilrichtung angegebene, dem Uhrzeiger entgegengesetzt

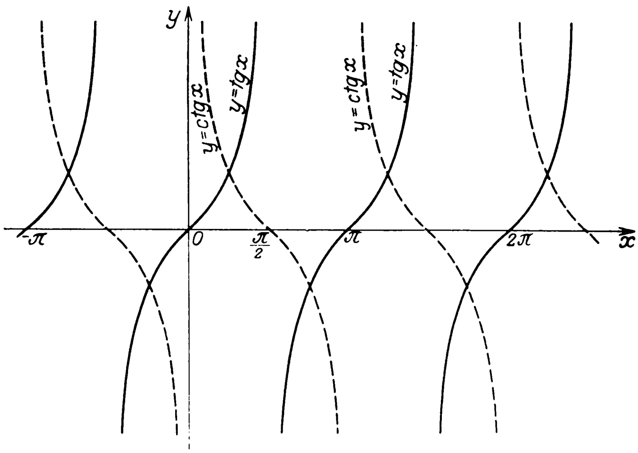



Fig. 15.

laufende verstanden wird. Die rechtwinkligen Koordinaten des Punktes $A$ liefern uns einfach die Funktionen $\cos x$ und $\sin x$. Die graphischen Bilder der Funktionen $\sin x$, $\cos x$, $\operatorname{tg} x$, $\operatorname{ctg} x$ werden durch Fig. 14 und 15 gegeben.

## 4. Exponentialfunktion und Logarithmus.

Man betrachtet neben den trigonometrischen Funktionen als weitere elementare transzendente Funktionen die Exponentialfunktion mit der positiven Basis $a$

$$y = a^x$$

und deren Umkehrfunktion, den Logarithmus zur Basis $a$

$$x = \log_a y\,.$$

Man pflegt auf der Schule über eigentümliche Schwierigkeiten in der Definition dieser Funktionen hinwegzugleiten, und auch wir wollen die genauere Diskussion dieser Funktionen noch ein wenig verschieben, bis wir feinere Hilfsmittel zur Hand haben (vgl. drittes Kapitel, § 6). Aber schon jetzt können wir sagen, wie wir diese Funktionen aufzufassen haben. Ist $x = \frac{p}{q}$ eine rationale Zahl ($p$ und $q > 0$ ganze Zahlen) — die Zahl $a$ setzten wir als positiv voraus —, so definieren wir $a^x$ als $\sqrt[q]{a^p} = a^{\frac{p}{q}}$, wo die Wurzel verabredungsgemäß positiv zu nehmen ist. Da die rationalen Werte von $x$ überall dicht liegen (vgl. S. 4), so liegt es nahe, die so definierten Werte von $a^x$ dadurch zu einer für alle, auch irrationale Werte von $x$ definierten stetigen Funktion, der „Exponentialfunktion", zu ergänzen, daß man für irrationale Werte von $x$ die Werte von $a^x$ einfach an die definierten Werte stetig anschließt und demgemäß unter $a^x$ diejenige stetige Funktion versteht, die für alle *rationalen* Werte von $x$ oben definiert wurde. Daß diese Festsetzung einwandfrei und nur auf eine Weise möglich ist, wollen wir für den Moment hinnehmen, indem wir uns dessen bewußt bleiben, daß wir später dafür eine Begründung nachzuholen haben[1]).

Die Funktion

$$x = \log_a y$$

läßt sich dann für $y > 0$ als Umkehrfunktion der Exponentialfunktion definieren.

## § 4. Funktionen einer ganzzahligen Veränderlichen.

Bisher haben wir die unabhängige Veränderliche als in einem ganzen Intervall stetig variabel betrachtet. Es kommen aber in der Mathematik auch zahlreiche Fälle vor, in denen eine Größe nur von einer ganzen Zahl, einer Nummer $n$, abhängt, welche die sämtlichen Werte $1, 2, 3, \ldots$ annehmen kann. Wir sprechen dann von einer Funktion einer ganzzahligen Veränderlichen, von einer *zahlentheoretischen Funktion* oder auch von einer Funktion eines *Index*, indem wir die ganzzahlige unabhängige Veränderliche, die Nummer, als Index bezeichnen. Am besten verstehen wir diese Begriffsbildung an der Hand von Beispielen.

1. Die Summe der ersten $n$ ganzen Zahlen

$$S_1(n) = 1 + 2 + 3 + 4 + \cdots + n = \frac{n(n+1)}{2}$$

---

[1]) Vgl. S. 54 und 139.

ist eine Funktion von $n$. Ebenso ist die Summe der ersten $n$ Quadratzahlen

$$S_2(n) = 1^2 + 2^2 + 3^2 + \cdots + n^2$$

eine Funktion der ganzzahligen Veränderlichen $n$[1]).

2. Andere einfache zahlentheoretische Funktionen sind der Ausdruck

$$n! = 1 \cdot 2 \cdots n$$

und die Binomialkoeffizienten

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}$$

bei festem $k$.

3. Jede ganze Zahl $n > 1$, die nicht Primzahl ist, läßt sich durch mehrere positive ganze Zahlen teilen, während die Primzahlen nur durch 1 und durch sich selbst teilbar sind. Offenbar können wir nun die Anzahl $T(n)$ dieser Teiler einer Zahl $n$ als Funktion der Zahl $n$ selbst ansehen. Für die ersten Zahlen ist diese zahlentheoretische Funktion durch die folgende Tabelle gegeben:

| $n =$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $T(n) =$ | 1 | 2 | 2 | 3 | 2 | 4 | 2 | 4 | 3 | 4 | 2 | 6 |

4. Eine viel kompliziertere zahlentheoretische Funktion ist die Anzahl $A(n)$ der Primzahlen, die unterhalb der Zahl $n$ liegen. Ihre nähere Untersuchung bildet eines der interessantesten und anziehendsten Probleme der höheren Zahlentheorie. Ich will hier nur das Haupt-

[1]) Man kann übrigens diese letzte Summe in folgender Weise leicht durch einen einfachen rationalen Ausdruck in $n$ darstellen. Wir gehen aus von der Formel

$$(\nu + 1)^3 - \nu^3 = 3\nu^2 + 3\nu + 1,$$

schreiben diese Gleichung für die Werte $\nu = 0, 1, 2, \ldots, n$ und addieren. Dann finden wir:

$$(n + 1)^3 = 3S_2 + 3S_1 + n + 1,$$

woraus durch Einsetzen von $S_1$ folgt

$$3S_2 = (n+1)\left\{(n+1)^2 - 1 - \frac{3}{2}n\right\} = (n+1)\left\{n^2 + \frac{1}{2}n\right\}$$

und somit

$$S_2 = \frac{n(n+1)(2n+1)}{6}.$$

Ebenso kann man die zahlentheoretischen Funktionen

$$S_3(n) = 1^3 + 2^3 + \cdots + n^3,$$
$$S_4(n) = 1^4 + 2^4 + \cdots + n^4,$$
$$\cdots\cdots\cdots\cdots$$

durch ein ganz ähnliches Verfahren wie oben als rationale Funktionen von $n$ ausdrücken.

resultat dieser Untersuchungen andeuten; es besagt, daß die Anzahl $A(n)$ für große Werte von $n$ angenähert durch die Funktion $\frac{n}{\log n}$ gegeben ist[1]), wenn man unter $\log n$ den Logarithmus zu der später zu definierenden „natürlichen Basis“ $e$ versteht (vgl. drittes Kapitel, § 6).

## § 5. Der Begriff des Grenzwertes einer Zahlenfolge. Beispiele.

Unter einer Zahlenfolge verstehen wir eine Aufeinanderfolge von unendlich vielen (nicht notwendig verschiedenen) Zahlen $a_1, a_2, a_3, \ldots, a_n, \ldots$, welche uns durch irgend eine Vorschrift gegeben sein können — im Grunde genommen also eine Funktion der ganzzahligen Veränderlichen $n$. Wir wollen uns zunächst an einer Reihe von Beispielen orientieren.

### 1. $a_n = \frac{1}{n}$.

Wir betrachten die Zahlenfolge

$$a_1 = 1, \quad a_2 = \frac{1}{2}, \quad a_3 = \frac{1}{3}, \quad \ldots, \quad a_n = \frac{1}{n}, \quad \ldots.$$

Keine Zahl der Folge ist Null; aber wir erkennen, daß, je größer der Index $n$ wird, die Zahl $a_n$ desto näher an Null liegen muß. Wenn wir also um die Zahl 0 ein Intervall abgrenzen, das wir so klein wählen mögen, wie wir wollen, so wird von einem bestimmten Index ab jede Zahl der Zahlenfolge $a_n$ in dieses Intervall hineinfallen. Diesen Sachverhalt bringen wir dadurch zum Ausdruck, daß wir sagen: Die Zahlen $a_n$ *streben* mit wachsendem $n$ gegen 0, oder sie besitzen den *Grenzwert* (*Limes*) 0, oder die Zahlenfolge $a_1, a_2, a_3, \ldots$ *konvergiert* gegen 0.

Anschaulich bedeutet dies bei der Darstellung der Zahlen durch Punkte auf der Zahlengeraden, daß die Punkte $\frac{1}{n}$ sich mit wachsendem $n$ immer dichter gegen den Grenzpunkt 0 andrängen.

Ganz ähnlich liegen die Verhältnisse bei der Folge

$$a_1 = 1, \quad a_2 = -\frac{1}{2}, \quad a_3 = \frac{1}{3}, \quad a_4 = -\frac{1}{4}, \quad \ldots, \quad a_n = \frac{(-1)^{n-1}}{n}, \quad \ldots$$

Auch hier streben die Zahlen $a_n$ mit wachsendem $n$ gegen Null; der Unterschied ist nur der, daß sie bald größer, bald kleiner als der Grenzwert 0 sind, daß sie, wie man sagt, um den Grenzwert herum *oszillieren*.

---

[1]) D. h. der Quotient der Zahl $A(n)$ durch die Zahl $\frac{n}{\log n}$ unterscheidet sich von 1 um beliebig wenig, wenn nur $n$ genügend groß ist.

Symbolisch pflegt man die Konvergenz der Zahlenfolge gegen 0 durch die Gleichung

$$\lim_{n\to\infty} a_n = 0$$

oder auch gelegentlich durch die Abkürzung

$$a_n \to 0$$

auszudrücken. Die Buchstaben lim sind eine Abkürzung für das Wort limes (Grenze).

$$\mathbf{2.}\ a_{2m} = \frac{1}{m};\quad a_{2m-1} = \frac{1}{2m}.$$

Bei den obigen Beispielen wird mit wachsendem $n$ die Differenz zwischen $a_n$ und dem Grenzwert absolut genommen immer kleiner. Daß dies nicht immer so zu sein braucht, zeigt das Beispiel der Zahlenfolge

$$a_1 = \frac{1}{2},\ a_2 = 1,\ a_3 = \frac{1}{4},\ a_4 = \frac{1}{2},\ a_5 = \frac{1}{6},\ a_6 = \frac{1}{3}, \ldots;$$

d. h. allgemein: für gerades $n = 2m$ sei $a_n = a_{2m} = \frac{1}{m}$, für ungerades $n = 2m - 1$ sei $a_n = a_{2m-1} = \frac{1}{2m}$. Auch diese Zahlenfolge hat einen Grenzwert, nämlich den Grenzwert Null; d. h. in jedes noch so kleine Intervall um die Zahl Null herum werden von einem gewissen $n$ an alle Zahlen $a_n$ hineinfallen; aber doch nicht mehr so, daß jede Zahl näher als die vorangehende an dem Grenzwerte Null liegt.

$$\mathbf{3.}\ a_n = \frac{n}{n+1}.$$

Wir betrachten die Folge

$$a_1 = \frac{1}{2},\quad a_2 = \frac{2}{3}, \ldots,\quad a_n = \frac{n}{n+1}, \ldots,$$

wobei der ganzzahlige Index $n$ alle Werte $1, 2, 3, \ldots$ durchläuft. Schreiben wir $a_n = 1 - \frac{1}{n+1}$, so erkennen wir ohne weiteres, daß mit wachsendem $n$ die Zahl $a_n$ sich immer mehr dem Grenzwert 1 nähern wird, in dem Sinn, daß in jedes noch so kleine um die Zahl 1 abgegrenzte Zahlenintervall alle auf ein gewisses $a_N$ folgenden Zahlen $a_n$ hineinfallen müssen. Man schreibt: $\lim\limits_{n\to\infty} a_n = 1$.

Ganz ähnlich liegt es z. B. bei der Zahlenfolge

$$a_n = \frac{n^2 - 1}{n^2 + n + 1}.$$

Auch diese Zahlenfolge strebt mit wachsendem $n$ einem Grenzwert, und zwar dem Wert 1 zu: $\lim\limits_{n\to\infty} a_n = 1$; man erkennt dies z. B. folgendermaßen. Wir schreiben

$$a_n = 1 - \frac{n+2}{n^2+n+1} = 1 - r_n$$

und brauchen nur noch zu beweisen, daß mit wachsendem $n$ die Zahlen $r_n$ gegen 0 streben. Nun ist, sobald $n$ größer als 2 gewählt wird, sicherlich $n + 2 < 2n$ und $n^2 + n + 1 > n^2$. Also gilt dann für den Rest

$$0 < r_n < \frac{2n}{n^2} = \frac{2}{n} \qquad (n > 2),$$

und daraus erkennt man sofort, daß dieser Rest $r_n$ bei wachsendem $n$ gegen 0 strebt. Unsere Betrachtung gibt uns gleichzeitig eine Abschätzung dafür, um wieviel die Zahl $a_n$ (für $n > 2$) ungünstigstenfalls sich von dem Grenzwerte 1 unterscheiden kann; dieser Unterschied wird sicherlich nicht mehr als $\frac{2}{n}$ betragen.

Der eben betrachtete Grenzübergang bringt die von vornherein sehr plausible Tatsache zum Ausdruck, daß im Zähler und Nenner des Bruches für $a_n$ die Glieder mit den höchsten Exponenten bei großen Werten von $n$ überwiegen und für die Bildung des Quotienten den Ausschlag geben.

### 4. $a_n = \sqrt[n]{p}$.

Es sei $p$ irgend eine positive Zahl. Dann betrachten wir die Zahlenfolge $a_1, a_2, a_3, \ldots, a_n, \ldots$, wobei

$$a_n = \sqrt[n]{p}$$

gesetzt ist. Wir behaupten, daß

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \sqrt[n]{p} = 1$$

ist. Um hierfür den Beweis zu führen, stützen wir uns am einfachsten auf einen auch für andere Zwecke nützlichen Hilfssatz: *Ist $h$ eine positive Größe, so gilt*

$$(1) \qquad (1+h)^n > 1 + nh \quad \text{für} \quad n > 1.$$

Der Beweis dieses Hilfssatzes ergibt sich unmittelbar aus dem binomischen Satze

$$(1+h)^n = 1 + nh + \frac{n(n-1)}{1\cdot 2} h^2 + \cdots,$$

da alle Glieder rechts positiv sind[1]). Aus dem binomischen Satz ergibt

[1]) Wenn wir den binomischen Satz nicht als bekannt voraussetzen wollen, können wir den Beweis für die obige Ungleichung folgendermaßen führen: Nehmen wir an, die Ungleichung

$$(1+h)^m > 1 + mh$$

sich übrigens noch eine zweite nützliche Hilfsungleichung für positives $h$:

$$(1+h)^n > 1 + \frac{n(n-1)}{2} h^2 \quad \text{für} \quad n > 1\,,$$

aus welcher, sobald $n \geqq 2$, also $n - 1 \geqq \frac{n}{2}$ ist, folgt:

$$(1+h)^n > \frac{n^2}{4} h^2. \tag{2}$$

Wir unterscheiden den Fall $p > 1$ und den Fall $p < 1$ (wenn $p = 1$ ist, wird $\sqrt[n]{p}$ für jeden Wert von $n$ selbst gleich 1 sein, und unsere Behauptung stellt dann eine Trivialität dar).

Ist zunächst $p > 1$, so wird auch $\sqrt[n]{p} > 1$ sein; wir setzen dann $\sqrt[n]{p} = 1 + h_n$, wo $h_n$ eine positive von $n$ abhängige Größe ist, und haben unter Berücksichtigung unserer obigen Ungleichung (1)

$$p = (1 + h_n)^n > 1 + n h_n,$$

woraus sofort folgt

$$0 < h_n < \frac{p-1}{n}.$$

Wir erkennen also, daß bei wachsendem $n$ die Größe $h_n$ gegen 0 streben muß, wodurch die behauptete Konvergenz der Größen $a_n$ gegen die Zahl 1 bewiesen ist. Gleichzeitig haben wir ein Mittel an der Hand, um uns eine Schätzung darüber zu verschaffen, wie nahe wir uns bei der Wahl einer Zahl $n$ schon an dem Grenzwert 1 befinden; die Abweichung der Zahl $a_n$ von 1 ist nämlich sicherlich nicht größer als $\frac{p-1}{n}$.

Ist $p < 1$, so wird $\sqrt[n]{p}$ ebenfalls kleiner als 1 sein und daher gleich $\frac{1}{1+h_n}$ gesetzt werden können, wo $h_n$ eine positive Zahl ist. Daraus folgt, wiederum unter Berücksichtigung unserer Ungleichung (1),

$$p = \frac{1}{(1+h_n)^n} < \frac{1}{1 + n h_n}.$$

(Wenn wir den Nenner verkleinern, vergrößern wir den Bruch.) Nunmehr folgt

$$1 + n h_n < \frac{1}{p}$$

---

wäre für einen gewissen Wert $m > 1$ schon bewiesen; dann multiplizieren wir rechts und links mit $1 + h$ und erhalten

$$(1+h)^{m+1} > (1 + mh)(1+h) = 1 + (m+1)h + mh^2.$$

Lassen wir rechts die positive Größe $mh^2$ fort, so steht unsere Ungleichung für den Exponenten $m + 1$ vor uns. Sie folgt also, wenn sie für den Exponenten $m$ gilt, auch für den Exponenten $m + 1$; da sie nun für den Exponenten $m = 2$ richtig ist, so besteht sie auch für den Exponenten $m = 3$, daher auch für den Exponenten $m = 4$ usw., also für jeden Exponenten. — Dies ist ein einfaches Beispiel für die sog. vollständige Induktion, eine oft angewandte Schlußweise.

und somit
$$h_n < \frac{\frac{1}{p} - 1}{n},$$
woraus wiederum folgt, daß $h_n$ bei wachsendem $n$ gegen 0 strebt. Als reziproker Wert einer gegen 1 strebenden Größe strebt $p$ also offenbar selbst gegen 1.

### 5. $a_n = \alpha^n$.

Wir betrachten bei festem $\alpha$ die Zahlenfolge $a_n = \alpha^n$, wo $n$ die Reihe der positiven ganzen Zahlen durchläuft.

Es sei zunächst $\alpha$ eine positive Zahl, die kleiner ist als 1. Dann können wir $\alpha = \frac{1}{1+h}$ mit positivem $h$ setzen und erhalten, wieder unter Benutzung der obigen Ungleichung (1),
$$a_n = \frac{1}{(1+h)^n} < \frac{1}{1+nh} < \frac{1}{nh} = \frac{1}{h} \cdot \frac{1}{n}.$$
Da $h$ und daher auch $\frac{1}{h}$ hier eine nur von $\alpha$ abhängige, d. h. gegenüber der Veränderlichkeit von $n$ feste Zahl ist, erkennen wir aus dieser Beziehung, daß $\alpha^n$ mit wachsendem $n$ dem Grenzwert 0 zustrebt:
$$\lim_{n\to\infty} \alpha^n = 0 \qquad (0 < \alpha < 1).$$
Dieselbe Beziehung gilt auch, wenn $\alpha$ Null oder negativ, aber größer als $-1$ ist. Man erkennt dies sofort, da jedenfalls $\lim_{n\to\infty} |\alpha|^n = 0$ ist.

Ist $\alpha = 1$, so wird offenbar $\alpha^n$ stets gleich 1 sein, und wir werden dann als Grenzwert von $\alpha^n$ die Zahl 1 anzusehen haben.

Ist $\alpha > 1$, so setzen wir $\alpha = 1 + h$ mit positivem $h$ und erkennen aus unserer Ungleichung unmittelbar, daß $\alpha^n$ mit wachsendem $n$ keinem bestimmten Grenzwert zustrebt, sondern über alle Grenzen wächst. Wir drücken diesen Sachverhalt aus, indem wir sagen: $\alpha^n$ strebt mit wachsendem $n$ gegen Unendlich, oder: $\alpha^n$ wird unendlich; in Zeichen
$$\lim_{n\to\infty} \alpha^n = \infty \qquad (\alpha > 1).$$
*Das Symbol $\infty$ bedeutet jedoch,* worauf ausdrücklich hingewiesen werden soll, *nicht etwa eine Zahl, mit der man rechnen kann, wie mit irgend einer anderen Zahl*; Gleichungen oder Aussagen, bei denen es heißt, daß eine Größe unendlich ist oder unendlich wird, haben niemals den Sinn einer Gleichung zwischen bestimmten Zahlen. Trotzdem aber ist eine solche Ausdrucksweise und die Einführung des Symbols $\infty$ äußerst bequem, wie wir im folgenden noch des öfteren sehen werden.

Ist $\alpha = -1$, so wird $\alpha^n$ keinem Grenzwert zustreben, sondern, wenn $n$ die Zahlenreihe durchläuft, zwischen den Werten $+1$ und $-1$ hin- und herpendeln. Ebenso wird, wenn $\alpha < -1$ ist, der Wert von $\alpha^n$ zwar absolut genommen über alle Grenzen wachsen, aber in seinem Vorzeichen abwechselnd positiv und negativ sein.

## 6. Zur geometrischen Veranschaulichung der Grenzwerte von $a^n$ und $\sqrt[n]{p}$.

Betrachten wir die Kurven $y = x^n$ bzw. die Kurven $y = x^{\frac{1}{n}} = \sqrt[n]{x}$ und beschränken uns dabei der Bequemlichkeit halber auf nicht negative Werte von $x$, so werden die obigen Grenzwerte durch Figur 16 und 17 veranschaulicht. Wir erkennen bei den Kurven $y = x^n$, daß diese sich bei wachsendem $n$ im Intervall von 0 bis 1 immer mehr der $x$-Achse nähern, während sie außerhalb dieses Intervalles immer steiler ansteigen und sich immer genauer einer Parallelen zur $y$-Achse anschmiegen. Alle diese Kurven aber laufen durch den Punkt mit den Koordinaten $x = 1$, $y = 1$ und durch den Nullpunkt hindurch.

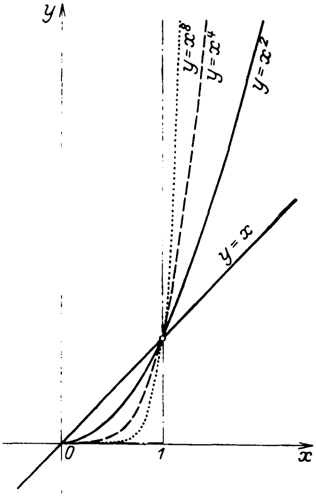



Fig. 16. $x^n$ für wachsende $n$.

Bei den Kurven $y = x^{\frac{1}{n}} = \sqrt[n]{x}$ nähert sich der Funktionswert $y$ für alle positiven Werte von $x$ immer mehr der 1, die Kurven also immer mehr der Parallelen im Abstande 1 zur $x$-Achse; andererseits müssen alle Kurven durch den Nullpunkt gehen. Die Kurven werden also in der Grenze sich dem Linienzug annähern, welcher aus der $y$-Achse zwischen den Ordinaten $y = 0$ und $y = 1$ einerseits, der Parallelen $y = 1$ zur $x$-Achse andererseits besteht. Beide Figuren stehen übrigens in engstem Zusammenhang miteinander, entsprechend der Tatsache, daß die Funktionen $y = \sqrt[n]{x}$ gerade die Umkehrfunktionen der $n$-ten Potenzen darstellen. Daraus können wir entnehmen, daß beide Figuren durch Umklappung um die Gerade $y = x$ ineinander übergehen.

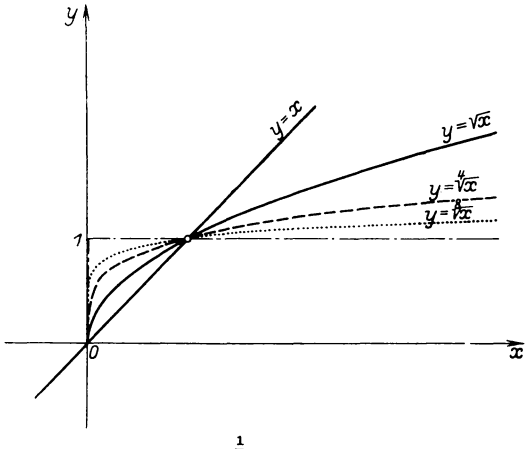



Fig. 17. $x^{\frac{1}{n}}$ für wachsende $n$.

## 7. Die geometrische Reihe.

Ein von der Schule her mehr oder weniger geläufiges Beispiel für einen Grenzwert liefert die *„geometrische Reihe“*

$$1 + q + q^2 + \cdots + q^{n-1} = S_n,$$

wobei die Zahl $q$ der *Quotient der Reihe* heißt. Der Wert dieser Summe läßt sich bekanntlich, wenn $q \neq 1$ ist, in folgender geschlossenen Form angeben:

$$S_n = \frac{1 - q^n}{1 - q},$$

wie man sofort erkennt, wenn man diese obige Summe $S_n$ mit $q$ multipliziert und die so entstehende Gleichung von der ursprünglichen Gleichung abzieht.

Es entsteht nun die Frage, was aus der Summe $S_n$ wird, wenn $n$ über alle Grenzen wächst. Das Resultat lautet: Die Summe $S_n$ hat einen bestimmten Grenzwert $S$, wenn $q$ zwischen $-1$ und $+1$ mit Ausschluß der Grenzen liegt, und zwar gilt dann

$$S = \lim_{n \to \infty} S_n = \frac{1}{1 - q}.$$

Um die Richtigkeit dieser Behauptung einzusehen, schreiben wir die Zahlen $S_n$ in die Form $S_n = \frac{1 - q^n}{1 - q} = \frac{1}{1 - q} - \frac{q^n}{1 - q}$; nun haben wir vorhin gezeigt, daß unter der Voraussetzung $|q| < 1$ die Größe $q^n$ — und damit auch $\frac{q^n}{1 - q}$ — mit wachsendem $n$ gegen 0 strebt, und daraus folgt sofort, daß die Zahl $S_n$ unter dieser Voraussetzung bei wachsendem $n$ den Grenzwert $\frac{1}{1 - q}$ besitzt, wie wir behauptet hatten.

Den Grenzübergang $1 + q + q^2 + \cdots + q^{n-1} \underset{n \to \infty}{\to} \frac{1}{1 - q}$ pflegt man auch dadurch zu kennzeichnen, daß man sagt, man könne die geometrische Reihe im Falle $|q| < 1$ ins Unendliche fortsetzen und *die Summe dieser unendlichen geometrischen Reihe sei der Ausdruck* $\frac{1}{1 - q}$.

Die Summen $S_n$ der endlichen geometrischen Reihe nennt man auch die *Teilsummen* oder *Partialsummen* der unendlichen geometrischen Reihe $1 + q + q^2 + \cdots$. (Man hat also die Zahlenfolge $S_1, S_2, \ldots, S_n, \ldots$ von der geometrischen Reihe scharf zu unterscheiden.)

Die Tatsache, daß die Teilsumme $S_n$ der geometrischen Reihe mit wachsendem $n$ gegen $S = \frac{1}{1 - q}$ konvergiert, drückt man auch durch die Redeweise aus: Die unendliche geometrische Reihe $1 + q + q^2 + \cdots$ *konvergiert* bei $|q| < 1$ gegen die Summe $S = \frac{1}{1 - q}$.

**8. $a_n = \sqrt[n]{n}$.**

Wir wollen zeigen, daß die Folge der Zahlen

$$a_1 = 1, \quad a_2 = \sqrt{2}, \quad a_3 = \sqrt[3]{3}, \quad \ldots, \quad a_n = \sqrt[n]{n}, \quad \ldots$$

mit wachsendem $n$ gegen 1 strebt:

$$\lim_{n\to\infty} \sqrt[n]{n} = 1.$$

Hier genügt die oben S. 22 benutzte Ungleichung (1) zum Beweise nicht mehr, sondern wir müssen uns, indem wir von vornherein $n \geqq 2$ voraussetzen, der Ungleichung (2) S. 23 bedienen. Wir können jedenfalls setzen

$$\sqrt[n]{n} = 1 + h_n,$$

wo $h_n$ eine von $n$ abhängige positive Zahl bedeutet, und es ist daher nach der Ungleichung (2)

$$n = (1 + h_n)^n > \frac{n^2}{4} h_n^2 \qquad (n \geqq 2),$$

woraus folgt

$$h_n^2 < \frac{4n}{n^2} = \frac{4}{n}.$$

Also ist gewiß $0 < h_n < \frac{2}{\sqrt{n}}$, und somit wird $h_n$ mit wachsendem $n$ gegen 0 streben. Dies aber bedeutet gerade unsere behauptete Limesrelation.

**9. $a_n = \sqrt{n+1} - \sqrt{n}$.**

Ich behaupte, daß

$$\lim_{n\to\infty} (\sqrt{n+1} - \sqrt{n}) = 0$$

ist.

Zu diesem Zwecke brauchen wir nur den zu untersuchenden Ausdruck in die Form

$$\sqrt{n+1} - \sqrt{n} = \frac{(\sqrt{n+1} - \sqrt{n})\,(\sqrt{n+1} + \sqrt{n})}{\sqrt{n+1} + \sqrt{n}} = \frac{1}{\sqrt{n+1} + \sqrt{n}}$$

zu setzen; an diesem Ausdruck erkennt man unmittelbar, daß er mit wachsendem $n$ gegen 0 strebt.

**10. $a_n = \frac{n}{2^n}$.**

Wir behaupten, daß bei wachsendem $n$ die Folge der Zahlen $a_n = \frac{n}{2^n}$ den Grenzwert 0 besitzt:

$$\lim_{n\to\infty} \frac{n}{2^n} = 0.$$

Wir brauchen zum Beweise nur zu beachten, daß für $n \geqq 2$ nach der Abschätzung (2)

$$2^n = (1+1)^n > \frac{1}{4} n^2$$

ist, woraus

$$a_n = \frac{n}{2^n} < \frac{4}{n}$$

folgt. Somit ist bewiesen, daß unsere Zahlenfolge den Grenzwert 0 hat.

## § 6. Genauere Erörterung des Grenzwertbegriffes.

### 1. Allgemeines.

Aus den Beispielen des vorangehenden Paragraphen abstrahieren wir den folgenden allgemeinen Begriff des Grenzwertes: *Wenn eine unendliche Folge von Zahlen $a_1, a_2, a_3, \ldots, a_n, \ldots$ gegeben ist und wenn es eine weitere Zahl $g$ gibt, derart, daß in jedes noch so kleine um $g$ abgegrenzte Intervall alle Zahlen $a_n$ mit Ausnahme von höchstens endlich vielen*[1]) *hineinfallen, so sagen wir: die Zahl $g$ ist der Grenzwert der Zahlenfolge $a_1, a_2, a_3, \ldots, a_n, \ldots$, oder: die Zahlenfolge $a_1, a_2, \ldots$ konvergiert gegen $g$*; in Zeichen: $\lim\limits_{n\to\infty} a_n = g$. Darunter ist, wie ausdrücklich bemerkt sei, auch z. B. der triviale Fall mit einbegriffen, daß alle Zahlen $a_n$ einander gleich sind und somit auch mit dem Grenzwert $g$ zusammenfallen.

In allen vorangehenden Beispielen lagen die Verhältnisse so, daß der Grenzwert der betrachteten Zahlenfolge wieder eine uns schon bekannte Zahl war. Wenn der Begriff des Grenzwertes uns nichts anderes lieferte als die Erkenntnis, daß wir gewisse bekannte Zahlen durch gewisse Folgen anderer bekannter Zahlen beliebig genau annähern können, so würde mit ihm nicht viel gewonnen sein. Die Fruchtbarkeit des Grenzwertbegriffes für die Analysis beruht aber wesentlich darauf, daß wir durch Grenzwerte bekannter Zahlen neue unmittelbar nicht mehr bekannte oder ausdrückbare Zahlen erfassen können.

Die ganze höhere Analysis bildet eine Kette von Beispielen für dieses Faktum, das uns in den späteren Kapiteln immer deutlicher werden wird. Die Darstellung der irrationalen Zahlen als Grenzwerte rationaler Zahlen kann man als ein erstes Beispiel auffassen. Weitere werden wir sogleich noch in diesem Paragraphen kennenlernen. Bevor wir uns diesen zuwenden, wollen wir jedoch noch einige allgemeine Bemerkungen voranschicken.

Wie erkennt man an einer gegebenen Zahlenfolge $a_1, a_2, a_3, \ldots, a_n, \ldots$ von vornherein, daß sie einem Grenzwert zustrebt, auch wenn man diesen Grenzwert nicht vorher angeben kann? Diese wichtige Frage

[1]) Statt „mit Ausnahme von höchstens endlich vielen" könnten wir auch „*von einem gewissen Werte des Index $n$ an*" sagen.

wird uns durch die *Konvergenzregel von Cauchy* beantwortet, welche folgendermaßen lautet: *Es werde in der Zahlenfolge* $a_1, a_2, \ldots, a_n, \ldots$ *der Ausdruck* $|a_n - a_m|$ *kleiner als eine beliebig klein gewählte positive Zahl* $\varepsilon$, *sobald man nur die Zahlen* $n$ *und* $m$ *beide hinreichend groß wählt, etwa größer als eine — im allgemeinen natürlich noch von* $\varepsilon$ *abhängige — Zahl* $N = N(\varepsilon)$, welche für $\varepsilon \to 0$ selbst über alle Grenzen streben wird; *dann folgt die Existenz eines Grenzwertes*

$$g = \lim_{n \to \infty} a_n.$$

Die Bedeutung des Cauchyschen Konvergenzkriteriums liegt darin, daß wir mit seiner Hilfe von dem Grenzwert einer Zahlenfolge auf Grund einer Betrachtung der Zahlenfolge selbst sprechen können, ohne vorher den Grenzwert anderweitig charakterisieren zu müssen.

Umgekehrt folgt aus der Existenz eines Grenzwertes $\lim\limits_{n \to \infty} a_n$ das Kleinwerden von $|a_n - a_m|$. Denn wenn die Zahlenfolge $a_1, a_2, \ldots$ gegen den Grenzwert $g$ strebt, so muß für beliebig kleines $\varepsilon > 0$ nach der Definition der Konvergenz $|g - a_n| < \frac{\varepsilon}{2}$ und $|g - a_m| < \frac{\varepsilon}{2}$ sein, wenn nur $n$ und $m$ groß genug sind; also $|a_n - a_m| = |(g - a_m) - (g - a_n)| \leqq |g - a_m| + |g - a_n| < \varepsilon$. Diese Beziehung drückt, da $\varepsilon$ beliebig klein gewählt werden kann, unsere Behauptung aus.

Die Aussage des Cauchyschen Kriteriums ist äußerst einfach und fast unmittelbar einleuchtend, wenn man sich die Zahlen auf der Zahlengeraden repräsentiert denkt; es besagt nämlich, daß eine Zahlenfolge sicher einen Grenzwert hat, wenn von einer gewissen Stelle $N$ ab die Werte der Zahlenfolge sich überhaupt nur noch in einem kleinen Spielraum ändern können, der beliebig klein wird, wenn $N$ hinreichend groß gewählt ist. Einen präzisen Beweis für unser Kriterium werden wir im Anhang dieses Kapitels, § 1, Nr. 2, besonders erbringen und für den Moment uns mit unserer anschaulichen Betrachtung begnügen.

Besonders einfach läßt sich die Frage nach der Konvergenz einer gegebenen Folge gegen einen Grenzwert entscheiden, wenn die Folge eine sog. *monotone Folge* ist, d. h. wenn entweder jede Zahl der Folge größer ist als die vorangehende (monoton zunehmende Folge) oder jede Zahl kleiner ist als die vorangehende (monoton abnehmende Folge). Es gilt der Satz: *Jede monoton zunehmende Folge, deren Zahlen nach oben beschränkt sind, d. h. unterhalb einer festen Zahl liegen, besitzt einen Grenzwert; ebenso besitzt jede monoton abnehmende Folge, deren Zahlen nicht unter eine feste untere Schranke sinken können, einen Grenzwert.* Auch diese beiden Sätze wollen wir vorläufig unter Hinweis auf die nähere Begründung im Anhang als anschaulich evident betrachten. Eine konvergente monoton zunehmende Folge kann natürlich nur gegen einen Grenzwert streben, der größer als jede

Zahl der Folge ist, während sich bei monoton abnehmenden Folgen die Zahlen gegen einen Grenzwert anhäufen, der kleiner ist als jede Zahl der Folge. So z. B. bilden die Zahlen $\frac{1}{n}$ eine monoton abnehmende Folge mit dem Grenzwert 0, die Zahlen $1 - \frac{1}{n}$ eine monoton zunehmende Folge mit dem Grenzwert 1.

Für manche Fälle ist es angenehmer, an Stelle der Forderung des monotonen Zunehmens einer Folge die schwächere Forderung zu stellen, daß die Glieder der Folge jedenfalls nicht abnehmen; mit anderen Worten, zu erlauben, daß aufeinanderfolgende Glieder der Folge einander gleich sind. Man spricht dann von einer monoton nicht abnehmenden Folge oder von einer monoton zunehmenden Folge im schwächeren Sinne. Unser Grenzwertsatz bleibt auch für solche Folgen bestehen, ebenso auch für Zahlenfolgen, welche monoton nicht zunehmen oder im schwächeren Sinne monoton abnehmen.

Es ist nützlich, sich zu vergegenwärtigen, *daß jede konvergente Zahlenfolge beschränkt ist*; d. h. zu jeder Zahlenfolge $a_1, a_2, a_3, \ldots$, für welche ein Grenzwert $\xi$ existiert, gibt es eine von $n$ unabhängige positive Zahl $M$, so daß für alle $a_n$ der Folge $|a_n| < M$ gilt.

Aus unseren Definitionen folgt dieser Satz leicht: Sicher gibt es nämlich einen Index $N$, so daß für $n > N$ immer $|a_n - \xi| < 1$ ist. Unter den $N$ Zahlen $|a_1 - \xi|, |a_2 - \xi|, \ldots, |a_N - \xi|$ sei $A$ die größte. Dann dürfen wir $M = |\xi| + A + 1$ setzen. Denn es ist sicher nach der Definition von $A$ die Ungleichung $|a_n - \xi| < A + 1$ für $n = 1, 2, \ldots, N$ erfüllt, während für $n > N$ gilt $|a_n - \xi| < 1 \leqq A + 1$.

Eine Zahlenfolge, welche nicht konvergiert, heißt *divergent*. Wachsen mit wachsendem $n$ die Zahlen $a_n$ positiv über alle Grenzen, so sprechen wir von einer nach $+\infty$ divergenten Folge und schreiben, wie schon früher gelegentlich: $\lim a_n = \infty$. Entsprechend schreiben wir $\lim a_n = -\infty$, wenn bei wachsendem $n$ die Zahlen $-a_n$ positiv über alle Grenzen streben. Divergenz einer Folge kann sich aber auch in anderer Weise ausprägen, wie z. B. bei der Folge $a_1 = -1$, $a_2 = +1$, $a_3 = -1$, $a_4 = +1, \ldots$, deren Glieder zwischen zwei verschiedenen Werten hin- und herpendeln[1]).

## 2. Rechnen mit Grenzwerten.

Endlich noch eine letzte Bemerkung über das Rechnen mit Grenzwerten. Aus dem Begriff des Grenzwertes folgt fast unmittel-

[1]) Eine weitere nützliche Bemerkung: An dem Konvergenzverhalten einer Folge ändert sich nichts, wenn wir endlich viele Zahlen $a_n$ daraus fortlassen. Wir machen im folgenden häufig davon Gebrauch, indem wir auch bei solchen Folgen von Konvergenz oder Divergenz reden, wo für endlich viele $n$ kein $a_n$ definiert ist.

bar, daß man die elementaren Rechenoperationen der Addition, Multiplikation, Subtraktion und Division mit Grenzwerten nach den folgenden Regeln ausführen kann:

Ist $a_1, a_2, \ldots$ eine Folge mit dem Grenzwert $a$ und $b_1, b_2, \ldots$ eine Folge mit dem Grenzwert $b$, so hat auch die Folge der Zahlen $c_n = a_n + b_n$ einen Grenzwert, und es ist

$$\lim_{n\to\infty} c_n = a + b.$$

Ebenso konvergiert die Folge der Zahlen $c_n = a_n b_n$ und genügt der Gleichung

$$\lim_{n\to\infty} c_n = a b.$$

Falls der Grenzwert $b$ von 0 verschieden ist, gilt weiter auch die Gleichung

$$\lim_{n\to\infty} c_n = \frac{a}{b},$$

wenn $c_n = \frac{a_n}{b_n}$ gesetzt wird. In Worten: Man kann die rationalen Rechenoperationen mit dem Prozesse der Grenzwertbildung *vertauschen; d. h. es ist für das Resultat gleichgültig, ob wir zunächst einen Grenzübergang und dann eine rationale Rechenoperation vornehmen oder umgekehrt.*

Zum Beweise dieser einfachen Regeln genügt es, ein Beispiel herauszugreifen, nach dessen Muster sich der Leser die weiteren Behauptungen leicht selbst klar machen kann. Wir betrachten etwa die Multiplikation von Grenzwerten. Die Beziehungen $a_n \to a$ und $b_n \to b$ besagen folgendes: Wenn wir eine noch so kleine Größe $\varepsilon$ vorgeben, so brauchen wir nur $n$ größer als $N$ zu nehmen, wobei $N = N(\varepsilon)$ eine zu $\varepsilon$ hinzuzubestimmende hinreichend große Zahl ist, und es wird sicherlich $|a - a_n| < \varepsilon$, sowie $|b - b_n| < \varepsilon$ sein. Schreiben wir $ab - a_n b_n = b(a - a_n) + a_n(b - b_n)$ und beachten, daß es eine von $n$ unabhängige positive Schranke $M$ geben muß, so daß $|a_n| < M$ ist, dann erhalten wir $|ab - a_n b_n| \leqq |b|\,|a - a_n| + |a_n|\,|b - b_n| < (|b| + M)\,\varepsilon$. Da die Größe $(|b| + M)\,\varepsilon$ mit $\varepsilon$ gegen 0 strebt, d. h. beliebig klein wird, wenn nur $\varepsilon$ hinreichend klein gewählt wird, so sehen wir, daß bei genügend großem $n$ tatsächlich die Zahlen $ab$ und $a_n b_n$ sich voneinander beliebig wenig unterscheiden; das aber ist gerade die durch die Gleichung $ab = \lim\limits_{n\to\infty} a_n b_n$ gemachte Aussage.

Auf Grund dieser Regeln bestimmen sich viele Grenzwerte besonders einfach; z. B. wird

$$\lim_{n\to\infty} \frac{n^2 - 1}{n^2 + n + 1} = \lim_{n\to\infty} \frac{1 - \frac{1}{n^2}}{1 + \frac{1}{n} + \frac{1}{n^2}} = 1,$$

da wir in dem zweiten Ausdruck sofort den Grenzübergang im Zähler und Nenner für sich vornehmen können.

Noch eine weitere einfache und selbstverständliche Regel verdient eine besondere Formulierung: Ist $\lim a_n = a$ und $\lim b_n = b$, ist ferner $a_n > b_n$ für jedes $n$, so gilt sicherlich $a \geqq b$. Wir dürfen aber keineswegs allgemein erwarten, daß $a > b$ sei; das lehrt uns das Beispiel der beiden Folgen $a_n = \frac{1}{n}$, $b_n = \frac{1}{2n}$, für welche $a = b = 0$ ist.

Wir wenden uns nun einigen neuen wichtigen Beispielen zu, die das oben geschilderte allgemeine Prinzip näher erläutern sollen.

## 3. Die Zahl $e$.

Als erstes Beispiel für die Erzeugung einer nicht von vornherein charakterisierbaren Zahl als Grenzwert einer Folge bekannter Zahlen betrachten wir die Summe

$$S_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!}.$$

Ich behaupte, daß diese Zahl $S_n$ mit wachsendem $n$ einem bestimmten Grenzwert zustrebt.

Um die Existenz des Grenzwertes zu beweisen, beachten wir, daß die Zahlen $S_n$ mit wachsendem $n$ monoton zunehmen. Andrerseits gilt für alle $n$

$$S_n < 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{n-1}} = 1 + \frac{1 - \frac{1}{2^n}}{1 - \frac{1}{2}} < 3.$$

Die $S_n$ besitzen also die obere Schranke 3 und haben somit als monoton wachsende Zahlenfolge einen Limes, den wir mit $e$ bezeichnen:

$$e = \lim_{n \to \infty} S_n.$$

Ich behaupte nun weiter, daß gegen die durch den obigen Grenzwert *definierte* Zahl $e$ auch die Zahlenfolge

$$T_n = \left(1 + \frac{1}{n}\right)^n$$

konvergiert.

Der Beweis ist einfach und zugleich lehrreich für das Operieren mit Grenzwerten. Nach dem binomischen Satze, den ich hier voraussetzen will, ist

$$T_n = \left(1 + \frac{1}{n}\right)^n = 1 + n \cdot \frac{1}{n} + \frac{n(n-1)}{2!}\frac{1}{n^2} + \cdots + \frac{n(n-1)(n-2)\ldots 1}{n!} \cdot \frac{1}{n^n}$$
$$= 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \cdots + \frac{1}{n!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{n-1}{n}\right).$$

Man erkennt hieraus sofort, daß einmal $T_n \leqq S_n$ ist und daß zweitens auch die $T_n$ eine monoton ansteigende Folge bilden [1]), woraus die Existenz des Grenzwertes $\lim\limits_{n\to\infty} T_n = T$ folgt. Um die Gleichung $T = e$ zu beweisen, beachten wir, daß jedenfalls

$$T_m > 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{m}\right) + \cdots + \frac{1}{n!}\left(1 - \frac{1}{m}\right)\cdots\left(1 - \frac{n-1}{m}\right)$$

ist, sobald $m > n$ wird. Lassen wir nun, indem wir $n$ festhalten, $m$ über alle Grenzen wachsen, so entsteht links der Grenzwert $T$, rechts gerade die Zahl $S_n$, und wir erhalten $T \geqq S_n$. Somit wird schließlich $T \geqq S_n \geqq T_n$. Nunmehr erst lassen wir $n$ wachsen, wobei $T_n$ gegen $T$ strebt und aus der hingeschriebenen Doppelungleichung sofort $T = \lim\limits_{n\to\infty} S_n = e$ folgt. Dies aber war unsere Behauptung.

Wir werden übrigens auf diese Zahl $e$ später (drittes Kapitel, § 6) noch von anderen Gesichtspunkten her geführt werden.

### 4. Die Zahl $\pi$ als Grenzwert.

Eine schon von der Schule her geläufige Grenzwertbetrachtung, die im Kern auf das klassische Altertum (Archimedes) zurückgeht, ist diejenige, durch welche man die Zahl $\pi$ definiert. Die geometrische Bedeutung von $\pi$ ist der Flächeninhalt des Kreises mit dem Radius 1. Die Existenz dieser Zahl $\pi$ entnehmen wir also der Anschauung, indem wir es als selbstverständlich ansehen, daß dieser Flächeninhalt sich durch eine (rationale oder irrationale) Zahl messen läßt, die wir dann eben einfach mit $\pi$ bezeichnen. Mit dieser Festsetzung aber ist wenig gewonnen, wenn man die Zahl wirklich mit einer gewissen Genauigkeit berechnen will. Man kann dann nicht anders vorgehen als so, daß man die Zahl durch einen Grenzprozeß darstellt, nämlich sie als Grenzwert einer Zahlenfolge wohlbekannter und leicht zu berechnender Zahlen auffaßt. Schon Archimedes hat in seiner Exhaustionsmethode diesen Weg beschritten, indem er den Kreis durch immer feiner sich ihm anschmiegende regelmäßige Polygone von wachsender Seitenzahl mehr und mehr ausschöpfte. Bezeichnen wir den Inhalt des dem Kreise einbeschriebenen $m$-Ecks mit $f_m$, so berechnet sich aus ihm der Inhalt des $2m$-Ecks durch die aus der Elementarmathematik bekannte Formel

$$f_{2m} = \frac{m}{2}\sqrt{2 - 2\sqrt{1 - \left(\frac{2 f_m}{m}\right)^2}}\,.$$

[1]) Wir erhalten nämlich $T_{n+1}$ aus $T_n$, indem wir die Faktoren $1 - \frac{1}{n}, 1 - \frac{2}{n}, \ldots$ durch die größeren $1 - \frac{1}{n+1}, 1 - \frac{2}{n+1}, \ldots$ ersetzen und ein letztes positives Glied hinzufügen.

Lassen wir $m$ nicht die Reihe aller Zahlen, sondern nur die Reihe der Potenzen von 2 durchlaufen, $m = 2^n$, bilden wir mit anderen Worten gerade diejenigen regelmäßigen Polygone, deren Ecken wir durch fortgesetzte Halbierung des Kreisumfanges erhalten, so wird der Flächeninhalt des Kreises durch den Grenzwert

$$\pi = \lim_{n\to\infty} f_{2^n}$$

gegeben sein. Diese Grenzwertdarstellung von $\pi$ ist tatsächlich eine Handhabe zur angenäherten numerischen Berechnung; denn wir können, ausgehend von dem Werte $f_4 = 2$, der Reihe nach die Zahlen unserer gegen $\pi$ strebenden Zahlenfolge berechnen.

Es sind dies übrigens Dinge, die dem Leser mehr oder weniger geläufig sein werden. Worauf ich aber schon hier hinweisen will, ist, daß die Berechnung von Flächeninhalten durch Ausschöpfung mittelst leicht berechenbarer geradlinig begrenzter Flächenstücke die Grundlage für den im nächsten Kapitel zu entwickelnden Integralbegriff gibt.

## 5. Das arithmetisch-geometrische Mittel.

Sind $\alpha$ und $\beta$ zwei positive Zahlen, von denen wir $\beta \leqq \alpha$ annehmen wollen, so ist immer ihr arithmetisches Mittel mindestens gleich dem geometrischen Mittel, d. h. es gilt immer

$$\frac{\alpha+\beta}{2} \geqq \sqrt{\alpha\beta};$$

das Gleichheitszeichen hat nur im Falle $\alpha = \beta$ statt. Der Beweis dieser Tatsache besteht in der Bemerkung, daß

$$\alpha + \beta - 2\sqrt{\alpha\beta} = (\sqrt{\alpha} - \sqrt{\beta})^2$$

als Quadrat nicht negativ ist und nur für $\alpha = \beta$ verschwindet.

Wir setzen nun $\alpha = a_0$ und $\beta = b_0$, wo $a_0 > b_0 > 0$. Dann bilden wir das arithmetische Mittel

$$a_1 = \frac{a_0 + b_0}{2}$$

und das geometrische Mittel

$$b_1 = \sqrt{a_0 b_0}\,.$$

Offenbar ist $b_0 < b_1 < a_1 < a_0$. Weiter bilden wir jetzt sowohl das arithmetische als auch das geometrische Mittel aus $a_1$ und $b_1$:

$$a_2 = \frac{a_1 + b_1}{2}, \qquad b_2 = \sqrt{a_1 b_1}$$

mit $b_0 < b_1 < b_2 < a_2 < a_1 < a_0$, und dann weiter

$$a_3 = \frac{a_2 + b_2}{2}, \qquad b_3 = \sqrt{a_2 b_2}$$

usw. Man erkennt sofort, daß die Zahlenfolge $a_n$ den Ungleichungen

$$a_0 > a_1 > a_2 > a_3 > \cdots > b_0,$$

die Zahlenfolge $b_n$ den Beziehungen

$$b_0 < b_1 < b_2 < b_3 < \cdots < a_0$$

genügt. Die $a_n$ bilden also eine absteigende, die $b_n$ eine ansteigende monotone Zahlenfolge. Da aber diese Zahlen sämtlich zwischen den Schranken $b_0$ und $a_0$ liegen, so existieren nach unseren obigen Betrachtungen die Grenzwerte

$$\lim_{n\to\infty} a_n = A \quad \text{und} \quad \lim_{n\to\infty} b_n = B.$$

Nun erkennt man leicht, daß $A = B$ sein muß; denn es ist

$$A = \lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{a_{n-1} + b_{n-1}}{2} = \frac{1}{2}\left(\lim_{n\to\infty} a_{n-1} + \lim_{n\to\infty} b_{n-1}\right) = \frac{A}{2} + \frac{B}{2},$$

also $A = B$.

Diesen gemeinsamen Grenzwert nennt man das *arithmetisch-geometrische Mittel* der Zahlen $a_0$ und $b_0$.

## § 7. Der Begriff des Grenzwertes bei stetigen Veränderlichen.

Wir haben bisher den Grenzwert von Zahlenfolgen, d. h. Funktionen $a_n$ einer ganzzahligen Veränderlichen $n$, betrachtet. Der Begriff des Grenzwertes tritt jedoch vielfach auch in Verbindung mit dem Begriff einer stetigen Veränderlichen $x$ und einer Funktion $f(x)$ bei folgender Frage auf:

Wir denken uns die Funktion durch eine Kurve repräsentiert und einen Punkt mit der Abszisse $x$ auf der Kurve so bewegt, daß die Abszisse sich immer mehr einer bestimmten Zahl $\xi$ nähert, — dabei kann die Bewegung auch hin und her laufen. Dann kann es der Fall sein, daß auch der Funktionswert $f(x)$ einem bestimmten Grenzwert zustrebt. Einen solchen Grenzübergang können wir dadurch präzis formulieren, daß wir ihn in folgender Weise auf Grenzübergänge mit Zahlenfolgen zurückführen.

Ist $x$ eine stetig veränderliche Größe und $\xi$ ein bestimmter Wert, so sagen wir: $x$ strebt gegen $\xi$, oder symbolisch ausgedrückt: $x \to \xi$, wenn $x$ eine Folge von Werten $x_1, x_2, x_3, \ldots, x_n, \ldots$ durchläuft, *welche alle von $\xi$ verschieden sind* und für welche $x_n \to \xi$ gilt. Ob die Wertefolge $x_n$ *monoton von rechts* oder *monoton von links* gegen $\xi$ strebt, ob sie um den Wert $\xi$ *oszilliert*, ist dabei ganz gleichgültig. Es sei nun $f(x)$ eine Funktion der stetigen Veränderlichen $x$. Dann sagen wir, daß für $x \to \xi$ der Funktionswert $f(x)$ gegen einen Grenzwert $g$ strebt, in Zeichen

$$\lim_{x\to\xi} f(x) = g,$$

wenn folgende Bedingung erfüllt ist: Für jede unserer Zahlenfolgen $x_1, x_2, \ldots, x_n, \ldots$ soll der Grenzwert der Zahlen $f(x_n)$ existieren, und dieser Grenzwert soll nicht von der speziellen Wahl der Folge abhängen[1]), sondern für alle Folgen dieselbe Zahl $g$ darstellen.

Eine andere, gleichwertige Formulierung ist die folgende: Grenzt man ein beliebig kleines Intervall $J_y$ um die Zahl $g$ ab, so fallen alle Funktionswerte $f(x)$ in dieses Intervall, wenn nur der von $\xi$ verschiedene Punkt $x$ hinreichend nahe an der Stelle $\xi$ liegt, d. h. innerhalb eines Intervalles $J_x$ um die Stelle $\xi$, dessen Länge von dem vorgegebenen Intervall $J_y$ um die Zahl $g$ noch abhängen kann und im allgemeinen um so kleiner sein wird, je kleiner $J_y$ ist.

Wir wollen uns diese abstrakte Definition an einigen einfachen Beispielen klar machen und betrachten zunächst die Funktion

$$f(x) = \frac{\sin x}{x}.$$

Wir behaupten, daß

$$\lim_{x \to 0} \frac{\sin x}{x} = 1$$

ist. Diese Behauptung können wir nicht etwa dadurch beweisen, daß wir den Grenzübergang in Zähler und Nenner für sich ausführen; denn Zähler und Nenner verschwinden für $x = 0$, und das Symbol $\frac{0}{0}$ als solches hat gar keinen Sinn. Wir schlagen zum Beweise dieser Limesgleichung folgenden Weg ein.

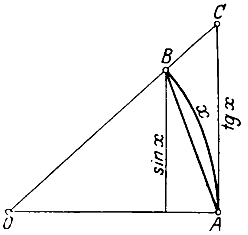


Fig. 18.

Aus Figur 18 entnehmen wir durch Vergleichen der Flächeninhalte der Dreiecke $OAB$ und $OAC$ und des Sektors $OAB$ für $0 < x < \frac{\pi}{2}$

$$\sin x < x < \operatorname{tg} x;$$

daraus folgt für $0 < |x| < \frac{\pi}{2}$:

$$1 < \frac{x}{\sin x} < \frac{1}{\cos x}.$$

Mithin ist der Quotient $\frac{\sin x}{x}$ zwischen den Grenzen 1 und $\cos x$ eingeschlossen. Mit $x \to 0$ strebt bekanntlich $\cos x$ gegen 1, und daraus

[1]) Diese zweite Forderung ist übrigens eine Konsequenz der ersten. D. h. wenn für jede Zahlenfolge $x_n$ mit $x_n \to \xi$ ein Grenzwert $g = \lim\limits_{n\to\infty} f(x_n)$ existiert, dann ist dieser Grenzwert derselbe für alle Folgen. Ist nämlich $x'_n$ eine zweite Zahlenfolge mit $x'_n \to \xi$ und dem Grenzwert $g' = \lim f(x'_n)$, so strebt auch die Zahlenfolge $x_1, x'_1, x_2, x'_2, \ldots, x_n, x'_n, \ldots$ gegen $\xi$, und daher existiert nach Voraussetzung auch ein Grenzwert $g''$ der Wertefolge $f(x_1), f(x'_1), \ldots, f(x_n), f(x'_n), \ldots$. Da dieser Grenzwert sich aber bei hinreichend großem $n$ sowohl von $f(x_n)$ als auch von $f(x'_n)$ nur beliebig wenig unterscheidet, so muß er sowohl mit $g$ als auch mit $g'$ identisch sein, woraus die behauptete Gleichheit von $g$ und $g'$ folgt.

folgt unmittelbar, daß der Quotient $\frac{\sin x}{x}$ sich von 1 nur beliebig wenig unterscheiden kann, wenn nur $x$ hinreichend nahe an 0 liegt. Dies aber ist gerade der Inhalt der behaupteten Limesgleichung.

Aus dem bewiesenen Resultat folgt die Relation

$$\lim_{x\to 0} \frac{\operatorname{tg} x}{x} = \lim_{x\to 0} \frac{\sin x}{x} \lim_{x\to 0} \frac{1}{\cos x} = 1$$

und die weitere

$$\lim_{x\to 0} \frac{1-\cos x}{x} = 0.$$

Diese letztere ergibt sich aus der für $0 < |x| < \frac{\pi}{2}$ gültigen Formel

$$\frac{1-\cos x}{x} = \frac{(1-\cos x)(1+\cos x)}{x(1+\cos x)} = \frac{1-\cos^2 x}{x(1+\cos x)} = \frac{\sin x}{x} \cdot \frac{1}{1+\cos x} \cdot \sin x.$$

Für $x \to 0$ strebt rechts der erste Faktor gegen 1, der zweite gegen $\frac{1}{2}$ und der dritte gegen 0, das Produkt also gegen 0, wie wir behaupteten.

Zum Schluß dieser Betrachtung über Grenzwerte bei stetigen Veränderlichen sei noch bemerkt, daß wir selbstverständlich auch Grenzwerte betrachten können, bei denen die stetige Veränderliche $x$ über alle Grenzen wächst. Z. B. ist ohne weiteres die Schreibweise verständlich:

$$\lim_{x\to\infty} \frac{x^2+1}{x^2-1} = \lim_{x\to\infty} \frac{1+\frac{1}{x^2}}{1-\frac{1}{x^2}} = 1.$$

Sie bedeutet, daß die Zahl links sich von 1 um beliebig wenig unterscheidet, sobald nur die Zahl $x$ hinreichend groß gewählt ist.

## § 8. Der Begriff der Stetigkeit.

### 1. Definitionen.

Wir haben schon in § 2 den Begriff der Stetigkeit an Beispielen veranschaulicht. Nunmehr sind wir mit Hilfe des Grenzwertbegriffes in der Lage, den Begriff der Stetigkeit vollständig zu präzisieren.

Das Bild einer in einem Intervall stetigen Funktion war für uns eine Kurve über diesem Intervall, die aus einem nicht zerreißenden Stück besteht; wir sagten auch, daß die Änderung der Funktion $y$ beliebig klein bleiben muß, wenn nur die Änderung der unabhängigen Veränderlichen $x$ sich auf ein hinreichend kleines Intervall beschränkt. Diesen Sachverhalt pflegt man etwas umständlicher, aber präziser folgendermaßen zu formulieren: Eine Funktion $f(x)$ heißt in dem Punkte $\xi$ *stetig*, wenn sie folgende Eigenschaft besitzt: Für die Stelle $\xi$ wird der Funktionswert $f(\xi)$ mit beliebig vorgegebener Genauigkeit $\varepsilon$

durch alle Funktionswerte $f(x)$ angenähert, für die $x$ hinreichend nahe an der Stelle $\xi$ liegt; d. h. ist $\varepsilon$ eine beliebig klein gewählte positive Zahl, so kann man zu ihr eine andere positive Zahl $\delta = \delta(\varepsilon)$ hinzubestimmen, derart, daß für alle Punkte, für welche $|x - \xi| < \delta$ ist, auch $|f(x) - f(\xi)| < \varepsilon$ wird. Mit noch anderen Worten: die Stetigkeitsforderung verlangt, daß für die Stelle $\xi$ die Limesgleichung

$$\lim_{x \to \xi} f(x) = f(\xi)$$

gilt. Der Funktionswert an der Stelle $\xi$ läßt sich als Grenzwert der Funktionswerte für die Zahlen jeder beliebigen Folge auffassen, welche gegen die Stelle $\xi$ konvergiert.

Es ist wichtig, zu beachten, daß unsere Forderung zweierlei verlangt, nämlich erstens die Existenz des Grenzwertes $\lim\limits_{x \to \xi} f(x)$ und zweitens die Übereinstimmung dieses Grenzwertes mit dem an der Stelle $\xi$ definierten Funktionswert $f(\xi)$.

Nachdem so die Stetigkeit *in einem Punkte* erklärt ist, definieren wir: Eine Funktion $f(x)$ heißt *in einem Intervalle stetig*, wenn sie in jedem Punkte des Intervalles stetig ist.

## 2. Unstetigkeitspunkte.

Man versteht den Begriff der Stetigkeit besser, wenn man sich ihn an seinem Gegenteil, an dem Begriff der Unstetigkeit erläutert. Die einfachste Art von Unstetigkeiten zeigen solche Stellen, bei denen die Funktion einen *Sprung* macht, d. h. bei denen die Funktionswerte sich verschiedenen, aber bestimmten Grenzwerten nähern, je nachdem, ob wir von rechts oder links der Sprungstelle zustreben; ob und wie in der Sprungstelle selbst der Funktionswert definiert ist, bleibt dahingestellt. Z. B. hat die durch die Gleichungen

$$f(x) = 0 \text{ für } x^2 > 1, \quad f(x) = 1 \text{ für } x^2 < 1, \quad f(x) = \frac{1}{2} \text{ für } x^2 = 1$$

definierte Funktion an den Stellen $\xi = 1$ und $\xi = -1$ Unstetigkeitspunkte. Die Grenzwerte von rechts und von links bei Annäherung an diese Stellen unterscheiden sich um 1 (und die Funktionswerte stimmen mit keinem der Grenzwerte überein, sondern sind deren arithmetisches Mittel).

Nebenbei bemerkt, wird uns unsere Funktion mit Hilfe des Grenzbegriffes durch den Ausdruck

$$f(x) = \lim_{n \to \infty} \frac{1}{1 + x^{2n}}$$

geliefert. Sobald nämlich $x^2 < 1$ ist, d. h. $x$ in dem Intervall

$$-1 < x < 1$$

liegt, wird $x^{2n}$ den Grenzwert 0 haben; die Funktion wird also den Wert 1 besitzen. Ist jedoch $x^2 > 1$, so wird $x^{2n}$ mit wachsendem $n$ über alle Grenzen wachsen; unsere Funktion wird also den Wert 0 haben. Endlich ist für $x^2 = 1$, d. h. für $x = +1$ und $x = -1$, der Funktionswert offenbar gleich $\frac{1}{2}$. Vgl. Fig. 19.

Weitere Kurven mit Sprüngen sind in Fig. 20a und 20b gezeichnet und repräsentieren Funktionen, die an den betreffenden Stellen Unstetigkeitspunkte haben.

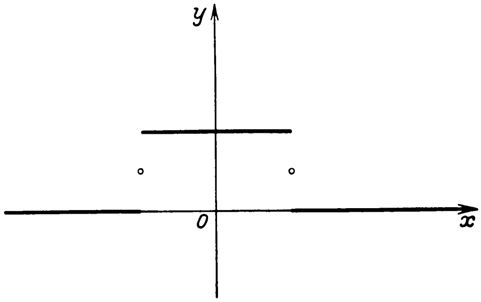


Fig. 19.

Während bei den Sprungstellen der Grenzwert von rechts wie der von links existiert, betrachten wir nun Unstetigkeitsstellen, in welchen diese Forderungen nicht erfüllt sind. Die wichtigsten derartigen Stellen sind die *Unendlichkeitsstellen.* Diese sind Stellen wie die Stelle $\xi = 0$ für die Funktion $\frac{1}{x}$ oder für die Funktion $\frac{1}{x^2}$, in denen der Funktionswert gar nicht definiert ist und bei welchen für $x \to \xi$ der Betrag $|f(x)|$ über alle

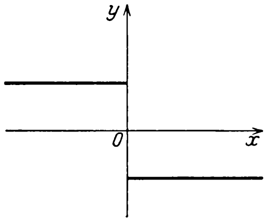


Fig. 20 a.

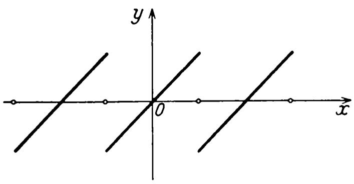


Fig. 20 b.

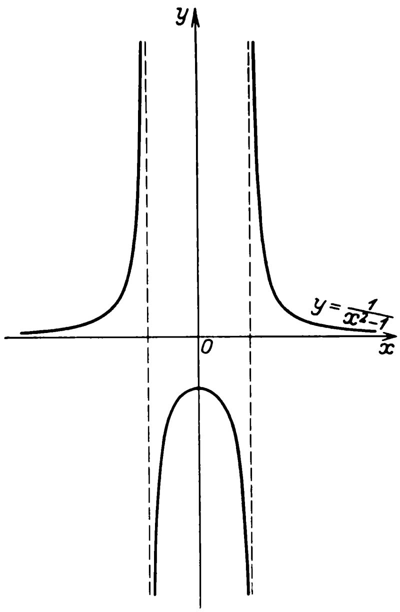


Fig. 21. Unendlichkeitsstellen.

Grenzen strebt. Bei der Funktion $\frac{1}{x}$ streben die Funktionswerte positiv bzw. negativ über alle Grenzen, wenn $x$ sich von rechts bzw. von links her dem Nullpunkte nähert. Dagegen hat die

Funktion $y = \frac{1}{x^2}$ für $x = 0$ eine Unendlichkeitsstelle, bei welcher von beiden Seiten her der Funktionswert positiv unendlich wird (vgl. Fig. 6 S. 11 und Fig. 12 S. 15). Die Funktion $y = \frac{1}{x^2 - 1}$, welche durch Figur 21 repräsentiert wird, hat sowohl an der Stelle $x = +1$ wie auch an der Stelle $x = -1$ eine Unendlichkeitsstelle.

Endlich erwähnen wir noch an der Hand eines Beispieles einen weiteren Typus von Unstetigkeiten, bei denen kein Grenzwert von rechts oder links existiert. Wir betrachten die Funktion

$$y = \sin \frac{1}{x}.$$

Die Funktion nimmt alle Werte zwischen $-1$ und $+1$ an, wenn die Zahl $\frac{1}{x}$ die Werte $\left(2n - \frac{1}{2}\right)\pi$ bis $\left(2n + \frac{1}{2}\right)\pi$ durchläuft, gleichgültig, welches die ganze Zahl $n$ ist. An den Stellen $x = \frac{2}{(4n+1)\pi}$ wird die Funktion den Wert 1, an den Stellen $x = \frac{2}{(4n-1)\pi}$ den Wert $-1$ annehmen. Man erkennt hieraus, daß die Funktion immer rascher zwischen den Werten $+1$ und $-1$ hin- und herpendelt, je mehr wir uns der Stelle $x = 0$ nähern, und daß in unmittelbarer Umgebung der Stelle $x = 0$ unendlich viele derartige *Oszillationen* stattfinden (vgl. Fig. 22). An der Stelle $x = 0$ selbst ist die Funktion gar nicht mehr definiert.

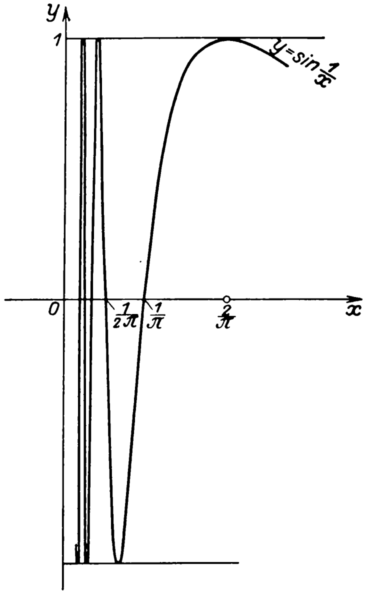


Fig. 22. Oszillierende Funktion mit Unstetigkeit.

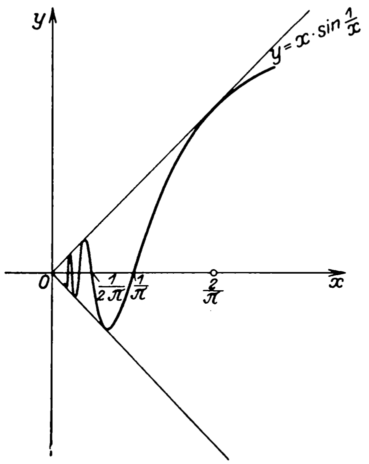


Fig. 23. Stetige oszillierende Funktion.

Es ist von Interesse, daß dagegen die Funktion $y = x \sin \frac{1}{x}$, deren Bild (vgl. Fig. 23) ebenfalls nebenstehend gezeichnet ist, an der Stelle

$x = 0$ stetig bleibt, wenn man ihr dort den Funktionswert 0 zuschreibt. Diese Stetigkeit wird dadurch erzwungen, daß der Faktor $x$ bei Annäherung an den Nullpunkt die Oszillationen des Sinus dämpft. Diese Funktion $y = x \sin \frac{1}{x}$ hat in der Nähe der Stelle $x = 0$ jedoch nicht mehr die Eigenschaft, nur endlich oft zwischen monotonem Wachsen und monotonem Abnehmen abzuwechseln; vielmehr oszilliert sie noch unendlich oft hin und her, wenn auch die Ausschläge bei diesen Oszillationen mit Annäherung an den Nullpunkt beliebig klein werden. Immerhin zeigt uns dieses Beispiel, daß der bloße Begriff der Stetigkeit noch allerlei merkwürdige und der naiven Anschauung fremdartige Möglichkeiten offen läßt.

Bei alledem möchte ich auf einen Umstand hinweisen, den man zur Präzisierung der Begriffe beachten muß. Es kann vorkommen, daß eine Funktion durch das ursprünglich gegebene Zuordnungsgesetz an einer bestimmten Stelle nicht definiert ist, wie z. B. die beiden zuletzt behandelten Funktionen an der Stelle $x = 0$. Bei dem letzteren Beispiel können wir nun an dieser Stelle durch eine neue Festsetzung, nämlich durch die Forderung $y = 0$ für $x = 0$ den Funktionswert an dieser Stelle so ergänzen, daß nach dieser Ergänzung die Funktion auch noch an dieser Stelle stetig bleibt; dazu ist nur nötig, daß die Grenzwerte von rechts und von links existieren und einander gleich sind; wir brauchen dann nur den Funktionswert an der Stelle gleich dem betreffenden Grenzwert zu setzen, um dort die Funktion stetig zu machen. Bei dem vorangehenden Beispiel $y = \sin \frac{1}{x}$ aber ist dies nicht möglich.

### 3. Sätze über stetige Funktionen.

Zum Schlusse seien noch folgende wichtigen allgemeinen Tatsachen angeführt, deren Beweis sich nach den Bemerkungen über das Rechnen mit Grenzwerten S. 31 von selbst versteht: *Summe, Differenz und Produkt von zwei stetigen Funktionen sind wieder stetig. Der Quotient zweier stetiger Funktionen ist dort stetig, wo der Nenner nicht verschwindet.*

Insbesondere ergibt sich hieraus sofort die Stetigkeit aller ganzen rationalen Funktionen und aller gebrochen rationalen Funktionen, abgesehen von den Nullstellen des Nenners. Daß die übrigen elementaren Funktionen, wie die trigonometrischen, stetig sind, wird sich bei unseren späteren Betrachtungen ganz von selbst mit ergeben (vgl. S. 53 und S. 76).

# Anhang zum ersten Kapitel.

## Vorbemerkungen.

In der griechischen Mathematik war weitgehend das Prinzip durchgeführt, alle Sätze in logisch bündiger Form dadurch zu beweisen, daß man sie auf ein System von möglichst wenigen nicht weiter zu beweisenden Axiomen zurückführte. Diese axiomatische Form der Darstellung, die zugleich einen Prüfstein für die Vollkommenheit der Untersuchung bildete, wurde zu Beginn der Neuzeit vorbildlich für Darstellungen in anderen Wissenszweigen; namentlich in der Philosophie haben Männer wie Descartes und Spinoza geglaubt, ihren Untersuchungen durch eine axiomatische oder, wie man es nannte, „geometrische" Darstellung größere Überzeugungskraft zu geben.

Aber anders war es mit der modernen Mathematik, die sich etwa gleichzeitig mit der neueren Philosophie zu entwickeln begann. In der Mathematik gab man vielfach sehr bald das antike Prinzip der axiomatischen Durchdringung des Stoffes auf. Anschauliche Evidenz in jedem einzelnen Falle wurde ein Hauptbeweismittel; ein gewisses gefühlsmäßiges, gelegentlich von mystischen Beimengungen nicht freies Operieren mit den neuen Begriffen (vor allem mit den ominösen „unendlich kleinen Größen") findet sich auch bei Forschern ersten Ranges. Eine Art blinder Glaube an die Allmacht der neuen Rechenmethoden riß die Forscher in ihren Betrachtungen auf Wege mit fort, die sie unter dem Drucke der Anforderungen voller Strenge niemals hätten gehen können. Kein Wunder, daß nur ein sicherer, genialer Instinkt vor groben Irrtümern schützen konnte.

Und doch ist es ein Glück, daß es so war und daß die später im 18. Jahrhundert einsetzende und sich im 19. Jahrhundert systematisch durchsetzende kritische Gegenströmung erst kam, als sie nicht mehr die Entwicklung hemmen, sondern nur noch ihre Früchte sichern und fördern konnte. Aber das Bedürfnis nach solcher kritischen Erfassung und Sicherung des Gewonnenen steigerte sich allmählich in einem Grade, daß seine Befriedigung mit Recht als eine der wichtigsten Leistungen in der Mathematik des 19. Jahrhunderts erscheinen muß.

Für die Integral- und Differentialrechnung ist hier besonders *Cauchy* zu nennen, welcher durch einwandfreie und klare Formulierung der Grundbegriffe das schon im 18. Jahrhundert begonnene Werk einer leicht faßlichen und von allen Unklarheiten des unendlich Kleinen befreiten Darstellung der höheren Analysis zu einer in vielen Zügen vorbildlichen Vollendung brachte.

Was hauptsächlich noch zu tun übrig blieb, war, in den Begründungen der Sätze und Methoden die anschaulichen Betrachtungen durch rein analytische zu ersetzen, die sich allein auf die Zahl und die mit Zahlen möglichen Operationen aufbauen, oder, wie man heute sagt, die Analysis zu *arithmetisieren*. In der Tat stellt die Berufung auf die geometrische Anschauung bei Beweisen in der Analysis ein für den kritisch geschulten Geist bedenkliches Element dar. Man braucht gar nicht auf die Frage nach der Genauigkeit oder Ungenauigkeit der Anschauung oder auf die Frage nach der Existenz einer „reinen Anschauung a priori" im Sinne Kants einzugehen, um zu erkennen, daß die naive anschauliche Betrachtung hinreichend viele Unbestimmtheiten enthält, die ihre Heranziehung zu vollständig strengen Beweisen in der Analysis verbieten. In den späteren Kapiteln wird uns das mehrfach deutlich vor Augen treten. Schon jetzt aber möchte ich darauf hinweisen, daß z. B. der Begriff der stetigen Kurve anschaulich sehr schwer zu erfassen ist. Eine stetige Kurve braucht keineswegs in jedem Punkte eine bestimmte Richtung zu haben; ja, es hat sich sogar gezeigt, daß es stetige Kurven gibt, die in keinem Punkte eine Richtung besitzen; ebenso gibt es stetige Kurven, für die nirgends der Begriff der Krümmung einen Sinn hat, und Kurven, denen man keine Länge zuschreiben kann. Angesichts solcher Tatsachen wird auch der Anfänger das Bedürfnis nach Arithmetisierung der Analysis anerkennen müssen[1]).

Wir wollen darüber aber nicht vergessen, daß eine jahrhundertelange glänzende, überaus erfolgreiche Entwicklung der Mathematik ohne eine Befriedigung dieses Bedürfnisses möglich war. Trotz aller Einwendungen bleibt die Anschauung doch immer die wichtigste lebendige Triebkraft für die mathematische Erfindung, und nur sie kann die Brücke von der Theorie zu den Anwendungen schlagen.

Ich will nun im Anschluß an Begriffsbildungen von *Bolzano* und *Weierstraß* diejenigen Gedankengänge entwickeln, welche die scharfe Begründung und Ergänzung der im ersten Kapitel nur unter Berufung auf die Anschauung formulierten Sätze geben.

Zum Schluß des Anhangs werden noch einige hiermit nicht zusammenhängende Ergänzungen zum ersten Kapitel folgen.

## § 1. Das Häufungsstellen-Prinzip und seine Anwendungen.

### 1. Das Häufungsstellen-Prinzip.

Bei der strengen Begründung der Analysis pflegt man das *Häufungsstellen-Prinzip* von *Weierstraß* an die Spitze zu stellen, ein Prinzip, welches vom Standpunkte der naiven Anschauung aus eine Selbstverständlichkeit formuliert, jedoch gerade als kurze Formulierung eines

[1]) Grundsätzlich muß man sich davon Rechenschaft geben, daß die scharfen mathematischen Begriffe immer weitgehende Idealisierungen der Vorstellungen darstellen, zu denen uns eine naive Anschauung führt, und daß daher die Fragen der letzten Grundlegung der Mathematik nicht durch Berufung auf die naive Anschauung beantwortet werden können. Erst in einer allmählichen Entwickelung, die bis in unsere Tage reicht und noch nicht zum Abschluß gekommen ist, haben die Grundlagenprobleme der Mathematik eine weitgehende Klärung erfahren.

immer wiederkehrenden Sachverhalts äußerst bequem zu verwenden ist wie eine Scheidemünze im täglichen Verkehr. Es lautet: *Wenn in einem endlichen Intervall unendlich viele Zahlen gegeben sind, so besitzen diese Zahlen mindestens eine Häufungsstelle $\xi$; d. h. es gibt mindestens einen Wert $\xi$ derart, daß in jedes auch noch so kleine um den Punkt $\xi$ abgegrenzte Intervall immer noch unendlich viele der gegebenen Zahlen hineinfallen.*

Um das Häufungsstellen-Prinzip rein arithmetisch zu beweisen, nehmen wir zunächst an, das gegebene Zahlenintervall sei das Intervall von 0 bis 1. Wir teilen es nun in zehn gleiche Teile durch die Punkte 0,1; 0,2; ...; 0,9. Dann muß jedenfalls in mindestens einem der zehn Teilintervalle noch eine unendliche Anzahl der gegebenen Zahlen liegen. Dieses Teilintervall oder, wenn es mehrere gibt, eines von ihnen möge das Intervall sein, das mit der Zahl $0,a_1$ beginnt. Dann teilen wir dieses Intervall wiederum in zehn Teile durch die Teilpunkte $0,a_1 1$; $0,a_1 2$; $0,a_1 3$; ...; $0,a_1 9$ und schließen nun wiederum, daß in mindestens einem dieser Teilintervalle noch unendlich viele Zahlen liegen müssen; ein solches Teilintervall möge mit dem Punkt $0,a_1 a_2$ beginnen. Indem wir dieses Teilintervall wiederum in zehn gleiche Teile einteilen, unsere Schlußweise wiederholen und dann immer weiter in derselben Art fortfahren, gelangen wir zu einer Folge von Ziffern $a_1$, $a_2, a_3, \ldots$, deren jede zwischen 0 und 9 liegt. Wir betrachten nun den Dezimalbruch

$$\xi = 0,a_1 a_2 a_3 \ldots$$

und erkennen unmittelbar, daß er eine Häufungsstelle unserer Zahlenmenge ist; denn jedes noch so kleine Zahlenintervall, in dessen Innern die Zahl $\xi$ liegt, enthält von einer gewissen Feinheit unserer Dezimalteilung ab alle jene oben gekennzeichneten Teilintervalle, in denen noch unendlich viele Zahlen der Zahlenmenge lagen.

Wenn das betrachtete Zahlenintervall nicht gerade das Intervall $0 \leqq x \leqq 1$ ist, sondern etwa das Intervall von $a$ bis $a + h$, so ändert sich nichts Wesentliches an unserer Betrachtung. Die fragliche Häufungsstelle wird uns dann einfach durch eine Zahl der Form

$$a + h \cdot 0,a_1 a_2 a_3 \ldots$$

dargestellt.

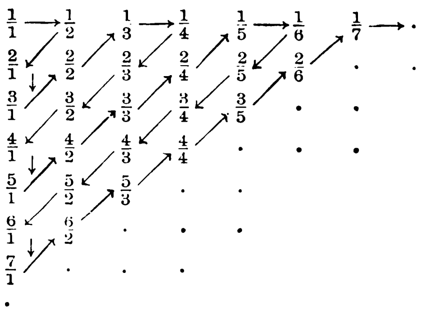
## 2. Grenzwerte von Zahlenfolgen. Beweis des Cauchyschen Konvergenzkriteriums.

Von dem gewonnenen Standpunkte aus erfährt der Begriff des Grenzwertes einer unendlichen Zahlenfolge $a_1, a_2, a_3, \ldots, a_n, \ldots$ eine neue Beleuchtung. Nehmen wir zunächst den Ausnahmefall vorweg, daß unendlich viele Zahlen der Folge einander gleich sind, so wollen wir diesen oder diese gemeinsamen Werte in Erweiterung unserer obigen

Definition ebenfalls als Häufungsstelle unserer Zahlenfolge bezeichnen. Befinden sich in der Folge unendlich viele verschiedene Zahlen und setzen wir voraus, daß die Zahlen $a_n$ dieser Folge „*beschränkt*“ sind, d. h. daß es eine Zahl $M$ gibt, für welche unabhängig von der Wahl des Index $n$ die Beziehung $|a_n| < M$ gilt, so bilden die Zahlen unserer Folge eine unendliche Zahlenmenge in einem endlichen Intervall, da sie ja alle zwischen $-M$ und $M$ liegen. Sie müssen also mindestens einen Häufungspunkt $\xi$ besitzen. Wenn es tatsächlich nur einen einzigen Häufungspunkt gibt, dann ist unsere Zahlenfolge offenbar konvergent mit genau demselben Grenzwert $\xi = \lim\limits_{n\to\infty} a_n$. *Wenn es dagegen mehrere Häufungspunkte gibt, so besitzt unsere Zahlenfolge keinen Grenzwert.* Die Existenz eines Limes und die Einzigkeit der Häufungsstelle einer beschränkten Zahlenfolge sind also gleichbedeutende Begriffe.

Der Fall der Nichtexistenz eines Limes ist durchaus als Regel zu betrachten. Z. B. besitzt die Zahlenfolge mit den Gliedern $a_{2n} = \frac{1}{n}$, $a_{2n-1} = 1 - \frac{1}{n}$ $(n = 1, 2, \ldots)$ die beiden Häufungsstellen 0 und 1.

Die Gesamtheit der positiven rationalen Zahlen läßt sich als eine Zahlenfolge auffassen, wobei allerdings die Ordnung nach der Größe vollständig zerstört wird. Zu einer solchen Anordnung als Zahlenfolge gelangen wir am einfachsten, indem wir die rationalen Zahlen nach dem beistehenden Schema anordnen, dieses Schema dann längs des eingezeichneten Linienzuges durchlaufen und dabei jeden Wert, der schon einmal vorgekommen ist $\left(\text{z.B.}\ \frac{2}{4}\right)$, nicht mehr berücksichtigen. Offenbar hat das System der rationalen Zahlen sämtliche rationalen und irrationalen Zahlen zu Häufungspunkten; es bildet also ein einfaches Beispiel für eine Zahlenfolge mit unendlich vielen Häufungsstellen.



Abzählung der rationalen Zahlen.

Wir können dem Häufungsstellen-Prinzip mit Hilfe des Begriffes der Konvergenz eine bemerkenswerte und für manche Anwendungen bequeme Formulierung geben: *Aus jeder beschränkten unendlichen Zahlenmenge läßt sich eine unendliche Teilfolge $a_1, a_2, a_3, \ldots$ herausgreifen, welche gegen einen Grenzpunkt $\xi$ konvergiert.* Hierzu brauchen wir nur eine Häufungsstelle $\xi$ der gegebenen unendlichen Zahlenmenge zu nehmen, dann eine Zahl $a_1$ aus der Zahlenmenge auszuwählen, deren Abstand von $\xi$ kleiner als $\frac{1}{10}$ ist; sodann eine weitere Zahl $a_2$ aus der Zahlenmenge, deren Abstand von $\xi$ kleiner als $\frac{1}{100}$ ist, eine dritte $a_3$, deren Abstand von $\xi$ kleiner als $\frac{1}{1000}$ ist, usw. Man erkennt, daß diese Zahlenfolge tatsächlich gegen den Grenzwert $\xi$ strebt.

Kehren wir nunmehr zu den konvergenten Zahlenfolgen, d. h. den beschränkten Folgen mit nur einer einzigen Häufungsstelle, zurück. Das früher in Kap. I, § 6 ausgesprochene Cauchysche Konvergenzkriterium wird nunmehr fast eine Selbstverständlichkeit. Setzen wir nämlich voraus, daß $|a_m - a_n|$ beliebig klein wird, wenn nur $m$ und $n$ hinreichend groß sind, dann liegen alle Zahlen $a_n$ in einem endlichen Intervall, besitzen also mindestens eine Häufungsstelle $\xi$. Gäbe es noch eine andere Häufungsstelle $\eta$, so müßte diese von $\xi$ einen Abstand $|\xi - \eta| = \alpha > 0$ haben. Beliebig nahe an $\xi$, etwa um weniger als $\frac{\alpha}{3}$ von $\xi$ entfernt, müßten noch unendlich viele Zahlen $a_n$ liegen, d. h. auch Werte $a_n$, für welche $n > N$ ist, wie groß wir auch $N$ vorgeben. Ebenso müssen in beliebiger Nähe der Häufungsstelle $\eta$, jedenfalls also auch um weniger als $\frac{\alpha}{3}$ von $\eta$ entfernt, noch unendlich viele Zahlen $a_m$ unserer Zahlenfolge liegen, also auch Zahlen $a_m$, deren Index $m > N$ ist. Für diese Werte $a_n$ und $a_m$ gilt aber nunmehr sicherlich $|a_m - a_n| > \frac{\alpha}{3}$, und diese Beziehung ist unvereinbar mit der Voraussetzung, daß $|a_m - a_n|$ bei hinreichend großem $N$ beliebig klein wird, sobald $m$ und $n$ gleichzeitig den Wert $N$ übersteigen. Mithin gibt es nicht zwei verschiedene Häufungsstellen, und das Cauchysche Kriterium ist bewiesen.

Ebenso einfach erkennen wir, daß *eine monoton zunehmende und eine monoton abnehmende beschränkte Zahlenfolge einen Grenzwert besitzen muß.* Denn ist im ersten Falle $\xi$ eine Häufungsstelle — eine solche gibt es ja sicher —, dann muß $\xi$ größer sein als jede Zahl der Zahlenfolge. Wäre nämlich eine Zahl $a_l$ der Zahlenfolge größer oder gleich $\xi$, dann würde für alle Zahlen $a_n$ mit $n > l + 1$ gelten $a_n > a_{l+1} > a_l \geqq \xi$; alle Zahlen der Zahlenfolge, höchstens mit Ausnahme der $l + 1$ ersten, würden also außerhalb eines Intervalles von der Breite $2|\xi - a_{l+1}|$ liegen, dessen Mittelpunkt der Punkt $\xi$ ist. Dies aber widerspricht der Voraussetzung, daß $\xi$ eine Häufungsstelle ist. Oberhalb $\xi$ können daher keine Zahlen und um so mehr auch keine Häufungsstellen der Folge liegen. Falls außer $\xi$ noch eine Häufungsstelle $\eta$ existiert, müßte also $\eta < \xi$ sein. Durch Wiederholung unserer Überlegung mit $\eta$ statt $\xi$ erhielten wir auch $\xi < \eta$ und damit einen Widerspruch. Eine ganz entsprechende Betrachtung gilt natürlich für monoton abnehmende Folgen.

Wir können übrigens, wie auf S. 30, unsere Aussage über monotone Folgen noch ergänzen, indem wir bei einer solchen Folge auch den Grenzfall zulassen, daß aufeinanderfolgende Zahlen der Folge einander gleich sind. Wir sprachen dann besser von einer monoton nicht abnehmenden bzw. nicht zunehmenden Folge. Der Satz von der Existenz eines Grenzwertes bleibt für solche Folgen bestehen.

## 3. Oberer und unterer Häufungspunkt, obere und untere Grenze einer Zahlenmenge.

Wenn wir in der Konstruktion, die uns in Nr. 1 zu einer Häufungsstelle $\xi$ geführt hat, die jeweilige Auswahl der Teilintervalle durch die Bedingung treffen, daß stets das letzte Teilintervall genommen werden soll, welches noch unendlich viele Punkte der Zahlenmenge enthält, so werden wir auf einen bestimmten Häufungspunkt $\beta$ geführt, den wir den „*oberen Häufungspunkt*" oder „*Limes superior*" der Zahlenmenge nennen und abgekürzt mit lim sup oder $\overline{\lim}$ bezeichnen. Er ist der am weitesten rechts liegende Häufungspunkt unserer Zahlenmenge; d. h. es können zwar noch unendlich viele Zahlen der Menge oberhalb $\beta$ liegen, aber, wie klein man auch die positive Zahl $\varepsilon$ wählt, nicht mehr unendlich viele oberhalb $\beta + \varepsilon$.

Wenn wir in der Konstruktion aus Nr. 1 stets das erste Teilintervall herausgreifen, in dem noch unendlich viele Punkte der Menge liegen, so werden wir auf einen Häufungspunkt $\alpha$ geführt, den „*unteren Häufungspunkt*" oder „*Limes inferior*" der Zahlenmenge (abgekürzt: lim inf oder $\underline{\lim}$), unterhalb dessen es keinen weiteren Häufungspunkt mehr gibt. Es können zwar noch unendlich viele Zahlen unterhalb $\alpha$ liegen, aber, wie klein auch die positive Zahl $\varepsilon$ gewählt wird, nicht mehr unter $\alpha - \varepsilon$. Den Beweis dieser Tatsachen kann ich dem Leser überlassen.

Der obere Häufungspunkt $\beta$ braucht ebenso wenig wie der untere $\alpha$ selbst zur Zahlenmenge zu gehören. Für die Menge der Zahlen $a_{2n} = \frac{1}{n}$; $a_{2n-1} = 1 - \frac{1}{n}$ ist z. B. $\alpha = 0$, $\beta = 1$, aber die Werte 0 und 1 befinden sich nicht in der gegebenen Zahlenmenge.

Bei unserem Beispiel gibt es oberhalb $\beta = 1$ keine Zahl der Menge. Man sagt, es ist in diesem Falle $\beta = 1$ auch die obere Grenze $G$ der Menge, indem man allgemein folgende Definition gibt: $G$ heißt *obere Grenze* einer Zahlenmenge, wenn es keine Zahl der Menge gibt, welche größer als $G$ ist, aber wenn bei jedem noch so klein gewählten positiven $\varepsilon$ immer mindestens eine Zahl der Menge vorhanden ist, die oberhalb $G - \varepsilon$ liegt. Entsprechend wird die *untere Grenze* $g$ definiert. Die obere Grenze kann, wie wir sehen, mit dem oberen Häufungspunkt zusammenfallen. Aber das Beispiel der Zahlenmenge $a_n = 1 + \frac{1}{n}$ $(n = 1, 2, \ldots)$ zeigt uns schon, daß dies nicht notwendig der Fall zu sein braucht. Hier ist $G = 2$ und $\beta = 1$.

Wir sehen jedenfalls, daß stets $G \geqq \beta$ sein muß, und erkennen ferner, daß die *obere Grenze, falls sie nicht mit dem oberen Häufungspunkt zusammenfällt, eine „isolierte" zur Zahlenmenge wirklich gehörende Zahl* sein muß. Entsprechend gilt für die untere Grenze $g$ stets $g \leqq \alpha$; es muß, wenn $g$ und $\alpha$ nicht zusammenfallen, $g$ eine *isolierte, zur Menge gehörende Zahl* sein.

## § 2. Sätze über stetige Funktionen.

### 1. Größter und kleinster Wert stetiger Funktionen.

Eine beschränkte unendliche Zahlenmenge muß eine obere Grenze $G$ und eine untere Grenze $g$ besitzen. Diese Zahlen $G$ und $g$ brauchen aber, wie wir sahen, nicht mehr selbst zu der Zahlenmenge zu gehören, oder, wie man sagt, die Zahlenmenge braucht keinen größten Wert und keinen kleinsten Wert zu besitzen. So besitzt z. B. die Zahlenfolge $1, \frac{1}{2}, \frac{1}{3}, \ldots$ die untere Grenze 0; aber 0 gehört nicht zur Zahlenmenge, und die Zahlenmenge enthält daher keine kleinste Zahl.

Angesichts dieser Tatsache ist der folgende Satz über stetige Funktionen durchaus nicht so selbstverständlich, wie er der naiven Anschauung nach erscheinen mag: *Jede in einem abgeschlossenen Intervall $a \leqq x \leqq b$ stetige Funktion $f(x)$ nimmt dort mindestens einmal einen größten und mindestens einmal einen kleinsten Wert an,* oder, wie man sagt, *sie besitzt einen größten und einen kleinsten Wert.*

Der Beweis ergibt sich leicht folgendermaßen. Jedenfalls stellt der Wertevorrat der stetigen Funktion $f(x)$ im Intervall $a \leqq x \leqq b$ eine beschränkte Zahlenmenge dar und besitzt daher eine obere Grenze $G$. Denn sonst gäbe es eine Folge von Zahlen $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ unseres Intervalles, für die $f(\xi_n)$ unbeschränkt wächst. Diese Folge hätte wenigstens einen Häufungspunkt $\xi^*$ im Intervall. Dann gäbe es aber in beliebiger Nähe von $\xi^*$ immer noch Zahlen $\xi_n$ unserer Folge, für die $|f(\xi_n) - f(\xi^*)| > 1$ (ja sogar beliebig groß) wäre, d. h. die Funktion wäre im Punkte $\xi^*$ unstetig. Da es also eine obere Grenze $G$ gibt, muß es auch eine Folge von Zahlen $x_1, x_2, \ldots, x_n, \ldots$ des Intervalles geben, derart, daß

$$\lim_{n \to \infty} f(x_n) = G$$

ist. Nach dem Häufungsstellen-Prinzip in der Fassung von S. 45 können wir aus der Folge der Zahlen $x_n$ eine Teilfolge auswählen, die gegen einen Grenzwert $\xi$ konvergiert; nennen wir sie wieder $\xi_1, \xi_2, \ldots, \xi_n, \ldots$, so daß also

$$\lim_{n \to \infty} \xi_n = \xi$$

ist. Es wird nun sicherlich auch

$$\lim_{n\to\infty} f(\xi_n) = G$$

sein. Andererseits ist wegen der vorausgesetzten Stetigkeit der Funktion $f(x)$, da $\xi$ zum Intervall gehört,

$$\lim_{n\to\infty} f(\xi_n) = f(\xi).$$

Somit ist $f(\xi) = G$. Der Wert $G$ wird also an einer bestimmten Stelle $\xi$ im Innern oder am Rande des Intervalles angenommen, was unsere Behauptung darstellt. Genau die entsprechende Betrachtung gilt natürlich für den kleinsten Wert.

Der Satz vom größten und kleinsten Wert stetiger Funktionen würde übrigens nicht allgemein richtig sein, wenn wir nicht ausdrücklich das Intervall als abgeschlossen voraussetzen, d. h. die Endpunkte in die Stetigkeitsvoraussetzung mit einbeziehen. Z. B. ist die Funktion $y = \frac{1}{x}$ in dem offenen Intervall $0 < x < \infty$ stetig; sie hat aber dort keinen größten Wert, sondern besitzt in der Nähe von $x = 0$ beliebig große Werte. Ebenso hat sie keinen kleinsten Wert, sondern wird für hinreichend großes $x$ beliebig klein, ohne den Wert 0 anzunehmen.

## 2. Die Gleichmäßigkeit der Stetigkeit.

Die Stetigkeit einer Funktion $f(x)$ in einem abgeschlossenen Intervall $a \leqq x \leqq b$ läßt, wie wir schon gesehen haben (vgl. S. 41) und wie wir noch weiter erkennen werden, einen so großen Spielraum für verschiedenartige unanschauliche Vorkommnisse, daß wir noch einige weitere, vom naiven Standpunkt aus sozusagen selbstverständliche Dinge über stetige Funktionen aus dem Begriff der Stetigkeit heraus logisch streng beweisen wollen. Unser Begriff der Stetigkeit besagt lediglich, daß aus der Beziehung $\lim\limits_{n\to\infty} x_n = \xi$ die Beziehung $\lim\limits_{n\to\infty} f(x_n) = f(\xi)$ folgt. Wir können dies auch so ausdrücken: Für jede Stelle $\xi$ gibt es zu jeder noch so kleinen Genauigkeitsschranke $\varepsilon > 0$ eine Zahl $\delta > 0$, derart, daß $|f(x) - f(\xi)| < \varepsilon$ wird, wenn nur $|x - \xi| < \delta$ ist, sobald alle betrachteten Zahlen $x$ in dem gegebenen Intervall $a \leqq x \leqq b$ liegen. Z. B. ist bei $c \neq 0$ für die Funktion $y = cx$ eine solche Zahl $\delta$ durch die Relation $\delta = \frac{\varepsilon}{|c|}$ gegeben; für die Funktion $y = x^2$ können wir uns folgendermaßen eine solche Zahl konstruieren: Wir nehmen z. B. an, daß $a = 0$ und $b = 1$ ist, und fragen: Wie nahe sollen wir $x$ an $\xi$ wählen, damit der Ausdruck $|x^2 - \xi^2| < \varepsilon$ wird? Zu dem Zweck schreiben wir $|x^2 - \xi^2| = |x - \xi|\,|x + \xi| \leqq |x - \xi|\,(1 + \xi)$. Wenn wir also $\delta \leqq \frac{\varepsilon}{1+\xi}$ wählen, so werden wir sicher sein, daß $|x^2 - \xi^2| < \varepsilon$ wird. Wir sehen an diesem Beispiel, daß die gewonnene Genauigkeitsgrenze $\delta$ nicht nur von $\varepsilon$ abhängt, sondern auch noch von der Stelle $\xi$

des Intervalles, an welcher wir die Stetigkeit der Funktion betrachten. Wir können nun aber, wenn wir darauf verzichten, die Schrankenbestimmung besonders günstig zu gestalten, diese Abhängigkeit der Größe $\delta$ von $\xi$ auch noch beseitigen, indem wir rechts $\xi$ durch die Größe 1 ersetzen, wodurch $\delta$ wegen $0 \leqq \xi \leqq 1$ höchstens verkleinert und schlimmstenfalls gleich $\frac{\varepsilon}{2}$ wird.

Es entsteht nun die Frage, ob etwas Ähnliches für jede in einem abgeschlossenen Intervall stetige Funktion gilt, d. h. ob man — unter Verzicht auf die jeweils günstigste Bestimmung von $\delta$ — für das ganze Intervall gleichzeitig oder, wie man besser sagt, *gleichmäßig in bezug auf* $\xi$ zu jedem $\varepsilon$ eine nur von $\varepsilon$ und nicht mehr von $\xi$ abhängige Schranke $\delta = \delta(\varepsilon)$ bestimmen kann, so daß die obige Beziehung

$$|f(x) - f(\xi)| < \varepsilon$$

gilt, sobald $|x - \xi| < \delta$ ist. In der Tat ist dies nun möglich lediglich auf Grund der allgemeinen Stetigkeitsdefinition und ohne weitere, speziellere Voraussetzungen. Diese Tatsache, auf die sich erst spät im 19. Jahrhundert die Aufmerksamkeit gerichtet hat, heißt der *Satz von der gleichmäßigen Stetigkeit stetiger Funktionen.*

Sein Beweis soll indirekt geführt werden. Wir werden also aus der Annahme, es gäbe eine in einem abgeschlossenen Intervall $a \leqq x \leqq b$ stetige, aber dort nicht gleichmäßig stetige Funktion $f(x)$, einen Widerspruch herleiten. Die gleichmäßige Stetigkeit bedeutet: Wenn man erreichen will, daß die Differenz $|f(u) - f(v)|$, wo $u$ und $v$ aus dem abgeschlossenen Intervall $a \leqq x \leqq b$ zu nehmen sind, kleiner wird als eine positive, beliebig klein wählbare Größe $\varepsilon$, so braucht man nur $u$ und $v$ nahe genug beieinander zu wählen, nämlich näher als um eine geeignet zu $\varepsilon$ hinzuzubestimmende Zahl $\delta = \delta(\varepsilon)$; *wo* ein solches Wertepaar $u, v$ im Intervall liegt, ist im übrigen gleichgültig. Wäre nun $f(x)$ nicht gleichmäßig stetig, so gäbe es danach eine feste positive (vielleicht sehr kleine) Zahl $\alpha$ und zu jeder Zahl $\delta_n$ einer beliebigen gegen 0 strebenden Folge $\delta_1, \delta_2, \delta_3, \ldots$ ein Wertepaar $u_n, v_n$ im Intervall, derart, daß für alle Werte von $n$ gilt $|u_n - v_n| < \delta_n$ und $|f(u_n) - f(v_n)| > \alpha$. Die Zahlen $u_n$ müssen nach dem Häufungsstellenprinzip eine Häufungsstelle $\xi$ besitzen und somit die Zahlen $v_n$ dieselbe Häufungsstelle. Grenzen wir um diese Häufungsstelle $\xi$ ein beliebig kleines Intervall $|x - \xi| < \delta$ ab, so werden also unendlich viele Wertepaare $u_n, v_n$ in dieses Intervall hineinfallen. Die „Schwankung" der Funktion $f(x)$ würde also innerhalb eines beliebigen solchen Intervalles noch immer größer als die feste Zahl $\alpha$ bleiben. Dies aber steht mit der vorausgesetzten Stetigkeit der Funktion $f(x)$ an der Stelle $\xi$ im Widerspruch; denn diese verlangt, daß für nahe an $\xi$ gelegene $x_1$ und $x_2$

$$|f(\xi) - f(x_1)| < \frac{\alpha}{2} \quad \text{und} \quad |f(\xi) - f(x_2)| < \frac{\alpha}{2},$$

also $|f(x_1) - f(x_2)| < \alpha$ ist. Damit ist die behauptete Gleichmäßigkeitseigenschaft bewiesen.

Bei unserem Beweis haben wir die Abgeschlossenheit des Intervalles wesentlich benützt[1]). In der Tat ist der Satz von der Gleichmäßigkeit der Stetigkeit für nicht abgeschlossene Intervalle auch gar nicht allgemein richtig. Z. B. ist die Funktion $\frac{1}{x}$ in dem offenen Intervall $0 < x \leqq 1$ stetig, aber nicht mehr gleichmäßig stetig. Denn wie klein man auch die Breite $\delta$ $(< 1)$ eines Intervalles wählt, so wird doch die Schwankung der Funktion in diesem Intervalle immer größer als eine feste Zahl, z. B. 1, sein, wenn das Intervall nur nahe genug am Nullpunkt, etwa als $\frac{\delta}{2} \leqq x \leqq \frac{3\delta}{2}$ gewählt wird. Die Ungleichmäßigkeit der Stetigkeit beruht natürlich darauf, daß in dem abgeschlossenen Intervall $0 \leqq x \leqq 1$ die Funktion eine Unstetigkeitsstelle im Anfangspunkt besitzt. Hätten wir die obige Funktion $y = x^2$ im ganzen (offenen) Intervalle $-\infty < x < \infty$ statt in einem abgeschlossenen Teilintervall betrachtet, wäre auch sie nicht gleichmäßig stetig gewesen.

### 3. Der Zwischenwertsatz.

Ein weiterer in der Analysis ständig benutzter Satz ist der folgende: *Eine in einem abgeschlossenen Intervall $a \leqq x \leqq b$ stetige Funktion $f(x)$, welche für $x = a$ negativ, für $x = b$ positiv ist (oder umgekehrt), nimmt im Intervall mindestens einmal den Wert 0 an.* Dieser für die geometrische Anschauung triviale Satz, welcher besagt, daß eine stetige Kurve, die am Anfang unterhalb, am Ende oberhalb der $x$-Achse liegt, dazwischen einmal die $x$-Achse schneiden muß, ist auch analytisch sehr einfach zu beweisen. Es gibt sicher im Intervall unendlich viele Punkte $x$, für welche $f(x) < 0$ ist — wegen der Stetigkeit der Funktion ist sie ja sicher in einem ganzen an den Anfangspunkt anschließenden Intervall negativ —. Die Menge dieser Punkte $x$ mit $f(x) < 0$ besitzt eine obere Häufungsstelle $\xi$, die im Innern des Intervalles liegen muß. Da in deren beliebiger Nähe Punkte $x$ mit $f(x) < 0$ liegen müssen, so ist jedenfalls wegen der Stetigkeit $f(\xi) \leqq 0$. Es kann aber nicht $f(\xi) < 0$ sein, da sonst auch in einer passend klein gewählten Umgebung von $\xi$ die Funktion $f(x)$ negativ wäre, also auch in Werten $x$, für die $x > \xi$ ist; dies widerspricht der Voraussetzung, daß $\xi$ die größte Häufungsstelle der Werte $x$ mit $f(x) < 0$ ist. Also gilt $f(\xi) = 0$, und unsere Behauptung ist bewiesen.

Wir können unseren Satz leicht ein wenig verallgemeinern: *Setzen wir voraus, daß $f(a) = \alpha$, $f(b) = \beta$ sei, und ist $\mu$ irgend ein Wert zwischen $\alpha$ und $\beta$, dann nimmt die stetige Funktion $f(x)$ in dem Intervall den Wert $\mu$*

[1]) Sonst brauchte nämlich der Häufungspunkt $\xi$ nicht zum Intervall zu gehören.

*mindestens einmal an.* Denn die stetige Funktion $\varphi(x) = f(x) - \mu$ wird an beiden Endpunkten des Intervalles verschiedene Vorzeichen haben, also im Intervall den Wert 0 annehmen.

## 4. Umkehrung einer stetigen monotonen Funktion.

Wenn die stetige Funktion $y = f(x)$ in dem Intervall $a \leqq x \leqq b$ monoton ist, so wird jeder Zwischenwert $\mu$ ein- und nur einmal angenommen; es gehört also, wenn $y$ das abgeschlossene Intervall zwischen den Werten $\alpha = f(a)$ und $\beta = f(b)$ durchläuft, zu jedem Wert von $y$ genau ein Wert von $x$. Wir können daher $x$ in diesem Intervall auch als eindeutige Funktion von $y$ betrachten, d. h. wir können die Funktion $y = f(x)$ eindeutig umkehren. Wir behaupten: Diese Umkehrfunktion $x = \varphi(y)$ ist wiederum eine stetige und monotone Funktion von $y$, wenn $y$ im Intervall zwischen $\alpha$ und $\beta$ variabel ist.

Der monotone Charakter der Umkehrfunktion $x = \varphi(y)$ versteht sich von selbst. Um den Beweis für ihre Stetigkeit streng zu führen, beachten wir, daß aus dem monotonen Charakter von $f(x)$ folgt: $|f(x_2) - f(x_1)| = |y_2 - y_1| > 0$, sobald $x_1$ und $x_2$ verschiedene Zahlen unseres Intervalles sind. Ist nun $h$ eine positive Zahl kleiner als $b - a$, so ist die Funktion $|f(x + h) - f(x)|$ im abgeschlossenen Intervalle $a \leqq x \leqq b - h$ stetig, besitzt also an einer Stelle $x = \xi$ einen kleinsten Wert $|f(\xi + h) - f(\xi)| = \alpha(h)$, welcher unserer obigen Bemerkung zufolge nicht Null ist[1]). Wir schließen hieraus: Wenn $x_1$ und $x_2$ zwei Stellen des Intervalles sind, für welche $|x_1 - x_2| \geqq h$ gilt, so wird $|f(x_1) - f(x_2)| \geqq \alpha(h)$. Daraus folgt aber sofort die Stetigkeit der Umkehrfunktion. Denn sobald $|y_1 - y_2|$ unter die positive Zahl $\alpha(h)$ sinkt, muß $|x_1 - x_2| < h$ sein; ist also eine Genauigkeitsschranke $\varepsilon$ vorgegeben, so brauchen wir nur $\delta = \alpha(\varepsilon)$ zu wählen, damit für alle Werte $y$, für die $|y_1 - y| < \delta$ ist, auch $|\varphi(y_1) - \varphi(y)| < \varepsilon$ wird.

## 5. Weitere Sätze über stetige Funktionen.

Ich überlasse dem Leser den Beweis der folgenden fast selbstverständlichen Tatsache: *Eine stetige Funktion von einer stetigen Funktion ist wieder stetig*; d. h. ist $\varphi(x)$ eine im Intervall $a \leqq x \leqq b$ stetige Funktion, deren Wertevorrat das Intervall $\alpha \leqq \varphi \leqq \beta$ ausfüllt, ist ferner $f(\varphi)$ eine in diesem letzteren Intervall stetige Funktion von $\varphi$, so ist $f(\varphi(x))$ auch eine im Intervall $a \leqq x \leqq b$ stetige Funktion von $x$. (Satz von der Stetigkeit der aus stetigen Funktionen zusammengesetzten Funktionen.)

Schon auf S. 41 wurde der Satz erwähnt, daß *Summe, Differenz, Produkt von stetigen Funktionen wieder stetig sind* und daß auch *der Quotient stetiger Funktionen stetig ist, solange der Nenner von* 0 *verschieden bleibt.*

---

1) Übrigens strebt $\alpha(h)$ wegen der Stetigkeit von $f(x)$ mit unbegrenzt abnehmendem $h$ selbst gegen Null.

## § 3. Bemerkungen über die elementaren Funktionen.

Im ersten Kapitel haben wir stillschweigend angenommen, daß die elementaren Funktionen stetig sind. Der Beweis dafür ist in der Tat nunmehr fast selbstverständlich. Zunächst ist $f(x) = x$ eine stetige Funktion, also auch $x^2 = x \cdot x$ als Produkt zweier stetiger Funktionen und ebenso jede Potenz von $x$ und daher jede ganze rationale Funktion als Summe von stetigen Funktionen; somit ist auch jede gebrochene rationale Funktion als Quotient stetiger Funktionen stetig in jedem Intervall, in welchem der Nenner nicht verschwindet.

Die $n$-te Wurzel aus $x$ ist als Umkehrfunktion der $n$-ten Potenz — die Funktion $x^n$ ist offenbar für $x > 0$ monoton und stetig — stetig, und also nach der Regel von der Stetigkeit zusammengesetzter Funktionen z. B. auch die $n$-te Wurzel aus einer rationalen Funktion.

Die Stetigkeit der trigonometrischen Funktionen, die dem Leser von der Schule her geläufig ist, könnten wir jetzt unschwer unter Benutzung der oben gegebenen Begriffsbildungen beweisen; ich unterlasse das aber hier, weil sich im nächsten Kapitel, § 3, diese Stetigkeit ganz von selbst als Folge der Differenzierbarkeit mit ergeben wird.

Einige Bemerkungen sind lediglich noch über die Definition und Stetigkeit der Exponentialfunktion $a^x$, der allgemeinen Potenz $x^\alpha$ und des Logarithmus nötig. Wir setzen wie im ersten Kapitel (§ 3) voraus, daß $a$ eine positive Zahl, etwa größer als 1 ist, und verstehen, wenn $r = \frac{p}{q}$ eine positive rationale Zahl bedeutet ($p$ und $q$ ganze Zahlen), unter $a^r = a^{\frac{p}{q}}$ den positiven Wert dieser Größe. Ist $\alpha$ irgend eine irrationale Zahl und bedeutet $r_1, r_2, \ldots, r_m, \ldots$ eine beliebige Folge rationaler Zahlen, welche gegen $\alpha$ streben, so behaupten wir, daß $\lim\limits_{m \to \infty} a^{r_m}$ existiert; wir nennen diesen Grenzwert dann $a^\alpha$.

Um die Existenz dieses Grenzwertes zu zeigen, brauchen wir nach dem Cauchyschen Konvergenzkriterium nur nachzuweisen, daß $|a^{r_n} - a^{r_m}|$ beliebig klein wird, sobald nur $n$ und $m$ hinreichend groß gewählt sind. Wir nehmen etwa an $r_n > r_m$, d. h. $r_n - r_m = \delta > 0$. Dann wird

$$a^{r_n} - a^{r_m} = a^{r_m}(a^\delta - 1).$$

Wir brauchen also, da $a^{r_m}$ beschränkt bleibt, nur noch zu zeigen, daß $|a^\delta - 1| = a^\delta - 1$ bei hinreichend großem $n$ und $m$ beliebig klein ist. Es ist aber $\delta$ eine rationale Zahl, die sicherlich, wenn $n$ und $m$ hinreichend groß sind, beliebig klein wird. Ist also $l$ eine beliebig große natürliche Zahl, so ist $\delta < \frac{1}{l}$, sobald $n$ und $m$ hinreichend groß sind. Nun ist[1]) wegen $\delta < \frac{1}{l}$ und $a > 1$

[1]) Diese Tatsache darf ich wohl aus der Schulmathematik als bekannt voraussetzen.

$$1 < a^\delta < a^{\frac{1}{l}},$$

und da $a^{\frac{1}{l}}$ mit wachsendem $l$ gegen 1 strebt (vgl. S. 22), folgt unmittelbar unsere Behauptung.

Ich kann es wiederum dem Leser überlassen, zu zeigen, daß die so auch für irrationale Werte $x$ definierte Funktion $a^x$ überall stetig ist, ferner, daß sie eine monotone Funktion von $x$ darstellt. Für negative Werte von $x$ wird diese Funktion naturgemäß durch die Gleichung

$$a^x = \frac{1}{a^{-x}}$$

erklärt. Sie nimmt dann, wenn $x$ von $-\infty$ bis $+\infty$ läuft, in monotoner Folge alle Werte zwischen 0 und $+\infty$ an. Daher besitzt sie eine stetige und monotone Umkehrfunktion, die wir als *Logarithmus zur Basis a* bezeichnen. — Ganz ähnlich kann man die Stetigkeit der allgemeinen Potenz $y = x^\alpha$ in $x$ beweisen, wobei $\alpha$ eine feste rationale oder irrationale Zahl, $x$ eine im Intervalle $0 < x < \infty$ laufende unabhängige Veränderliche ist.

Die hier skizzierte „elementar-mathematische Betrachtung" der Exponentialfunktion, des Logarithmus und der Potenz $x^\alpha$, wie sie zum großen Teil schon von der Schule her geläufig ist, werden wir im dritten Kapitel, § 6, durch eine prinzipiell viel einfachere Betrachtungsweise ersetzen.

## § 4. Polarkoordinaten.

Wir haben im ersten Kapitel den Funktionsbegriff an die Spitze gestellt und ihn durch das geometrische Bild einer Kurve veranschaulicht. Es ist jedoch nützlich, daran zu erinnern[1]), daß die analytische Geometrie den umgekehrten Weg einschlägt, nämlich von einer geometrisch gegebenen Kurve ausgeht und dann diese Kurve durch eine Funktion darstellt, etwa durch eine Funktion, welche die eine Koordinate eines Kurvenpunktes vermittelst der anderen ausdrückt. Diese Auffassung führt ganz naturgemäß dazu, außer den bisher von uns im ersten Kapitel allein benutzten rechtwinkligen Koordinaten auch noch andere zu betrachten, deren Einführung der geometrisch gegebenen Kurve besser angepaßt sein kann. Das wichtigste Beispiel dafür sind die *Polarkoordinaten* $r$, $\vartheta$, welche mit den rechtwinkeligen Koordinaten $x$, $y$ eines Punktes $P$ durch die Gleichungen

$$x = r\cos\vartheta, \quad y = r\sin\vartheta, \quad r^2 = x^2 + y^2, \quad \operatorname{tg}\vartheta = \frac{y}{x}$$

verbunden sind und deren geometrische Bedeutung durch Figur 24 klar wird.

[1]) Vgl. auch S. 9.

Betrachten wir etwa die *Lemniskate*: sie ist geometrisch definiert als der Ort aller Punkte $P$, für welche das Produkt der Abstände $r_1$ und $r_2$ von den festen Punkten $F_1$ und $F_2$ mit den rechtwinkeligen

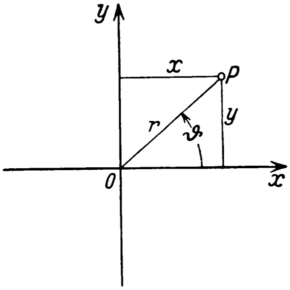



Fig. 24. Polarkoordinaten.

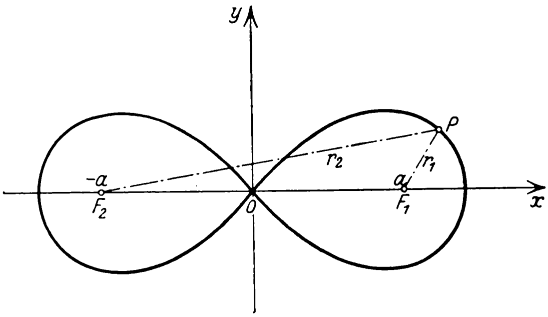



Fig. 25. Lemniskate.

Koordinaten $x = a$, $y = 0$ bzw. $x = -a$, $y = 0$ den Wert $a^2$ besitzt (vgl. Fig. 25). Da

$$r_1^2 = (x-a)^2 + y^2, \quad r_2^2 = (x+a)^2 + y^2$$

ist, so ergibt sich nach einfacher Umrechnung als Lemniskatengleichung

$$(x^2+y^2)^2 - 2a^2(x^2-y^2) = 0.$$

Führen wir nun Polarkoordinaten ein, so erhalten wir

$$r^4 - 2a^2r^2(\cos^2\vartheta - \sin^2\vartheta) = 0$$

oder schließlich nach Weglassung des Faktors $r^2$ gemäß einer einfachen trigonometrischen Formel

$$r^2 = 2a^2\cos 2\vartheta.$$

Wir sehen also, daß die Gleichung der Lemniskate in Polarkoordinaten einfacher wird als in rechtwinkligen.

## § 5. Bemerkungen über komplexe Zahlen.

Die Gesamtheit der reellen Zahlen wird für uns die Grundlage der Betrachtung abgeben. Ich möchte jedoch schon an dieser Stelle im Hinblick auf die Anwendungen im achten, neunten und zehnten Kapitel daran erinnern, daß die Aufgaben der Algebra noch zu einer umfassenden Erweiterung des Zahlbegriffes geführt haben, nämlich zur Einführung der *komplexen Zahlen*. Ebenso wie die Erweiterung des Zahlbereichs der natürlichen Zahlen zu dem aller reellen Zahlen dem Bestreben entspringt, Ausnahmeerscheinungen zu beseitigen bzw. gewisse Operationen, wie Division, Subtraktion, Zuordnung von Zahlen und Strecken ausnahmslos möglich zu machen, wird man zur Einführung der komplexen Zahlen durch die Forderung genötigt, jeder quadratischen Gleichung und sogar jeder algebraischen Gleichung eine Lösung zuschreiben zu können. Will man z. B. erreichen, daß die Gleichung

$$x^2 + 1 = 0$$

Wurzeln besitzt, so ist man gezwungen, neue Symbole $+i$ und $-i$ als Wurzeln dieser Gleichung einzuführen (und dadurch erreicht man, wie in der Algebra gezeigt wird, gleichzeitig die Auflösbarkeit aller algebraischen Gleichungen)[1].

Sind $a$ und $b$ gewöhnliche reelle Zahlen, so bedeutet die ***komplexe Zahl*** $c = a + ib$ ein Zahlenpaar $(a, b)$, wobei das Rechnen mit solchen Zahlenpaaren einfach nach folgender allgemeinen Regel zu erfolgen hat: Man addiert, multipliziert, dividiert komplexe Zahlen (unter denen für $b = 0$ als Spezialfall auch die reellen Zahlen enthalten sind), indem man das Symbol $i$ wie eine unbestimmte Rechengröße behandelt, aber stets alle Ausdrücke durch die Benutzung der Beziehung $i^2 = -1$ so vereinfacht, daß höhere Potenzen dieser Größe als die ersten nicht stehen bleiben und daß wieder ein Ausdruck der Form $a + bi$ entsteht.

Ich darf voraussetzen, daß der Leser eine gewisse Bekanntschaft mit diesen komplexen Zahlen mitbringt. Trotzdem will ich eine besonders wichtige Beziehung hier noch hervorheben, indem ich an die geometrische bzw. trigonometrische Darstellung der komplexen Zahlen anknüpfe. Ist $c = x + iy$ eine solche Zahl, so repräsentiert man sie in einem rechtwinkligen Koordinatensystem durch den Punkt $P$ mit den Koordinaten $x$ und $y$. Führen wir nun (vgl. Fig. 24, S. 55) durch die Gleichungen $x = r\cos\vartheta$, $y = r\sin\vartheta$ an Stelle der rechtwinkligen Koordinaten Polarkoordinaten $r$ und $\vartheta$ ein, wobei wir $r \geqq 0$ und $-\pi < \vartheta \leqq \pi$ wählen, so wird $r = \sqrt{x^2 + y^2}$ der Abstand des Punktes $P$ vom Koordinatenanfangspunkte $O$, und $\vartheta$ ist der Winkel zwischen der positiven $x$-Achse und dem Strahl $OP$. Die komplexe Zahl $c$ stellt sich dann in der Form dar

$$c = r(\cos\vartheta + i\sin\vartheta).$$

Wir nennen den Winkel $\vartheta$ den zur Zahl $c$ gehörigen ***Winkel*** oder ***Arcus***, die Größe $r$ ihren ***absoluten Betrag***, für den man auch $|c|$ schreibt. Zur „***konjugiert komplexen***“ Zahl $\bar{c} = x - iy$ gehört offenbar derselbe absolute Betrag $r$, aber, außer im Falle eines negativ-reellen $c$, der Winkel $-\vartheta$. Es ist

$$r^2 = |c|^2 = c\bar{c} = x^2 + y^2.$$

Mit Hilfe der obigen trigonometrischen Darstellung nimmt die Multiplikation der komplexen Zahlen eine besonders einfache Form an. Es ist nämlich

[1] Die Tatsache, daß jede algebraische Gleichung reelle oder komplexe Wurzeln besitzt, ist die Aussage des „Fundamentalsatzes“ der Algebra.

$$c \cdot c' = r(\cos\vartheta + i\sin\vartheta)\cdot r'(\cos\vartheta' + i\sin\vartheta')$$
$$= rr'\cdot((\cos\vartheta\cos\vartheta' - \sin\vartheta\sin\vartheta') + i(\cos\vartheta\sin\vartheta' + \sin\vartheta\cos\vartheta'));$$

wenn man die bekannten Additionsgesetze für die trigonometrischen Funktionen beachtet, ergibt sich hieraus

$$c \cdot c' = rr'\cdot(\cos(\vartheta + \vartheta') + i\sin(\vartheta + \vartheta')).$$

Man multipliziert also komplexe Zahlen, indem man ihre absoluten Beträge multipliziert und ihre Winkel addiert. Die merkwürdige Formel

$$(\cos\vartheta + i\sin\vartheta)(\cos\vartheta' + i\sin\vartheta') = \cos(\vartheta + \vartheta') + i\sin(\vartheta + \vartheta')$$

pflegt man die *Formel von Moivre* zu nennen; sie führt im übrigen unmittelbar zu der Beziehung

$$(\cos\vartheta + i\sin\vartheta)^n = \cos n\vartheta + i\sin n\vartheta,$$

welche uns z. B. gestattet, die Gleichung $x^n = 1$ für ganzzahliges positives $n$ aufzulösen, indem man sofort ihre Wurzeln

$$\varepsilon_1 = \varepsilon = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}, \qquad \varepsilon_2 = \varepsilon^2 = \cos\frac{4\pi}{n} + i\sin\frac{4\pi}{n}, \ \ldots,$$
$$\varepsilon_{n-1} = \varepsilon^{n-1} = \cos\frac{(n-1)2\pi}{n} + i\sin\frac{(n-1)2\pi}{n}, \qquad \varepsilon_n = \varepsilon^n = 1$$

hinschreiben kann.

Denkt man sich übrigens den Ausdruck auf der linken Seite der Gleichung $(\cos\vartheta + i\sin\vartheta)^n = \cos n\vartheta + i\sin n\vartheta$ nach dem binomischen Satz entwickelt, so liefert uns die Gleichung, wenn wir Reelles und Imaginäres trennen, eine Darstellung von $\cos n\vartheta$ und $\sin n\vartheta$ durch Potenzen und Potenzprodukte von $\cos\vartheta$ und $\sin\vartheta$.

Zweites Kapitel.

# Grundbegriffe der Integral- und Differentialrechnung.

Unter den Grenzwertbildungen der Analysis spielen zwei eine besonders wichtige Rolle, nicht nur weil sie immer wieder in den verschiedensten Zusammenhängen auftreten, sondern vor allem weil sie miteinander in einer engen Wechselbeziehung stehen. Diese beiden Grenzwertbildungen, das *Integral* und der *Differentialquotient*, wurden an Hand vereinzelter Beispiele schon seit langer Zeit, z. T. sogar schon im klassischen Altertum betrachtet; aber erst die Tatsache, daß man ihren engen gegenseitigen Zusammenhang erkannte und sie, gestützt darauf, zur Grundlage ganz neuer methodischer Rechenverfahren machte, bildet den Beginn der eigentlichen systematischen Integral- und Differentialrechnung. Das Verdienst, diese Entwicklung angebahnt zu haben, gebührt gleichmäßig den zwei großen Geistern des 17. Jahrhunderts *Newton* und *Leibniz*, die, wie man heute weiß, ihre Entdeckungen unabhängig voneinander machten. Wenn vielleicht Newton in seinen Untersuchungen zu größerer begrifflicher Klarheit durchdrang, so haben sich doch die Leibnizschen Bezeichnungen und Rechenmethoden in höherem Grade durchgesetzt als die Newtonschen; noch heute bilden diese formalen Seiten der Leibnizschen Gedankenentwicklung ein unentbehrliches Element in der Theorie.

## § 1. Das bestimmte Integral.

Als ersten Punkt behandeln wir das Integral, welches aus sachlichen und historischen Gründen bei jeder Darstellung der höheren Analysis viel mehr im Vordergrund stehen müßte, als dies einer auf Zufälligkeiten beruhenden und bis heute fortwirkenden Lehrtradition entspricht. Das Integral tritt uns ursprünglich entgegen bei dem Problem, den Flächeninhalt eines krummlinig begrenzten Stückes der Ebene auszumessen. Eine verfeinerte Betrachtung führt dann sofort zu einer Loslösung des Integralbegriffes von der naiven anschaulichen Vorstellung des Flächeninhaltes und zu einer rein auf dem Zahlbegriff aufgebauten analytischen Fassung. Die Bedeutung dieser analytischen Integraldefinition werden wir nicht nur darin erkennen, daß sie allein uns volle begriffliche Klarheit liefert, sondern ebenso sehr auch in der Möglichkeit zu mannigfachen, über die Flächeninhaltsbestimmung weit hinausgehenden Anwendungen.

Wir beginnen mit einer anschaulichen Betrachtung.

### 1. Das Integral als Flächeninhalt.

Ist eine in einem Intervall stetige positive Funktion $f(x)$ gegeben und sind $x = a$ und $x = b$ $(a < b)$ zwei Werte in diesem Intervall, so denken wir uns die Funktion durch eine Kurve repräsentiert und betrachten den Inhalt der Fläche, welche von der Kurve einerseits, den beiden geraden Linien $x = a$ und $x = b$ andererseits und schließlich dem Stück der $x$-Achse zwischen den Punkten $a$ und $b$ begrenzt ist (Fig. 26).

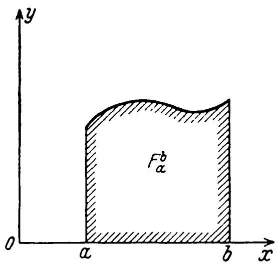



Fig. 26.

Daß es einen bestimmten Sinn hat, von diesem Flächeninhalt zu sprechen, ist eine von der Anschauung inspirierte Annahme, welche wir hier ausdrücklich als Voraussetzung formulieren. Wir nennen diesen Flächeninhalt $F_a^b$ das *bestimmte Integral der Funktion $f(x)$ zwischen den Grenzen $a$ und $b$.* Wenn wir versuchen, diesen Flächeninhalt wirklich durch eine Maßzahl zu charakterisieren, so werden wir davon ausgehen müssen, daß wir zwar nicht Flächenstücke mit krummlinigen Rändern, wohl aber geradlinig begrenzte Polygone ausmessen können, indem wir sie in Dreiecke bzw. Rechtecke zerlegen. Dies ist im allgemeinen exakt bei unserem Flächenstück nicht möglich. Wohl aber liegt es sehr nahe, den Flächeninhalt in folgender Weise als Grenzwert einer Summe von Rechtecksinhalten aufzufassen. Teilen wir die Strecke zwischen den beiden Punkten $a$ und $b$ der Abszissenachse in $n$ gleiche Teile und errichten in allen Teilpunkten die Ordinaten bis zur Kurve, so wird die ganze Fläche ebenfalls in $n$ Streifen zerlegt. Jeden dieser Streifen können wir zwar im allgemeinen ebensowenig direkt mit Hilfe der Funktion $f(x)$ berechnen wie den genannten Flächeninhalt; aber wenn wir, wie in Figur 27 ersichtlich, in jedem Teilintervall jeweils den kleinsten und den größten Funktionswert von $f(x)$ aufsuchen und den betreffenden Streifen einmal durch ein Rechteck ersetzen, dessen Höhe gleich dem kleinsten Funktionswert ist, das andere Mal durch ein Rechteck, dessen Höhe gleich dem größten Funktionswert ist, so erhalten wir zwei treppenförmige Figuren; in der einen ist der treppenförmige Linienzug ausgezogen, in der anderen punktiert. Die erste treppenförmige Figur besitzt offenbar einen Flächeninhalt, der höchstens so groß ist wie der zu bestimmende Flächeninhalt $F_a^b$; die zweite einen Flächeninhalt, der mindestens so groß ist wie $F_a^b$. Bezeichnen wir die Summe der Flächeninhalte der in der einen bzw. in der anderen Art gebildeten Rechtecke mit $\underline{F_n}$ („*Untersumme*") bzw. $\overline{F_n}$ („*Obersumme*"), so gilt die Beziehung

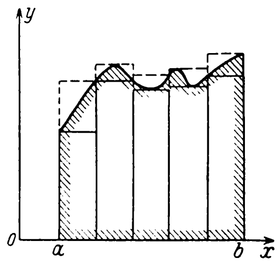


Fig. 27. Obersumme und Untersumme.

$$\underline{F_n} \leqq F_a^b \leqq \overline{F_n}.$$

Machen wir nun die Einteilung feiner und feiner, d. h. lassen wir $n$ über alle Grenzen wachsen, so entnehmen wir der Anschauung, daß die beiden Größen $\overline{F_n}$ und $\underline{F_n}$ sich einander immer mehr nähern und dem gemeinsamen Grenzwert $F_a^b$ zustreben werden. Wir können also unser Integral als den Grenzwert

$$F_a^b = \lim_{n \to \infty} \underline{F_n} = \lim_{n \to \infty} \overline{F_n}$$

auffassen.

Die anschauliche Betrachtung zeigt uns sofort die Möglichkeit einer Verallgemeinerung. Es war keineswegs notwendig, die $n$ Teilintervalle einander gleich lang zu machen. Sie dürfen vielmehr verschiedene Längen besitzen, wenn wir nur voraussetzen, daß bei wachsendem $n$ die Länge des längsten der Teilintervalle gegen 0 strebt.

## 2. Die analytische Definition des Integrales.

Während wir soeben das bestimmte Integral als eine durch den Flächeninhalt gegebene, gewissermaßen von vornherein bekannte Zahl angesehen haben, die wir nur hinterher als Grenzwert darstellen konnten, wollen wir jetzt das Verhältnis umkehren. Wir wollen uns nicht mehr auf den Standpunkt stellen, daß wir aus der Anschauung schon wüßten, daß und wie einer krummen stetigen Linie in der obigen Weise ein Flächeninhalt zugeordnet werden kann; sondern wir wollen umgekehrt von rein analytisch gebildeten Summen, wie den vorhin definierten Ober- und Untersummen, ausgehen und dann beweisen, daß diese Summen einem bestimmten Grenzwert zustreben. Diesen Grenzwert betrachten wir als Definition des Integrals bzw. des Flächeninhaltes. Wir werden dabei ganz von selbst auf diejenigen formalen Bezeichnungen geführt, welche seit Leibniz für die Integralrechnung üblich geworden sind.

Es sei $f(x)$ eine positive stetige Funktion im Intervall $a \leqq x \leqq b$. Wir denken uns das Intervall von der Länge $b - a$ durch $n - 1$ Teilpunkte $x_1, x_2, \ldots, x_{n-1}$ in $n$ beliebige gleiche oder ungleiche Teilintervalle geteilt und außerdem $x_0 = a$, $x_n = b$ gesetzt; in jedem Teilintervall wählen wir einen ganz beliebigen Punkt, der im Innern oder am Rande des Intervalls liegen kann, etwa im ersten Intervall den Punkt $\xi_1$, im zweiten den Punkt $\xi_2, \ldots$, im $n$-ten den Punkt $\xi_n$. Wir betrachten nunmehr statt der stetigen Funktion $f(x)$ eine unstetige Funktion (Treppenfunktion), die im ersten Teilintervall den konstanten Wert $f(\xi_1)$, im zweiten Teilintervall den konstanten Wert $f(\xi_2), \ldots$, im $n$-ten den konstanten Wert $f(\xi_n)$ besitzt. Das Bild dieser Treppenfunktion definiert auf die in der Figur 28 angegebenen Art eine Reihe von Rechtecken, deren Inhaltssumme durch den Ausdruck

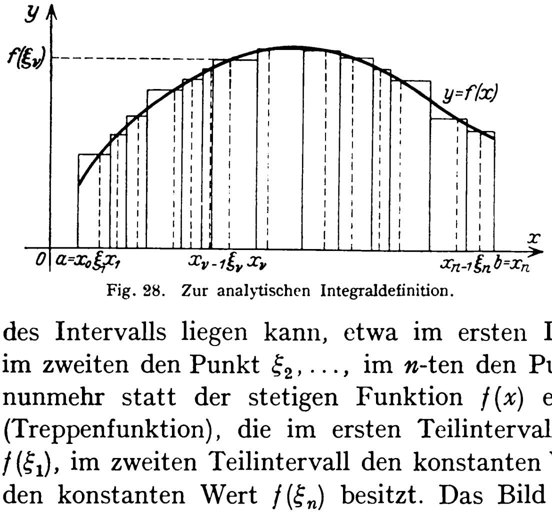


Fig. 28. Zur analytischen Integraldefinition.

$$F_n = (x_1 - x_0)\,f(\xi_1) + (x_2 - x_1)\,f(\xi_2) + \cdots + (x_n - x_{n-1})\,f(\xi_n)$$

gegeben ist. Abgekürzt pflegt man sie unter Verwendung des Summenzeichens $\Sigma$ in der Form

$$F_n = \sum_{\nu=1}^{n} (x_\nu - x_{\nu-1})\,f(\xi_\nu)$$

zu schreiben oder schließlich, wenn wir zur weiteren Abkürzung den Ausdruck

$$\Delta x_\nu = x_\nu - x_{\nu-1}$$

einführen, in der Form

$$F_n = \sum_{\nu=1}^{n} f(\xi_\nu)\, \Delta x_\nu .$$

— Das Symbol $\Delta$ ist hier nicht etwa ein Faktor, sondern bedeutet „*Differenz*". — Unsere grundlegende Behauptung lautet nun folgendermaßen: *Lassen wir die Anzahl der Teilpunkte über alle Grenzen wachsen und dabei zugleich die Länge $\Delta x$ des längsten Teilintervalles gegen* 0 *streben, so strebt die obige Summe einem Grenzwerte zu. Der Grenzwert ist unabhängig davon, wie wir die Teilpunkte und die Zwischenwerte $\xi_1$, $\xi_2, \ldots, \xi_n$ in den einzelnen Teilintervallen gewählt haben.*

Diesen Grenzwert nennen wir das ***bestimmte Integral*** der Funktion $f(x)$ zwischen den Grenzen $a$ und $b$ und betrachten ihn, wie schon erwähnt, geradezu als Definition des Flächeninhaltes[1]).

Wir können und müssen diesen Satz von der Existenz des Integrales einer stetigen Funktion rein analytisch ohne Berufung auf die Anschauung beweisen. Ich will den Beweis jedoch hier noch übergehen und erst im Anhang am Schluß des Kapitels nachholen, nachdem der Leser durch den Erfolg unserer Begriffsbildung ein größeres Interesse an ihrer exakten Fundierung gewonnen hat. Einstweilen wollen wir uns damit begnügen, daß die anschauliche Betrachtung aus Nr. 1 uns den Satz außerordentlich plausibel macht.

### 3. Ergänzungen, Bezeichnungen und Grundregeln für das bestimmte Integral.

Die eben gegebene Definition des Integrals als Grenzwert einer Summe hat Leibniz dazu veranlaßt, das Integral durch folgendes Symbol auszudrücken:

$$\int_a^b f(x)\, dx .$$

Das Integralzeichen ist dabei entstanden durch Stilisierung eines Summenzeichens, welches die Gestalt eines lateinischen $S$ hatte. Der Grenzübergang von der Intervalleinteilung in endliche Differenzen $\Delta x_\nu$ zu verschwindenden Differenzen ist angedeutet, indem statt des Symboles $\Delta$ das Symbol $d$ geschrieben ist. Wir müssen uns aber davor hüten, hier etwa das $dx$ als „unendlich kleine Größe" und das Integral als „Summe aus unendlich vielen unendlich kleinen Summanden" aufzufassen; eine solche Auffassung würde jedes klaren Sinnes entbehren; was in ihr einem richtigen sachgemäßen Gefühl entspricht, wird eben gerade durch den oben ausgeführten Grenzübergang präzisiert.

---

[1]) Man kann natürlich den Flächeninhaltsbegriff auch rein geometrisch definieren und dann die Äquivalenz einer solchen geometrischen Definition mit der obigen Grenzwertdefinition beweisen. Vgl. fünftes Kapitel, § 2, Nr. 1.

In unseren obigen Figuren hatten wir einmal vorausgesetzt, daß die Funktion $f(x)$, der „*Integrand*", in dem ganzen Intervall positiv ist, und zweitens, daß $b > a$ ist. Die Formel, welche das Integral als Grenzwert einer Summe definiert, ist aber von jeder solchen Voraussetzung unabhängig. Ist zunächst $f(x)$ in unserem Intervall oder in einem Teile desselben negativ, so sind eben einfach in unserer Summe die betreffenden Faktoren $f(\xi_\nu)$ negativ. Wir werden dem betreffenden von der Kurve begrenzten Flächenstück dann naturgemäß einen negativen Flächeninhalt zuschreiben, was ja mit dem von der analytischen Geometrie her bekannten Vorzeichenprinzip durchaus im Einklang steht. Der Gesamtflächeninhalt, der von einem Kurvenstück begrenzt ist, wird sich also im allgemeinen aus positiven und negativen Summanden zusammensetzen, je nach der Anzahl der Kurventeile, die oberhalb oder unterhalb der $x$-Achse verlaufen [1]).

Lassen wir auch noch die Voraussetzung $a < b$ fallen und nehmen zunächst an, es sei $a > b$, so können wir unsere arithmetische Integraldefinition doch beibehalten; nur werden die Differenzen $\Delta x_\nu$, wenn wir das Intervall von $a$ bis $b$ durchlaufen, negativ werden. Wir werden so ohne weiteres zu der für beliebiges $a$ und $b$ gültigen Relation

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$$

geführt. Ihr entsprechend setzen wir fest:

$$\int_a^a f(x)\,dx = 0.$$

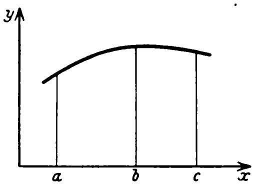


Fig. 29.

Ebenso ergibt sich aus unserer Integraldefinition ohne weiteres (vgl. Fig. 29) die **grundlegende Beziehung**

$$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx$$

für $a < b < c$. Wegen der obigen Beziehungen gilt diese Gleichung ohne weiteres auch bei beliebiger Lage der drei Punkte $a$, $b$, $c$.

Zu einer wichtigen einfachen **Grundregel** über das Integral gelangen wir, wenn wir die Funktion $cf(x)$ betrachten, wobei $c$ eine Konstante ist. Aus der Definition des Integrales ergibt sich unmittelbar

$$\int_a^b c\,f(x)\,dx = c\int_a^b f(x)\,dx.$$

Weiter erwähne ich noch die folgende **Summenregel**: Wenn

$$f(x) = \varphi(x) + \psi(x)$$

[1]) Über den durch beliebige geschlossene Kurven begrenzten Flächeninhalt vgl. fünftes Kapitel, § 2.

ist, so wird

$$\int_a^b f(x)\,dx = \int_a^b \varphi(x)\,dx + \int_a^b \psi(x)\,dx\,.$$

Ihr Beweis ist ebenfalls sehr einfach.

Endlich noch eine an sich selbstverständliche, aber doch für die Handhabung wichtige Bemerkung über die Bezeichnung der „*Integrationsvariablen*". Wir haben unser Integral in der Form $\int_a^b f(x)\,dx$ geschrieben. Für die Auswertung des Integrals ist es belanglos, ob wir die Abszissen des Koordinatensystemes, d. h. die unabhängige Veränderliche mit $x$ oder irgendwie anders bezeichnen. Es ist daher vollständig gleichgültig, wie wir diese Integrationsvariable nennen, und wir können ebensogut wie $\int_a^b f(x)\,dx$ auch $\int_a^b f(t)\,dt$ oder $\int_a^b f(u)\,du$ oder ähnlich schreiben.

## § 2. Beispiele.

Wir können den Grenzprozeß, den unsere Integraldefinition vorschreibt, nun tatsächlich in vielen Fällen bis zu einer vollständigen Berechnung des fraglichen Flächeninhaltes durchführen und wollen dies an einer Reihe von Beispielen erläutern, wobei wir uns, außer in Nr. 5, nur der Ober- oder Untersummen bedienen werden[1]).

### 1. Erstes Beispiel.

Wir wollen zunächst die Funktionen $f(x) = x^n$ betrachten, wo $n$ eine ganze Zahl $\geqq 0$ ist. Für $n = 0$, d. h. für $f(x) = 1$ ist das Ergebnis ohne jeden Grenzübergang so selbstverständlich, daß ich es einfach hinschreiben kann:

$$\int_a^b 1\,dx = \int_a^b dx = b - a\,.$$

Auch für die Funktion $f(x) = x$ ist die Integration geometrisch eine Trivialität. Das Integral der Funktion $f(x) = x$

$$\int_a^b x\,dx$$

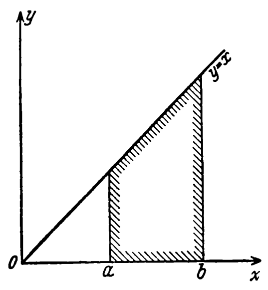


Fig. 30.

ist einfach der Flächeninhalt des in Figur 30 angedeuteten Trapezes und besitzt somit nach einer elementaren Formel den Wert

$$(b-a)\,\frac{b+a}{2} = \frac{b^2-a^2}{2}\,.$$

[1]) Ich überlasse es dem Leser als nützliche Übungsaufgabe, sich in den nachfolgenden Beispielen selbst davon zu überzeugen, daß tatsächlich bei Benutzung von Obersummen und Untersummen derselbe Grenzwert entsteht.

Bestätigen wir, daß dieses Resultat auch auf Grund unserer Grenzwertbildung herauskommt! Dabei können wir uns zur Berechnung der Grenzwerte auf die Diskussion der Obersummen oder Untersummen beschränken. Wir teilen das Intervall von $a$ bis $b$ in $n$ gleiche Teile durch die Teilpunkte

$$a + h, \quad a + 2h, \quad \ldots, \quad a + (n-1)h,$$

wobei $h = \frac{b-a}{n}$ ist. Das Integral muß dann der Grenzwert der folgenden Untersumme (bei $b > a$) bzw. Obersumme (bei $b < a$) sein:

$$\begin{aligned} &h\{a + (a+h) + (a+2h) + \cdots + (a + (n-1)h)\} \\ &= h\{na + h + 2h + \cdots + (n-1)h\}. \end{aligned}$$

Nun ist

$$1 + 2 + \cdots + (n-1) = \frac{(n-1)n}{2},$$

und unser Ausdruck geht daher über in

$$nh\left\{a + h\frac{n-1}{2}\right\} = (b-a)\left(a + \frac{b-a}{2}\frac{n-1}{n}\right).$$

Bei wachsendem $n$ strebt die rechte Seite offenbar gerade gegen den Grenzwert

$$(b-a)\left(a + \frac{b-a}{2}\right) = \frac{b^2 - a^2}{2},$$

was bewiesen werden sollte.

## 2. Zweites Beispiel.

Nicht ganz so einfach ist das Beispiel der Funktion $f(x) = x^2$ oder, geometrisch gesprochen, die Bestimmung des von einem Parabelsegment, einem Stück $x$-Achse und zwei Ordinaten begrenzten Flächenstückes. Wir berechnen etwa das Integral

$$\int_0^b x^2\,dx$$

für $b \geqq 0$ (vgl. Fig. 31) und teilen das Intervall $0 \leqq x \leqq b$ in $n$ gleiche Teile von der Länge $h = \frac{b}{n}$; dann ist der gesuchte Flächeninhalt der Parabel der Grenzwert des folgenden Ausdruckes (Obersumme):

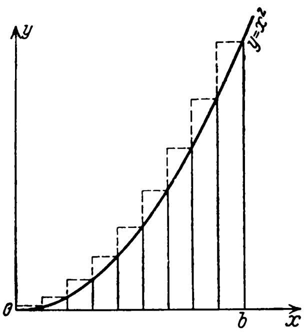


Fig. 31.

$$\begin{aligned} h(h^2 + 2^2h^2 + 3^2h^2 + \cdots + n^2h^2) &= h^3(1^2 + 2^2 + \cdots + n^2) \\ &= \frac{b^3}{n^3}(1^2 + 2^2 + \cdots + n^2). \end{aligned}$$

Die Summe in der Klammer haben wir aber auf S. 19 (Anm. 1) durch die folgende Gleichung ausgedrückt:

$$1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

Setzen wir diesen Wert oben ein, so erhält unsere Summe nach einer selbstverständlichen Umformung die Gestalt

$$\frac{b^3}{6}\left(1 + \frac{1}{n}\right)\left(2 + \frac{1}{n}\right).$$

Bei unbegrenzt wachsendem $n$ ergibt sich also als Grenzwert $\frac{b^3}{3}$, und wir erhalten die gesuchte Integralformel

$$\int_0^b x^2\,dx = \frac{b^3}{3}.$$

Hieraus ergibt sich sofort nach unseren obigen Relationen die allgemeinere Formel

$$\int_a^b x^2\,dx = \int_0^b x^2\,dx - \int_0^a x^2\,dx = \frac{b^3 - a^3}{3}.$$

### 3. Integration von $x^\alpha$ bei beliebigem positiven ganzzahligen $\alpha$.

Als drittes Beispiel betrachten wir die Integration der Potenz

$$y = f(x) = x^\alpha,$$

wo $\alpha$ eine beliebige ganze positive Zahl sei. Für die Ausführung des Integrales

$$\int_a^b x^\alpha\,dx,$$

wobei wir $0 < a < b$ voraussetzen, würde es unpraktisch sein, das Intervall in $n$ gleiche Teile einzuteilen[1]). Der Grenzübergang vollzieht sich aber sehr einfach, wenn man eine Einteilung in „geometrischer Progression" folgendermaßen vornimmt: Wir setzen $\sqrt[n]{\frac{b}{a}} = q$ und teilen das Intervall durch die Punkte

$$a, \quad aq, \quad aq^2, \quad \ldots, \quad aq^{n-1}, \quad aq^n = b.$$

[1]) Wir müßten die Ausführung des Integrales dann auf die Bestimmung der Grenzwerte von $\frac{1}{n^{\alpha+1}}(1^\alpha + 2^\alpha + \cdots + n^\alpha)$ für $n \to \infty$ stützen, was der Leser nach den Bemerkungen von S. 19, Anm. 1 selbst durchführen mag.

Das gesuchte Integral ist dann der Grenzwert der folgenden Summe:

$$a^\alpha(aq-a)+(aq)^\alpha(aq^2-aq)+(aq^2)^\alpha(aq^3-aq^2)+\cdots+(aq^{n-1})^\alpha(aq^n-aq^{n-1})$$
$$=a^{\alpha+1}(q-1)\{1+q^{\alpha+1}+q^{2(\alpha+1)}+q^{3(\alpha+1)}+\cdots+q^{(n-1)(\alpha+1)}\},$$

für unbegrenzt zunehmendes $n$.

In der letzten Klammer steht eine geometrische Reihe mit dem Quotienten $q^{\alpha+1} \neq 1$. Summieren wir sie, so erhalten wir für den gesamten Ausdruck den Wert

$$a^{\alpha+1}(q-1)\frac{q^{n(\alpha+1)}-1}{q^{\alpha+1}-1}.$$

Setzen wir hier $q=\left(\frac{b}{a}\right)^{\frac{1}{n}}$ ein, so ergibt sich für unsere Summe der Ausdruck

$$(b^{\alpha+1}-a^{\alpha+1})\frac{q-1}{q^{\alpha+1}-1}.$$

Lassen wir nun $n$ über alle Grenzen wachsen, so behält der erste Faktor seinen Wert; der zweite Faktor, den wir wegen $q \neq 1$ nach der eben benutzten Formel für die geometrische Reihe in die Gestalt

$$\frac{1}{q^\alpha+q^{\alpha-1}+\cdots+1}$$

setzen können, wird, da $q$ wegen $q=\left(\frac{b}{a}\right)^{\frac{1}{n}}$ mit $n\to\infty$ gegen 1 strebt, den Grenzwert $\frac{1}{\alpha+1}$ haben, und somit ergibt sich schließlich der gesuchte Wert unseres Integrales durch die Formel

$$\int_a^b x^\alpha\,dx=\frac{1}{\alpha+1}(b^{\alpha+1}-a^{\alpha+1}).$$

Wir werden später (vgl. § 4) sehen, daß wir diese zwar im Prinzip einfache, aber in der Durchführung doch etwas umständliche Rechnung vollständig vermeiden können, wenn wir unser Integrationsproblem in einem etwas größeren methodischen Zusammenhang anfassen.

### 4. Integration von $x^\alpha$ für beliebiges rationales $\alpha \neq -1$.

Das gewonnene Ergebnis läßt sich ohne wesentliche Komplikation der Überlegungen erheblich verallgemeinern. Es sei $\alpha=\frac{r}{s}$ eine rationale positive Zahl, $r$ und $s$ ganze positive Zahlen; dann ändert sich an unserer eben durchgeführten Bestimmung des Integrales nichts als die Bestimmung des Grenzwertes von $\frac{q-1}{q^{\alpha+1}-1}$ für $q\to 1$. Dieser Ausdruck ist jetzt einfach $\frac{q-1}{q^{\frac{r+s}{s}}-1}$. Setzen wir $q^{\frac{1}{s}}=\tau\,(\neq 1)$, so wird $\tau$ zugleich mit $q$ gegen 1 streben, und wir haben also den Grenzwert von $\frac{\tau^s - 1}{\tau^{r+s} - 1}$ für $\tau \to 1$ aufzusuchen. Dieser Grenzwert ergibt sich, wenn wir Zähler und Nenner zunächst durch $\tau - 1$ dividieren und die vorhin angewandte algebraische Umformung sowohl für den Zähler als auch für den Nenner vornehmen, einfach als

$$\lim_{\tau \to 1} \frac{\tau^{s-1} + \tau^{s-2} + \cdots + 1}{\tau^{r+s-1} + \tau^{r+s-2} + \cdots + 1}$$

und läßt sich unmittelbar bestimmen, indem wir im Zähler und Nenner, die beide in $\tau$ stetig sind, einfach $\tau = 1$ einsetzen; so ergibt sich der Wert $\frac{s}{r+s} = \frac{1}{\alpha + 1}$, so daß wir auch für beliebiges positives rationales $\alpha$ die Integralformel

$$\int_a^b x^\alpha \, dx = \frac{1}{\alpha + 1} (b^{\alpha+1} - a^{\alpha+1})$$

erhalten. Diese Formel bleibt aber auch für negatives rationales $\alpha$ gültig, wenn wir nur $\alpha \neq -1$ annehmen, wo offenbar die oben angewandte Summationsformel für die geometrische Reihe ihren Sinn verliert. Um bei negativem $\alpha = -\frac{r}{s}$ die Grenzwertbestimmung von $\frac{q-1}{q^{\alpha+1} - 1}$ vorzunehmen, setzen wir $q^{-\frac{1}{s}} = \tau$; demgemäß ist $q = \tau^{-s}$ und $q^{\alpha+1} = q^{-\frac{r}{s}+1} = q^{-\frac{r-s}{s}} = \tau^{r-s}$, und wir haben den Grenzwert von

$$\frac{\tau^{-s} - 1}{\tau^{r-s} - 1} = \frac{1 - \tau^s}{\tau^r - \tau^s}$$

zu bestimmen. Ich kann es dem Leser überlassen, zu beweisen, daß dieser Grenzwert wiederum gleich $\frac{1}{\alpha+1}$ ist, und somit ergibt sich ganz allgemein für positives oder negatives rationales $\alpha$, mit Ausnahme des Wertes $\alpha = -1$, die Integralformel

$$\int_a^b x^\alpha \, dx = \frac{1}{\alpha + 1} (b^{\alpha+1} - a^{\alpha+1}).$$

Die Gestalt der rechten Seite zeigt uns im übrigen deutlich, daß die Formel versagt, wenn $\alpha = -1$ ist, weil im Zähler und Nenner des Quotienten dann Null stehen würde.

Es liegt nahe, zu vermuten, daß sich der Gültigkeitsbereich unserer letzten Formel auch auf irrationale Werte von $\alpha$ ausdehnt. In der Tat werden wir dies in § 7 durch einen einfachen Grenzübergang bestätigen.

### 5. Integration von $\sin x$ und $\cos x$.

Als letztes Beispiel, das wir ebenfalls mit Hilfe eines speziellen Kunstgriffes behandeln wollen, betrachten wir die Funktion $f(x) = \sin x$. Wir werden das Integral

$$\int_a^b \sin x\, dx$$

als Grenzwert der folgenden Summe auffassen:

$$S_h = h\,(\sin(a+h) + \sin(a+2h) + \cdots + \sin(a+nh)),$$

wo $h = \frac{b-a}{n}$ gesetzt ist. Multiplizieren wir die Klammer rechts mit $2\sin\frac{h}{2}$ und berücksichtigen die bekannte trigonometrische Formel

$$2\sin u \sin v = \cos(u-v) - \cos(u+v),$$

so erhalten wir, wenn $h$ kein Vielfaches von $2\pi$ ist, die Formel

$$\begin{aligned} S_h &= \frac{h}{2\sin\frac{h}{2}} \left\{ \begin{aligned} &\cos\left(a+\frac{h}{2}\right) - \cos\left(a+\frac{3}{2}h\right) + \cos\left(a+\frac{3}{2}h\right) - \cos\left(a+\frac{5}{2}h\right) \\ &\quad + \cdots + \cos\left(a+\frac{2n-1}{2}h\right) - \cos\left(a+\frac{2n+1}{2}h\right) \end{aligned} \right\} \\ &= \frac{h}{2\sin\frac{h}{2}} \left\{ \cos\left(a+\frac{h}{2}\right) - \cos\left(a+\frac{2n+1}{2}h\right) \right\}. \end{aligned}$$

Das Integral wird also wegen $a + nh = b$ der Grenzwert von

$$\frac{h}{2\sin\frac{h}{2}} \left\{ \cos\left(a+\frac{h}{2}\right) - \cos\left(b+\frac{h}{2}\right) \right\}$$

für $h \to 0$.

Nun wissen wir aus dem ersten Kapitel, daß bei abnehmendem $h$ der Ausdruck $\frac{h}{2} \cdot \frac{1}{\sin\frac{h}{2}}$ gegen 1 strebt. Der gesuchte Grenzwert wird also einfach $\cos a - \cos b$, und wir erhalten die Integralformel

$$\int_a^b \sin x\, dx = -(\cos b - \cos a).$$

Ganz ebenso erhalten wir, wie der Leser selbst überlegen mag, die Integralformel

$$\int_a^b \cos x\, dx = \sin b - \sin a.$$

Fast jedes der behandelten Beispiele haben wir auf Grund einer besonderen Überlegung, eines besonderen Kunstgriffes durchführen müssen. Es wird nun gerade der wesentliche Punkt in der methodischen Integral-und Differentialrechnung sein, daß an Stelle solcher spezieller Kunstgriffe Überlegungen allgemeinen Charakters treten, die ganz zwangsläufig zu den gewünschten Resultaten führen. Um zu diesen Methoden zu gelangen, müssen wir uns dem anderen Grundbegriffe der höheren Analysis, dem des Differentialquotienten, zuwenden.

## § 3. Die Ableitung oder der Differentialquotient.

Ebenso wie der Integralbegriff ist auch der Begriff der Ableitung oder des Differentialquotienten anschaulichen Ursprunges. Seine Quellen sind einmal die Aufgabe, an eine gegebene Kurve in einem Punkte die *Tangente* zu legen, und sodann die Aufgabe, für den Begriff der *Geschwindigkeit* einer beliebigen Bewegung eine präzise Definition zu finden.

### 1. Differentialquotient und Kurventangente.

Ich knüpfe zunächst an das Tangentenproblem an. Ist $P$ ein Punkt auf einer gegebenen krummen Linie (vgl. Fig. 32), so werden wir der naiven Anschauung gemäß die Kurventangente in $P$ durch folgenden geometrischen Grenzübergang kennzeichnen: Wir betrachten außer dem Punkte $P$ einen zweiten Punkt $P_1$ auf der Kurve und legen durch die beiden Punkte $P$, $P_1$ eine gerade Linie, eine Sekante der Kurve. Lassen wir dann den Punkt $P_1$ längs der Kurve in den Punkt $P$ hineinrücken, so strebt diese Sekante einer Grenzlage zu, welche ganz unabhängig davon ist, ob $P_1$ von links oder von rechts in $P$ hineinrückt. Diese Grenzlage der Sekante ist die Tangente, und daß tatsächlich eine solche bestimmte Grenzlage der Sekante existiert, ist gleichbedeutend mit der Annahme, daß die Kurve im Punkte $P$ eine bestimmte Tangente oder eine bestimmte Richtung besitzt. (Das Wort „Annahme" soll uns dabei andeuten, daß wirklich eine Voraussetzung zugrunde liegt, die selbst bei stetigen Kurven nicht immer erfüllt zu sein braucht — z. B. an jeder Ecke einer Kurve nicht gilt.)

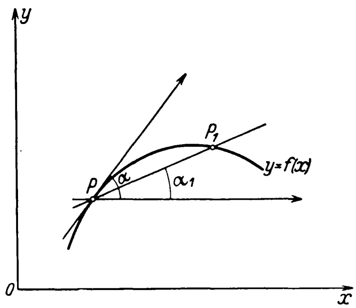


Fig. 32. Sehne und Tangente.

Sobald wir nun die Kurve durch eine Funktion $y = f(x)$ dargestellt haben, entsteht die Aufgabe, den geometrischen Grenzübergang analytisch mit Hilfe dieser Funktion $f(x)$ darzustellen. Verstehen wir unter dem Winkel, den die positive $x$-Achse mit einer gerichteten Geraden $g$ bildet, denjenigen Winkel, um den man die positive $x$-Achse in positivem Sinne[1]) drehen muß, bis sie zum ersten Male der Geraden $g$

[1]) D. h. in dem Sinne, bei welchem sie durch eine Drehung um $\frac{\pi}{2}$ in die positive $y$-Achse übergeht.

parallel wird, und bezeichnen wir den Winkel, den die Sekante $PP_1$ mit der positiven $x$-Achse bildet, mit $\alpha_1$, den Winkel, den entsprechend die Tangente mit der positiven $x$-Achse bildet, mit $\alpha$ (vgl. Fig. 32), dann wird offenbar, wenn man zunächst von dem Fall einer wagerechten Tangente absieht, in unmittelbar verständlicher Bezeichnung

$$\lim_{P_1 \to P} \alpha_1 = \alpha .$$

Wenn $x, y = f(x)$ und $x_1, y_1 = f(x_1)$ die Koordinaten der Punkte $P$ bzw. $P_1$ bedeuten, so erhalten wir unmittelbar

$$\operatorname{tg} \alpha_1 = \frac{y_1 - y}{x_1 - x} = \frac{f(x_1) - f(x)}{x_1 - x} \text{ [1]};$$

und somit stellt sich unser Grenzübergang durch die Gleichung

$$\lim_{x_1 \to x} \frac{f(x_1) - f(x)}{x_1 - x} = \operatorname{tg} \alpha$$

dar. Den Ausdruck

$$\frac{f(x_1) - f(x)}{x_1 - x} = \frac{y_1 - y}{x_1 - x} = \frac{\Delta y}{\Delta x}$$

bezeichnen wir als den *Differenzenquotienten* der Funktion $y = f(x)$, indem wir durch die Symbole $\Delta y$ und $\Delta x$ die Differenzen der Funktion $y = f(x)$ und der unabhängigen Variabeln $x$ bezeichnen. ($\Delta$ bedeutet also hier, wie auch in § 1, nicht einen Faktor, sondern eine Abkürzung für Differenz.) Es ist also der Tangens des Richtungswinkels der Kurve gleich dem Grenzwert, den der Differenzenquotient unserer Funktion erhält, wenn $x_1$ gegen $x$ strebt.

Diesen Grenzwert nennen wir die *Ableitung* oder den *Differentialquotienten* der Funktion $y = f(x)$ an der Stelle $x$ und bezeichnen ihn nach *Lagrange* durch das Symbol $y' = f'(x)$ oder nach *Leibniz* durch das Symbol $\frac{dy}{dx}$ oder $\frac{df(x)}{dx}$ oder $\frac{d}{dx} f(x)$. Die Leibnizsche Schreibweise werden wir ihrer Bedeutung nach in Nr. 7 noch näher diskutieren; hier weise ich darauf hin, daß wir durch die Bezeichnung $f'(x)$ zum Ausdruck gebracht haben, daß die Ableitung wiederum eine Funktion von $x$ ist, entsprechend der Tatsache, daß sie für jeden Wert $x$ des betrachteten Intervalles zu bilden ist. Ich schreibe noch einmal die Definitionsgleichung der Ableitung hin:

$$f'(x) = \lim_{x_1 \to x} \frac{f(x_1) - f(x)}{x_1 - x}$$

---

[1]) Damit diese Gleichungen einen Sinn haben, müssen wir $0 < | x - x_1 | < \delta$ mit hinreichend kleinem $\delta$ voraussetzen. Entsprechende Voraussetzungen machen wir bei Vorbereitungen von Grenzübergängen im folgenden oft stillschweigend.

oder

$$\frac{dy}{dx} = \frac{df(x)}{dx} = f'(x) = \lim_{x_1 \to x} \frac{f(x_1) - f(x)}{x_1 - x} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h},$$

wobei zuletzt $x_1$ durch $x + h$ ersetzt wurde[1]).

Man kann den Differentialquotienten nicht etwa unmittelbar bilden, indem man in dem Ausdruck des Differenzenquotienten einfach $x_1 = x$, d. h. Zähler und Nenner gleich Null setzt, was zu dem sinnlosen Symbol $\frac{0}{0}$ führen würde. Vielmehr beruht die wirkliche Ausführung des Grenzüberganges im Einzelfalle immer auf gewissen Vorbereitungen (Umformung des Differenzenquotienten). Ist z. B. $f(x) = x^2$, so wird $\frac{f(x_1) - f(x)}{x_1 - x} = \frac{x_1^2 - x^2}{x_1 - x} = x_1 + x$ [2]). Nach dieser Umformung steht rechts eine Funktion von $x_1$, die auch für $x_1 = x$ einen Sinn hat und dort stetig ist, und wir dürfen nun den Grenzübergang $x_1 \to x$ einfach ausführen, indem wir rechts $x_1 = x$ setzen, wobei wir sofort erhalten:

$$f'(x) = \frac{d(x^2)}{dx} = 2x.$$

Die Ausführung eines solchen Überganges, d. h. die wirkliche Bildung des Differentialquotienten, nennt man die *Differentiation* oder das *Differenzieren* der Funktion $f(x)$. Wir werden im Folgenden sehen, daß und wie man diesen Prozeß des Differenzierens bei allen wichtigen Funktionen tatsächlich durchführen kann.

Es ist nun von großer Bedeutung, daß die Aufgabe der Differentiation einer gegebenen Funktion $f(x)$ losgelöst von der geometrischen Anschauung der Tangente einen bestimmten Sinn besitzt. Ganz ebenso wie wir bei der Integraldefinition uns von der ursprünglichen geometrischen Anschauung des Flächeninhaltes losgelöst und im Gegenteil den Flächeninhaltsbegriff auf die Integraldefinition gestützt hatten, werden wir jetzt unabhängig von der geometrischen Deutung einer Funktion $y = f(x)$ durch eine Kurve als die Ableitung der Funktion $y = f(x)$ die neue Funktion $y' = f'(x)$ gemäß der obigen Gleichung definieren, vorausgesetzt, daß der Grenzwert des Differenzenquotienten existiert oder, wie man auch sagt, daß die Funktion *differenzierbar* ist. Wir werden diese Voraussetzung der Differenzierbarkeit im übrigen stillschweigend immer dort machen, wo nichts besonderes Gegenteiliges gesagt ist[3]). Man muß beachten, daß für die Differenzierbarkeit der Funktion $f(x)$ an der Stelle $x$ der obige Grenzwert von $\frac{f(x+h) - f(x)}{h}$ für $h \to 0$ existieren muß, un-

---

[1]) Gelegentlich findet man in der Literatur, insbesondere in der englischen, auch die auf *Cauchy* zurückgehende Bezeichnung $Df(x)$ („*Derivierte*" von $f$).

[2]) Vgl. auch S. 70, Anm. 1.

[3]) Beispiele für Fälle, wo die Voraussetzung nicht erfüllt ist, werde ich später geben (vgl. Nr. 5).

abhängig davon, wie $h$ gegen Null strebt, ob durch positive oder durch negative Werte oder ohne Vorzeichenbeschränkung.

Nachdem wir den Differentialquotienten $f'(x)$ gewonnen haben, können wir der Kurve als Tangentenrichtung im Punkte $(x, y)$ diejenige Richtung zuschreiben, für welche der Winkel $\alpha$ mit der positiven $x$-Achse durch die Gleichung $\operatorname{tg} \alpha = f'(x)$ gegeben ist. Wir vermeiden so die Schwierigkeiten, welche sich aus der Unbestimmtheit der geometrischen Anschauung ergeben, indem wir die geometrische Definition auf die analytische stützen und nicht umgekehrt.

Selbstverständlich ist immer die Veranschaulichung des Differentialquotienten durch die Kurventangente ein wichtiges Hilfsmittel für

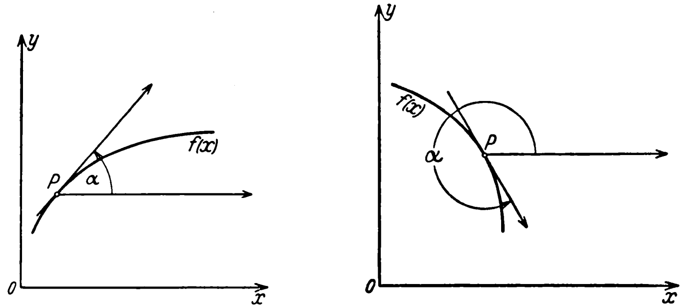



Fig. 33. Tangente bei steigender und fallender Funktion.

das Verständnis auch bei rein analytischen Überlegungen. So halten wir von vornherein die anschauliche Vorstellung fest: *Wenn* $f'(x)$ *positiv ist und man die Kurve in Richtung wachsender* $x$ *durchläuft, so weist die Tangente nach oben; es steigt also die Kurve an der betreffenden Stelle bei wachsendem* $x$ *an; wenn dagegen* $f'(x)$ *negativ ist, weist die Tangente nach unten, die Kurve fällt also.* Vgl. Fig. 33. Analytisch ergibt sich diese Tatsache aus der Bemerkung, daß der Grenzwert von $\frac{f(x+h)-f(x)}{h}$ nur dann positiv sein kann, wenn die Funktion an der Stelle $x$ wächst, d. h., wenn wenigstens für hinreichend nahe an 0 gelegenes positives bzw. negatives $h$ auch $f(x+h) - f(x)$ positiv bzw. negativ ist. (Entsprechendes gilt natürlich für negatives $f'(x)$.)

## 2. Der Differentialquotient als Geschwindigkeit.

Ebenso wie die naive Anschauung uns ein unmittelbares Gefühl für den Begriff der Richtung und damit den der Tangente einer in einem bestimmten Sinne durchlaufenen Kurve gibt, fordert sie von uns auch, einer Bewegung eine Geschwindigkeit zuzuschreiben. Die Definition dieser Geschwindigkeit führt uns ebenfalls auf genau denselben Grenzübergang, den wir soeben als Differentiation bezeichnet haben.

Betrachten wir z. B. die Bewegung eines Punktes auf einer geraden Linie, auf welcher die Lage des Punktes durch die Koordinate $y$, seinen mit einem Vorzeichen versehenen Abstand von einem festen Anfangspunkt, gemessen wird. Die Bewegung ist gegeben, wenn wir $y = f(t)$ als Funktion der Zeit $t$ kennen. Ist diese Funktion $f(t)$ eine ganze lineare Funktion $f(t) = ct + b$, dann nennen wir die Bewegung eine gleichförmige Bewegung mit der Geschwindigkeit $c$ und können für beliebige verschiedene Werte $t$ und $t_1$ schreiben

$$c = \frac{f(t_1) - f(t)}{t_1 - t}.$$

Die Geschwindigkeit ist also der Differenzenquotient der Funktion $ct + b$, und dieser Differenzenquotient ist gänzlich unabhängig davon, welche beiden Zeitmomente $t_1$ und $t$ wir herausgreifen. Was aber werden wir unter der Geschwindigkeit der Bewegung im Zeitmoment $t$ zu verstehen haben, wenn die Bewegung nicht mehr gleichförmig ist?

Wir betrachten, um zu dieser Definition zu gelangen, den Differenzenquotienten $\frac{f(t_1) - f(t)}{t_1 - t}$, den wir als Durchschnittsgeschwindigkeit im Zeitintervall zwischen $t_1$ und $t$ bezeichnen wollen. Wenn nun die Durchschnittsgeschwindigkeit einem bestimmten Grenzwert zustrebt, sobald wir den Zeitmoment $t_1$ immer mehr an den Zeitmoment $t$ heranrücken lassen, so werden wir diesen Grenzwert als die Geschwindigkeit im Zeitmoment $t$ zu definieren haben. Mit anderen Worten: *Die Geschwindigkeit im Zeitmoment t wird durch den Differentialquotienten*

$$f'(t) = \lim_{t_1 \to t} \frac{f(t_1) - f(t)}{t_1 - t}$$

*gemessen.*

Wir sehen an dieser neuen Bedeutung des Differentialquotienten, welche an und für sich mit dem Tangentenproblem nichts mehr zu tun hat, daß es tatsächlich sachgemäß ist, den Grenzübergang des Differenzierens als analytische Operation unabhängig von der geometrischen Anschauung zu definieren. Die Differenzierbarkeit der Bewegungsfunktion $y = f(t)$ ist auch hier wieder eine Voraussetzung, die wir stets stillschweigend machen werden und die durchaus notwendig dafür ist, daß der Geschwindigkeitsbegriff einen Sinn erhält.

Ein einfaches Beispiel für den Zusammenhang zwischen Bewegung und Geschwindigkeit gibt uns der freie Fall. Gehen wir von dem durch Beobachtungen gewonnenen Fallgesetz aus, daß die beim freien Fall von einem Massenpunkt in der Zeit $t$ durchfallene Strecke proportional der Größe $t^2$ ist, daß also diese Strecke durch eine Funktion der Form

$$y = f(t) = at^2$$

mit konstantem $a$ dargestellt wird, so gewinnen wir sofort, ähnlich wie in Nr. 1, für die Geschwindigkeit den Ausdruck $f'(t) = 2at$, welcher uns zeigt, daß die Geschwindigkeit beim freien Fall proportional der Zeit wächst.

### 3. Beispiele.

Wir wollen nun die Differentiation von Funktionen an einer Reihe von Beispielen wirklich durchführen und betrachten als erstes die Funktion $y = f(x) = c$, wo $c$ eine Konstante ist. Es wird $f(x+h) - f(x) = c - c = 0$, also auch $\lim\limits_{h \to 0} \frac{f(x+h) - f(x)}{h} = 0$; d. h. *der Differentialquotient einer Konstanten ist Null.*

Für eine ganze lineare Funktion $y = cx + b$ ergibt sich

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{ch}{h} = c.$$

Weiter differenzieren wir die Funktion

$$y = f(x) = x^\alpha,$$

wobei wir zunächst voraussetzen, daß $\alpha$ eine positive ganze Zahl ist. Im Falle $x_1 \neq x$ wird

$$\frac{f(x_1) - f(x)}{x_1 - x} = \frac{x_1^\alpha - x^\alpha}{x_1 - x},$$

und die rechte Seite ist nach einer bekannten elementaren Formel der Algebra (vgl. auch § 2, Nr. 3) gerade gleich $x_1^{\alpha-1} + x_1^{\alpha-2} x + \cdots + x^{\alpha-1}$. Nach dieser Umformung können wir wegen der Stetigkeit des erhaltenen Ausdruckes unmittelbar den Grenzübergang $x_1 \to x$ ausführen, indem wir in dem letzten Ausdruck einfach überall die Größe $x_1$ durch $x$ ersetzen, so daß jeder Summand gleich $x^{\alpha-1}$ wird. Da die Anzahl der Summanden genau $\alpha$ ist, so erhalten wir

$$y' = f'(x) = \frac{d(x^\alpha)}{dx} = \alpha x^{\alpha-1}.$$

Genau zu demselben Resultat gelangen wir, wenn $\alpha = -\beta$ eine negative ganze Zahl ist; dabei machen wir aber ausdrücklich die Voraussetzung, daß $x$ von 0 verschieden ist. Es wird dann

$$\frac{f(x_1) - f(x)}{x_1 - x} = \frac{\frac{1}{x_1^\beta} - \frac{1}{x^\beta}}{x_1 - x} = -\frac{x^\beta - x_1^\beta}{x - x_1} \cdot \frac{1}{x^\beta x_1^\beta} = -\frac{x^{\beta-1} + x^{\beta-2} x_1 + \cdots + x_1^{\beta-1}}{x_1^\beta x^\beta}.$$

Nunmehr können wir wiederum den Grenzübergang $x_1 \to x$ ausführen, indem wir überall $x_1$ durch $x$ ersetzen. Wir erhalten dann genau wie oben für den Grenzwert den Ausdruck

$$y' = -\beta \frac{x^{\beta-1}}{x^{2\beta}} = -\beta x^{-\beta-1}.$$

Mithin wird auch für *negatives ganzzahliges* $\alpha = -\beta$ für die Funktion $y = x^\alpha$ der Differentialquotient

$$y' = \alpha x^{\alpha-1}.$$

Endlich wollen wir dieselbe Formel für den Fall beweisen, daß $x$ positiv und $\alpha$ eine beliebige rationale Zahl ist, etwa $\alpha = \frac{p}{q}$, wo $p$

und $q$ ganze rationale Zahlen sind, die wir etwa beide als positiv voraussetzen. (Wenn eine von ihnen negativ wäre, so würde sich in der Betrachtung nichts Wesentliches ändern; für $\alpha = 0$ aber ist uns das Ergebnis schon vorhin begegnet.) Es ist jetzt

$$\frac{f(x_1) - f(x)}{x_1 - x} = \frac{x_1^{\frac{p}{q}} - x^{\frac{p}{q}}}{x_1 - x}.$$

Setzen wir $x^{\frac{1}{q}} = \xi$ und $x_1^{\frac{1}{q}} = \xi_1$, so wird

$$\frac{f(x_1) - f(x)}{x_1 - x} = \frac{\xi_1^p - \xi^p}{\xi_1^q - \xi^q} = \frac{\xi_1^{p-1} + \xi_1^{p-2}\xi + \cdots + \xi^{p-1}}{\xi_1^{q-1} + \xi_1^{q-2}\xi + \cdots + \xi^{q-1}}.$$

Nach dieser letzten Umformung können wir den Grenzübergang $x_1 \to x$ oder, was jetzt auf dasselbe herauskommt, $\xi_1 \to \xi$ wieder unmittelbar vornehmen und erhalten für den Grenzwert den Ausdruck

$$y' = \frac{p}{q}\frac{\xi^{p-1}}{\xi^{q-1}} = \frac{p}{q}\xi^{p-q} = \frac{p}{q}x^{\frac{p-q}{q}} = \frac{p}{q}x^{\frac{p}{q}-1}$$

oder schließlich

$$f'(x) = y' = \alpha x^{\alpha-1},$$

also formal genau dasselbe Resultat wie oben. Ich überlasse es dem Leser, selbst zu bestätigen, daß auch für negative rationale Exponenten $\alpha$ dieselbe Differentiationsformel besteht. Im übrigen kommen wir auf die Differentiation der Potenz später noch einmal in mehr systematischem Zusammenhange zurück.

Schließlich behandeln wir als weiteres Beispiel die Differentiation der trigonometrischen Funktionen $\sin x$ und $\cos x$. Es wird unter Benutzung einer elementaren trigonometrischen Formel

$$\frac{\sin(x+h) - \sin x}{h} = \frac{\sin x \cos h + \cos x \sin h - \sin x}{h}$$
$$= \sin x \frac{\cos h - 1}{h} + \cos x \frac{\sin h}{h}.$$

Nun wissen wir aus dem ersten Kapitel, § 7, daß

$$\lim_{h\to 0}\frac{\sin h}{h} = 1, \quad \lim_{h\to 0}\frac{\cos h - 1}{h} = 0$$

wird. Somit erhalten wir unmittelbar als den gesuchten Differentialquotienten

$$y' = \frac{d\sin x}{dx} = \cos x.$$

Ganz ähnlich erfolgt die Differentiation der Funktion

$$y = \cos x.$$

Es wird

$$\frac{\cos(x+h) - \cos x}{h} = \cos x \frac{\cos h - 1}{h} - \sin x \frac{\sin h}{h},$$

und der Grenzübergang $h \to 0$ liefert uns sofort die Ableitung

$$y' = \frac{d \cos x}{d x} = -\sin x.$$

## 4. Einige Grundregeln für die Differentiation.

Ebenso wie beim Integral ergeben sich auch für den Differentialquotienten einige einfache Grundregeln unmittelbar aus der Definition: Ist $\varphi(x) = f(x) + g(x)$, so wird $\varphi'(x) = f'(x) + g'(x)$. Ist $\psi(x) = c f(x)$ ($c$ eine Konstante), so wird $\psi'(x) = c f'(x)$. Denn es ist ja

$$\frac{\varphi(x+h) - \varphi(x)}{h} = \frac{f(x+h) - f(x)}{h} + \frac{g(x+h) - g(x)}{h}$$

bzw.
$$\frac{\psi(x+h) - \psi(x)}{h} = c\,\frac{f(x+h) - f(x)}{h},$$

und nunmehr ergeben sich unsere Behauptungen unmittelbar durch Grenzübergang.

Nach diesen Regeln ist z. B. der Differentialquotient der Funktion $\varphi(x) = f(x) + a x + b$ bei konstantem $a$ und $b$ durch die Gleichung

$$\varphi'(x) = f'(x) + a$$

gegeben.

## 5. Differenzierbarkeit und Stetigkeit der Funktionen.

Es ist nützlich, zu bemerken, daß wir von einer Funktion, die wir differenzieren können, die Stetigkeit niemals besonders zu beweisen brauchen. *Vielmehr ist die Stetigkeit einer Funktion eine Folge ihrer Differenzierbarkeit.* Wenn nämlich der Differenzenquotient $\frac{f(x+h) - f(x)}{h}$ einem bestimmten Grenzwert zustrebt, sobald die Größe $h$ gegen 0 rückt, so muß notwendig mit $h$ zugleich auch der Zähler des Bruches, d. h. die Zahl $f(x+h) - f(x)$ gegen 0 streben; und durch diese Tatsache drückt sich gerade die Stetigkeit der Funktion $f(x)$ an der betreffenden Stelle $x$ aus.

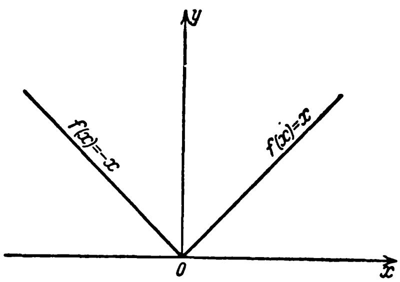


Fig. 34. $f(x) = |x|$.

Es ist aber keineswegs umgekehrt so, daß eine stetige Funktion sich auch überall differenzieren läßt. Das einfachste Gegenbeispiel gegen eine solche Annahme gibt uns die Funktion $f(x) = |x|$, d. h. $f(x) = -x$ für $x \leqq 0$, $f(x) = x$ für $x \geqq 0$, welche in Figur 34 gezeichnet ist. An der Stelle $x = 0$ ist diese Funktion zwar stetig, hat aber keinen Differentialquotienten. Der Grenzwert von $\frac{f(x+h) - f(x)}{h}$ wird nämlich gleich 1, wenn $h$ durch positive Werte gegen Null strebt, und gleich

$-1$, wenn $h$ durch negative Werte gegen Null strebt. Man sagt: unsere Funktion besitzt an der Stelle $x = 0$ verschiedenen vorderen und hinteren Differentialquotienten, indem man allgemein als *vorderen bzw. hinteren Differentialquotienten* den Grenzwert von $\frac{f(x+h) - f(x)}{h}$ definiert, wenn das eine Mal $h$ durch positive Werte, das zweite Mal $h$ durch negative Werte gegen Null rückt. Differenzierbarkeit einer Funktion verlangt dann neben der Existenz die Gleichheit von vorderem und hinterem Differentialquotienten. Die Verschiedenheit der beiden bedeutet geometrisch das Auftreten einer *Ecke* der Kurve.

Als weitere Beispiele von Stellen, wo eine stetige Funktion nicht differenzierbar ist, betrachten wir die *Unendlichkeitsstellen des Differentialquotienten*, d. h. solche Stellen, wo weder der hintere noch der vordere Differentialquotient existiert, sondern wo für $h \to 0$ der Differenzenquotient $\frac{f(x+h) - f(x)}{h}$ über alle Grenzen wächst. Beispielsweise ist die Funktion $y = f(x) = \sqrt[3]{x} = x^{\frac{1}{3}}$ für alle Werte von $x$ definiert und stetig — übrigens eine ungerade Funktion —. Ihr Differentialquotient ist für $x \neq 0$ nach Nr. 3 durch $y' = \frac{1}{3} x^{-\frac{2}{3}}$ gegeben. An der Stelle $x = 0$ wird $\frac{f(x+h) - f(x)}{h} = \frac{h^{\frac{1}{3}}}{h} = h^{-\frac{2}{3}}$, und wir erkennen hieraus sofort, daß für $h \to 0$ kein Grenzwert existiert, vielmehr der Differenzenquotient gegen $\infty$ strebt. Die Funktion besitzt dort, wie man kurz sagt, einen *unendlich großen Differentialquotienten* oder den Differentialquotienten $\infty$. Geometrisch bedeutet dies das Auftreten einer *senkrechten Tangente* an die Kurve bei glattem Verlauf (vgl. Fig. 35).

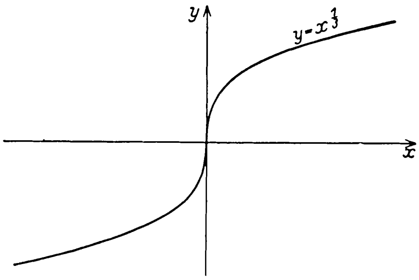


Fig. 35.

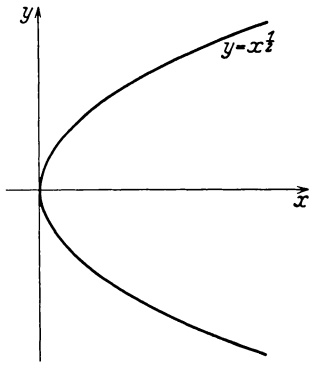


Fig. 36.

Auch die Funktion $y = f(x) = \sqrt{x}$, die ja für $x \geqq 0$ definiert und stetig ist, wird im Punkte $x = 0$ nicht mehr differenzierbar sein. Wir sind hier, da $y$ für negatives $x$ nicht definiert ist, auf die Bildung des vorderen Differentialquotienten angewiesen; wegen $\frac{f(h) - f(0)}{h} = \frac{1}{\sqrt{h}}$

wird auch dieser unendlich sein: Die Kurve wird im Nullpunkt die $y$-Achse berühren (vgl. Fig. 36).

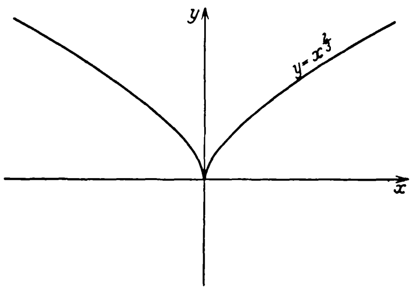


Fig. 37. Spitze.

Endlich erkennen wir in der Funktion $y = \sqrt[3]{x^2} = x^{\frac{2}{3}}$ das Beispiel eines Falles, wo der vordere Differentialquotient für die Stelle $x = 0$ positiv unendlich, der hintere dagegen negativ unendlich wird, wie aus der Beziehung

$$\frac{f(h) - f(0)}{h} = \frac{1}{\sqrt[3]{h}}$$

sofort hervorgeht. Die stetige Kurve $y = x^{\frac{2}{3}}$ selbst, eine sog. *Neilsche Parabel*, setzt im Nullpunkt des Koordinatensystemes senkrecht zur $x$-Achse mit einer *Spitze* auf (Fig. 37).

### 6. Höhere Ableitungen und ihre Bedeutung.

Der Differentialquotient $f'(x)$ einer Funktion ist selbst wieder eine Funktion von $x$, deren Bild man als die „*Differentialkurve*" der gegebenen Kurve bezeichnet. Beispielsweise wird die Differentialkurve der Parabel $y = x^2$ eine gerade Linie, dargestellt durch die Funktion $y = 2x$, und die Differentialkurve der Sinuskurve $y = \sin x$ die Cosinuskurve $y = \cos x$, sowie die Differentialkurve der Kurve $y = \cos x$ die Kurve $y = -\sin x$. (Diese letzteren Kurven gehen, wie Figur 38 zeigt, durch Parallelverschiebungen in Richtung der $x$-Achse auseinander hervor.)

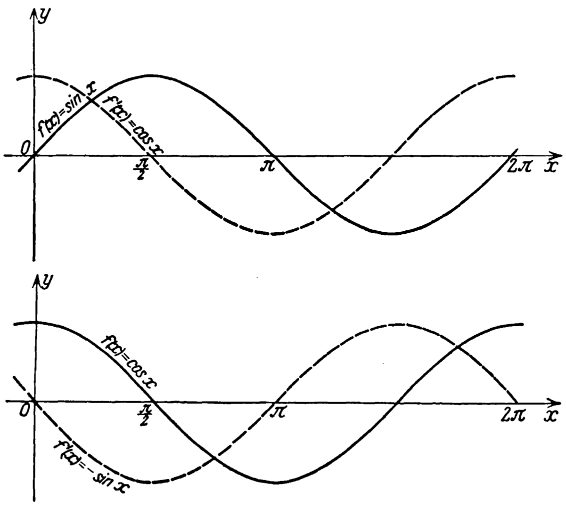


Fig. 38. Differentialkurven von $\sin x$ und $\cos x$.

Es liegt nun nahe, zu diesen Differentialkurven wieder die Differentialkurven zu bilden, d. h. den Differentialquotienten der Funktion $f'(x) = \varphi(x)$ aufzusuchen. Diesen Differentialquotienten

$$\varphi'(x) = \lim_{h \to 0} \frac{f'(x+h) - f'(x)}{h}$$

nennen wir dann, vorausgesetzt, daß er wirklich existiert, den *zweiten Differentialquotienten* oder die *zweite Ableitung* der Funktion $f(x)$ und bezeichnen ihn mit $f''(x)$.

Ebenso können wir den Differentialquotienten von $f''(x)$ zu bilden versuchen, den sog. dritten Differentialquotienten von $f(x)$, den wir dann mit $f'''(x)$ bezeichnen. Bei den meisten wichtigen Funktionen hindert uns nichts, diesen Prozeß beliebig weit fortgesetzt zu denken und auf diese Weise einen *n-ten Differentialquotienten* oder eine *n-te Ableitung* $f^{(n)}(x)$ zu definieren. Zuweilen wird es zweckmäßig sein, als den 0-ten Differentialquotienten die Funktion $f(x)$ selbst zu bezeichnen.

Wenn man die unabhängige Veränderliche $x$ als Zeit $t$ deutet und wie oben durch die Funktion $f(t)$ die Bewegung eines Punktes repräsentiert, so ergibt sich die physikalische Bedeutung des zweiten Differentialquotienten oder der zweiten Ableitung als die Geschwindigkeit, mit der sich die Geschwindigkeit $f'(t)$ ändert, oder, wie man sagt, die *Beschleunigung*. Die geometrische Bedeutung des zweiten Differentialquotienten werden wir später noch eingehend diskutieren. Schon hier aber erkennen wir unmittelbar folgendes: An einer Stelle, an welcher $f''(x)$ positiv ist, wird bei wachsendem $x$ auch $f'(x)$ wachsen; ist dagegen $f''(x)$ negativ, so wird $f'(x)$ abnehmen.

### 7. Differentialquotienten und Differenzenquotienten; Bezeichnungen von Leibniz.

Die Tatsache, daß bei unseren die Ableitung definierenden Grenzübergängen die Differenz $\Delta x$ gegen 0 strebt, pflegt man auch durch die Formulierung auszudrücken: Die Größe $\Delta x$ *wird* unendlich klein. Man will mit dieser Bezeichnung andeuten, daß man den Grenzübergang als einen Prozeß empfindet, während dessen die Größe $\Delta x$ niemals Null ist, aber sich doch der Null beliebig nähert. Im Anschluß an *Leibniz* hat man symbolisch den Grenzübergang bei der Bildung des Differentialquotienten, das Differenzieren, dadurch zum Ausdruck gebracht, daß man das Symbol $\Delta$ durch das Symbol $d$ ersetzte, so daß wir dieses Leibnizsche Symbol durch die Gleichung

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

definieren können. Seit Leibniz hat man für die Ableitung, welche ja als Grenzwert eines Differenzenquotienten definiert ist, das Wort *Differentialquotient* eingeführt. Wenn wir aber den Sinn der Differentialrechnung klar erfassen wollen, müssen wir uns davor hüten, in dem Differentialquotienten einen Quotienten zweier tatsächlich „unendlich kleiner Größen" zu erblicken. Es ist durchaus so, daß wir stets den Differenzenquotienten $\frac{\Delta y}{\Delta x}$ bilden müssen mit Differenzen $\Delta x$, die von 0 verschieden sind. Nach Bildung dieses Differenzenquotienten muß man auf Grund von Umformungen oder sonst irgendwie den Grenzübergang ausgeführt denken. Man hat aber kein Recht dazu, erst so etwas wie einen Grenzübergang von $\Delta x$ zu einer unendlich kleinen Größe, welche doch nicht 0 ist, auszuführen, etwa von $\Delta x$ bzw. $\Delta y$ zu $dx$ bzw. $dy$, und sodann den Quotienten dieser „unendlich kleinen Größen" zu bilden. Eine solche Auffassung des Differentialquotienten ist mit der Forderung mathematischer Klarheit der Begriffe unvereinbar, ja völlig sinnlos. Sie hat unzweifelhaft für manchen naiven Menschen einen gewissen Reiz, eben den Reiz des Geheimnisvollen, der mit dem Wort „unendlich" stets verbunden ist, und in den Anfängen der Differentialrechnung, insbesondere auch bei Leibniz selbst, mischt sich eine solche unklar-mystische Auffassung mit einer durchaus klaren Erfassung des Grenzüberganges. Zwar hat dieser Nebel, der über den Grundlagen der neuen Wissenschaft schwebte, Leibniz, wie seine großen Nachfolger, nicht an dem Auffinden der richtigen Wege gehindert. Aber wir sind dadurch nicht von der Pflicht entbunden, in der Begründung der Differential- und Integralrechnung jede solche nebelhafte Vorstellung zu vermeiden.

Der Grund, warum die Leibnizsche Bezeichnung nicht nur eine solche Anziehungskraft ausübt, sondern tatsächlich von äußerstem Nutzen und großer Geschmeidigkeit ist, liegt darin, daß man in vielen Rechnungen und formalen Umformungen mit den Symbolen $dy$ und $dx$ so umgehen kann, als ob sie einfach Rechengrößen wären wie irgendwelche Zahlen. Viele Rechnungen gewinnen auf diese Weise, wenn sie sich prinzipiell stets auch ohne diese Symbolik durchführen lassen, wesentlich an Übersichtlichkeit. Wir werden diese Tatsache im folgenden immer wieder bestätigt finden und die Berechtigung begründen, von ihr ausgiebig Gebrauch zu machen, wenn wir uns nur des symbolischen Charakters der Zeichen $dy$ und $dx$ bewußt bleiben.

Auch für den zweiten und höheren Differentialquotienten hat Leibniz eine Bezeichnung von suggestiver Kraft und von großer praktischer Nützlichkeit geschaffen. Er faßt nämlich den zweiten Differentialquotienten als Grenzwert des „zweiten Differenzenquotienten" auf folgende Art auf. Wir betrachten neben dem Argument $x$ die Argumente $x_1 = x + h$ und $x_2 = x + 2h$. Dann verstehen wir unter dem zweiten Differenzenquotienten den ersten Differenzenquotienten des ersten Differenzenquotienten, d. h. den Ausdruck

$$\frac{1}{h}\left(\frac{y_2 - y_1}{h} - \frac{y_1 - y}{h}\right) = \frac{1}{h^2}(y_2 - 2y_1 + y),$$

wo $y = f(x)$, $y_1 = f(x_1)$, $y_2 = f(x_2)$ gesetzt wird. Schreiben wir noch $h = \Delta x$ und $y_2 - y_1 = \Delta y_1$, $y_1 - y = \Delta y$, so werden wir sinngemäß den Ausdruck in der letzten Klammer als die Differenz der Differenz von $y$ oder die *zweite Differenz* von $y$ bezeichnen und symbolisch schreiben

$$y_2 - 2y_1 + y = \Delta\Delta y = \Delta^2 y \text{ [1]}.$$

[1] $\Delta\Delta = \Delta^2$ bedeutet also hier nicht ein Quadrat, sondern nur ein Symbol für „Differenz von Differenz" oder „zweite Differenz".

Der zweite Differenzenquotient wird in dieser symbolischen Schreibweise dann einfach $\frac{\Delta^2 y}{(\Delta x)^2}$ geschrieben werden, wobei im Nenner wirklich das Quadrat von $\Delta x$ steht, während im Zähler die Zahl 2 symbolisch die Wiederholung des Differenzenprozesses bezeichnet. Diese Schreibweise des Differenzenquotienten[1]) veranlaßte Leibniz dazu, auch für den zweiten und ebenso für die höheren Differentialquotienten die symbolische Bezeichnung

$$y'' = f''(x) = \frac{d^2 y}{d x^2}, \qquad y''' = f'''(x) = \frac{d^3 y}{d x^3} \quad \text{usw.}$$

einzuführen, eine Bezeichnung, die sich ebenfalls, wie wir später sehen werden, formal sehr bewährt.

## 8. Der Mittelwertsatz.

Zwischen dem Differentialquotienten $\frac{dy}{dx} = f'(x)$ und dem Differenzenquotienten besteht eine einfache, für viele Zwecke wichtige Beziehung, der *Mittelwertsatz*, zu dem wir folgendermaßen gelangen: Wir betrachten den Differenzenquotienten

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = \frac{\Delta f}{\Delta x}$$

einer Funktion $f(x)$, von der wir voraussetzen, daß sie überall in dem betrachteten Intervall einen Differentialquotienten besitzt, deren Kurve also überall eine bestimmte Tangente haben soll. Unser Differenzenquotient wird uns in Figur 39 durch die Richtung der Sekante repräsentiert, und zwar als der Tangens des in der Figur angegebenen Winkels $\alpha$. Denken wir uns nun diese Sekante parallel zu sich verschoben, so wird dabei mindestens einmal eine Lage eintreten, bei welcher diese Gerade die Kurve in einem Zwischenpunkte zwischen den Abszissen $x_1$ und $x_2$ berührt, nämlich in einem solchen Punkte des Kurvenbogens, der von der Sekante möglichst weit entfernt ist. Es wird

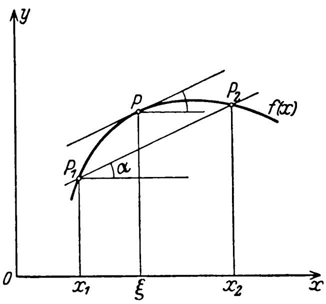


Fig. 39. Mittelwertsatz.

[1]) Ich möchte nicht einen Hinweis darauf unterlassen, daß die Darstellbarkeit des zweiten Differentialquotienten als Grenzwert des oben hingeschriebenen zweiten Differenzenquotienten eines Beweises bedarf. Denn wir haben vorher die zweite Ableitung nicht auf diese Weise definiert, sondern als Grenzwert des ersten Differenzenquotienten der ersten Ableitung. Den Nachweis, daß beide Definitionen bei stetiger zweiter Ableitung äquivalent sind, kann ich hier um so eher übergehen, als er sich später im Anhang zum sechsten Kapitel von selbst ergeben und im übrigen für uns keine Bedeutung gewinnen wird.

also einen solchen Zwischenwert $\xi$ geben, daß

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = f'(\xi)$$

wird. Diese Tatsache heißt der *Mittelwertsatz der Differentialrechnung.* Wir können ihm noch eine etwas andere Form geben, indem wir beachten, daß die Zahl $\xi$ sich in der Gestalt $\xi = x_1 + \vartheta(x_2 - x_1)$ darstellen läßt, wobei $\vartheta$ eine gewisse Zahl zwischen 0 und 1 bedeutet, die man bei Anwendungen des Mittelwertsatzes oft nicht näher bestimmen kann, aber — wie sich herausstellen wird — auch nicht näher zu kennen braucht. Der Mittelwertsatz lautet dann, genau formuliert: *Wenn $f(x)$ in dem abgeschlossenen Intervall $x_1 \leqq x \leqq x_2$ stetig und wenigstens an jeder Stelle des offenen Intervalles $x_1 < x < x_2$ differenzierbar ist, so gibt es mindestens einen Wert $\vartheta$ zwischen* 0 *und* 1*:*

$$0 < \vartheta < 1,$$

*derart, daß*

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(x_1 + \vartheta(x_2 - x_1))$$

*ist.*

Ersetzen wir $x_1$ durch $x$ und $x_2$ durch $x + h$, so können wir den Mittelwertsatz auch durch die Formel

$$\frac{f(x+h) - f(x)}{h} = f'(\xi) = f'(x + \vartheta h), \quad x < \xi < x + h$$

repräsentieren.

Daß man zwar die Stetigkeit von $f(x)$ bis in die Randpunkte des Intervalles hinein voraussetzen muß, aber die Existenz der Ableitung in den Randpunkten nicht vorauszusetzen braucht, ist eine zunächst unscheinbare Bemerkung, die jedoch für manche Anwendungen nützlich sein wird.

Wenn die Ableitung an einer Stelle im Innern nicht mehr existiert, so braucht übrigens der Mittelwertsatz nicht mehr richtig zu sein, wie bei dem obigen Beispiel $f(x) = |x|$ auf Seite 76.

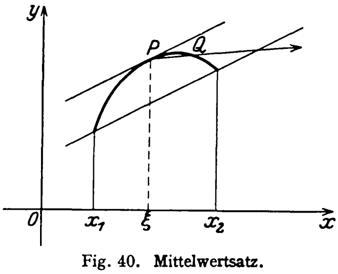


Fig. 40. Mittelwertsatz.

Unsere anschauliche Begründung können wir durch folgende Betrachtung ergänzen: Auf der Kurve gibt es sicherlich mindestens einen Punkt $P$, der eine größtmögliche Entfernung von der Sehne hat (vgl. Fig. 40). In diesem Punkte hat die Kurve nach Voraussetzung eine bestimmte Tangente, und diese Tangente muß der Sehne parallel sein; denn die Tangente ist die Grenzlage einer Sekante und entsteht, wenn wir $P$ mit einem Punkt $Q$ des Kurvenbogens verbinden und $Q$ gegen $P$ rücken lassen. Da nach Voraussetzung $P$ nicht weniger weit von der Sehne entfernt ist als $Q$, muß jedesmal die Verlängerung der Strecke $PQ$ in Richtung von $P$ nach $Q$ die Sehne schneiden oder ihr parallel laufen; und da dies der Fall sein muß, auf welcher Seite von $P$ der Punkt $Q$ auch liegt, so ist dies nur möglich, wenn die Grenzlage parallel zu der Sehne ist, welche die Endpunkte der Ordinaten in $x_1$ und $x_2$ verbindet. Für die oben mit $\xi$ bezeichnete Zahl kann also einfach die Abszisse des Kurvenpunktes $P$ genommen werden.

Den strengen Beweis des Mittelwertsatzes pflegt man in folgender Form zu führen. Man schickt den *Hilfssatz von Rolle* voraus, der einen Spezialfall des allgemeinen Mittelwertsatzes darstellt: „Wenn eine Funktion $\varphi(x)$ im abgeschlossenen Intervall $x_1 \leqq x \leqq x_2$ stetig ist, wenn ferner $\varphi(x_1) = 0$ und $\varphi(x_2) = 0$ ist, wenn endlich $\varphi(x)$ im Innern differenzierbar ist, so gibt es sicherlich mindestens eine Stelle $\xi$ im Innern des Intervalles, für welche $\varphi'(\xi) = 0$ ist." In der Tat muß es gewiß im Innern mindestens einen Punkt $\xi$ geben, für welchen die Funktion $\varphi(x)$ einen größten oder einen kleinsten Wert annimmt (vgl. erstes Kapitel, Anhang, § 2); wir setzen etwa voraus, daß $\xi$ eine solche Stelle sei, für welche $\varphi(\xi)$ nicht kleiner ist als alle andern im Intervall angenommenen Funktionswerte $\varphi(x)$. Dann ist für kleines $|h|$ sicher $\varphi(\xi) - \varphi(\xi + h) \geqq 0$. Lassen wir nun in dem Ausdruck $\frac{\varphi(\xi + h) - \varphi(\xi)}{h}$ die Größe $h$ von positiven Werten her gegen Null streben, so folgt durch Grenzübergang $\varphi'(\xi) \leqq 0$. Da aber ebenso, wenn $h$ durch negative Werte gegen Null strebt, $\frac{\varphi(\xi + h) - \varphi(\xi)}{h} \geqq 0$ ist und also der Grenzwert nicht negativ sein kann, so muß $\varphi'(\xi) = 0$ sein, womit die Behauptung bewiesen ist.

Den Rolleschen Satz wenden wir nun auf die Funktion[1])

$$\varphi(x) = f(x) - f(x_1) - \frac{x - x_1}{x_2 - x_1}(f(x_2) - f(x_1))$$

an, welche offenbar den Bedingungen $\varphi(x_1) = \varphi(x_2) = 0$ genügt und die Form $\varphi(x) = f(x) + ax + b$ mit den konstanten Koeffizienten $a = -\frac{f(x_2) - f(x_1)}{x_2 - x_1}$ und $b$ besitzt. Nach Nr. 4 ist $\varphi'(x) = f'(x) + a$ und somit wegen des Rolleschen Satzes für einen geeigneten Zwischenwert $\xi$

$$0 = \varphi'(\xi) = f'(\xi) + a,$$

was mit der Aussage des Mittelwertsatzes gleichbedeutend ist.

---

[1]) Diese stellt übrigens, wie der Leser leicht selbst überlegen kann, bis auf einen von $x$ unabhängigen Faktor den Abstand des Kurvenpunktes von der Sekante dar.

## 9. Angenäherte Darstellung beliebiger Funktionen durch lineare. — Differentiale.

Die Definitionsgleichung $\lim\limits_{h\to 0} \frac{f(x+h)-f(x)}{h} = f'(x)$ des Differentialquotienten ist gleichbedeutend damit, daß

$$f(x+h) - f(x) = f'(x)h + \varepsilon h$$

oder

$$f(x+\Delta x) = y + \Delta y = f(x) + f'(x)\Delta x + \varepsilon \Delta x$$

wird, wobei $\varepsilon$ eine mit $h = \Delta x$ zugleich gegen Null strebende Größe ist. Denken wir uns für den Augenblick die Stelle $x$ festgehalten und den Zuwachs $h = \Delta x$ veränderlich, so ist durch diese Formeln der Zuwachs der Funktion, d. h. die Größe $\Delta y$, in zwei Summanden zerlegt, nämlich einen zu $h$ proportionalen „linearen" Anteil $hf'(x)$ und einen „Fehler", welcher im Verhältnis zu $h$ um so kleiner wird, je kleiner der Zuwachs $h$ selbst bleibt. Die Funktion $f(x+h)$ wird also bei festem $x$ in ihrer Abhängigkeit von $h$ durch den linearen Anteil $f(x) + hf'(x)$ um so besser dargestellt, auf ein je kleineres Intervall um den Punkt $x$ herum wir uns beschränken. Dieser angenäherten Darstellung der Funktion $f(x+h)$ durch die angegebene lineare Funktion von $h$ entspricht geometrisch die Ersetzung der Kurve durch ihre Tangente im Punkte $x$. Wir kommen auf die praktische Anwendung dieser Überlegung zur Ausführung von Näherungsrechnungen noch später im siebenten Kapitel zurück.

Hier will ich nur beiläufig die Bemerkung anschließen, daß man an diese genäherte Darstellung des Zuwachses $\Delta y$ durch den linearen Ausdruck $hf'(x)$ eine logisch einwandfreie Definition des Begriffes „Differential" anschließen kann, wie es nach dem Vorgange von Leibniz vor allem Cauchy getan hat.

Wenn auch der Begriff eines Differentials als unendlich kleine Größe keinen Sinn hat und es deshalb unsinnig ist, den Differentialquotienten als den Quotienten zweier solcher Größen zu definieren, so kann man dennoch eine Deutung der Gleichung $f'(x) = \frac{dy}{dx}$ erstreben, bei welcher der Ausdruck $\frac{dy}{dx}$ nicht bloß symbolisch zu verstehen ist, sondern als wirklicher Quotient zweier Größen $dy$ und $dx$. Man definiert zu diesem Zwecke zunächst die Ableitung $f'(x)$ durch unseren Grenzübergang, denkt sich $x$ festgehalten und betrachtet den Zuwachs $h = \Delta x$ als Veränderliche. Diese Größe nennt man jetzt das *Differential von* $x$ und schreibt $h = dx$. Nunmehr definieren wir als *Differential der Funktion* $y$ den Ausdruck $dy = y'dx = hf'(x)$, also wiederum eine Zahl, die mit unendlich kleinen Größen nichts zu tun hat. Jetzt ist wirklich der Differentialquotient $y' = f'(x)$ der

Quotient der beiden Differentiale $dy$ und $dx$; aber in dieser Aussage steckt nichts Merkwürdiges mehr, sondern eine bloße Tautologie, eine Umschreibung der Wortdefinition. Das Differential $dy$ ist also der lineare Anteil des Zuwachses $\Delta y$ (vgl. Fig. 41).

Wir werden von diesen Differentialen zunächst keinen Gebrauch machen. Ich möchte jedoch der Vollständigkeit halber noch erwähnen, daß man nun auch zweite und höhere Differentiale bilden kann. Denkt man sich nämlich jetzt $h$ irgendwie gewählt, und zwar denselben Wert $h$ für verschiedene $x$, so wird $dy = hf'(x)$ eine Funktion von $x$ sein, von der man wieder das Differential bilden kann. Diese Größe wird man *zweites Differential von* $y$ nennen und durch das Symbol $d^2y = d^2f(x)$ bezeichnen. Sie ist der lineare Anteil des Zuwachses $hf'(x+h) - hf'(x)$, d. h. es wird $d^2y = h^2f''(x)$. Ebenso kann man natürlich beliebig weiter gehen und gelangt zu dritten, vierten Differentialen von $y$ usw., welche durch die Ausdrücke $h^3f'''(x)$, $h^4f^{\mathrm{IV}}(x), \ldots$ definiert werden können.

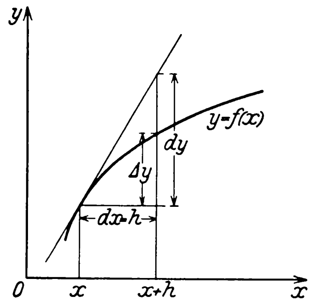


Fig. 41. Differential $dy$.

## 10. Bemerkungen über die Anwendungen unserer Begriffe in der Naturwissenschaft.

In den Anwendungen der Mathematik auf Naturerscheinungen haben wir es niemals mit scharf definierten Größen zu tun. Ob eine Länge exakt gleich einem Meter ist, das ist eine Frage, welche durch kein Experiment entschieden werden kann und welche infolgedessen keinen „physikalischen Sinn" hat. Ebensowenig hat es einen unmittelbaren physikalischen Sinn, von einem materiellen Stabe zu sagen, seine Länge sei rational oder sie sei irrational; wir werden sie immer mit jeder wünschenswerten Genauigkeit durch rationale Zahlen messen können, und was hauptsächlich interessiert, ist, ob wir bei einer solchen Messung durch rationale Zahlen mit verhältnismäßig kleinen Nennern auskommen oder nicht. Ebenso wie die Frage nach Rationalität oder Irrationalität in der strengen Bedeutung der „Präzisionsmathematik" keinen physikalischen Sinn hat, wird auch sonst in den Anwendungen die wirkliche Durchführung von Grenzübergängen gewöhnlich nur eine mathematische Idealisierung darstellen. Die Bedeutung solcher Idealisierungen für die Anwendungen beruht vor allem darin, daß durch sie alle analytischen Ausdrücke wesentlich einfacher und handlicher werden. Z. B. ist es außerordentlich viel einfacher und bequemer, mit dem Begriff der Momentangeschwindigkeit zu arbeiten, die Funktion nur eines bestimmten Zeitmomentes ist, als mit dem Begriff der Durchschnittsgeschwindigkeit zwischen zwei verschiedenen Zeitmomenten. Jede rationelle Naturbetrachtung wäre ohne derartige mathematische Idealisierungen zu hoffnungsloser Komplikation verurteilt und müßte in den ersten Anfängen stecken bleiben. Ich will mich jedoch hier nicht in eine Betrachtung über das Verhältnis der Mathematik zur Wirklichkeit einlassen; ich möchte lediglich für unsere gegenwärtige Begriffsbildung hervorheben, daß man bei den Anwendungen das Recht hat, einen Differenzenquotienten an die Stelle eines Differentialquotienten zu setzen und umgekehrt, sobald die betrachteten Differenzen nur klein genug sind, um bei der Annäherung eine genügende Genauigkeit zu gewährleisten. Der Physiker oder der Biologe oder der Techniker, oder wer sonst mit diesen Begriffen praktisch zu tun hat, wird dann das Recht haben, innerhalb seiner Genauigkeitsgrenzen den Differenzenquotienten mit dem Differentialquotienten zu identifizieren. Er wird dann für den Zuwachs $h$ der unabhängigen Veränderlichen (nach der Ausdrucksweise von Nr. 9 das Differential $h = dx$) den Zuwachs $\Delta y = f(x + h) - f(x)$ um so genauer durch die Größe $hf'(x)$ (das Differential $dy$) darstellen können, je kleiner $h$ ist. Solange er bei dieser Ersetzung innerhalb der Genauigkeitsgrenzen bleibt, die ihm bei der Aufgabe jeweils gesteckt sind, pflegt er die Zuwächse $dx = h$ und $dy = hf'(x)$ als „unendlich kleine Größen" zu bezeichnen. Diese *„physikalisch unendlich kleinen" Größen* haben einen präzisen Sinn. Es sind durchaus endliche, von Null verschiedene Größen, nur für die betreffende Betrachtung klein genug gewählt, z. B. kleiner als der Bruchteil einer Wellenlänge oder kleiner als der Abstand zweier Elektronen im Atom oder dgl., allgemein kleiner als der verlangte Grad der Genauigkeit.

## § 4. Das unbestimmte Integral, die primitive Funktion und die Fundamentalsätze der Differential- und Integralrechnung.

Wie schon erwähnt, ist gerade der Zusammenhang zwischen dem Problem der Integration und dem Problem der Differentiation der Angelpunkt der Differential- und Integralrechnung. Diesen Zusammenhang wollen wir nunmehr entwickeln.

### 1. Das Integral als Funktion der oberen Grenze.

Das bestimmte Integral einer Funktion $f(x)$ hängt in seinem Wert von der Wahl der beiden Integrationsgrenzen $a$ und $b$ ab. Es ist eine Funktion sowohl der unteren Grenze $a$ als auch der oberen Grenze $b$. Um diese Abhängigkeit näher zu studieren, stellen wir uns zunächst einmal die untere Grenze $a$ als eine bestimmte feste Zahl vor, bezeichnen die Integrationsvariable nicht mehr mit $x$, sondern mit $u$, was ja doch vollständig gleichgültig ist (vgl. S. 63), und bezeichnen die obere Grenze statt mit $b$ mit $x$, um anzudeuten, daß wir die obere Grenze variabel lassen und den Wert des Integrales als Funktion der oberen Grenze untersuchen wollen. Wir setzen demgemäß

$$\int_a^x f(u)\,du = \Phi(x).$$

Diese Funktion $\Phi(x)$ nennen wir *ein unbestimmtes Integral* der Funktion $f(x)$. Wenn wir von einem und nicht von dem unbestimmten Integral sprechen, so deuten wir damit an, daß wir statt der unteren Grenze $a$ auch irgend eine andere hätten wählen und so einen anderen Wert für das Integral erhalten können. Geometrisch wird für jeden Wert von $x$ das unbestimmte Integral durch den aus Figur 42 ersichtlichen Flächeninhalt gegeben, der zur Kurve $y = f(u)$ gehört und von den Ordinaten $u = a$ und $u = x$ begrenzt wird; das Vorzeichen des Flächeninhaltes ist dabei natürlich nach den früher (§ 1, Nr. 3) angegebenen Regeln zu bestimmen.

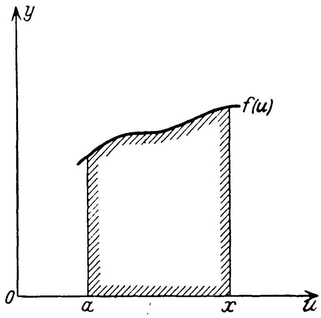


Fig. 42.

Wenn wir statt der unteren Grenze $a$ eine andere untere Grenze $\alpha$ wählen, so erhalten wir das unbestimmte Integral

$$\Psi(x) = \int_\alpha^x f(u)\,du.$$

Die Differenz $\Psi(x) - \Phi(x)$ wird offenbar durch das Integral

$$\int_\alpha^a f(u)\,du$$

gegeben, ist also, da $\alpha$ und $a$ als jeweils fest gegebene Zahlen zu betrachten sind, eine Konstante. Es ist also

$$\Psi(x) = \Phi(x) + \text{const.};$$

d. h. *unsere verschiedenen unbestimmten Integrale derselben Funktion unterscheiden sich lediglich durch eine additive Konstante.*

Man könnte auch ebenso das Integral als Funktion der unteren Grenze betrachten, indem man die Funktion

$$\varphi(x) = \int_x^b f(u)\,du$$

einführt, wobei die Zahl $b$ fest gewählt ist. Auch hier unterscheiden sich zwei Funktionen mit verschiedenen oberen Grenzen $b$ und $\beta$ lediglich um eine additive Konstante, nämlich $\int_b^\beta f(u)\,du$.

## 2. Der Differentialquotient des unbestimmten Integrales.

Wir wollen nun das unbestimmte Integral $\Phi(x)$ nach der Variablen $x$ differenzieren. Das Ergebnis dieser Differentiation wird der folgende Satz sein:

*Das unbestimmte Integral*

$$\Phi(x) = \int_a^x f(u)\,du$$

*einer stetigen Funktion $f(x)$ besitzt stets einen Differentialquotienten $\Phi'(x)$, und zwar ist*

$$\Phi'(x) = f(x),$$

*d. h. die Differentiation des unbestimmten Integrales der gegebenen stetigen Funktion liefert wieder diese Funktion.*

*Diese Tatsache bildet den Kernpunkt der ganzen Differential- und Integralrechnung.* Ihr Beweis ergibt sich außerordentlich einfach an Hand der Bedeutung des Integrales als Flächeninhalt. Wir bilden

$$\frac{\Phi(x+h) - \Phi(x)}{h}$$

und deuten den Zähler

$$\Phi(x+h) - \Phi(x) = \int_a^{x+h} f(u)\,du - \int_a^x f(u)\,du = \int_x^{x+h} f(u)\,du$$

als Flächeninhalt von der Ordinate, die zu $x$ gehört, bis zur Ordinate, die zu $x + h$ gehört.

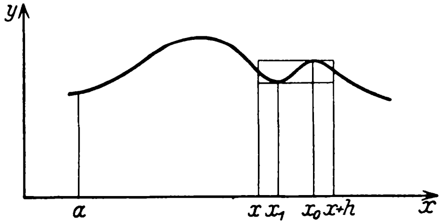



Fig. 43. Differentiation des unbestimmten Integrales.

Ist $x_0$ in dem Intervall zwischen $x$ und $x+h$ eine Stelle, für welche $f(x)$ den größten, und $x_1$ eine Stelle, für die $f(x)$ den kleinsten Wert annimmt (vgl. Fig. 43), so wird der fragliche Flächeninhalt zwischen den Werten $hf(x_0)$ und $hf(x_1)$ liegen, welche die Inhalte von Rechtecken mit diesem Intervall als Grundlinie und den Höhen $f(x_0)$ bzw. $f(x_1)$ darstellen. Analytisch ausgedrückt ist also sicherlich

$$f(x_0) \geqq \frac{\Phi(x+h) - \Phi(x)}{h} \geqq f(x_1),$$

wie man auch ohne Berufung auf die geometrische Deutung des Integrales aus der Definition des Integrales als Grenzwert folgendermaßen entnimmt[1]). Es sei

$$\int_x^{x+h} f(u)\,du = \lim_{n\to\infty} \sum_{\nu=1}^{n} f(u_\nu)\,\Delta u_\nu,$$

wobei $u_0 = x, u_1, \ldots, u_n = x + h$ Teilpunkte des Intervalles von $x$ bis $x + h$

[1]) Vergleiche hierzu übrigens auch die späteren Betrachtungen, S. 102.

bedeuten und der größte der absoluten Beträge der Differenzen $\Delta u_\nu = u_\nu - u_{\nu-1}$ mit wachsendem $n$ gegen Null strebt. Dann wird $\frac{\Delta u_\nu}{h}$ sicherlich positiv sein, gleichviel ob $h$ positiv ist oder negativ. Da nun $f(x_0) \geqq f(u_\nu) \geqq f(x_1)$ gilt und da die Summe der Größen $\Delta u_\nu$ gleich $h$ ist, so folgt

$$f(x_0) \geqq \frac{1}{h} \sum f(u_\nu)\, \Delta u_\nu \geqq f(x_1)$$

und somit durch Grenzübergang $n \to \infty$ zum Integral die oben behauptete Relation für $\frac{1}{h}\int\limits_x^{x+h} f(u)\,du = \frac{\Phi(x+h) - \Phi(x)}{h}$.

Wenn jetzt $h$ gegen 0 strebt, müssen gleichzeitig $f(x_0)$ und $f(x_1)$ wegen der Stetigkeit der Funktion gegen den Wert $f(x)$ streben. Wir erkennen also unmittelbar, daß

$$\Phi'(x) = \lim_{h \to 0} \frac{\Phi(x+h) - \Phi(x)}{h} = f(x)$$

ist, eine Gleichung, die genau den Inhalt unserer Behauptung ausmacht.

Aus der so bewiesenen Differenzierbarkeit der Funktion $\Phi(x)$ folgt mit Rücksicht auf § 3, Nr. 5 sofort: *Das Integral einer stetigen Funktion $f(x)$ ist wiederum eine stetige Funktion der oberen Grenze.*

Als Ergänzung zu unserem Ergebnis sei bemerkt, daß, wenn man das bestimmte Integral nicht als Funktion der oberen Grenze, sondern als Funktion $\varphi(x)$ der unteren Grenze auffaßt, der Differentialquotient nicht gleich $f(x)$, sondern gleich $-f(x)$ wird; in Formeln: Wird

$$\varphi(x) = \int\limits_x^b f(u)\,du$$

gesetzt, so gilt

$$\varphi'(x) = -f(x).$$

Der Beweis dieser Tatsache folgt sofort aus der Bemerkung, daß

$$\int\limits_x^b f(u)\,du = -\int\limits_b^x f(u)\,du$$

ist.

### 3. Die primitive Funktion (Stammfunktion); allgemeine Definition des unbestimmten Integrales.

Der eben bewiesene Satz zeigt uns, daß das unbestimmte Integral $\Phi(x)$ ganz von selbst die Lösung des folgenden Problemes gibt: *Zu einer gegebenen Funktion $f(x)$ ist eine Funktion $F(x)$ zu bestimmen, derart, daß*

$$F'(x) = f(x)$$

*wird.* Dieses Problem verlangt von uns, den *Prozeß der Differentiation umzukehren.* Es ist eines der typischen Umkehrungsprobleme, wie sie in der Mathematik an vielen Stellen vorkommen und wie wir sie in dieser Vorlesung schon mehrfach als fruchtbares mathematisches Erzeugungsprinzip kennengelernt haben. (Z. B. geschah die erste Erweiterung des natürlichen Zahlbegriffes unter dem Drucke der Forderung, gewisse einfache Rechenprozesse umzukehren. Die Bildung der Umkehrfunktionen führte uns zu neuartigen Funktionen und wird uns zu neuen Funktionen führen.)

Eine solche Funktion $F(x)$, für welche $F'(x) = f(x)$ ist, nennen wir eine zur Funktion $f(x)$ gehörige *primitive Funktion* oder *Stammfunktion*; man will mit dieser Bezeichnung andeuten, daß aus ihr die Funktion $f(x)$ durch Ableitung oder Differentiation entsteht.

Dieses Problem der Umkehrung der Differentiation oder der Auffindung einer primitiven Funktion zu $f(x)$ trägt zunächst einen ganz andersartigen Charakter als das Problem der Integration. Das Resultat aus Nr. 2 sagt uns nun jedoch: *Jedes unbestimmte Integral $\Phi(x)$ der Funktion $f(x)$ ist eine zu $f(x)$ primitive Funktion.*

Wir haben mit diesem Ergebnis aber das Problem der Auffindung primitiver Funktionen noch nicht vollständig gelöst. Denn wir wissen noch nicht, ob wir damit alle Lösungen der Aufgabe gefunden haben. Die Frage nach der Gesamtheit aller primitiven Funktionen wird nun durch den folgenden, gelegentlich ebenfalls als Fundamentalsatz der Differential- und Integralrechnung bezeichneten Satz beantwortet: *Es ist*

$$F_1(x) - F_2(x) = c,$$

*d. h. die Differenz zweier verschiedener primitiver Funktionen $F_1(x)$ und $F_2(x)$ zu $f(x)$ ist stets eine Konstante. Wir erhalten also zu einer beliebigen primitiven Funktion $F(x)$ alle anderen in der Gestalt*

$$F(x) + c,$$

*bei geeigneter Wahl der Konstanten $c$. Umgekehrt stellt der Ausdruck $F_1(x) = F(x) + c$ für jeden Wert der Konstanten $c$ eine primitive Funktion zu $f(x)$ dar.*

Daß für jede beliebige Wahl dieser Konstanten die Funktion $F(x) + c$ zugleich mit der Funktion $F(x)$ eine primitive Funktion zu $f(x)$ wird, ist fast selbstverständlich. Denn es ist (vgl. auch § 3, Nr. 4)

$$\frac{(F(x+h)+c) - (F(x)+c)}{h} = \frac{F(x+h) - F(x)}{h},$$

und da der Grenzwert der rechten Seite bei abnehmendem $h$ nach Voraussetzung gleich $f(x)$ wird, so gilt dies auch für den Grenzwert der linken Seite, d. h. es ist

$$\frac{d}{dx}(F(x) + c) = f(x) = F'(x).$$

Wir brauchen nun zum Beweise unserer obigen Behauptung lediglich zu zeigen, daß die Differenz zweier beliebiger primitiver Funktionen $F_1(x)$ und $F_2(x)$ auch stets eine Konstante ist. Zu dem Zwecke betrachten wir die Differenz

$$F_1(x) - F_2(x) = G(x)$$

und bilden

$$G'(x) = \lim_{h \to 0} \left\{ \frac{F_1(x+h) - F_1(x)}{h} - \frac{F_2(x+h) - F_2(x)}{h} \right\}.$$

Beide Ausdrücke auf der rechten Seite haben bei abnehmendem $h$ nach Voraussetzung den Grenzwert $f(x)$; d. h. für jeden betrachteten Wert von $x$ ist $G'(x) = 0$. Eine Funktion aber, deren Ableitung durchweg 0 ist, muß durch eine Kurve repräsentiert werden, deren Tangente überall parallel zur $x$-Achse verläuft, d. h. sie muß selbst eine Konstante sein; es ist also $G(x) = c$, wie wir behaupteten. Ohne Berufung auf die Anschauung leitet man diese letzte Tatsache mit Hilfe des Mittelwertsatzes folgendermaßen ab: Es wird

$$G(x_2) - G(x_1) = (x_2 - x_1)\, G'(\xi); \quad x_1 < \xi < x_2.$$

Da die Ableitung $G'(x)$ nach Voraussetzung für jeden betrachteten Argumentwert, also z. B. auch für den Wert $\xi$ verschwindet, so folgt sofort $G(x_2) = G(x_1)$, und da $x_1$ und $x_2$ beliebige Werte von $x$ in dem betrachteten Intervall sind, so muß $G(x)$ dort konstant sein.

Man kann also jetzt, wenn man das Resultat von Nr. 2 heranzieht, sagen: *Jede primitive Funktion $F(x)$ zu $f(x)$ läßt sich in der Form*

$$F(x) = c + \Phi(x) = c + \int_a^x f(u)\, du$$

*darstellen, wo $c$ und $a$ Konstante sind, und umgekehrt stellt unser Ausdruck bei beliebig fest gewähltem $a$ und beliebigem Wert von $c$ eine primitive Funktion dar.*

Es liegt nahe, zu vermuten, daß man die Konstante $c$ allgemein weglassen kann, da ja durch Änderung der unteren Integrationsgrenze $a$ sich das unbestimmte Integral um eine additive Konstante verändert. Man würde jedoch bei Weglassung von $c$ in vielen Fällen tatsächlich nicht alle primitiven Funktionen erhalten, wie schon das Beispiel der Funktion $f(x) = 0$ zeigt. Hier ist das unbestimmte Integral $\Phi(x)$ aus Nr. 1 immer Null, unabhängig von der unteren Grenze $a$; primitive Funktion zu $f(x) = 0$ ist aber jede beliebige Konstante. Ein zweites Beispiel liefert die Funktion $f(x) = \sqrt{x}$; sie ist nur für nicht negative Werte von $x$ definiert, und wir wollen uns auf den Funktionszweig $f(x) \geqq 0$ beschränken. Es wird dann

$$\Phi(x) = \frac{2}{3} x^{\frac{3}{2}} - \frac{2}{3} a^{\frac{3}{2}},$$

und wir sehen, daß, wie wir auch immer die untere Grenze $a$ wählen, stets das unbestimmte Integral $\Phi(x)$ aus $\frac{2}{3} x^{\frac{3}{2}}$ durch Addition einer

Konstanten $\leqq 0$, nämlich der Konstanten $-\frac{2}{3} a^{\frac{3}{2}}$ entsteht, während z. B. auch $\frac{2}{3} x^{\frac{3}{2}} + 1$ eine primitive Funktion zu $\sqrt{x}$ ist. Daher können wir in dem allgemeinen Ausdruck der primitiven Funktion die additive Konstante nicht entbehren. Der gefundene Zusammenhang legt es nahe, nunmehr den Begriff des unbestimmten Integrales ein wenig zu erweitern. Wir wollen nämlich von nun an jeden Ausdruck der Form $c + \Phi(x) = c + \int_a^x f(u)\, du$ als ein unbestimmtes Integral von $f(x)$ bezeichnen. Mit anderen Worten, *wir wollen keinen Unterschied zwischen den primitiven Funktionen und den unbestimmten Integralen mehr machen.* Trotzdem aber ist es für das Verständnis der Zusammenhänge unbedingt erforderlich, sich darüber klar zu sein, daß zunächst Integration und Umkehrung der Differentiation zwei vollständig voneinander verschiedene Dinge sind und daß erst die Erkenntnis ihres Zusammenhanges uns das Recht gibt, für die primitive Funktion auch das Wort „unbestimmtes Integral" zu verwenden.

Es ist üblich, das unbestimmte Integral durch eine an und für sich vielleicht nicht ganz klare Bezeichnungsweise symbolisch darzustellen. Man schreibt nämlich

$$F(x) = c + \int_a^x f(u)\, du = \int f(x)\, dx,$$

d. h. man läßt die obere Grenze $x$ fort, ebenso die untere Grenze $a$ und die additive Konstante $c$ und schreibt für die Integrationsvariable den Buchstaben $x$, den man konsequenterweise eigentlich vermeiden sollte, um eine Verwechselung mit der oberen Grenze $x$, der unabhängigen Veränderlichen in $F(x)$, auszuschließen. Bei der Bezeichnung $\int f(x)\, dx$ für das unbestimmte Integral muß man sich stets der damit verbundenen Unbestimmtheit bewußt bleiben, nämlich der Tatsache, daß dieses Symbol immer nur e i n unbestimmtes Integral bedeutet.

### 4. Die Verwendung der primitiven Funktion zur Ausführung bestimmter Integrale.

Nehmen wir an, daß wir irgend eine primitive Funktion $F(x) = \int f(x)\, dx$ zu $f(x)$ kennen und daß wir das bestimmte Integral $\int_a^b f(u)\, du$ berechnen wollen. Dann wissen wir, daß das unbestimmte Integral

$$\Phi(x) = \int_a^x f(u)\, du,$$

weil es auch eine primitive Funktion zu $f(x)$ ist, sich von $F(x)$ nur um eine additive Konstante unterscheiden kann. Es ist also

$$\Phi(x) = F(x) + c,$$

und die additive Konstante $c$ bestimmt sich sofort durch Berücksichtigung der Tatsache, daß das unbestimmte Integral $\Phi(x) = \int\limits_a^x f(u)\,du$ für $x = a$ den Wert 0 haben muß. Es ist also $0 = \Phi(a) = F(a) + c$ und daher $c = -F(a)$ und $\Phi(x) = F(x) - F(a)$, speziell also für $x = b$

$$\int\limits_a^b f(u)\,du = F(b) - F(a),$$

und somit ergibt sich die wichtige Regel: *Man erhält das bestimmte Integral der Funktion $f(x)$ zwischen den Grenzen $a$ und $b$, indem man mit irgend einer primitiven Funktion $F(x)$ die Differenz $F(b) - F(a)$ bildet.*

Wir können die eben gefundene Regel unter Benutzung der Beziehung $F'(x) = f(x)$ auch folgendermaßen schreiben:

$$F(b) - F(a) = \int\limits_a^b F'(x)\,dx = \int\limits_a^b \frac{dF(x)}{dx}\,dx.$$

Diese Formel läßt sich nun sehr einfach direkt verstehen und beweisen.

Ersetzen wir rechts den Differentialquotienten $\frac{dF}{dx}$ durch den Differenzenquotienten $\frac{\Delta F}{\Delta x}$ und das Symbol $dx$ durch $\Delta x$, das Integral durch die Summe $\sum \frac{\Delta F}{\Delta x_\nu} \Delta x_\nu$, so erhalten wir gerade die Summe der Differenzen $\Delta F$, und es ist klar, daß diese Summe unabhängig von der Intervalleinteilung den Wert $F(b) - F(a)$ besitzt. Andererseits aber ist sie gleich dem Integral. Nach dem Mittelwertsatz ist nämlich $\frac{\Delta F}{\Delta x_\nu} = F'(\xi_\nu)$, wo $\xi_\nu$ einen Zwischenwert zwischen $x_\nu$ und $x_{\nu+1}$ bedeutet. Unsere Summe ist also gleich $\sum \Delta x_\nu F'(\xi_\nu)$; sie geht also zufolge der Integraldefinition bei Verfeinerung der Einteilung tatsächlich in das Integral $\int\limits_a^b F'(x)\,dx$ über, womit unsere Formel bestätigt ist.

Bei der Handhabung unserer Regel bedient man sich oft des Zeichens $|$, um die Differenz $F(b) - F(a)$ auszudrücken. Man schreibt nämlich

$$\int\limits_a^b f(x)\,dx = F(b) - F(a) = F(x)\Big|_a^b$$

und deutet also mit dem Strich an, daß in dem davorstehenden Ausdruck für $x$ einmal der Wert $b$, dann der Wert $a$ einzusetzen ist und schließlich die Differenz der beiden so entstehenden Zahlen zu bilden ist.

## 5. Einige Beispiele.

Wir sind ohne weiteres in der Lage, den dargelegten Zusammenhang zwischen bestimmtem Integral, unbestimmtem Integral und Differentialquotienten an einer Reihe einfacher Beispiele zu verfolgen. Vermöge des Satzes von Nr. 2 können wir aus jeder der in § 2 direkt bewiesenen Integralformeln eine Differentiationsformel ableiten.

In § 2 waren wir für beliebige von $-1$ verschiedene rationale $\alpha$ und positive $a$, $b$ zu der Integralformel

$$\int_a^b x^\alpha dx = \frac{1}{\alpha+1}(b^{\alpha+1} - a^{\alpha+1})$$

gelangt, die wir, indem wir die Integrationsvariable mit $u$ und die obere Grenze mit $x$ bezeichnen, in die Form

$$\int_a^x u^\alpha du = \frac{1}{\alpha+1}(x^{\alpha+1} - a^{\alpha+1})$$

setzen können. Daraus folgt nach unserem Fundamentalsatz, daß die rechte Seite eine primitive Funktion zum Integranden ist, d. h. daß die Differentiationsformel

$$\frac{d}{dx} x^{\alpha+1} = (\alpha+1)\, x^\alpha$$

gilt, und zwar für alle rationalen Werte von $\alpha$ mit Ausnahme des Wertes $\alpha = -1$ und für alle $x > 0$. Durch Einsetzen zeigt sich übrigens, daß *diese* Formel auch für $\alpha = -1$ gilt, wenn wieder $x > 0$ ist. Das gewonnene Resultat stimmt genau mit dem überein, das wir in § 3 durch direkte Durchführung der Differentiation erhalten haben. Wir hätten uns also, gestützt auf den Fundamentalsatz, die besondere Mühe dieser Differentiation sparen können, nachdem wir einmal die Integration ausgeführt haben.

Weiter folgt aus der in § 2 bewiesenen Integralformel

$$\int_a^x \cos u\, du = \sin x - \sin a$$

die Differentiationsformel $\frac{d}{dx}\sin x = \cos x$ in Übereinstimmung mit dem in § 3 gewonnenen Ergebnis.

Wir können aber auch umgekehrt jede direkt bewiesene Differentiationsformel $F'(x) = f(x)$ als Verknüpfung zwischen primitiver Funktion $F(x)$ und abgeleiteter Funktion $f(x)$ ansehen, also als Formel der unbestimmten Integration auffassen und dann gemäß Nr. 4 hieraus das bestimmte Integral von $f(x)$ erhalten. — Gerade von dieser letzten Methode macht man, wie wir im Kapitel IV sehen werden, außerordentlich häufig Gebrauch. — Insbesondere können wir von den Resultaten des § 3 ausgehen und aus ihnen vermöge des Fundamentalsatzes die Integralformeln des § 2 beweisen. Es ist z. B. nach § 3 $\frac{d}{dx} x^{\alpha+1} = (\alpha + 1)\, x^{\alpha}$. Also ist $\frac{x^{\alpha+1}}{\alpha+1}$ bei $\alpha \neq -1$ eine primitive Funktion oder ein unbestimmtes Integral zu $x^{\alpha}$, und somit ergibt sich jetzt nach Nr. 4 die obige Integralformel.

## § 5. Einfachste Methoden zur graphischen Integration.

Ein unbestimmtes Integral oder eine primitive Funktion zu $f(x)$ ist eine Funktion $y = F(x)$, die man zweckmäßigerweise nicht nur durch Flächeninhalte veranschaulichen wird, sondern für die man auch ebensogut wie für jede Funktion eine graphische Darstellung durch eine Kurve geben kann. Aus der Definition unserer Begriffe ergibt sich unmittelbar die Möglichkeit, diese Kurve in einfacher Weise genähert zu konstruieren und sich so ein anschauliches Bild von der Integralfunktion zu machen.

Man muß zunächst bedenken, daß diese letztere Kurve nicht eindeutig, sondern wegen der willkürlichen additiven Konstanten nur bis auf eine willkürliche Parallelverschiebung in Richtung der $y$-Achse bestimmt ist. Man kann also von der Integralkurve noch verlangen, daß sie durch einen bestimmten, willkürlich vorgeschriebenen Punkt geht, z. B., wenn $x = 1$ dem Definitionsintervall von $f(x)$ angehört, durch den Punkt mit den Koordinaten $x = 1$, $y = 0$. Dann ist die Integralkurve durch die Forderung bestimmt, daß ihre Richtung für jeden Wert von $x$ durch den zugehörigen Wert von $f(x)$ gegeben wird. Um eine angenäherte Konstruktion einer diesen Bedingungen genügenden Kurve zu geben, versucht man, statt der krummen, durch $y = F(x)$ dargestellten Linie einen aus geraden Stücken bestehenden Polygonzug zu konstruieren, dessen Eckpunkte über vorgegebenen Teilpunkten der $x$-Achse liegen und dessen Seiten jeweils ein Stück weit die Richtung der gesuchten Kurve annähern. Hierzu teile man das betrachtete Intervall der $x$-Achse durch die Punkte $x = 1, x_1, x_2, \ldots$ in eine gewisse Anzahl von Teilen ein, welche übrigens nicht notwendig gleich sein müssen, und errichte in jedem Teilpunkt die Parallele zur $y$-Achse. Dann lege man (vgl. Fig. 44) durch den Punkt $x = 1$, $y = 0$ die gerade Linie, deren Richtungstangens gerade gleich $f(1)$ ist; durch den Schnittpunkt dieser Geraden mit der Linie $x = x_1$ lege man die Gerade mit dem Richtungstangens $f(x_1)$, schneide sie wieder mit der Linie $x = x_2$ und setze dies Verfahren fort. Um die Geraden mit den vorgegebenen Richtungen prak-

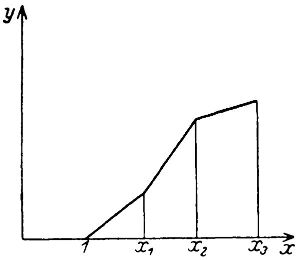


Fig. 44. Graphische Integration.

tisch zu konstruieren, errichte man in jedem Teilpunkt die Ordinate der Kurve $y = f(x)$. Diese Ordinate projiziere man auf irgend eine Parallele zur $y$-Achse, z. B. auf diese selbst. Dann erhält man die

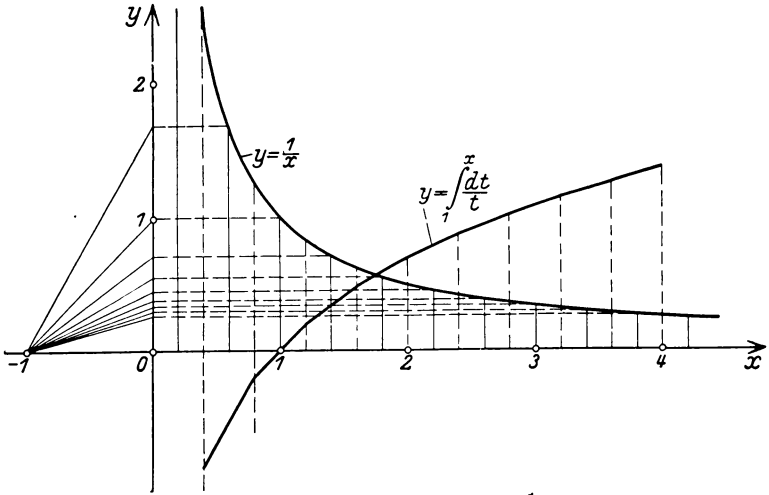



Fig. 45. Graphische Integration von $\frac{1}{x}$.

Richtung der Integralkurve, indem man — bei der letzten Annahme — den Punkt mit den Koordinaten $x = 0$ und $y = f(x)$ mit dem Punkt $x = -1$, $y = 0$ verbindet (vgl. die Figuren 45 und 46). Indem man diese Richtungen parallel zu sich selbst überträgt, konstruiert man einen Polygonzug, dessen Eckpunkte über den gegebenen Teilpunkten der $x$-Achse liegen und dessen Richtungen jeweils in den Anfangspunkten mit den dortigen Richtungen der Integralkurven übereinstimmen. Macht man die Intervalleinteilung hinreichend fein, so werden die Polygone eine beliebig genaue Annäherung an die Integralkurve darstellen. Man kann häufig die Genauigkeit der Konstruktion noch etwas verbessern, indem man für jede Polygonseite diejenige Richtung wählt, welche nicht zum Anfangspunkt des betreffenden Intervalles, sondern zu dessen Mittelpunkt gehört (vgl. die Figuren)[1].

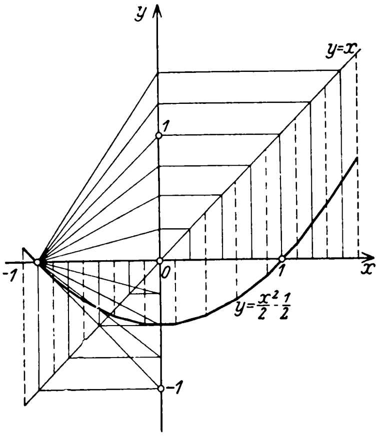



Fig. 46. Graphische Integration von $x$.

[1]) Beiläufig möchte ich erwähnen, daß die graphische Integration, d. h. die zeichnerische Ermittlung einer primitiven Funktion $F(x)$ zu einer durch eine Kurve gegebenen Funktion $f(x)$, auch durch mechanische Apparate, die sog.

In Figur 46 ist die geschilderte Konstruktion für die Funktion $f(x) = x$ durchgeführt. Wir erhalten so durch graphische Integration als Integralkurve die Parabel $y = \frac{x^2}{2} - \frac{1}{2}$. Weiter ist in Figur 45 für die Funktion $f(x) = \frac{1}{x}$ die Integralfunktion gezeichnet, eine Funktion, die wir später (vgl. drittes Kapitel, § 6) noch eingehend untersuchen werden — sie wird sich nämlich als die Logarithmusfunktion erweisen. Endlich rate ich dem Leser, selbst weitere Beispiele durchzuführen, etwa die graphische Integration der Funktionen $\sin x$ und $\cos x$.

## § 6. Weitere Bemerkungen über den Zusammenhang zwischen dem Integral und dem Differentialquotienten.

Bevor wir zu einer systematischen Verfolgung der im § 4 gefundenen Zusammenhänge übergehen, will ich diese noch von einem anderen Gesichtspunkte aus beleuchten, der in engster Beziehung mit der anschaulichen Vorstellung von Massendichte und anderen physikalischen Begriffen steht.

### 1. Die Massenverteilung und Dichte; Gesamtquantität und spezifische Quantität.

Denken wir uns auf einer geraden Linie, der $x$-Achse, irgendwelche Massen verteilt, und zwar stetig, aber nicht notwendig gleichförmig. Stellen wir uns etwa eine vertikale Luftsäule vor, welche über einem Flächenstück von der Größe 1 steht — als $x$-Achse nehmen wir dabei die senkrecht nach oben weisende Gerade, als Nullpunkt $x = 0$ den Punkt am Erdboden. Die Gesamtmasse zwischen zwei Abszissen $x = x_1$ und $x = x_2$ wird durch eine sog. Summenfunktion $F(x)$ folgendermaßen charakterisiert. Man geht von dem Anfangspunkt der Massenverteilung längs der Geraden, dem Punkte $x = 0$, aus und versteht unter $F(x)$ die Gesamtmasse zwischen der Abszisse 0 und der Abszisse $x$. Der Massenzuwachs von der Abszisse $x_1$ bis zur Abszisse $x_2$ wird dann einfach durch

$$F(x_2) - F(x_1)$$

gegeben, wobei diesem Zuwachs ein Vorzeichen zugeschrieben wird, welches sich bei Vertauschung von $x_1$ und $x_2$ ändert.

---

*Integraphen*, ausgeführt werden kann, bei denen ein Stift längs der gegebenen Kurve $y = f(x)$ geführt wird und ein Zeichenstift automatisch eine der Kurven $y = F(x)$ beschreibt, für welche $F'(x) = f(x)$ ist. Die unbestimmt bleibende Integrationskonstante drückt sich in einer gewissen Willkürlichkeit bei der Wahl der Anfangsstellung des Apparates aus. Näheres vgl. z. B. bei *A. Galle*, Mathematische Instrumente, Leipzig 1912, IX. Abschnitt.

Der Durchschnittswert der Masse pro Längeneinheit in dem Intervall $x_1$ bis $x_2$ ist

$$\frac{F(x_2) - F(x_1)}{x_2 - x_1}.$$

Nehmen wir an, daß die Funktion $F(x)$ differenzierbar ist, so strebt für $x_2 \to x_1$ dieser Wert gegen die Ableitung $F'(x_1)$. Und diese Größe ist gerade das, was man als *spezifische Masse* oder *Dichte* unserer Verteilung an der Stelle $x_1$ zu bezeichnen pflegt — sie ist natürlich im allgemeinen mit dem Ort veränderlich —. Zwischen der Dichte $f(x)$ und der Summenfunktion $F(x)$ besteht also einfach die Beziehung

$$F(x) = \int_0^x f(u)\,du; \quad f(x) = F'(x).$$

*Die Summenfunktion ist eine primitive Funktion der Dichte,* oder gleichbedeutend: *Die Masse ist das Integral der Dichte*; und umgekehrt: *Die Dichte ist der Differentialquotient der Summenfunktion.*

Genau das entsprechende Verhältnis finden wir auch sonst vielfach in der Physik vor. Bezeichnen wir z. B. die gesamte Wärmemenge, die man braucht, um eine gegebene Einheit eines Stoffes von der Temperatur $t_0$ auf die Temperatur $t$ zu bringen, mit $Q(t)$, so wird für eine Temperaturerhöhung von der Temperatur $t_1$ auf die Temperatur $t_2$ die Wärmemenge

$$Q(t_2) - Q(t_1)$$

gebraucht. Zwischen $t_1$ und $t_2$ ist die durchschnittliche für die Erhöhung um die Temperatureinheit aufgewandte Wärmemenge

$$\frac{Q(t_2) - Q(t_1)}{t_2 - t_1};$$

wenn wir wiederum die Differenzierbarkeit der Funktion $Q(t)$ voraussetzen, erhalten wir in der Grenze eine Funktion

$$q(t) = \lim_{t_1 \to t} \frac{Q(t) - Q(t_1)}{t - t_1},$$

die wir als *spezifische Wärme* des Stoffes bezeichnen. Diese spezifische Wärme ist allgemein als eine Funktion der Temperatur $t$ anzusehen.

Auch hier besteht zwischen spezifischer Wärme und gesamter Wärmemenge die für das Verhältnis des Integrals zum Differentialquotienten charakteristische Relation

$$\int_a^b q(t)\,dt = Q(b) - Q(a).$$

Ganz ähnliche Verhältnisse treffen wir überall dort, wo man einer gesamten Quantität eine spezifische Quantität gegenüberstellen kann, wie der elektrischen Ladung die Dichte der elektrischen Ladung, einer Gesamtkraft auf eine Fläche die Kraftdichte oder den Druck in jedem Punkt usw.

In der Natur liegt die Sache gewöhnlich so, daß uns von vornherein nicht etwa die spezifischen Größen, d. h. die Dichte, sondern vielmehr die Gesamtquantität gegeben ist; daß also die Integrale das Primäre sind (was ja auch durch die Bezeichnung „primitive Funktion" zum Ausdruck kommt) und daß erst ein Grenzübergang, nämlich die Differentiation uns zu den spezifischen Größen führt.

Beiläufig bemerke ich übrigens, daß bei unserer Auffassung, wenn wir die Massen als von Natur positiv ansehen, die Summenfunktion $F(x)$ eine mit $x$ monoton zunehmende Funktion sein muß und daß dementsprechend die spezifische Größe, die Dichte $f(x)$, positiv ist. An und für sich aber hindert uns nichts, auch die Vorstellung negativer Massen heranzuziehen (wie z. B. negative Elektrizität); dann sind unsere Summenfunktionen $F(x)$ durch keine Monotoniebedingung mehr eingeschränkt.

## 2. Gesichtspunkte der Anwendungen.

Das Verhältnis der primitiven Summenfunktion zu der Verteilungsdichte wird vielleicht noch deutlicher, wenn man sich wieder vergegenwärtigt, daß vom Standpunkte der physikalischen Tatsachen aus sowohl der Grenzübergang bei der Integration als auch der bei der Differentiation eine Idealisierung darstellt und daß diesen Grenzübergängen exakt in der Natur nichts entspricht, daß man vielmehr an Stelle des Integrals in der physikalischen Wirklichkeit genau genommen nur eine Summe sehr vieler kleiner Summanden und an Stelle eines Differentialquotienten einen Differenzenquotienten aus sehr kleinen Größen bilden muß. Die Größen $\Delta x$ bleiben von 0 verschieden; der Grenzübergang $\Delta x \to 0$ hat lediglich die mathematische Bedeutung einer weitgehenden Vereinfachung, bei welcher dennoch die Genauigkeit der mathematischen Darstellung der Wirklichkeit nicht merklich beeinträchtigt wird.

Beispielsweise werden wir uns entsprechend der atomistischen Struktur der Materie die Masse in einer Luftsäule nicht wirklich stetig als Funktion von $x$ verteilt denken können; wir wollen uns vielmehr (was allerdings wiederum eine idealisierende Vereinfachung ist) vorstellen, daß die Masse in Form von sehr vielen, sehr dicht beieinanderliegenden, aber punktförmigen Molekülen über die $x$-Achse verteilt sei. Dann wird die Summenfunktion $F(x)$ keine stetige Funktion sein; sie wird vielmehr in dem Intervall zwischen zwei Molekülen einen konstanten Wert haben und plötzlich springen, wenn wir mit der Variablen den Ort eines neuen Moleküles passieren; die Größe dieses Sprunges ist gerade gleich der Masse des Moleküles, die durchschnittliche Entfernung zweier Moleküle wird dabei nach den Ergebnissen der Atomtheorie die Größenordnung von $10^{-8}$ cm haben. Wenn wir unsere Größen $\Delta x$ zwar sehr klein, aber doch noch groß genug gegenüber der Länge $10^{-8}$ cm machen, etwa $10^{-4}$ cm, dann liegen in jedem Intervall zwischen $x$ und $x + \Delta x$ noch so viele Moleküle, nämlich der Größenordnung nach $10^4$, daß man auch von ihnen noch den Eindruck einer stetigen Verteilung behält und sich das Recht nimmt, diese Verteilung durch eine stetige und differenzierbare Funktion zu idealisieren. Als Verteilungsdichte wird man dann einfach den Differenzenquotienten $\frac{\Delta F(x)}{\Delta x} = \frac{F(x+\Delta x) - F(x)}{\Delta x}$ betrachten; selbstverständlich ist es dabei eine wesentliche physikalische Annahme, daß die Bildung dieses Differenzenquotienten nicht zu merklich verschiedenen Ergebnissen führt, wenn wir die Größe $\Delta x$ innerhalb gewisser Grenzen verändern, etwa innerhalb der Grenzen $10^{-4}$ cm und $10^{-5}$ cm.

Man sieht unmittelbar, daß diese Dichte mit der Gesamtverteilung durch die jetzt selbstverständliche Gleichung

$$F(x) = \sum \frac{\Delta F(x)}{\Delta x} \Delta x$$

zusammenhängt.

Es ist vielleicht angebracht, noch ein weiteres Beispiel für die Begriffe: Summenfunktion und Verteilungsdichte zu besprechen. In der Statistik, z. B. in der kinetischen Theorie der Materie oder in der statistischen Biologie, treten diese Begriffe häufig in einer Form auf, bei der der Charakter der mathematischen Idealisierung besonders deutlich wird. Betrachten wir etwa die Moleküle eines in einem Gefäß eingeschlossenen Gases. Ihre Anzahl sei $N$. Die Anzahl derjenigen, welche eine Geschwindigkeit kleiner als $x$ haben, sei $N\Phi(x)$. Dann bedeutet $\Phi(x)$ das Verhältnis der Anzahl derjenigen Moleküle, welche sich mit einer zwischen 0 und $x$ gelegenen Geschwindigkeit bewegen, zur Gesamtzahl (die ganze Betrachtung bezieht sich auf einen bestimmten Zeitmoment). Diese Summenfunktion ist natürlich nicht stetig, sondern sie wird wieder stückweise konstant sein und jeweils um die Zahl $\frac{1}{N}$ sprunghaft anwachsen, wenn bei wachsendem $x$ gerade ein $x$-Wert passiert wird, dem ein Molekül mit der Geschwindigkeit $x$ entspricht.

Die mathematische Idealisierung, die wir vornehmen, ist nun die, daß wir die Zahl $N$ in der Idee über alle Grenzen wachsen lassen. Wir nehmen an, daß bei diesem Grenzübergang $N \to \infty$ die Summenfunktion $\Phi(x)$ gegen eine bestimmte stetige Grenzfunktion $F(x)$ strebt. Daß dies wirklich der Fall ist (d. h. daß man mit hinreichender Genauigkeit $\Phi(x)$ durch eine solche stetige Funktion $F(x)$ ersetzen darf), stellt selbstverständlich eine wesentliche physikalische Annahme dar; und eine weitere solche Annahme ist es, wenn wir voraussetzen, daß diese Summenfunktion $F(x)$ einen Differentialquotienten $F'(x)=f(x)$ besitzt, den wir dann als Verteilungsdichte bezeichnen werden. Die Summenfunktion hängt mit der Verteilungsdichte durch die Gleichungen

$$F(x) = \int_0^x f(u)\,du; \quad F(b) - F(a) = \int_a^b f(x)\,dx$$

zusammen.

Man nennt diese Verteilungsdichte gelegentlich auch die *spezifische Wahrscheinlichkeit* dafür, daß ein Molekül die Geschwindigkeit $x$ besitzt. Die eben durchgeführte Idealisierung, welche in der auf *Maxwell* zurückgehenden kinetischen Gastheorie eine große Rolle spielt, stellt sich mathematisch in genau derselben Form bei zahlreichen Fragen der mathematischen Statistik ein.

## § 7. Integralabschätzungen und Mittelwertsatz der Integralrechnung.

Wir schließen dieses Kapitel mit einigen Betrachtungen über eine Frage von allgemeiner Bedeutung, deren ganze Wichtigkeit allerdings erst etwas später zutage treten wird. Es handelt sich um die Abschätzung von Integralen.

### 1. Der Mittelwertsatz der Integralrechnung.

Die erste und einfachste dieser Abschätzungsregeln lautet: Wenn die stetige Funktion $f(x)$ in einem Intervall $a \leqq x \leqq b$ überall nicht negativ ist, d. h. überall entweder positiv oder Null ist, so ist auch das bestimmte Integral

$$\int_a^b f(x)\,dx$$

nicht negativ. Ebenso ist dieses Integral nicht positiv, wenn die Funktion in dem Intervall nirgends positiv ist. Der Beweis dieser Tatsache ergibt sich ohne weiteres aus der Definition des Integrales.

Hieraus entspringt der folgende Satz: Wenn in dem Intervalle $a \leqq x \leqq b$ überall

$$f(x) \geqq g(x)$$

ist, so ist auch

$$\int_a^b f(x)\,dx \geqq \int_a^b g(x)\,dx.$$

Denn das Integral der Differenz $f(x) - g(x)$ ist nach der ersten Bemerkung nicht negativ, und nach unserer Summenregel (S. 62f.) ist

$$\int_a^b (f(x) - g(x))\,dx = \int_a^b f(x)\,dx - \int_a^b g(x)\,dx.$$

Es sei nun $M$ der größte, $m$ der kleinste Wert der Funktion $f(x)$ in dem Intervall $a$ bis $b$. Dann ist die Funktion $M - f(x)$ in dem Intervall nicht negativ, und dasselbe gilt für die Funktion $f(x) - m$. Es ergibt sich also nach den obigen Bemerkungen sofort die doppelte Ungleichung:

$$\int_a^b m\,dx \leqq \int_a^b f(x)\,dx \leqq \int_a^b M\,dx,$$

und da $\int_a^b m\,dx = m\int_a^b dx = m\,(b-a)$ und ebenso $\int_a^b M\,dx = M\,(b-a)$ ist, folgt

$$m\,(b-a) \leqq \int_a^b f(x)\,dx \leqq M\,(b-a).$$

Das betrachtete Integral läßt sich daher darstellen als das Produkt von $b - a$ mit einem Werte $\mu$ zwischen $m$ und $M$:

$$\int_a^b f(x)\,dx = \mu\,(b-a), \qquad m \leqq \mu \leqq M.$$

Die genaue Größe dieses Zwischenwertes $\mu$ brauchen wir nicht allgemein anzugeben. Wir können aber aussagen, daß er an einer gewissen Stelle $\xi$ des Intervalles $a < \xi < b$ von der Funktion angenommen wird, da eine stetige Funktion in einem Intervalle alle Werte zwischen ihrem größten und ihrem kleinsten Wert annimmt (vgl. S. 51f.). Wie beim Mittelwertsatz der Differentialrechnung ist bei vielen Anwendungen die genaue Angabe des Wertes $\xi$ unwichtig. Wir dürfen also $\mu = f(\xi)$ setzen, wo $\xi$ ein solcher Zwischenwert ist, und können dann schreiben

$$\int_a^b f(x)\,dx = (b-a)\,f(\xi), \qquad a < \xi < b.$$

In dieser letzteren Formulierung bezeichnen wir unsere Abschätzungsformel als *Mittelwertsatz der Integralrechnung*. Wir können diesen Mittelwertsatz noch ein wenig verallgemeinern, indem wir an Stelle des Integranden $f(x)$ einen *Integranden der Form* $f(x)\,p(x)$ betrachten, *wobei* $p(x)$ *eine beliebige positive Funktion bedeutet, die wir, ebenso wie* $f(x)$, *als stetig annehmen*. Da $m\,p(x) \leqq f(x)\,p(x) \leqq M\,p(x)$ ist, so erhalten wir sofort *die Beziehung*

$$m\int_a^b p(x)\,dx \leqq \int_a^b f(x)\,p(x)\,dx \leqq M\int_a^b p(x)\,dx$$

*oder, in eine Gleichung zusammengefaßt,*

$$\int_a^b f(x)\,p(x)\,dx = f(\xi)\int_a^b p(x)\,dx,$$

*wo* $\xi$ *wieder ein Mittelwert ist.*

**2. Anwendungen. Die Integration von $x^{\alpha}$ für beliebiges irrationales $\alpha$.**

Der Mittelwertsatz bzw. die mit ihm gleichbedeutenden Integralabschätzungen liefern uns sofort eine sehr wichtige Einsicht in eine anschaulich übrigens sofort einleuchtende Tatsache: Der Wert eines Integrales ändert sich nur wenig, wenn die Funktion selbst überall wenig geändert wird. Präzis ausgedrückt: Wenn zwei Funktionen $f(x)$ und $g(x)$ sich in dem ganzen Intervall $a \leqq x \leqq b$ absolut genommen um weniger als eine Zahl $\varepsilon$ unterscheiden, so unterscheiden sich ihre Integrale voneinander absolut genommen um weniger als $\varepsilon(b-a)$. In Formeln: Wenn in unserem Intervall $|f(x) - g(x)| < \varepsilon$ gilt, so ist

$$\left|\int_a^b f(x)\,dx - \int_a^b g(x)\,dx\right| < \varepsilon(b-a)$$

oder, was damit gleichbedeutend ist,

$$-\varepsilon(b-a) + \int_a^b g(x)\,dx < \int_a^b f(x)\,dx < \int_a^b g(x)\,dx + \varepsilon(b-a).$$

Figur 47 zeigt uns diese Tatsache sehr deutlich. Wir ziehen zu der Kurve $y = f(x)$ die „Parallelkurven" $y = f(x) + \varepsilon$ und $y = f(x) - \varepsilon$. In dem durch diese beiden „Parallelkurven" gebildeten Streifen verläuft nach Voraussetzung die Funktion $g(x)$. Es wird hieraus sofort klar, daß die Flächeninhalte, die von den Kurven $f(x)$ und $g(x)$ begrenzt sind, sich um weniger als den halben Streifeninhalt voneinander unterscheiden, und der Streifeninhalt ist gerade die Zahl

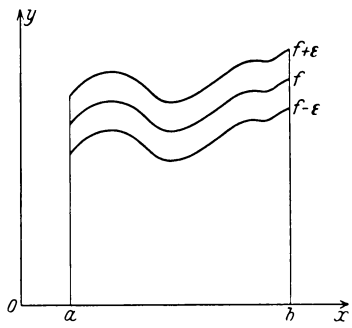


Fig. 47. Stetigkeit eines Integrales.

$$\int_a^b (f(x) + \varepsilon)\,dx - \int_a^b (f(x) - \varepsilon)\,dx = 2\varepsilon(b-a).$$

Ohne Berufung auf die Anschauung folgt (analog zu den Betrachtungen in Nr. 1) aus den Ungleichungen $-\varepsilon + g(x) < f(x) < \varepsilon + g(x)$ die Beziehung

$$\int_a^b (-\varepsilon + g(x))\,dx < \int_a^b f(x)\,dx < \int_a^b (g(x) + \varepsilon)\,dx,$$

welche sofort infolge der Grundregeln über Integrale die Gestalt

$$-\varepsilon(b-a) + \int_a^b g(x)\,dx < \int_a^b f(x)\,dx < \int_a^b g(x)\,dx + \varepsilon(b-a)$$

annimmt — wir haben dabei nur das Integral einer Summe durch die entsprechende Summe der Integrale ersetzt und beachtet, daß

$$\int_a^b \varepsilon\, dx = \varepsilon(b-a)$$

ist.

Zur Erläuterung der eben bewiesenen Tatsache will ich zeigen, daß man mit ihrer Hilfe imstande ist, die Funktion $x^\alpha$ für irgend einen irrationalen Wert von $\alpha$ zu integrieren, genauer, das bestimmte Integral $\int_a^b x^\alpha\, dx$ zu bilden. Wir setzen dabei voraus, daß $0 < a < b$ ist.

Den Exponenten $\alpha$ stellen wir als Grenzwert einer Folge von rationalen Zahlen $\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$ dar, $\alpha = \lim_{n\to\infty} \alpha_n$, wobei wir voraussetzen dürfen, daß keiner dieser Werte $\alpha_n$ mit $-1$ zusammenfällt; denn $\alpha$ selbst ist ja von $-1$ verschieden. Für die Potenz $x^\alpha$ benutzen wir dann die Definition

$$x^\alpha = \lim_{n\to\infty} x^{\alpha_n}$$

und beachten folgendes: Wie klein auch immer wir die positive Zahl $\varepsilon$ wählen, stets können wir den Index $n$ so groß hinzubestimmen, daß in dem ganzen Intervalle $a \leqq x \leqq b$ gilt: $|x^\alpha - x^{\alpha_n}| < \varepsilon$ [1]).

Nunmehr brauchen wir nur unsere obige Beziehung auf die Funktionen $f(x) = x^\alpha$ und $g(x) = x^{\alpha_n}$ anzuwenden und erhalten

$$-\varepsilon(b-a) + \int_a^b x^{\alpha_n}\, dx < \int_a^b x^\alpha\, dx < \int_a^b x^{\alpha_n}\, dx + \varepsilon(b-a).$$

Rechts und links aber können wir die Integrale gemäß dem Resultat aus § 2, Nr. 4 ausführen und erhalten

$$-\varepsilon(b-a) + \frac{1}{\alpha_n+1}(b^{\alpha_n+1} - a^{\alpha_n+1}) < \int_a^b x^\alpha\, dx < \frac{1}{\alpha_n+1}(b^{\alpha_n+1} - a^{\alpha_n+1}) + \varepsilon(b-a).$$

Lassen wir nunmehr die Zahl $\varepsilon$ immer kleiner werden und gegen 0 streben, so müssen wir $n$ immer größer und größer wählen; da die Zahlen $\alpha_n$

---

1) Der Beweis ergibt sich, wenn wir der Kürze halber voraussetzen, daß $\alpha_n > \alpha > 0$, also $\delta_n = \alpha_n - \alpha > 0$ ist, sehr einfach folgendermaßen: Unter Berücksichtigung des monotonen Charakters von $x^\alpha$ wird

$$|x^\alpha - x^{\alpha_n}| = x^\alpha\,|1 - x^{\delta_n}| \leq b^\alpha\,(|1 - a^{\delta_n}| + |1 - b^{\delta_n}|).$$

Da $\lim_{n\to\infty} a^{\delta_n} = \lim_{n\to\infty} b^{\delta_n} = 1$, also $\lim |1 - a^{\delta_n}| = \lim |1 - b^{\delta_n}| = 0$ ist, so strebt die von der Lage von $x$ im Intervall gar nicht mehr abhängige Zahl $\varepsilon_n = b^\alpha(|1 - a^{\delta_n}| + |1 - b^{\delta_n}|)$ mit wachsendem $n$ gegen Null, woraus die Behauptung folgt.

gegen $\alpha$ und $a^{\alpha_n}$ bzw. $b^{\alpha_n}$ gegen $a^{\alpha}$ bzw. $b^{\alpha}$ konvergieren, erhalten wir unmittelbar das Resultat

$$\int_a^b x^{\alpha}\,dx = \frac{1}{\alpha+1}\left(b^{\alpha+1} - a^{\alpha+1}\right).$$

Mit anderen Worten: *Auch für irrationale Exponenten $\alpha$ gilt dieselbe Integralformel wie für rationale.* Beiläufig bemerkt folgt hiernach aus dem Fundamentalsatz von § 4 für positives $x$, daß auch für irrationale $\alpha$ die Differentiationsformel

$$\frac{d}{dx}\, x^{\alpha+1} = (\alpha + 1)\, x^{\alpha}$$

gilt.

# Anhang zum zweiten Kapitel.

## § 1. Die Existenz des bestimmten Integrales einer stetigen Funktion.

Ich bin noch den analytischen Beweis der Tatsache schuldig, daß das bestimmte Integral einer stetigen Funktion $f(x)$ zwischen den Grenzen $a$ und $b$ $(a < b)$ stets existiert. Hierzu betrachten wir in der Bezeichnung des zweiten Kapitels, § 1, die Summe

$$F_n = \sum_{\nu=1}^{n} f(\xi_\nu)\,\Delta x_\nu.$$

Es ist sicherlich

$$\underline{F_n} = \sum_{\nu=1}^{n} f(v_\nu)\,\Delta x_\nu \leqq F_n \leqq \sum_{\nu=1}^{n} f(u_\nu)\,\Delta x_\nu = \overline{F_n},$$

wo $f(v_\nu)$ den kleinsten, $f(u_\nu)$ den größten Funktionswert im $\nu$-ten Teilintervall bedeutet. Die Aufgabe ist, zu beweisen, daß $F_n$ gegen einen bestimmten, von der speziellen Art der Intervalleinteilung und der Wahl der $\xi_\nu$ unabhängigen Grenzwert konvergiert, wenn bei wachsendem $n$ die Länge des längsten Teilintervalles $\Delta x_\nu$ gegen 0 strebt. Offenbar ist es zu diesem Zwecke notwendig und hinreichend, zu zeigen, daß die Ausdrücke $\underline{F_n}$ und $\overline{F_n}$ beide gegen einen und denselben Grenzwert konvergieren.

Wenn die Intervalleinteilung so fein gewählt ist, daß in jedem Intervall die Schwankung $|f(u_\nu) - f(v_\nu)|$ der stetigen Funktion kleiner als eine beliebig klein vorgegebene Zahl $\varepsilon$ wird (was wegen der Gleichmäßigkeit der Stetigkeit der Funktion $f(x)$ stets möglich ist), dann ist gewiß

$$0 \leqq \overline{F_n} - \underline{F_n} = \sum_{\nu=1}^{n} \Delta x_\nu \left(f(u_\nu) - f(v_\nu)\right) < \varepsilon\,(b - a),$$

und wir sehen also, daß diese Differenz bei wachsendem $n$ gegen 0 strebt, so daß wir uns damit begnügen können, z. B. die Konvergenz von $\overline{F}_n$ zu beweisen. Diese Konvergenz ist bewiesen, sobald gezeigt ist, daß $|\overline{F}_n - \overline{F}_m|$ beliebig klein wird, wenn nur die beiden entsprechenden Einteilungen, die wir als „Einteilung $n$" bzw. „Einteilung $m$" bezeichnen wollen, einen bestimmten Grad der Feinheit übersteigen. Dieser „Grad der Feinheit" soll dadurch charakterisiert werden, daß für beide Einteilungen die Schwankung der Funktion $f(x)$ in jedem Teilintervalle kleiner als $\varepsilon$ bleibt ($\varepsilon > 0$). Dann gehen wir zu einer dritten Einteilung über, deren Teilpunkte wir erhalten, wenn wir sämtliche Teilpunkte der beiden Einteilungen, die zu den Zahlen $n$ und $m$ gehören, zusammennehmen. Diese Teilung, welche $l$ Teilpunkte haben möge, bezeichnen wir durch den Index $l$ und betrachten einen zugehörigen Wert $\overline{F}_l$. Wir werden nun den Ausdruck $|\overline{F}_n - \overline{F}_m|$ abschätzen, indem wir erst für die Ausdrücke $|\overline{F}_n - \overline{F}_l|$ und $|\overline{F}_m - \overline{F}_l|$ Abschätzungen gewinnen. Ich behaupte, daß folgende beiden Beziehungen gelten:

$$\underline{F}_n \leqq \overline{F}_l \leqq \overline{F}_n \quad \text{und} \quad \underline{F}_m \leqq \overline{F}_l \leqq \overline{F}_m .$$

Der Beweis folgt sofort aus der Bedeutung unserer Ausdrücke. Betrachten wir etwa das $\nu$-te Teilintervall der Einteilung $n$, dann entsprechen diesem Teilintervall ein oder mehrere Teilintervalle der Einteilung $l$; sämtliche zu diesen Intervallen gehörigen Summanden bestehen aus zwei Faktoren, deren erster eine Differenz $\varDelta x$ und deren zweiter sicherlich nicht größer als $f(u_\nu)$ und nicht kleiner als $f(v_\nu)$ wird. Die Summe der Differenzen $\varDelta x$ bei der feineren Einteilung $l$, die zu dem betrachteten $\nu$-ten Intervall der gröberen Einteilung $n$ gehören, ist aber gerade $\varDelta x_\nu$. Wir sehen also, daß der betreffende Beitrag zu dem Ausdruck $\overline{F}_l$ zwischen den Grenzen $\varDelta x_\nu f(u_\nu)$ und $\varDelta x_\nu f(v_\nu)$ liegt. Indem wir jetzt die Summe über sämtliche $n$ Teilintervalle bilden, erhalten wir sofort die erste der beiden obigen Ungleichungen, und die zweite ergibt sich genau so, wenn wir statt von der Einteilung $n$ von der Einteilung $m$ ausgehen.

Wir haben vorhin gesehen, daß $\overline{F}_n - \underline{F}_n < \varepsilon(b-a)$ ist; ebenso gilt $\overline{F}_m - \underline{F}_m < \varepsilon(b-a)$. Nach den eben bewiesenen Ungleichungen für $\overline{F}_l$ folgt also $0 \leqq \overline{F}_n - \overline{F}_l < \varepsilon(b-a)$ und $0 \leqq \overline{F}_m - \overline{F}_l < \varepsilon(b-a)$. Mithin gilt sicherlich auch

$$|\overline{F}_n - \overline{F}_m| = |(\overline{F}_n - \overline{F}_l) - (\overline{F}_m - \overline{F}_l)| < 2\varepsilon(b-a).$$

Diese Beziehung lehrt, da $\varepsilon$ beliebig klein vorgegeben werden kann, tatsächlich nach dem Cauchyschen Konvergenzkriterium die Konvergenz unserer Zahlen $\overline{F}_n$. Gleichzeitig erkennen wir aus unserer Betrachtung unmittelbar, daß der Grenzwert von der Art der Einteilung vollständig unabhängig ist [1]).

Der Existenzbeweis für das bestimmte Integral einer stetigen Funktion ist damit geführt.

Unser Beweisverfahren lehrt noch mehr. Es zeigt uns, daß man in vielen Fällen auch durch etwas allgemeinere Grenzwertbildungen zum Integral geführt wird. Wenn z. B. $f(x) = \varphi(x)\,\psi(x)$ ist und das Integrationsintervall von $a$ bis $b$ durch die Teilpunkte $x_\nu$ in $n$ Teile geteilt ist, so betrachten wir jetzt statt der Summe $\sum f(\xi_\nu)\,\Delta x_\nu$ die allgemeinere Summe

$$\sum \varphi(\xi'_\nu)\,\psi(\xi''_\nu)\,\Delta x_\nu\,,$$

wo nunmehr $\xi'_\nu$ und $\xi''_\nu$ zwei nicht notwendig zusammenfallende Punkte des $\nu$-ten Teilintervalles sind. Auch diese Summe wird bei wachsendem $n$ gegen das Integral

$$\int_a^b f(x)\,dx = \int_a^b \varphi(x)\,\psi(x)\,dx$$

streben, wenn dabei die größte Intervallänge gegen Null konvergiert.

Entsprechendes gilt für alle analog gebildeten Summen; z. B. wird die Summe

$$\sum_{\nu=1}^{n} \sqrt{\varphi(\xi'_\nu)^2 + \psi(\xi''_\nu)^2}\,\Delta x_\nu$$

gegen das Integral

$$\int_a^b \sqrt{\varphi(x)^2 + \psi(x)^2}\,dx$$

konvergieren. Der Beweis für diese Tatsachen verläuft vollständig ebenso wie der oben geführte und braucht daher nicht besonders ausgeführt zu werden.

## § 2. Zusammenhang des Mittelwertsatzes der Differentialrechnung mit dem Mittelwertsatz der Integralrechnung.

Zwischen dem Mittelwertsatz der Differentialrechnung und dem der Integralrechnung besteht eine einfache Beziehung, welche durch den Fundamentalsatz gegeben wird und die ich als lehrreiches Beispiel für seine Handhabung hier ausführen will. Nehmen wir zunächst den Mittelwertsatz der Integralrechnung in seiner speziellen Form

$$\int_a^b f(x)\,dx = (b-a)\,f(\xi)\,,$$

[1]) Vgl. hierzu Anm. 1 auf S. 36.

dann erhalten wir, wenn wir $\int f(x)\,dx = F(x)$, also $f(x) = F'(x)$ setzen, unmittelbar als Ausdruck des eben hingeschriebenen Mittelwertsatzes die Gleichung

$$F(b) - F(a) = (b - a)\,F'(\xi)$$

oder

$$\frac{F(b) - F(a)}{b - a} = F'(\xi)\,.$$

Hierbei können wir für $F(x)$ offenbar eine beliebige Funktion wählen, deren erste Ableitung $F'(x) = f(x)$ stetig ist, und damit ist der Mittelwertsatz der Differentialrechnung für solche Funktionen bewiesen.

Die allgemeinere Fassung des Mittelwertsatzes der Integralrechnung

$$\int_a^b f(x)\,p(x)\,dx = f(\xi)\int_a^b p(x)\,dx\,,$$

wo $p(x)$ eine in unserem Intervall positive, $f(x)$ eine dort beliebige stetige Funktion bedeutet, führt entsprechend zu einem verallgemeinerten Mittelwertsatz der Differentialrechnung. Setzen wir

$$\int f(x)\,p(x)\,dx = F(x), \quad \text{d. h.} \quad f(x)\,p(x) = F'(x);$$
$$\int p(x)\,dx = G(x), \quad \text{d. h.} \quad p(x) = G'(x),$$

so nimmt die obige Mittelwertformel die Gestalt an

$$F(b) - F(a) = (G(b) - G(a))\,f(\xi)$$

oder, da $f(x) = \frac{F'(x)}{G'(x)}$ ist, für $a \neq b$

$$\frac{F(b) - F(a)}{G(b) - G(a)} = \frac{F'(\xi)}{G'(\xi)}\,.$$

Diese Formel, in welcher $\xi$ wieder einen gewissen Zwischenwert zwischen $a$ und $b$ bedeutet, nennt man den *verallgemeinerten Mittelwertsatz der Differentialrechnung*. Für seine Gültigkeit braucht offenbar lediglich vorausgesetzt zu werden, daß $F$ und $G$ stetige Funktionen mit stetigen Differentialquotienten sind und daß überdies $G'(x)$ überall positiv (oder auch überall negativ) ist. Man kann nämlich unter diesen Voraussetzungen die ganze Betrachtung unmittelbar umkehren.

Es sei schließlich darauf hingewiesen, daß wir bei den hier gegebenen Herleitungen für die Mittelwertsätze der Differentialrechnung engere Voraussetzungen machen müssen, als es an sich nötig ist. (Vgl. hierzu zweites Kapitel, § 3, Nr. 8 und später S. 165.)

Drittes Kapitel.

# Differential- und Integralrechnung der elementaren Funktionen.

## § 1. Die einfachsten Differentiationsregeln und ihre Anwendungen.

In der höheren Analysis und ihren Anwendungen ist die Sachlage gewöhnlich die, daß die Probleme des Integrierens wichtiger sind als die des Differenzierens, daß jedoch die Differentiation weit weniger Schwierigkeiten bietet. Demgemäß erscheint es im Aufbau der Integral- und Differentialrechnung als das naturgemäße Verfahren, zunächst möglichst weite Klassen von Funktionen differenzieren zu lernen und die gewonnenen Resultate vermöge der Fundamentalsätze des § 4 vom zweiten Kapitel zur Lösung von Integrationsproblemen nutzbar zu machen. Die Durchführung dieses Programmes wird die Aufgabe der nächsten Paragraphen sein. Wir wollen dabei gewissermaßen noch einmal von vorn anfangen, indem wir die wichtigsten Differentiationen und Integrationen ohne Berufung auf die Resultate des vorigen Kapitels in systematischem Zusammenhange ausführen. Dabei spielen einige Differentiationsregeln, von denen wir die ersten übrigens schon kennengelernt haben (zweites Kapitel, § 3, Nr. 4), eine wichtige Rolle.

### 1. Differentiationsregeln.

Wir setzen voraus, daß in dem betrachteten Intervalle die Funktionen $f(x)$ und $g(x)$ differenzierbar sind; dann lauten unsere Regeln:

1. Regel. *Multiplikation mit einer Konstanten.* Ist $c$ eine Konstante und $\varphi(x) = c f(x)$, so ist $\varphi(x)$ differenzierbar, und es wird

$$\varphi'(x) = c f'(x).$$

Der Beweis folgt unmittelbar aus der Beziehung

$$\frac{\varphi(x+h) - \varphi(x)}{h} = c\,\frac{f(x+h) - f(x)}{h}$$

und dem Grenzübergang $h \to 0$.

2. Regel. *Differentialquotient einer Summe.* Ist $\varphi(x) = f(x) + g(x)$, so ist $\varphi(x)$ differenzierbar, und es wird

$$\varphi'(x) = f'(x) + g'(x),$$

d. h. wir können die Prozesse der Differentiation und Addition miteinander vertauschen. Dasselbe gilt bei einer Summe aus einer beliebigen Anzahl, etwa $n$ Summanden

$$\varphi(x) = \sum_{\nu=1}^{n} f_\nu(x).$$

Es ist dann

$$\varphi'(x) = \sum_{\nu=1}^{n} f_\nu'(x).$$

Den Beweis kann ich nach dem zweiten Kapitel, § 3, als fast selbstverständlich übergehen.

3. Regel. *Differentialquotient eines Produktes.* Ist

$$\varphi(x) = f(x)\, g(x),$$

so ist $\varphi(x)$ differenzierbar, und es wird

$$\varphi'(x) = f(x)\, g'(x) + g(x)\, f'(x).$$

Der Beweis folgt aus der Gleichung

$$\frac{\varphi(x+h) - \varphi(x)}{h} = \frac{f(x+h)\, g(x+h) - f(x)\, g(x)}{h}$$

$$= \frac{f(x+h)\, g(x+h) - f(x+h)\, g(x) + f(x+h)\, g(x) - f(x)\, g(x)}{h}$$

$$= f(x+h)\,\frac{g(x+h) - g(x)}{h} + g(x)\,\frac{f(x+h) - f(x)}{h}.$$

In dem letzten Ausdruck läßt sich der Grenzübergang $h \to 0$ ohne weiteres ausführen und führt unmittelbar zu der behaupteten Formel.

Diese Formel gewinnt eine noch übersichtlichere Gestalt, wenn wir sie durch $\varphi(x) = f(x)\, g(x)$ dividieren[1]). Wir erhalten dann

$$\frac{\varphi'(x)}{\varphi(x)} = \frac{f'(x)}{f(x)} + \frac{g'(x)}{g(x)}.$$

Wendet man diese Produktregel mehrfach an, so erhält man für den Differentialquotienten des Produkts von mehreren, etwa $n$ Funktionen einen Ausdruck aus $n$ Summanden, deren jeder aus dem Produkt der Ableitung eines der Faktoren mit den übrigen Faktoren des ursprünglichen Produktes besteht. Als Formel geschrieben:

$$\varphi'(x) = \frac{d}{dx}\left(f_1(x)\, f_2(x) \cdots f_n(x)\right) = f_1'(x)\, f_2(x) \cdots f_n(x)$$

$$+ f_1(x)\, f_2'(x)\, f_3(x) \cdots f_n(x) + \cdots + f_1(x)\, f_2(x) \cdots f_n'(x) = \sum_{\nu=1}^{n} f_\nu'(x)\,\frac{\varphi(x)}{f_\nu(x)} \;{}^{1})$$

oder übersichtlicher nach Division durch $\varphi(x) = f_1(x)\, f_2(x) \cdots f_n(x)$

$$\frac{\varphi'(x)}{\varphi(x)} = \frac{f_1'(x)}{f_1(x)} + \frac{f_2'(x)}{f_2(x)} + \cdots + \frac{f_n'(x)}{f_n(x)} = \sum_{\nu=1}^{n} \frac{f_\nu'(x)}{f_\nu(x)} \;{}^{1}).$$

4. Regel. *Differentialquotient eines Quotienten.* Für den Quotienten

$$\varphi(x) = \frac{f(x)}{g(x)}$$

---

[1]) Wobei wir allerdings annehmen müssen, daß $\varphi(x)$ nirgends verschwindet.

gilt die folgende Differentiationsregel: An allen Stellen, an denen $g(x)$ nicht verschwindet, ist $\varphi(x)$ differenzierbar, und es ist

$$\varphi'(x) = \frac{g(x)\,f'(x) - g'(x)\,f(x)}{g(x)^2}$$

oder bei $\varphi(x) \neq 0$

$$\frac{\varphi'(x)}{\varphi(x)} = \frac{f'(x)}{f(x)} - \frac{g'(x)}{g(x)}.$$

Setzen wir die Differenzierbarkeit von $\varphi(x)$ schon voraus, so können wir nach der Produktregel aus $f(x) = \varphi(x)\,g(x)$ schließen:

$$f'(x) = \varphi(x)\,g'(x) + g(x)\,\varphi'(x).$$

Setzt man hier rechts für $\varphi(x)$ den Quotienten $\frac{f(x)}{g(x)}$ ein und rechnet aus der gewonnenen Gleichung $\varphi'(x)$ aus, so erhält man ohne weiteres die obige Regel. Um jedoch die Differenzierbarkeit von $\varphi(x)$ gleichzeitig mit zu beweisen, geben wir dem Beweise folgende Wendung. Wir schreiben:

$$\frac{\varphi(x+h) - \varphi(x)}{h} = \frac{\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}}{h} = \frac{g(x)\,\frac{f(x+h) - f(x)}{h} - \frac{g(x+h) - g(x)}{h}\,f(x)}{g(x)\,g(x+h)}.$$

Lassen wir jetzt $h$ gegen Null streben, so erhalten wir das behauptete Resultat; denn nach Voraussetzung existieren die Grenzwerte der beiden nach Ausführung der Division sich ergebenden Summanden rechts und sind gleich $\frac{g(x)\,f'(x)}{g(x)^2}$ bzw. gleich $\frac{g'(x)\,f(x)}{g(x)^2}$, womit dann auch die Existenz des Grenzwertes der linken Seite und die Differentiationsformel folgt.

### 2. Differentiation der rationalen Funktionen.

Wir leiten zunächst noch einmal für jedes ganzzahlige positive $n$ die Differentiationsformel

$$\frac{d}{dx}\,x^n = n\,x^{n-1}$$

ab, indem wir uns auf die Produktregel stützen. Wir fassen $x^n$ auf als ein Produkt von $n$ Faktoren: $x^n = x \cdots x$ und erhalten daraus

$$\frac{d}{dx}\,x^n = 1 \cdot x^{n-1} + 1 \cdot x^{n-1} + \cdots + 1 \cdot x^{n-1} = n\,x^{n-1}.$$

Der zweite Differentialquotient der Funktion $x^n$ ergibt sich nunmehr unter Benutzung der eben gewonnenen Formel und der ersten Differentiationsregel sofort:

$$\frac{d^2}{dx^2}\,x^n = n\,(n-1)\,x^{n-2},$$

und ebenso folgt hieraus

$$\frac{d^3}{dx^3}x^n = n(n-1)(n-2)\,x^{n-3},$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\frac{d^n}{dx^n}x^n = 1\cdot 2\cdots n = n!$$

Man erkennt, daß der $(n+1)$-te Differentialquotient von $x^n$ überall verschwindet.

Mit der Differentiation der Potenzen ist auf Grund unserer beiden ersten Regeln sofort auch die Differentiation einer beliebigen ganzen rationalen Funktion

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

möglich. Es wird einfach

$$y' = a_1 + 2a_2 x + 3a_3 x^2 + \cdots + n a_n x^{n-1}$$

und weiter

$$y'' = 2a_2 + 3\cdot 2a_3 x + 4\cdot 3a_4 x^2 + \cdots + n\cdot(n-1)\,a_n x^{n-2},\ \text{usw.}$$

Die Differentiation der gebrochenen rationalen Funktion ergibt sich nun mit Hilfe der Quotientenregel. Speziell wollen wir noch einmal die Differentiationsformel für die Funktion $x^n$ ableiten, wenn $n = -m$ eine ganze negative Zahl ist. Die Anwendung der Quotientenregel ergibt, unter Berücksichtigung der Tatsache, daß der Differentialquotient des Zählers gleich Null ist, das Resultat

$$\frac{d}{dx}\left(\frac{1}{x^m}\right) = -\frac{m x^{m-1}}{x^{2m}} = -\frac{m}{x^{m+1}}$$

oder, wenn wir $m = -n$ einsetzen,

$$\frac{d}{dx}x^n = n x^{n-1}$$

in genauer formaler Übereinstimmung mit der Formel für positives $n$ und ebenso wie diese im Einklang mit dem früher (zweites Kapitel, § 3, Nr. 3) gewonnenen Resultat.

### 3. Differentiation der trigonometrischen Funktionen.

Für die trigonometrischen Funktionen $\sin x$ und $\cos x$ hatten wir schon früher (S. 75f.) die Differentiationsformeln

$$\frac{d}{dx}\sin x = \cos x \quad \text{und} \quad \frac{d}{dx}\cos x = -\sin x$$

gewonnen. Wir können nunmehr mit Hilfe der Quotientenregel auch die Funktionen

$$y = \operatorname{tg} x = \frac{\sin x}{\cos x} \quad \text{und} \quad y = \operatorname{ctg} x = \frac{\cos x}{\sin x}$$

differenzieren. Der Differentialquotient der ersten Funktion wird nach dieser Regel

$$y' = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x},$$

und wir erhalten das Resultat

$$\frac{d}{dx}\operatorname{tg} x = 1 + \operatorname{tg}^2 x = \frac{1}{\cos^2 x}.$$

Ganz ebenso ergibt sich

$$\frac{d}{dx}\operatorname{ctg} x = -(1 + \operatorname{ctg}^2 x) = -\frac{1}{\sin^2 x}.$$

## § 2. Die entsprechenden Integralformeln.

### 1. Allgemeine Integrationsregeln.

Der Fundamentalsatz des § 4 vom zweiten Kapitel bzw. die Definition des unbestimmten Integrales eröffnet uns die Möglichkeit, jeder Differentiationsformel eine entsprechende äquivalente Integralformel gegenüberzustellen. Mit den ersten drei Differentiationsregeln des vorigen Paragraphen sind offenbar die folgenden z. T. schon im zweiten Kapitel, § 1, erwähnten Integrationsregeln vollständig äquivalent.

*Multiplikation mit einer Konstanten:* Ist $c$ eine Konstante, so ist

$$\int c f(x)\, dx = c \int f(x)\, dx.$$

*Integration einer Summe:* Es ist stets

$$\int (f(x) + g(x))\, dx = \int f(x)\, dx + \int g(x)\, dx.$$

Der dritten Differentiationsregel entspricht die *Regel der Produktintegration oder Teilintegration (partielle Integration).* Durch Integration erhalten wir aus der Produktregel

$$\int (f(x)\, g(x))'\, dx = \int f(x)\, g'(x)\, dx + \int g(x)\, f'(x)\, dx.$$

Das unbestimmte Integral auf der linken Seite ist offenbar (bis auf eine additive Konstante) $f(x)\, g(x)$, und wir können daher die Regel der Produktintegration in der folgenden Gestalt schreiben:

$$\int f(x)\, g'(x)\, dx = f(x)\, g(x) - \int g(x)\, f'(x)\, dx.$$

Ich habe diese Integralformel nur der Vollständigkeit halber als Gegenstück zu der Produktregel der Differentiation angegeben; von Wichtigkeit wird sie für uns erst im nächsten Kapitel werden (siehe dort § 4).

### 2. Integration der einfachsten Funktionen.

Den oben gewonnenen Differentiationsformeln für spezielle Funktionen stellen wir nun die inhaltlich gleichbedeutenden Integrationsformeln gegenüber. Die Formel

$$\frac{d}{dx} x^n = n x^{n-1}$$

lautet als Integralformel

$$\int x^{n-1}\, dx = \frac{x^n}{n}, \qquad n \neq 0.$$

Denn diese Formel besagt nichts anderes, als daß der Differentialquotient der rechten Seite gleich dem links unter dem Integralzeichen stehenden Integranden ist. — Ersetzen wir $n$ durch $n+1$, so gewinnen wir die Integralformel

$$\int x^n\,dx = \frac{1}{n+1}x^{n+1}, \qquad n \neq -1.$$

Diese Formel gilt für jeden ganzzahligen Exponenten $n$ (bei $n<0$ natürlich nur für $x \neq 0$), ausgenommen für $n = -1$, wo der Nenner $n+1$ verschwinden würde. Dieser Ausnahmefall $n = -1$ soll uns später im § 6 noch ausführlicher beschäftigen.

Der Fundamentalsatz der Integralrechnung erlaubt uns sofort, unsere Integralformeln zur Flächeninhaltsbestimmung, d. h. zur Berechnung b e s t i m m t e r  I n t e g r a l e, auszunützen. Wir erhalten unmittelbar nach Kap. II, § 4, Nr. 4

$$\int_a^b x^n\,dx = \frac{1}{n+1}(b^{n+1} - a^{n+1}), \qquad n \neq -1,$$

wobei wir, wenn $n$ negativ ist, $a$ und $b$ beide als gleichen Vorzeichens voraussetzen, weil sonst der Integrand im Integrationsintervall unstetig wird.

Den Differentiationsformeln für $\sin x$, $\cos x$, $\operatorname{tg} x$ und $\operatorname{ctg} x$ entsprechen als Integralformeln die folgenden:

$$\int \cos x\,dx = \sin x, \qquad \int \sin x\,dx = -\cos x,$$

$$\int \frac{1}{\cos^2 x}\,dx = \operatorname{tg} x, \qquad \int \frac{1}{\sin^2 x}\,dx = -\operatorname{ctg} x.$$

Aus diesen Formeln erhalten wir sofort nach der Grundregel vom zweiten Kapitel, § 4, die Werte des bestimmten Integrales zwischen irgendwelchen Grenzen, z. B.

$$\int_a^b \cos x\,dx = \sin x\,\Big|_a^b = \sin b - \sin a.$$

Daß es mit Hilfe der ersten beiden Integrationsregeln nunmehr möglich ist, beliebige ganze rationale Funktionen, sowie mit beliebigen konstanten Koeffizienten gebildete lineare Kombinationen der übrigen hier integrierten Funktionen zu integrieren, brauche ich hier kaum noch besonders zu betonen. Ich möchte aber noch auf den folgenden Punkt hinweisen. Integrationsregeln und Differentiationsregeln müssen nach dem Fundamentalsatz einander äquivalent sein; man kann daher auch die allgemeinen Integrationsregeln dieses Paragraphen zuerst beweisen und aus ihnen die Differentiationsregeln des vorangehenden Paragraphen ablesen. Ich möchte dem Leser raten, diesen Gedanken allein durchzuführen.

# § 3. Die Umkehrfunktion und ihr Differentialquotient.

## 1. Die allgemeine Differentiationsformel.

Wir haben früher (S. 14 und 52) erkannt, daß eine stetige Funktion $y = f(x)$ sich in einem Intervall, in welchem sie monoton ist, eindeutig umkehren läßt. Genauer: *Ist $a \leqq x \leqq b$ ein Intervall, in welchem die stetige Funktion $y = f(x)$ monoton verläuft und wird $f(a) = \alpha$ und $f(b) = \beta$ gesetzt, so ist $x$ eine in dem Intervall zwischen $\alpha$ und $\beta$ eindeutig bestimmte stetige und monotone Funktion $x = \varphi(y)$ von $y$.*

Der Begriff des Differentialquotienten gibt uns ein sehr einfaches Mittel, um den monotonen Charakter und damit die Umkehrbarkeit einer Funktion zu erkennen. Eine differenzierbare Funktion ist nämlich sicherlich stets dann monoton zunehmend, wenn in dem betreffenden Intervall überall $f'(x) > 0$ ist, und entsprechend monoton abnehmend, wenn dort überall $f'(x) < 0$ ist. Diese anschaulich einleuchtende Tatsache folgt analytisch sofort aus dem Mittelwertsatz $f(x+h) - f(x) = hf'(x + \vartheta h)$ $(0 < \vartheta < 1)$. Ist $f'(x)$ überall positiv bzw. negativ und liegt $x$ sowie $x + h$ im Intervalle, so wird also $f(x+h) - f(x)$ stets das gleiche bzw. das entgegengesetzte Vorzeichen wie $h$ besitzen.

Wir zeigen nun folgenden Satz: *Ist $y = f(x)$ im Intervall $a < x < b$ differenzierbar und dort entweder überall $f'(x) > 0$ oder überall $f'(x) < 0$, dann besitzt in diesem Intervall auch die Umkehrfunktion $x = \varphi(y)$ einen Differentialquotienten, und zwischen dem Differentialquotienten der gegebenen Funktion $y = f(x)$ und dem der Umkehrfunktion $x = \varphi(y)$ besteht für zusammengehörige Werte $x, y$ die Beziehung $f'(x) \cdot \varphi'(y) = 1$, welche wir auch*

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}$$

*schreiben können.*

In der letzten Schreibweise kommt wieder die Schmiegsamkeit der Leibnizschen Bezeichnung zum Ausdruck. Es ist eben tatsächlich so, als ob die Symbole $dy$ und $dx$ Rechengrößen wären, mit denen man operieren kann wie mit wirklichen Zahlen. Entsprechend einfach ergibt sich auch der Beweis dieser Formel, wenn man die Differentialquotienten als Grenzwerte von Differenzenquotienten auffaßt:

$$y' = f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{x_1 \to x} \frac{y_1 - y}{x_1 - x},$$

wobei $x$ und $y = f(x)$ bzw. $x_1$ und $y_1 = f(x_1)$ zusammengehörige Wertepaare bedeuten. Da nach Voraussetzung der erste dieser Grenzwerte nicht Null ist und da die Gleichung $\lim \Delta x = 0$ mit der Gleichung $\lim \Delta y = 0$ wegen der Stetigkeit von $y = f(x)$ und $x = \varphi(y)$ gleichbedeutend ist, also auch die Beziehungen $y_1 \to y$ und $x_1 \to x$ gleichbedeutend sind, so existiert auch der Grenzwert

$$\lim_{x_1 \to x} \frac{x_1 - x}{y_1 - y} = \lim_{y_1 \to y} \frac{x_1 - x}{y_1 - y}$$

und ist gleich $\frac{1}{f'(x)}$. Andererseits bedeutet dieser Grenzwert nach Definition den Differentialquotienten $\varphi'(y)$ der Umkehrfunktion $\varphi(y)$, und damit ist unsere Formel bewiesen.

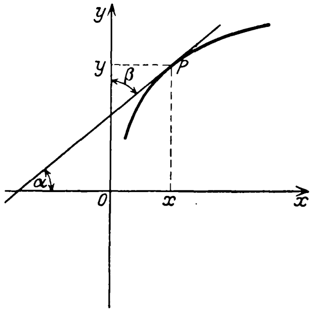


Fig. 48. Zur Differentiation der Umkehrfunktion.

Übrigens hat diese Formel eine einfache geometrische Bedeutung, welche an Figur 48 unmittelbar klar wird. Die Tangente an die Kurve $y = f(x)$ oder $x = \varphi(y)$ bildet mit der positiven $x$-Achse einen Winkel $\alpha$, mit der positiven $y$-Achse einen Winkel $\beta$, und zufolge der geometrischen Bedeutung der Differentialquotienten ist

$$f'(x) = \operatorname{tg}\alpha\,, \quad \varphi'(y) = \operatorname{tg}\beta\,.$$

Da aber die Winkel $\alpha$ und $\beta$ sich gerade zu $\frac{\pi}{2}$ ergänzen, so ist $\operatorname{tg}\alpha \operatorname{tg}\beta = 1$, und diese Relation drückt genau unsere Differentiationsformel aus.

Wir haben bisher ausdrücklich vorausgesetzt, daß entweder $f'(x) > 0$ oder $f'(x) < 0$ sei, daß also nirgends $f'(x) = 0$ wird. Was geschieht nun, wenn diese Gleichung doch besteht? Gilt überall in einem Intervalle $f'(x) = 0$, so ist die Funktion dort konstant, also gewiß nicht umkehrbar, weil demselben Werte von $y$

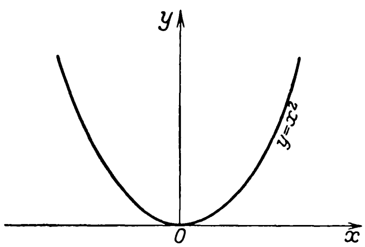


Fig. 49.

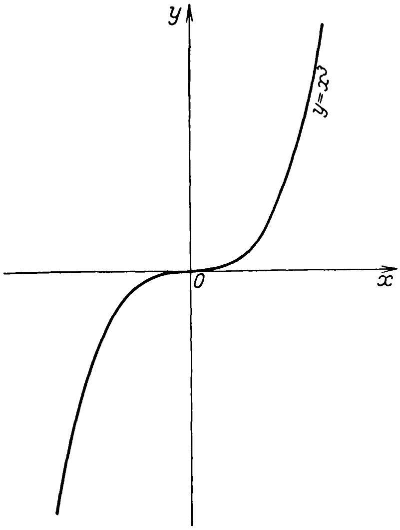


Fig. 50.

alle Werte $x$ dieses Intervalles entsprechen müßten. Gilt jedoch die Gleichung $f'(x) = 0$ nur an isolierten Stellen und wird $f'(x)$ der Einfachheit halber stetig angenommen, so hat man zu unterscheiden, ob beim Durchgang durch diese Stellen $f'(x)$ sein Zeichen wechselt oder nicht. Im ersten Falle trennt diese Stelle ein Intervall, wo die Funktion monoton zunimmt, von einem solchen, wo sie monoton abnimmt. In der Umgebung einer solchen Stelle ist eine eindeutige Umkehrbarkeit nicht vorhanden. Im zweiten Falle stört die Nullstelle der Ableitung den monotonen Charakter der Funktion $y = f(x)$ nicht, also auch nicht die eindeutige Umkehrbarkeit. Nur wird die Umkehrfunktion an der betreffenden Stelle nicht mehr differenzierbar sein; vielmehr wird dort ihre Ableitung unendlich werden.

Beispiele für beide Vorkommnisse liefern die Funktionen $y = x^2$ bzw. $y = x^3$ an der Stelle $x = 0$. Die Figuren 49 und 50 veranschaulichen uns das Verhalten der beiden Funktionen beim Durchgang durch den Nullpunkt und zeigen zugleich, daß man die Funktion $y = x^3$ in der Umgebung des Nullpunktes eindeutig umkehren kann, die Funktion $y = x^2$ aber nicht.

## 2. Die Umkehrfunktionen der Potenzen und der trigonometrischen Funktionen.

Das einfachste Beispiel bieten die Funktionen $y = x^n$ für ganzes positives $n$ und, wie wir zunächst allgemein voraussetzen wollen, positives $x$. Es ist dann $y'$ immer positiv, so daß wir überall für positives $y$ in eindeutiger Weise eine positive Umkehrfunktion

$$x = \sqrt[n]{y} = y^{\frac{1}{n}}$$

bilden können. Der Differentialquotient dieser Umkehrfunktion ergibt sich nach der obigen allgemeinen Regel nunmehr unmittelbar durch die folgende Rechnung:

$$\frac{d y^{\frac{1}{n}}}{d y} = \frac{d x}{d y} = \frac{1}{\frac{d y}{d x}} = \frac{1}{n x^{n-1}} = \frac{1}{n} \frac{1}{y^{\frac{n-1}{n}}} = \frac{1}{n} y^{\frac{1}{n}-1},$$

und wir können als Schlußresultat, indem wir jetzt wieder die unabhängige Variable mit $x$ bezeichnen, schreiben

$$\frac{d \sqrt[n]{x}}{d x} = \frac{d}{d x}\left(x^{\frac{1}{n}}\right) = \frac{1}{n} x^{\frac{1}{n}-1},$$

übrigens in Übereinstimmung mit dem in Kap. II, § 3, Nr. 3 direkt abgeleiteten Ergebnis.

Eine besondere Beachtung verlangt der Grenzpunkt $x = 0$. Nähert sich $x$ von positiven Werten her der Null, so wird $\frac{d x^{\frac{1}{n}}}{d x}$ für $n > 1$ offenbar über alle Grenzen wachsen, was der Tatsache entspricht, daß die Ableitung der $n$-ten Potenz $f(x) = x^n$ für $n > 1$ im Nullpunkt verschwindet. Geometrisch bedeutet dies, daß die Kurven $y = x^{\frac{1}{n}}$ für $n > 1$ im Nullpunkte alle die $y$-Achse berühren (vgl. Fig. 17 von S. 25).

Ergänzend sei bemerkt, daß wir für ungerades $n$ die Voraussetzung $x > 0$ aufgeben und die Funktionen $y = x^n$ für alle Werte von $x$ betrachten dürfen, ohne daß ihr monotoner Charakter und ihre eindeutige Umkehrbarkeit durch die Funktion $x = y^{\frac{1}{n}}$ verloren geht. Auch die Differentiationsformel $\frac{d}{dy} y^{\frac{1}{n}} = \frac{1}{n} y^{\frac{1}{n}-1}$ bleibt dann für negative $y$ bestehen; für $x=0$, $n>1$ wird $\frac{d(x^n)}{dx} = 0$, was einem unendlich großen Differentialquotienten $\frac{dx}{dy}$ an der Stelle $y = 0$ bei der Umkehrfunktion entspricht.

Um die trigonometrischen Funktionen umzukehren, betrachten wir nochmals die Kurven von $\sin x$, $\cos x$, $\operatorname{tg} x$, $\operatorname{ctg} x$. Wir erkennen sofort aus Figur 14 und 15 (S. 17), daß für jede dieser Funktionen ein bestimmtes Intervall ausgewählt werden muß, wenn man von einer eindeutigen Umkehrung sprechen will; denn die Parallelen $y = c$ zur $x$-Achse schneiden diese Kurven, wenn überhaupt, in unendlich vielen Punkten.

Für $y = \sin x$ wird die Ableitung $y' = \cos x$ z. B. im Intervall $-\frac{\pi}{2} < x < \frac{\pi}{2}$ positiv sein. In diesem Intervall können wir also den Sinus eindeutig umkehren; wir schreiben diese Umkehrfunktion des Sinus in der Form

$$x = \arcsin y$$

(gesprochen arcussinus $y$; man meint damit den Bogen, dessen Sinus den Wert $y$ hat). Diese Funktion läuft monoton von $-\frac{\pi}{2}$ bis $+\frac{\pi}{2}$, wenn $y$ ebenso das Intervall von $-1$ bis $+1$ durchläuft. Will man besonders betonen, daß man die Umkehrung des Sinus gerade für dieses Intervall meint, so spricht man von dem *Hauptwert* des Arcussinus. Führt man die Umkehrung für ein anderes Intervall aus, in welchem $\sin x$ monoton verläuft, z. B. das Intervall $\frac{\pi}{2} < x < \frac{3\pi}{2}$, so erhält man „einen andern *Zweig*“ des Arcussinus; ohne genaue Angabe des Intervalles, in dem die Funktionswerte liegen sollen, ist der Arcussinus eine *mehrdeutige* und zwar *unendlich vieldeutige Funktion.*

Allgemein kommt die Mehrdeutigkeit von $\arcsin y$ dadurch zum Ausdruck, daß zu demselben Wert $y$ des Sinus zugleich mit dem Bogen $x$ auch der Bogen $2k\pi + x$, sowie der Bogen $(2k+1)\pi - x$ gehört, wie auch immer die Zahl $k$ als ganzzahliger Wert gewählt wird (vgl. Fig. 51).

Die Differentiation der Funktion $x = \arcsin y$ vollzieht sich nunmehr ohne weiteres vermöge unserer allgemeinen Regel durch die folgende kleine Rechnung:

$$\frac{dx}{dy} = \frac{1}{y'} = \frac{1}{\cos x} = \frac{1}{\sqrt{1-\sin^2 x}} = \frac{1}{\sqrt{1-y^2}},$$

wobei die Quadratwurzel positiv zu nehmen ist, wenn wir uns in dem zuerst angegebenen Intervall befinden[1]).

Schreibt man schließlich wieder für die unabhängige Veränderliche $x$

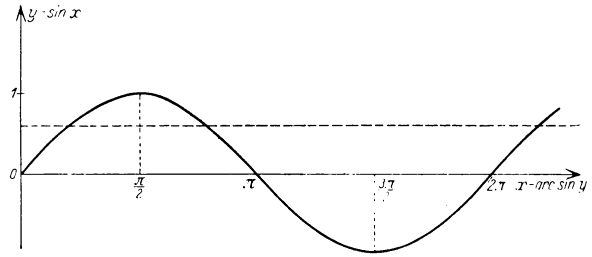



Fig. 51. Umkehrung von sin $x$.

statt $y$, so erhält man die Differentiationsformel der Funktion arc sin $x$ in folgender Form:

$$\frac{d}{dx} \operatorname{arc} \sin x = \frac{1}{\sqrt{1-x^2}}.$$

Dabei ist vorausgesetzt, daß $\operatorname{arc}\sin x$ zwischen $-\frac{\pi}{2}$ und $+\frac{\pi}{2}$ liegt, und das Vorzeichen der Quadratwurzel ist positiv zu wählen.

Genau ebenso ergibt sich für die Umkehrfunktion von $y = \cos x$, die wir mit arc cos $x$ bezeichnen, die Differentiationsformel

$$\frac{d}{dx} \operatorname{arc} \cos x = -\frac{1}{\sqrt{1-x^2}}.$$

Hier ist das Vorzeichen der Wurzel positiv zu nehmen, wenn man

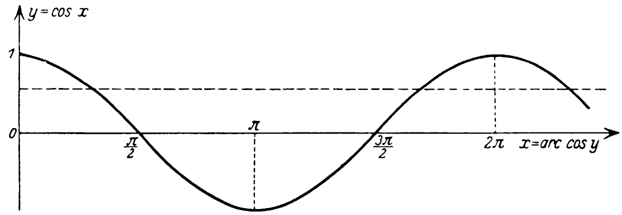



Fig. 52. Umkehrung von cos $x$.

die Werte von $\operatorname{arc}\cos x$ im Intervall zwischen 0 und $\pi$ (nicht wie beim $\operatorname{arc}\sin x$ zwischen $-\frac{\pi}{2}$ und $+\frac{\pi}{2}$) wählt (vgl. Fig. 52).

Noch ein Wort ist über die Endpunkte $x = -1$ und $x = +1$ des Intervalles für $x$ zu sagen. Die Differentialquotienten werden bei Annäherung an diese Endpunkte unendlich, entsprechend der Tatsache, daß die Kurven, welche den arc sin bzw. den arc cos darstellen, an diesen Stellen vertikale Tangenten besitzen müssen.

[1]) Wenn wir statt dessen das Intervall $\frac{\pi}{2} < x < \frac{3\pi}{2}$ zugrunde legten, was der Ersetzung von $x$ durch $x + \pi$ entspricht, so hätten wir das negative Vorzeichen der Wurzel zu wählen, weil dort $\cos x$ negativ ist.

Analog können wir die Umkehrfunktion des tg und ctg behandeln. Die Funktion $y = \operatorname{tg} x$, deren Differentialquotient $\frac{1}{\cos^2 x}$ für $x \neq \frac{\pi}{2} + k\pi$ überall positiv ist, läßt sich im Intervalle $-\frac{\pi}{2} < x < \frac{\pi}{2}$ eindeutig umkehren. Wir nennen diese Umkehrfunktion $x = \operatorname{arc\,tg} y$ bzw. (unter Vertauschung der Buchstaben $x$ und $y$) $y = \operatorname{arc\,tg} x$ und erkennen sofort aus Figur 53, daß die ursprünglich, d. h. ohne genaue Festlegung eines Intervalls für die Werte der Funktion $\operatorname{arc\,tg} x$, vorhandene Mehrdeutigkeit bei der Umkehrung zum Ausdruck kommt, indem wir zu jedem $x$ an Stelle eines Wertes $y$ in unserem Intervall auch den Wert $y + k\pi$ für beliebiges ganzzahliges $k$ hätten wählen können. Für die Funktion $y = \operatorname{ctg} x$ wird die Umkehrfunktion $x = \operatorname{arc\,ctg} y$ bzw. (unter Vertauschung von $x$ und $y$) $y = \operatorname{arc\,ctg} x$ eindeutig bestimmt sein, wenn wir für deren Werte das Intervall von 0 bis $\pi$ festlegen; die Mehrdeutigkeit von $\operatorname{arc\,ctg} x$ ist im übrigen dieselbe wie bei $\operatorname{arc\,tg} x$.

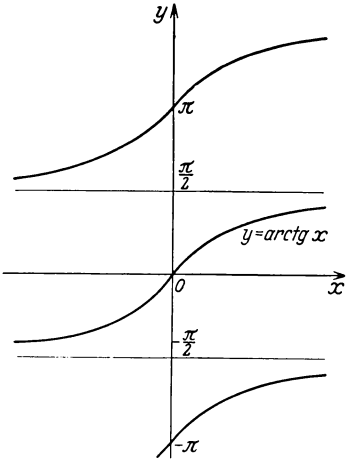


Fig. 53. Arcustangens.

Die Differentiationsformeln ergeben sich wieder in der folgenden Art:

$$x = \operatorname{arc\,tg} y, \quad \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \cos^2 x = \frac{1}{1 + \operatorname{tg}^2 x} = \frac{1}{1 + y^2};$$

$$x = \operatorname{arc\,ctg} y, \quad \frac{dx}{dy} = -\sin^2 x = -\frac{1}{1 + \operatorname{ctg}^2 x} = -\frac{1}{1 + y^2};$$

oder, wenn wir schließlich wieder die unabhängige Veränderliche mit $x$ bezeichnen,

$$\frac{d}{dx} \operatorname{arc\,tg} x = \frac{1}{1 + x^2},$$

$$\frac{d}{dx} \operatorname{arc\,ctg} x = -\frac{1}{1 + x^2}.$$

### 3. Die zugehörigen Integralformeln.

In der Sprache der unbestimmten Integration ausgedrückt, lauten unsere eben abgeleiteten Formeln folgendermaßen:

$$\int \frac{1}{\sqrt{1 - x^2}} dx = \arcsin x, \qquad \int \frac{1}{\sqrt{1 - x^2}} dx = -\arccos x,$$

$$\int \frac{1}{1 + x^2} dx = \operatorname{arc\,tg} x, \qquad \int \frac{1}{1 + x^2} dx = -\operatorname{arc\,ctg} x.$$

Zwischen den ersten beiden Formeln und zwischen den letzten beiden, die den unbestimmten Integralen zwei anscheinend ganz verschiedene Funktionen zuordnen, besteht kein Widerspruch. Wir müssen bedenken, daß im unbestimmten Integral eine beliebige additive Konstante verfügbar bleibt. Ziehen wir aus dieser Konstanten den Bestandteil $\frac{\pi}{2}$ heraus und beachten, daß $\frac{\pi}{2} - \operatorname{arc}\cos x = \operatorname{arc}\sin x$ ist und ebenso $\frac{\pi}{2} - \operatorname{arc}\operatorname{ctg} x = \operatorname{arc}\operatorname{tg} x$, so klärt sich diese formale Unstimmigkeit sofort auf. Diese Unbestimmtheit beruht also lediglich darauf, daß das unbestimmte Integral nicht *eine* bestimmte Funktion, sondern eine ganze Schar von Funktionen bedeutet, welche sich untereinander durch beliebige additive Konstanten unterscheiden. Eine Gleichung für ein unbestimmtes Integral legt nicht *den* Wert, sondern nur *einen* Wert desselben fest. Es wäre, wie schon erwähnt, korrekter, diese Tatsache stets dadurch zum Ausdruck zu bringen, daß man die unbestimmten Integrationskonstanten besonders mitschreibt, also statt

$$\int f(x)\,dx = F(x)$$

stets

$$\int f(x)\,dx = F(x) + c\,.$$

Aus Bequemlichkeitsgründen aber pflegt man von dieser umständlicheren Schreibweise abzusehen; man muß sich um so mehr der Unbestimmtheit, die mit der abgekürzten Schreibweise verbunden ist, stets bewußt bleiben (vgl. auch zweites Kapitel, S. 92).

Aus den Formeln der unbestimmten Integration folgen sofort bestimmte Integralformeln, wenn man gemäß dem zweiten Kapitel, § 4, Nr. 4, die Integrationsgrenzen einführt. Speziell ist

$$\int_a^b \frac{dx}{1+x^2} = \operatorname{arc}\operatorname{tg} x \Big|_a^b = \operatorname{arc}\operatorname{tg} b - \operatorname{arc}\operatorname{tg} a\,.$$

Setzen wir $a = 0$, $b = 1$ und beachten, daß $\operatorname{tg} 0 = 0$ und $\operatorname{tg}\frac{\pi}{4} = 1$ ist, so gewinnen wir die merkwürdige Formel

$$\frac{\pi}{4} = \int_0^1 \frac{1}{1+x^2}\,dx\,.$$

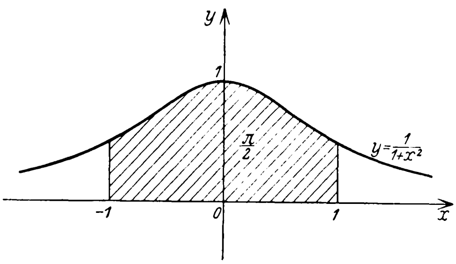



Fig. 54. $\frac{\pi}{2}$ als Flächeninhalt.

*Die Zahl $\pi$, die ursprünglich aus der Betrachtung des Kreises heraus entstand, wird durch diese Formel mit der rationalen Funktion $\frac{1}{1+x^2}$ in eine sehr einfache Verbindung gebracht und als Flächeninhalt gedeutet, welcher in der durch Figur 54 angegebenen Weise definiert ist.*

## § 4. Die Differentiation der zusammengesetzten Funktionen.

### 1. Die Kettenregel.

Die früheren Differentiationsregeln gestatten uns, jede Funktion zu differenzieren, die sich durch Funktionen mit bekannten Differentialquotienten linear gebrochen ausdrücken läßt. Wir können aber noch einen ganz wesentlichen Schritt weitergehen und alle diejenigen Funktionen differenzieren, welche sich durch *Zusammensetzen* aus Funktionen mit bekannten Differentialquotienten ableiten lassen. Es sei $\varphi(x)$ eine in einem Intervall $a \leqq x \leqq b$ differenzierbare Funktion, die dort alle Werte eines Intervalles $\alpha \leqq \varphi \leqq \beta$ annimmt. Dann wollen wir eine zweite differenzierbare Funktion $g(\varphi)$ der unabhängigen Veränderlichen $\varphi$ betrachten, bei der diese Veränderliche das letzte Intervall von $\alpha$ bis $\beta$ durchläuft, und nunmehr $g(\varphi) = g(\varphi(x)) = f(x)$ als Funktion von $x$ in dem Intervall $a \leqq x \leqq b$ auffassen. Es wird dann die Funktion $f(x) = g(\varphi(x))$ als eine aus den Funktionen $g$ und $\varphi$ *zusammengesetzte* Funktion der unabhängigen Veränderlichen $x$ im Intervalle $a \leqq x \leqq b$ bezeichnet.

Ist z. B. $\varphi(x) = 1 - x^2$ und $g(\varphi) = \sqrt{\varphi}$, so heißt diese zusammengesetzte Funktion einfach $f(x) = \sqrt{1 - x^2}$. Als Intervall $a \leqq x \leqq b$ werden wir hier das Intervall $0 \leqq x \leqq 1$ nehmen. Der Wertevorrat der Funktion $\varphi(x)$ füllt dabei ebenfalls das gesamte Intervall $0 \leqq \varphi \leqq 1$ aus; es ist also die zusammengesetzte Funktion $f(x) = \sqrt{1 - x^2}$ in dem Intervall $0 \leqq x \leqq 1$ definiert.

Ein anderes Beispiel für die Zusammensetzung von Funktionen ist die Funktion $f(x) = \sqrt{1 + x^2}$, wobei die Zusammensetzung durch die Gleichungen

$$\varphi(x) = 1 + x^2, \qquad g(\varphi) = \sqrt{\varphi}$$

gekennzeichnet wird und wobei der Funktionswert $\varphi(x)$ die Gesamtheit aller positiven Zahlen $\geqq 1$ durchläuft und dementsprechend die Funktion $f(x) = g(\varphi(x))$ für alle Werte von $x$ gebildet werden kann.

Bei der Zusammensetzung von Funktionen in dieser Art hat man natürlich immer zu beachten, daß man sich auf solche Intervalle $a \leqq x \leqq b$ beschränkt, bei welchen die zusammengesetzte Funktion in dem ganzen zugehörigen Intervall definiert ist. Z. B. existiert die zusammengesetzte Funktion $\sqrt{1 - x^2}$ nur in dem Gebiet $-1 \leqq x \leqq 1$ der unabhängigen Veränderlichen $x$, nicht aber in dem Gebiet $1 < x \leqq 2$, weil der Wertevorrat der Funktion $\varphi(x)$ dort aus negativen Zahlen besteht, für welche die Funktion $g(\varphi)$ nicht definiert ist.

Ebenso wie man zwei Funktionen miteinander zusammensetzen kann, können und müssen natürlich auch Funktionen betrachtet werden, bei denen ein solcher Zusammensetzungsprozeß öfter vorgenommen wird.

Z. B. die Funktion

$$\sqrt{1 + \operatorname{arc\,tg} x^2},$$

welche durch den Zusammensetzungsprozeß

$$\varphi(x) = x^2, \quad \psi(\varphi) = 1 + \operatorname{arc\,tg} \varphi, \quad g(\psi) = \sqrt{\psi(\varphi)} = f(x)$$

erklärt ist.

Für die Differentiation der zusammengesetzten Funktionen gilt nun folgender grundlegende Satz, die *Kettenregel der Differentialrechnung*: *Die Funktion $f(x) = g(\varphi(x))$ ist differenzierbar, und ihr Differentialquotient wird gegeben durch die Gleichung*

$$f'(x) = g'(\varphi) \cdot \varphi'(x),$$

oder in der Leibnizschen Bezeichnung

$$\frac{dy}{dx} = \frac{dy}{d\varphi} \cdot \frac{d\varphi}{dx}.$$

In Worten: *Man erhält die Ableitung der zusammengesetzten Funktion als das Produkt der Ableitungen der bei der Zusammensetzung benutzten Funktionen.* Der Beweis dieser Formel gelingt sehr einfach, wenn wir auf die Bedeutung des Differentialquotienten zurückgreifen. Zu beliebigem $\Delta x \neq 0$ und zugehörigen Werten von $\Delta\varphi$ und $\Delta g$ gibt es zwei mit $\Delta x$ gegen 0 strebende Größen $\varepsilon$ und $\eta$ mit

$$\Delta g = g'(\varphi)\,\Delta\varphi + \varepsilon\,\Delta\varphi \quad \text{und} \quad \Delta\varphi = \varphi'(x)\,\Delta x + \eta\,\Delta x;$$

man braucht nur $\eta$ aus der zweiten und im Falle $\Delta\varphi \neq 0$ auch $\varepsilon$ aus der ersten Gleichung auszurechnen, im Falle $\Delta\varphi = 0$ aber $\varepsilon = 0$ zu wählen. Setzen wir nun in die erste der beiden Gleichungen den Ausdruck für $\Delta\varphi$ aus der zweiten ein, so ergibt sich

$$\Delta g = g'(\varphi)\,\varphi'(x)\,\Delta x + (\eta\, g'(\varphi) + \varepsilon\,\varphi'(x) + \varepsilon\,\eta)\,\Delta x$$

oder

$$\frac{\Delta g}{\Delta x} = g'(\varphi)\,\varphi'(x) + (\eta\, g'(\varphi) + \varepsilon\,\varphi'(x) + \varepsilon\,\eta).$$

In dieser Gleichung aber können wir nunmehr $\Delta x$ gegen 0 streben lassen und erhalten sofort das behauptete Ergebnis, da die Klammer auf der rechten Seite für $\Delta x \to 0$ selbst gegen 0 strebt. Mithin besitzt auch die linke Seite unserer Gleichung einen Grenzwert $f'(x)$, und dieser ist gleich dem ersten Glied der rechten Seite, wie behauptet wurde[1]).

---

[1]) Wir hätten die Regel auch beweisen können, indem wir den Grenzübergang $\Delta x \to 0$ und den daraus folgenden $\Delta\varphi \to 0$ in der Beziehung $\frac{\Delta g}{\Delta x} = \frac{\Delta g}{\Delta\varphi} \cdot \frac{\Delta\varphi}{\Delta x}$ ausführen. Der im Text gegebene Weg ist jedoch vorzuziehen, weil er Sonderbetrachtungen für den Fall $\varphi'(x) = 0$ vermeidet.

Unsere Regel dehnt sich, wenn wir sie mehrfach hintereinander anwenden, unmittelbar auf den Fall aus, daß eine Funktion $f(x)$ durch *Zusammensetzung von mehr als zwei Funktionen* entsteht. Ist z. B.

$$y = g(u), \qquad u = \varphi(v), \qquad v = \psi(x),$$

so können wir $y = f(x)$ als Funktion von $x$ auffassen; ihre Differentiation vollzieht sich nach der Regel

$$\frac{dy}{dx} = y' = g'(u)\,\varphi'(v)\,\psi'(x) = \frac{dy}{du}\cdot\frac{du}{dv}\cdot\frac{dv}{dx},$$

und analog liegen die Dinge, wenn wir eine beliebige Anzahl von Funktionen miteinander zusammensetzen. Den Beweis dafür kann ich dem Leser selbst überlassen.

## 2. Beispiele.

Als einfachstes Beispiel betrachten wir die Funktion $y = x^\alpha$, wo $\alpha = \frac{p}{q}$ gesetzt ist und $q$ eine positive ganze Zahl, $p$ eine positive oder negative ganze Zahl, $\alpha$ also eine beliebige positive oder negative rationale Zahl bedeuten darf. Es sei $x > 0$. Nach der Kettenregel ergibt sich für

$$y = \varphi^p, \qquad \varphi = x^{\frac{1}{q}}$$

die Formel

$$y' = p\,\varphi^{p-1}\cdot\frac{1}{q}\,x^{\frac{1}{q}-1} = \frac{p}{q}\,x^{\frac{p}{q}-1},$$

so daß wir für beliebige rationale $\alpha$ die Differentiationsformel

$$\frac{d}{dx}x^\alpha = \alpha\,x^{\alpha-1}$$

erhalten, in Übereinstimmung mit der schon früher in Kap. II, § 3 auf anderem Wege gegebenen Ableitung.

Als zweites Beispiel betrachten wir

$$y = \sqrt{1-x^2} \qquad \text{oder} \qquad y = \sqrt{\varphi},$$

wo $\varphi = 1 - x^2$ und $-1 < x < 1$ ist. Die Kettenregel liefert

$$y' = \frac{1}{2\cdot\sqrt{\varphi}}\cdot(-2x) = -\frac{x}{\sqrt{1-x^2}}.$$

Weitere Beispiele sind durch die folgenden kleinen Rechnungen gegeben.

1. $$y = \arcsin\sqrt{1-x^2},$$

$$\frac{dy}{dx} = \frac{1}{\sqrt{1-(1-x^2)}}\cdot\frac{d\sqrt{1-x^2}}{dx} = \frac{1}{|x|}\cdot\frac{-x}{\sqrt{1-x^2}} = \mp\frac{1}{\sqrt{1-x^2}}.$$

2.
$$y=\sqrt{\frac{1+x}{1-x}},$$
$$\frac{dy}{dx}=\frac{1}{2\sqrt{\frac{1+x}{1-x}}}\cdot\frac{d\left(\frac{1+x}{1-x}\right)}{dx}=\frac{\sqrt{1-x}}{2\sqrt{1+x}}\cdot\frac{2}{(1-x)^2}$$
$$=\frac{1}{(1+x)^{\frac{1}{2}}(1-x)^{\frac{3}{2}}}.$$

Auch die Kettenregel der Differentialrechnung können wir in die Gestalt einer Integrationsformel setzen, entsprechend der Tatsache, daß jeder Differentiationsformel eine Integrationsformel als vollständiges Äquivalent entspricht. Wir wollen jedoch diese Formel an dieser Stelle übergehen, da wir sie zunächst noch nicht brauchen werden und ohnehin an einer späteren Stelle (viertes Kapitel, § 2) ausführlich behandeln müssen.

### 3. Nochmals Integration und Differentiation von $x^\alpha$ für irrationales $\alpha$.

Entsprechend der elementaren Definition der Potenz $x^\alpha$ für irrationales $\alpha$ durch die Gleichung
$$x^\alpha=\lim x^{r_n},$$
wo die Zahlen $r_n$ eine Folge rationaler Zahlen mit dem Limes $\alpha$ bedeuten, könnte man versucht sein, die Differentiation von $x^\alpha$ durch die direkte Ausführung des Grenzüberganges in der Differentiationsformel
$$\frac{d}{dx}x^{r_n}=r_n x^{r_n-1}$$
durchzuführen. Man müßte hierzu das Recht haben, aus der Beziehung $x^{r_n}\to x^\alpha$ auch auf die Gleichung $\frac{d}{dx}x^{r_n}\to\frac{d}{dx}x^\alpha$ zu schließen. Es steht aber einem solchen Grenzübergang ein schwerwiegendes prinzipielles Bedenken entgegen. Man kann nämlich in beliebig naher Nachbarschaft einer Kurve andere Kurven ziehen, deren Richtung von der Richtung der ursprünglichen Kurve an beliebigen Stellen beliebig viel abweicht; z. B. kann man eine gerade Linie durch einen beliebig nahe an ihr verlaufenden Wellenzug annähern, dessen Wellen gegenüber dieser geraden Linie immer wieder eine Steilheit von 45° erreichen (vgl. Fig. 55). Mit anderen Worten, es lehrt uns dieses Beispiel: Man darf aus der Tatsache, daß zwei Funktionen sich überall nur wenig voneinander unterscheiden, nicht ohne weiteres schließen, daß auch ihre

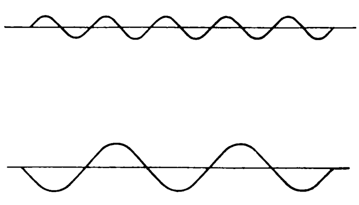


Fig. 55. Approximation einer Geraden durch Wellenlinien.

Differentialquotienten überall nahe aneinander liegen; und dieser Einwand verbietet uns, den so naheliegenden Grenzübergang ohne besondere Rechtfertigung auszuführen.

Ganz anders aber als bei dem Differentialquotienten liegen die Verhältnisse beim Integral. Hier haben wir im zweiten Kapitel, § 7, Nr. 2, erkannt, daß die Integrale zweier Funktionen, die sich in einem ganzen Intervall von $a$ bis $b$ überall um weniger als $\varepsilon$ unterscheiden, selbst um weniger als $\varepsilon(b-a)$ voneinander verschieden sind. Daraus haben wir an jener Stelle mit Hilfe des Fundamentalsatzes die Gültigkeit der Differentiationsformel

$$\frac{1}{\alpha+1}\frac{d}{dx}x^{\alpha+1} = x^{\alpha}$$

oder bei Ersetzung von $\alpha + 1$ durch $\alpha$

$$\frac{d}{dx}x^{\alpha} = \alpha x^{\alpha-1}$$

auch für irrationales $\alpha$ abgeleitet, so daß also auf diesem indirekten Wege schließlich doch die oben angegebene Beziehung $\frac{d}{dx}x^{r_n} \to \frac{d}{dx}x^{\alpha}$ ihre Bestätigung gefunden hat.

Jene Betrachtung war für uns ein charakteristisches Beispiel für das Ineinandergreifen von Differentialrechnung und Integralrechnung. Wir werden es jedoch aus prinzipiellen Gründen vorziehen, in § 6 an Stelle der elementaren Definition von $x^{\alpha}$ eine andere, innerlich einfachere zu setzen, welche uns von neuem und ohne Umwege zu dem gewonnenen Resultat hinführen wird.

## § 5. Maxima und Minima.

Nachdem wir zu einer gewissen Beherrschung des Problemes der Differentiation der elementaren Funktionen und der aus ihnen zusammengesetzten gelangt sind, bietet sich uns die Möglichkeit zu mannigfachen Anwendungen. Ich möchte die elementarste dieser Anwendungen, die Theorie der Maxima und Minima einer Funktion, im Zusammenhang mit einer geometrischen Diskussion der zweiten Ableitung, schon an dieser Stelle besprechen, um dann im nächsten Paragraphen den Faden der allgemeinen Theorie wieder aufzunehmen.

### 1. Allgemeine Vorbemerkungen über die geometrische Bedeutung der Differentialquotienten.

Nach Definition des Differentialquotienten $\frac{d}{dx}f(x)$ einer Funktion $f(x)$ gibt dieser die Steigung der Kurve $y = f(x)$ an. Diese Steigung können wir wiederum durch eine Kurve $y' = \frac{d}{dx}f(x) = f'(x)$ repräsentieren, die *Differentialkurve* zu der gegebenen. Die Steigung dieser letzteren

Kurve wird durch den Differentialquotienten $\frac{d}{dx} f'(x) = \frac{d^2}{dx^2} f(x) = f''(x)$, den zweiten Differentialquotienten von $f(x)$, gegeben usw. Ist der zweite Differentialquotient $f''(x)$ an einer Stelle $x$ — und infolge der hier vorausgesetzten Stetigkeit dann auch in einer gewissen Nachbarschaft dieser Stelle — positiv, so bedeutet dies, daß beim Passieren dieser Stelle in Richtung wachsender $x$-Werte der Differentialquotient der Kurve $y = f(x)$ zunimmt. Die Kurve wird also ihre konvexe Seite nach der Richtung abnehmender $y$-Werte zeigen. Umgekehrt ist es, wenn $f''(x)$

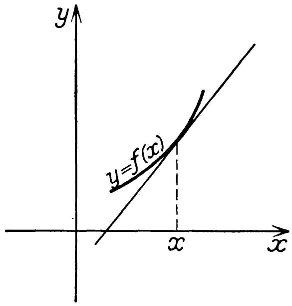



Fig. 56a. $f'' > 0$.

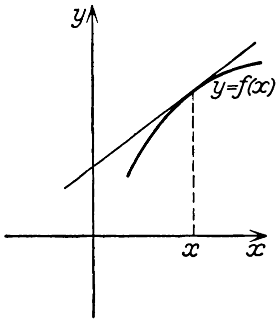



Fig. 56b. $f'' < 0$.

negativ ist. Im ersten Falle wird also die Kurve in der Umgebung der Stelle $x$ oberhalb ihrer Tangente, im zweiten Falle unterhalb ihrer Tangente verlaufen (vgl. Fig. 56a und b).

Eine besondere Beachtung verlangen nur die Stellen, wo $f''(x) = 0$ ist. Im allgemeinen wird beim Durchgang durch eine solche Stelle die zweite Ableitung $f''(x)$ ihr Vorzeichen wechseln. Dann wird ein solcher Punkt eine Übergangsstelle zwischen den beiden oben gezeichneten Typen sein; d. h. die Tangente in diesem Punkte wird auf der einen Seite oberhalb, auf der anderen Seite unterhalb verlaufen, also die Kurve nicht nur berühren, sondern zugleich durchschneiden (vgl. Fig. 57). Ein solcher Punkt heißt *Wendepunkt*, die zugehörige Tangente *Wendetangente* der Kurve. Das einfachste Beispiel gibt uns die Funktion $y = x^3$, die kubische Parabel, für welche im Punkt $x = 0$ die $x$-Achse selbst eine Wendetangente ist. Ein anderes Beispiel liefert die Funktion $f(x) = \sin x$, für welche $f'(x) = \frac{d}{dx} \sin x = \cos x$, $f''(x) = \frac{d^2}{dx^2} \sin x = - \sin x$ wird. Es ist hiernach $f'(0) = 1$ und

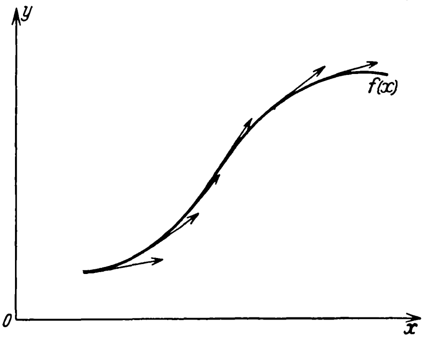



Fig. 57. Wendepunkt.

$f''(0) = 0$; die Sinuskurve hat also, da $f''(x)$ für $x = 0$ das Zeichen wechselt, im Nullpunkt eine um $45^0$ geneigte Wendetangente.

Man muß jedoch beachten, daß es auch Stellen geben kann, bei denen zwar $f''(x) = 0$ ist, bei denen jedoch die Tangente die Kurve nicht schneidet, sondern ganz auf einer Seite der Kurve bleibt. Dies gilt z. B. für die Kurve $y = x^4$, welche ganz oberhalb der $x$-Achse liegt und für welche ebenfalls für $x = 0$ die zweite Ableitung $f''(x)$ verschwindet.

## 2. Maxima und Minima.

Wir sagen, daß eine stetige Funktion oder eine Kurve $y = f(x)$ an einer Stelle $x = \xi$ ein *Maximum* bzw. ein *Minimum* besitzt, wenn wenigstens in irgend einer Umgebung dieser Stelle $x = \xi$ die Funktionswerte $f(x)$ für $x \neq \xi$ sämtlich kleiner bzw. größer als $f(\xi)$ sind. Unter einer *Umgebung* einer Stelle $\xi$ verstehen wir dabei ein Intervall $\alpha \leqq x \leqq \beta$, welches den Punkt $\xi$ im Innern enthält. Geometrisch gesprochen sind also solche Maxima und Minima Wellenberge bzw. Wellentäler der Kurve. Ein Blick auf Figur 58 zeigt uns, daß sehr wohl der Wert des Maximums an einer Stelle $P_5$ kleiner sein kann als der Wert des Minimums an einer anderen Stelle $P_2$; dem Begriff des Maximums und Minimums haftet also zunächst stets etwas Relatives, nämlich die Beziehung auf eine jeweils abzugrenzende Umgebung an.

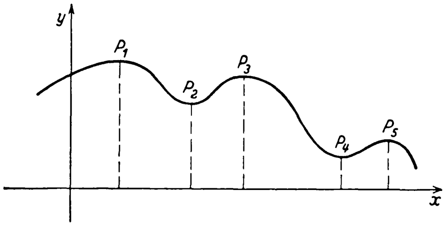


Fig. 58. Maxima und Minima.

Will man den wirklich kleinsten Wert oder den wirklich größten Wert der Funktion feststellen, so muß man noch besondere Betrachtungen anstellen, um zu entscheiden, ob und wie man ihn aus den Minima oder Maxima aussondern kann.

Für uns kommt es zunächst darauf an, die (relativen) Maxima oder Minima oder, wie wir mit einem neutralen Wort sagen wollen, die *relativen Extrema* (Extremum heißt äußerster Wert) einer Funktion oder Kurve zu finden. Diese Aufgabe, welche von der Geometrie, der Mechanik und der Physik sehr häufig gestellt wird und welche auch sonst in zahlreichen Anwendungen immer wieder auftritt, hat einen der ersten Anlässe zur Ausbildung der Differential- und Integralrechnung im 17. Jahrhundert gebildet.

Man erkennt sofort, daß bei vorausgesetzter Differenzierbarkeit für eine solche Extremumstelle $\xi$ die Tangente an die Kurve horizontal verlaufen muß. Es ist also die Bedingung

$$f'(\xi) = 0$$

eine notwendige Bedingung für das Vorliegen eines Extremums;

durch Auflösung dieser Gleichung nach der Unbekannten $\xi$ erhalten wir die Stellen, für welche ein solches Extremum eintreten kann. Unsere Bedingung ist aber keineswegs etwa eine hinreichende Bedingung für ein Extremum; es kann nämlich Stellen geben, für welche die Ableitung verschwindet, d. h. die Tangente horizontal verläuft, für welche die Kurve aber weder ein Maximum noch ein Minimum besitzt. Dieser Fall tritt dann ein, wenn die Kurve an der betreffenden Stelle eine horizontale sie durchschneidende Wendetangente besitzt, wie in dem obigen Beispiel der Funktion $y = x^3$ an der Stelle $x = 0$.

Hat man aber einmal eine Stelle $\xi$ gefunden, für welche $f'(\xi)$ verschwindet, so kann man sofort schließen, daß die Stelle eine Maximumstelle ist, wenn $f''(\xi) < 0$ wird, eine Minimumstelle, wenn $f''(\xi) > 0$ ist; denn im ersten Falle liegt in der Umgebung dieses Punktes die Kurve ganz unterhalb der Tangente, im zweiten ganz oberhalb.

Statt uns bei der Ableitung unserer notwendigen Bedingungen auf die Anschauung zu berufen, hätten wir natürlich leicht auch einen rein analytischen Beweis geben können (vgl. hierzu die ganz entsprechenden Betrachtungen zum Rolleschen Satz auf S. 83). Ist die Stelle $\xi$ eine Stelle des Maximums, so muß bei hinreichend kleinem von 0 verschiedenem $|h|$ der Ausdruck $f(\xi) - f(\xi + h)$ positiv sein. Also wird der Quotient $\frac{f(\xi + h) - f(\xi)}{h}$ positiv oder negativ, je nachdem $h$ negativ oder positiv ist. Wenn $h$ von negativen Werten her gegen 0 strebt, kann also der Grenzwert dieses Quotienten nicht negativ sein, während er, wenn $h$ von positiven Werten her gegen 0 strebt, nicht positiv sein kann. Da beide Grenzwerte wegen der vorausgesetzten Existenz der Ableitung einander gleich und zwar gleich $f'(\xi)$ sein müssen, so können sie nur den Wert Null haben; d. h. es muß $f'(\xi) = 0$ sein. Eine entsprechende Schlußweise gilt für ein Minimum.

Wir können übrigens notwendige und hinreichende Bedingungen für das Eintreten eines Maximums oder Minimums auch formulieren und analytisch beweisen, ohne die zweite Ableitung heranzuziehen. *Es besteht an einer Stelle $\xi$ dann und nur dann ein Minimum oder ein Maximum, wenn die Ableitung $f'(x)$ beim Durchgang durch diese Stelle ihr Vorzeichen wechselt; und zwar liegt ein Minimum vor, wenn die Ableitung links von $\xi$ negativ, rechts positiv ist, andernfalls ein Maximum.*

Zum Beweise bilden wir unter Benutzung des Mittelwertsatzes $f(\xi + h) - f(\xi) = h f'(\xi + \vartheta h)$. Ein Extremum liegt dann und nur dann vor, wenn $f(\xi + h) - f(\xi)$ bei hinreichend klein gewähltem $|h|$ dasselbe Vorzeichen hat, gleichviel ob $h$ positiv oder negativ ist; und hieraus ergibt sich unmittelbar die Richtigkeit unserer Behauptung.

Gleichzeitig erkennen wir, daß tatsächlich der Funktionswert $f(\xi)$ den kleinsten bzw. den größten Wert in jedem den Punkt $\xi$ enthaltenden Intervalle darstellt, in welchem kein weiterer Zeichenwechsel der Ableitung mehr eintritt.

Da der Mittelwertsatz, auf den allein wir diese Betrachtung gestützt haben, auch dann noch anwendbar bleibt, wenn $f(x)$ an der Stelle $\xi$ selbst nicht mehr differenzierbar ist, wohl aber an allen andern betrachteten Stellen, wenn also z. B. die Kurve $y = f(x)$ für $x = \xi$ eine Ecke oder Spitze hat, so sind wir auf diese Weise noch zu einem etwas allgemeineren Resultat gelangt: Wenn die Funktion $f(x)$ in einem die Stelle $\xi$ enthaltenden Intervalle stetig und, höchstens mit Ausnahme dieser Stelle, differenzierbar ist, so hat sie an dieser Stelle dann und nur dann ein Extremum, wenn dort zwei Intervalle mit verschiedenem Vorzeichen von $f'(x)$ getrennt werden. Beispielsweise hat die Funktion $y = |x|$ an der Stelle $x = 0$ ein Minimum, weil $y' > 0$ für $x > 0$ und $y' < 0$ für $x < 0$ gilt (vgl. Fig. 34, S. 76). Ebenso hat die Funktion $y = \sqrt[3]{x^2}$ an dieser Stelle ein Minimum, obwohl dort der Differentialquotient $\frac{2}{3} x^{-\frac{1}{3}}$ unendlich wird (vgl. Fig. 37, S. 78).

Im übrigen möchte ich zur allgemeinen Theorie der Maxima und Minima noch folgende Bemerkung machen: Die Aufsuchung der Maxima und Minima ist nicht ohne weiteres gleichbedeutend mit der Aufsuchung der größten und kleinsten Funktionswerte in einem abgeschlossenen Intervalle. Bei einer monotonen Funktion werden diese größten und kleinsten Werte am Rande des Intervalles angenommen und sind dann keine Maxima und Minima in unserem Sinne mehr — denn dieser Begriff bezog sich immer auf eine volle Umgebung der betreffenden Stelle. So hat z. B. die Funktion $f(x) = x$ im Intervall $0 \leqq x \leqq 1$ an der Stelle 0 ihren kleinsten und an der Stelle 1 ihren größten Wert, und Entsprechendes gilt für jede monotone Funktion. Die Funktion $y = \operatorname{arc\,tg} x$ mit der Ableitung $y' = \frac{1}{1+x^2}$ ist für $-\infty < x < \infty$ monoton und besitzt in diesem offenen Intervall weder ein Maximum oder Minimum noch einen größten oder kleinsten Wert.

Will man nach Aufsuchung der Nullstellen von $f'(x)$ sicher gehen, daß man damit auch die Stellen eines kleinsten oder größten Wertes gefunden hat, so wird man sich oft des folgenden Kriteriums bedienen: *Eine Nullstelle $\xi$ von $f'(x)$ gibt sicherlich dann für ein ganzes Intervall den kleinsten oder größten Wert von $f(x)$, wenn in dem ganzen Intervall $f''(x) > 0$ bzw. $f''(x) < 0$ ist.* Es ist nämlich dann wegen des Mittelwertsatzes, wenn außer $\xi$ auch noch $\xi + h$ zum Intervall gehört,

$$f'(\xi + h) = f'(\xi + h) - f'(\xi) = h f''(\xi + \vartheta h).$$

Es hat also die Ableitung $f'(x)$ an der Stelle $x = \xi + h$ das Vorzeichen von $h$ oder das entgegengesetzte, je nachdem $f''(x) > 0$ oder $f''(x) < 0$ vorausgesetzt wurde; hieraus aber folgt die Behauptung nach der Bemerkung am Schluß von S. 129.

### 3. Beispiele für Maxima und Minima.

1. Aufgabe. Unter allen Rechtecken gegebenen Flächeninhaltes ist eines mit kleinstem Umfang zu suchen.

Der Flächeninhalt sei $a^2$, die eine Seite $x$ (wobei wir für $x$ das Intervall $0 < x < \infty$ zu betrachten haben), dann ist die andere Seite $\frac{a^2}{x}$, und der halbe Umfang wird gegeben durch

$$f(x) = x + \frac{a^2}{x}.$$

Es wird

$$f'(x) = 1 - \frac{a^2}{x^2}, \qquad f''(x) = \frac{2a^2}{x^3}.$$

Die Gleichung $f'(\xi) = 0$ hat die einzige positive Wurzel $\xi = a$. Für diese wird $f''(x)$ positiv (ebenso wie für jedes positive $x$); sie liefert also den gesuchten kleinsten Wert, und wir erhalten das sehr plausible Resultat, daß von allen betrachteten Rechtecken das Quadrat den kleinsten Umfang besitzt.

2. Aufgabe. Unter allen Dreiecken mit gegebener Grundlinie und gegebenem Flächeninhalt ist dasjenige von kleinstem Umfang zu bestimmen.

Um diese Aufgabe zu lösen, machen wir die gegebene Grundlinie $AB$ zu einem Teil der $x$-Achse eines Koordinatensystems, den Mittelpunkt $O$ von $AB$ zum Nullpunkt. Ist $C$ die Spitze des betreffenden Dreiecks, $h$ seine — fest gegebene — Höhe und $x, h$ die Koordinaten der Spitze, so wird die Summe der beiden zu bestimmenden Seiten $AC$ und $BC$ des Dreiecks offenbar, wenn die Länge der gegebenen Grundlinie $2a$ ist, durch

$$f(x) = \sqrt{(x+a)^2 + h^2} + \sqrt{(x-a)^2 + h^2}$$

gegeben, und wir erhalten weiter:

$$f'(x) = \frac{x+a}{\sqrt{(x+a)^2+h^2}} + \frac{x-a}{\sqrt{(x-a)^2+h^2}},$$

$$f''(x) = \frac{-(x+a)^2}{\left(\sqrt{(x+a)^2+h^2}\right)^3} + \frac{1}{\sqrt{(x+a)^2+h^2}} + \frac{-(x-a)^2}{\left(\sqrt{(x-a)^2+h^2}\right)^3} + \frac{1}{\sqrt{(x-a)^2+h^2}}$$

$$= \frac{h^2}{\left(\sqrt{(x+a)^2+h^2}\right)^3} + \frac{h^2}{\left(\sqrt{(x-a)^2+h^2}\right)^3}.$$

Man erkennt sofort, daß erstens $f'(0)$ verschwindet, daß zweitens $f''(x)$ stets positiv ist; daher wird die Stelle $x = 0$ tatsächlich ein Minimum liefern; da $f''(x) > 0$ ist, nimmt $f'(x)$ immer zu, und es kann also an keiner andern Stelle $f'(x) = 0$ sein, so daß die Stelle $x = 0$ tatsächlich den kleinsten Wert geben muß. Dieser kleinste Wert wird also durch das gleichschenkelige Dreieck geliefert.

Ganz ähnlich erkennt man, daß unter allen Dreiecken gegebenen Umfangs und gegebener Grundlinie das gleichschenkelige den größten Flächeninhalt besitzt.

3. Aufgabe. Kleinste Entfernungssumme eines Punktes einer Geraden von zwei gegebenen festen Punkten.

Es seien eine gerade Linie und zwei feste Punkte $A$, $B$ auf derselben Seite der Geraden gegeben. Auf der geraden Linie ist ein Punkt $P$ gesucht, so daß die Entfernungssumme $PA + PB$ möglichst klein wird.

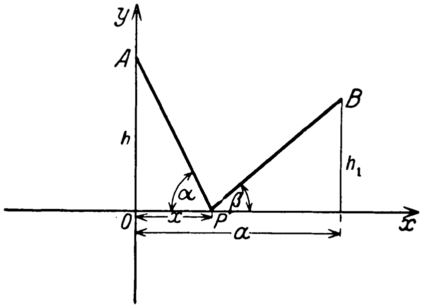



Fig. 59. Reflexionsgesetz.

Wir machen die gegebene Gerade zur $x$-Achse eines Koordinatensystems und wählen die Bezeichnungen von Figur 59. Dann wird die gesuchte Entfernungssumme gegeben durch

$$f(x) = \sqrt{x^2 + h^2} + \sqrt{(x-a)^2 + h_1^2},$$

und wir erhalten weiter

$$f'(x) = \frac{x}{\sqrt{x^2+h^2}} + \frac{x-a}{\sqrt{(x-a)^2+h_1^2}},$$

$$f''(x) = \frac{-x^2}{(\sqrt{x^2+h^2})^3} + \frac{1}{\sqrt{x^2+h^2}} + \frac{-(x-a)^2}{(\sqrt{(x-a)^2+h_1^2})^3} + \frac{1}{\sqrt{(x-a)^2+h_1^2}}$$

$$= \frac{h^2}{(\sqrt{x^2+h^2})^3} + \frac{h_1^2}{(\sqrt{(x-a)^2+h_1^2})^3}.$$

Die Gleichung $f'(\xi) = 0$ liefert uns also

$$\frac{\xi}{\sqrt{\xi^2+h^2}} = \frac{a-\xi}{\sqrt{(\xi-a)^2+h_1^2}},$$

$$\cos\alpha = \cos\beta,$$

und dies bedeutet, daß die beiden Geraden $PA$ und $PB$ mit der gegebenen Geraden gleiche Winkel bilden müssen. Daß wir damit tatsächlich den kleinsten Wert erhalten, zeigt uns das positive Vorzeichen von $f''(x)$.

Die Lösung dieser Aufgabe steht in engstem Zusammenhang mit dem Spiegelungsgesetz der Optik. Nach einem wichtigen Prinzip der Optik, dem sog. *Fermatschen Prinzip der kürzesten Lichtzeit*, wird die Bahn eines Lichtstrahls durch die Eigenschaft charakterisiert, daß die Zeit, die das Licht braucht, um von einer Stelle $A$ zu einer Stelle $B$ unter gewissen Bedingungen zu kommen, möglichst kurz sein muß. Wird dem Lichtstrahl die Bedingung auferlegt, auf seinem Wege von $A$ nach $B$ einen Punkt der gegebenen Geraden (etwa eines Spiegels) zu passieren, so sehen wir, daß die kürzeste Lichtzeit von einem solchen Strahl erreicht wird, bei dem „Einfallswinkel" gleich „Ausfallswinkel" ist.

4. Aufgabe. Brechungsgesetz.

Es seien zwei Punkte $A$ und $B$ auf verschiedenen Seiten der $x$-Achse gegeben. Welcher Weg führt in möglichst kurzer Zeit von $A$ nach $B$, wenn die Geschwindigkeit auf der einen Seite der $x$-Achse $c_1$, die Geschwindigkeit auf der anderen Seite der $x$-Achse $c_2$ ist?

Es ist klar, daß dieser „kürzeste" Weg aus zwei geradlinigen Strecken bestehen muß, welche in einem Punkte $P$ der $x$-Achse zusammenstoßen. Mit den Bezeichnungen von Figur 60 erhält man für die Länge der Strecken $PA$ und $PB$ die beiden Ausdrücke

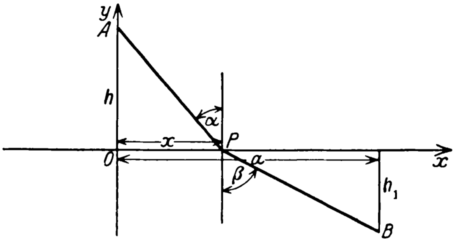


Fig. 60. Brechungsgesetz.

$\sqrt{h^2+x^2}$ bzw. $\sqrt{h_1^2+(a-x)^2}$, und wir erhalten die Zeit, die zu dem Wege benötigt wird, indem wir die beiden durchmessenen Wege jeweils durch die entsprechende Geschwindigkeit dividieren. Die für den Weg gebrauchte Zeit ergibt sich also als

$$f(x) = \frac{1}{c_1}\sqrt{h^2+x^2} + \frac{1}{c_2}\sqrt{h_1^2+(a-x)^2}.$$

Durch Differentiation erhalten wir hieraus:

$$f'(x) = \frac{1}{c_1}\frac{x}{\sqrt{h^2+x^2}} - \frac{1}{c_2}\frac{a-x}{\sqrt{h_1^2+(a-x)^2}},$$

$$f''(x) = \frac{1}{c_1}\frac{\sqrt{h^2+x^2}-\frac{x^2}{\sqrt{h^2+x^2}}}{(\sqrt{h^2+x^2})^2} + \frac{1}{c_2}\frac{\sqrt{h_1^2+(a-x)^2}-\frac{(a-x)^2}{\sqrt{h_1^2+(a-x)^2}}}{(\sqrt{h_1^2+(a-x)^2})^2}$$

$$= \frac{1}{c_1}\frac{h^2}{(\sqrt{h^2+x^2})^3} + \frac{1}{c_2}\frac{h_1^2}{(\sqrt{h_1^2+(a-x)^2})^3}.$$

Die Gleichung $f'(x) = 0$, d. h. $\frac{1}{c_1}\frac{x}{\sqrt{h^2+x^2}} = \frac{1}{c_2}\frac{a-x}{\sqrt{h_1^2+(a-x)^2}}$ ist, wie man aus der Figur leicht abliest, gleichbedeutend mit der Bedingung $\frac{1}{c_1}\sin\alpha = \frac{1}{c_2}\sin\beta$ oder

$$\frac{\sin\alpha}{\sin\beta} = \frac{c_1}{c_2}.$$

Ich kann es dem Leser überlassen, zu beweisen, daß es nur eine einzige Stelle gibt, welche dieser Bedingung genügt, und daß diese Stelle wirklich den kleinsten Wert liefert. Die physikalische Bedeutung unseres Beispieles ergibt sich wiederum aus dem optischen Prinzip der kürzesten Lichtzeit. Ein Lichtstrahl beschreibt zwischen zwei Punkten diejenige Bahn, für die er die kürzeste Zeit braucht. Bedeuten $c_1$ und $c_2$ die

Lichtgeschwindigkeit auf beiden Seiten einer Trennungsebene zweier optischer Medien, so wird dieser Lichtweg gemäß unserem Ergebnisse verlaufen; unser Ergebnis liefert also das *Snelliussche Brechungsgesetz.*

## § 6. Logarithmus und Exponentialfunktion.

In dem systematischen Zusammenhange der Integral- und Differentialrechnung ergibt sich ganz von selbst ein bequemer Zugang zu der Exponentialfunktion und dem Logarithmus. Wenn wir diese Funktionen auch schon früher behandelt haben, so wollen wir sie nunmehr, ohne von den früheren Definitionen und den auf ihnen beruhenden Überlegungen Gebrauch zu machen, hier von neuem definieren und ihre Theorie entwickeln. Wir gehen dabei von dem Logarithmus aus und gewinnen erst aus ihm durch Umkehrung die Exponentialfunktion.

### 1. Definition des Logarithmus. Differentiationsformel.

Wir haben gesehen, daß das unbestimmte Integral der Potenz $x^n$ für ganzzahliges $n$ im allgemeinen wieder zu einer Potenz führt; die einzige Ausnahme fanden wir bei der Funktion $\frac{1}{x}$, welche nicht als Differentialquotient einer der von uns bisher behandelten Funktionen erschien. Es liegt daher nahe, zu vermuten, daß das unbestimmte Integral der Funktion $\frac{1}{x}$ eine neuartige Funktion darstellen wird, und dieser Vermutung nachgehend, wollen wir im Folgenden die Funktion

$$y = \int_1^x \frac{d\xi}{\xi} = f(x)$$

für $x > 0$ untersuchen. Wir *bezeichnen* sie als den *Logarithmus von* $x$, genauer als den *natürlichen* Logarithmus von $x$, und schreiben sie $y = \log x$ oder auch $y = \log \operatorname{nat} x$. Die Integrationsvariable haben wir, um eine Verwechslung mit der oberen Grenze $x$ auszuschließen, mit $\xi$ bezeichnet.

Daß wir als untere Grenze des Integrals die Zahl 1 genommen haben, ist eine zunächst willkürliche Festsetzung, die sich jedoch bald als zweckmäßig erweisen wird.

Daß die hier definierte Logarithmusfunktion mit dem früher auf „elementare Art“ definierten Logarithmus übereinstimmt, wird sich im Laufe der folgenden Betrachtung ganz von selbst ergeben. Die Ergebnisse der folgenden Überlegungen sind aber, worauf ich noch einmal hinweise, von den früher gewonnenen unabhängig.

Geometrisch bedeutet unsere Logarithmusfunktion den in Figur 61 schraffiert gezeichneten Flächeninhalt, welcher von der gleichseitigen Hyperbel $y = \frac{1}{\xi}$ und der $\xi$-Achse einerseits und den Geraden $\xi = 1$ und $\xi = x$ andererseits begrenzt ist. Er ist dabei positiv zu rechnen, wenn $x > 1$, negativ, wenn $x < 1$ wird. Für $x = 1$ verschwindet der Flächeninhalt, und wir haben daher $\log 1 = 0$.

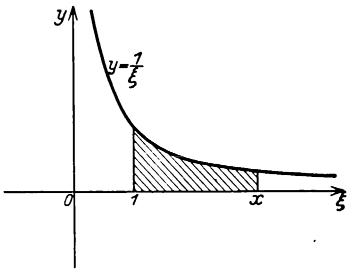


Fig. 61. $\log x$ als Flächeninhalt.

Gemäß der obigen Definition wird die Differentiation des Logarithmus durch die Formel

$$\frac{d \log x}{dx} = \frac{1}{x}$$

gegeben.

Es sei ausdrücklich hervorgehoben, daß wir für das Argument $x$ stets voraussetzen, es sei positiv; den Logarithmus von 0 oder gar von negativen Werten gemäß unserer obigen Formel zu bilden, verhindert uns die Tatsache, daß der Integrand $\frac{1}{\xi}$ für $\xi = 0$ unendlich wird. Dagegen kann man natürlich, wenn man nur als untere Grenze eine negative Zahl, etwa $-1$ wählt, das Integral mit einer negativen oberen Grenze $x$ bilden, d. h. den Ausdruck

$$\int_{-1}^{x} \frac{d\xi}{\xi} \qquad (x < 0)$$

betrachten. Auf Grund der Bedeutung des Integrals als Grenzwert einer Summe oder als Flächeninhalt erkennt man, daß für $x < 0$

$$\int_{-1}^{x} \frac{d\xi}{\xi} = \int_{1}^{-x} \frac{d\xi}{\xi} = \int_{1}^{|x|} \frac{d\xi}{\xi} = \log|x|$$

ist. Wir können demgemäß allgemein als *Formel der unbestimmten Integration*

$$\int \frac{dx}{x} = \log|x|$$

schreiben.

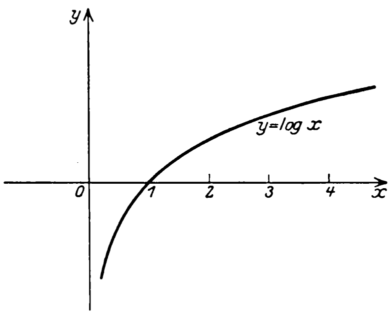


Fig. 62.

Den Logarithmus wird man sich natürlich auch durch eine Kurve veranschaulichen können. Diese Kurve, *die Logarithmuskurve*, ist in Figur 62 gezeichnet. Wie man zu ihrer graphischen Konstruktion gelangen kann, ergibt sich aus den Ausführungen im zweiten Kapitel, § 5.

## 2. Das Additionstheorem.

Der so definierte Logarithmus genügt dem folgenden fundamentalen Gesetz:

$$\log(a\,b) = \log a + \log b.$$

Der Beweis dieses *Additionstheoremes* ergibt sich sehr einfach aus der Integraldefinition und ihrer geometrischen Bedeutung. Es ist nämlich

$$\log(a\,b) = \int\limits_1^{ab} \frac{d\xi}{\xi} = \int\limits_1^{a} \frac{d\xi}{\xi} + \int\limits_a^{ab} \frac{d\xi}{\xi},$$

und es braucht daher nur noch gezeigt zu werden, daß

$$\int\limits_a^{ab} \frac{d\xi}{\xi} = \int\limits_1^{b} \frac{d\xi}{\xi}$$

ist.

Wir können schreiben:

$$\int\limits_1^{b} \frac{d\xi}{\xi} = \lim_{n\to\infty} \sum_{\nu=0}^{n-1} \frac{\Delta\xi_\nu}{\xi_\nu},$$

wobei

$$\xi_\nu = 1 + \nu h \quad \text{mit} \quad h = \frac{b-1}{n} \quad \text{und} \quad \Delta\xi_\nu = \xi_{\nu+1} - \xi_\nu$$

ist. Setzen wir die Summe rechts in die Form

$$\sum_{\nu=0}^{n-1} \frac{a\,\Delta\xi_\nu}{a\,\xi_\nu} = \sum_{\nu=0}^{n-1} \frac{\Delta\eta_\nu}{\eta_\nu}$$

mit

$$\eta_\nu = a\,\xi_\nu, \quad \text{d. h.} \quad \eta_\nu = a + \nu\frac{a\,b - a}{n}, \quad \text{und} \quad \Delta\eta_\nu = \eta_{\nu+1} - \eta_\nu,$$

so erkennen wir sofort, daß sie für $n \to \infty$ in das Integral $\int\limits_a^{ab} \frac{d\xi}{\xi}$ übergeht, und damit ist die behauptete Gleichheit der Integrale bewiesen, welche eine bemerkenswerte geometrische Tatsache über die von der Hyperbel $y = \frac{1}{x}$ begrenzten Flächeninhalte ausspricht.

Aus dem Additionstheorem des Logarithmus folgt nun für beliebige positive Zahlen $a_1, a_2, \ldots, a_n$ die Gleichung

$$\log(a_1\,a_2 \cdots a_n) = \log a_1 + \log a_2 + \cdots + \log a_n.$$

Sind speziell alle Zahlen $a_1, a_2, \ldots, a_n$ gleich einer und derselben Zahl $a$, so ergibt sich

$$\log a^n = n \log a.$$

Ebenso folgt

$$\log a + \log\frac{1}{a} = \log 1 = 0.$$

Es ist also
$$\log a = -\log \frac{1}{a}.$$
Setzen wir weiter $\sqrt[n]{a} = \alpha$, so folgt $\log a = n \log \alpha$ oder
$$\log \sqrt[n]{a} = \log a^{\frac{1}{n}} = \frac{1}{n} \log a.$$

Hieraus aber folgt durch nochmalige Anwendung des Additionstheoremes, wenn $m$ eine natürliche Zahl ist,
$$\frac{m}{n} \log a = \log \sqrt[n]{a^m} = \log a^{\frac{m}{n}}.$$

Die Gleichung
$$\log a^r = r \log a$$
ist somit für alle positiven rationalen Werte von $r$ bewiesen und offenbar auch für $r = 0$ richtig. Für negative rationale $r$ gilt sie gleichfalls; denn es wird
$$\log a^r = \log \frac{1}{a^{-r}} = -\log a^{-r} = r \log a.$$

### 3. Monotoner Charakter und Wertevorrat des Logarithmus.

Offenbar wächst der Wert der Funktion $\log x$, sobald $x$ wächst, und nimmt entsprechend ab, sobald $x$ abnimmt: der Logarithmus ist eine monotone Funktion.

Da die Ableitung $\frac{1}{x}$ bei wachsendem $x$ immer kleiner wird, wird das Anwachsen der Funktion mit wachsendem $x$ immer schwächer vor sich gehen. Trotzdem aber nähert sich die Funktion $\log x$ für unbegrenzt wachsendes $x$ nicht etwa einem bestimmten positiven Grenzwert, sondern sie wird unendlich, d. h. zu jeder noch so großen positiven Zahl $A$ gibt es Werte von $x$, so daß $\log x > A$ wird. Diese Tatsache folgt sehr leicht aus dem Additionstheorem. Es ist nämlich $\log 2^n = n \log 2$. Da $\log 2$ eine positive Zahl ist, so wird für $x = 2^n$ bei hinreichend großem $n$ die Funktion $\log x$ offenbar positiv beliebig groß werden.

Da $\log \frac{1}{2^n} = -n \log 2$ ist, so erkennen wir, daß $\log x$, sobald $x$ von positiven Werten her gegen Null strebt, in negativem Sinne über alle Grenzen wächst.

Zusammenfassend können wir sagen:

Die Funktion $\log x$ ist eine monotone Funktion, welche alle Werte zwischen $-\infty$ und $+\infty$ annimmt, wenn die unabhängige Veränderliche $x$ das Kontinuum der positiven Zahlen durchwandert.

### 4. Die Umkehrfunktion des Logarithmus (Exponentialfunktion).

Da die Funktion $y = \log x$ $(x > 0)$ eine monotone Funktion von $x$ ist, die sämtliche reellen Werte annimmt, so ist die Umkehrfunktion, die wir zunächst mit $x = E(y)$ bezeichnen, eine eindeutige monotone, für jeden Wert von $y$ definierte Funktion; sie ist differenzierbar, weil $\log x$ selbst differenzierbar ist (s. § 3, S. 115). Indem wir die Bezeichnung der abhängigen und unabhängigen Veränderlichen vertauschen, wollen wir diese Funktion $E(x)$ näher studieren. Zunächst erkennen wir, daß sie für jeden Wert von $x$ positiv sein muß. Ferner sehen wir, daß

$$E(0) = 1$$

sein muß; denn diese Gleichung ist gleichbedeutend damit, daß der Logarithmus von 1 den Wert 0 hat.

Zweitens folgt aus dem Additionstheorem für den Logarithmus sofort das „*Multiplikationstheorem*"

$$E(\alpha)\,E(\beta) = E(\alpha + \beta).$$

Zum Beweise brauchen wir nur zu beachten, daß die Gleichungen

$$E(\alpha) = a, \quad E(\beta) = b, \quad E(\alpha + \beta) = c$$

gleichbedeutend sind mit

$$\alpha = \log a, \quad \beta = \log b, \quad \alpha + \beta = \log c.$$

Da (nach dem Additionstheorem für den Logarithmus) $\alpha + \beta = \log ab$ ist, so muß $c = ab$ sein, womit wir das Multiplikationstheorem bewiesen haben.

Aus diesem Theorem ergibt sich eine Grundeigenschaft der Funktion $y = E(x)$, welcher wir die Berechtigung entnehmen werden, unsere Funktion als *Exponentialfunktion* zu bezeichnen und symbolisch in der Form

$$y = e^x$$

zu schreiben. Um zu dieser Eigenschaft zu gelangen, bedenken wir, daß es eine Zahl — wir wollen sie $e$ nennen[1]) — geben muß, für welche

$$\log e = 1$$

ist. Gleichbedeutend damit ist die Definition

$$E(1) = e.$$

Wegen des Multiplikationstheorems der Funktion $E(x)$ schließt man nun analog wie oben beim Logarithmus für jedes ganzzahlige positive $n$

$$E(n) = e^n$$

und ebenso bei ganzzahligen positiven $n$ und $m$

$$E\left(\frac{m}{n}\right) = e^{\frac{m}{n}},$$

was wir übrigens auch direkt aus dem Additionstheorem des Logarithmus hätten entnehmen können.

[1]) Ihre Identität mit der im ersten Kapitel betrachteten Zahl $e$ werden wir in Nr. 6 dieses Paragraphen nachweisen.

Die so für positive rationale Zahlen $r$ bewiesene Gleichung $E(r) = e^r$ gilt wegen der Gleichung

$$E(r)\,E(-r) = E(0) = 1$$

auch für negative rationale Zahlen $r$.

Es ist also die Funktion $E(x)$ eine für alle Werte von $x$ stetige Funktion, welche für rationale Werte von $x$ mit $e^x$ übereinstimmt. Diese Tatsache rechtfertigt es, auch für beliebige irrationale Werte von $x$ unsere Funktion mit $e^x$ zu bezeichnen[1]. (Man beachte dabei, daß die Stetigkeit der Funktion $e^x$ bei dieser Definition als Umkehrfunktion einer stetigen monotonen Funktion von vornherein feststeht, während sie bei der elementaren Definition bewiesen werden muß.)

Die Differentiation der Exponentialfunktion wird durch die Formel

$$\frac{d}{dx} e^x = e^x \quad \text{oder} \quad y' = y$$

geleistet. *Diese Formel drückt die wichtige Tatsache aus, daß die Exponentialfunktion bei der Differentiation sich reproduziert.* Ihr Beweis ist äußerst einfach. Es ist nämlich $x = \log y$, woraus mit Rücksicht auf die Differentiationsformel für den Logarithmus folgt $\frac{dx}{dy} = \frac{1}{y}$, also nach der Regel für inverse Funktionen

$$\frac{dy}{dx} = y = e^x,$$

wie behauptet war.

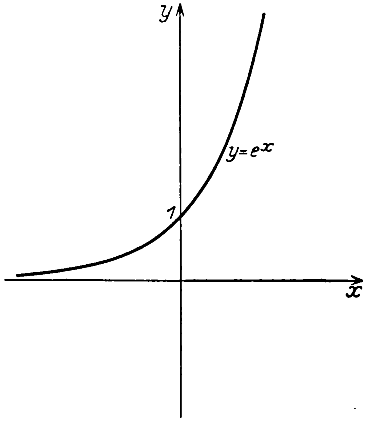


Fig. 63. Die Exponentialfunktion.

Die Kurve, durch welche man die Exponentialfunktion $e^x$ darstellt, die sog. Exponentialkurve, ergibt sich einfach durch Spiegelung der Logarithmuskurve an der Winkelhalbierenden des positiven Quadranten. Sie ist in Figur 63 gezeichnet.

## 5. Die allgemeine Exponentialfunktion $a^x$ und die allgemeine Potenz $x^\alpha$.

Die *Exponentialfunktion* $a^x$ für eine beliebige positive Grundzahl $a$ *definieren* wir jetzt einfach durch die Gleichung

$$y = a^x = e^{x \log a},$$

[1]) Nimmt man das aus Nr. 6 hervorgehende Ergebnis der Identität unserer Zahl $e$ mit der früher so bezeichneten Zahl vorweg, so ist nunmehr gezeigt, daß unsere jetzt gegebene Definition dieselbe Exponentialfunktion mit der Basis $e$ liefert, welche früher elementar durch Potenzieren definiert wurde. Denn nach jener elementaren Definition hat man für einen irrationalen Wert von $x$ die Funktion $e^x$ durch den Grenzwert der Ausdrücke $e^{x_n}$ erklärt, wobei $x_n$ eine Folge von rationalen Zahlen mit dem Grenzwert $x$ durchläuft.

was wegen

$$e^{\log a} = a$$

mit der elementaren Definition übereinstimmt. Es ergibt sich nun ohne weiteres mit Hilfe der Kettenregel

$$\frac{d}{dx} a^x = \frac{d}{dx} e^{x \log a} = e^{x \log a} \cdot \log a,$$

$$\frac{d}{dx} a^x = a^x \log a.$$

Die Umkehrfunktion der Exponentialfunktion $y = a^x$ bezeichnet man als den *Logarithmus zur Basis a* und schreibt

$$x = \underset{a}{\log} y,$$

während man den oben eingeführten Logarithmus, wo eine Unterscheidung nötig ist, als natürlichen Logarithmus bezeichnet.

Man entnimmt unmittelbar aus der Definition die Relation

$$x \log a = \log a \cdot \underset{a}{\log} y = \log y,$$

welche uns zeigt, daß man den Logarithmus einer Zahl $y$ für eine beliebige positive Basis $a \neq 1$ aus dem natürlichen Logarithmus erhält, indem man diesen mit dem reziproken natürlichen Logarithmus von $a$, dem *Modul* des Logarithmensystems der Basis $a$[1]), multipliziert.

Anstatt unserer früheren Definition für die *allgemeine Potenz* $x^\alpha$ $(x > 0)$ ist also nunmehr diese Potenz durch die Gleichung

$$x^\alpha = e^{\alpha \log x}$$

definiert.

Die Differentiationsregel für die Potenz $x^\alpha$ ergibt sich nach der Kettenregel aus dieser Definition unmittelbar; denn es ist

$$\frac{d}{dx} x^\alpha = e^{\alpha \log x} \frac{\alpha}{x} = \alpha x^{\alpha - 1},$$

in Übereinstimmung mit unserem früheren Resultat (vgl. S. 105).

## 6. Exponentialfunktion und Logarithmus dargestellt durch Grenzwerte.

Unsere Betrachtungen erlauben uns, für die eingeführten Größen in einfacher Weise wichtige Grenzwertbeziehungen aufzustellen. Wir gehen aus von der Differentiationsformel für die Funktion $f(x) = \log x$

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \frac{1}{x}.$$

---

[1]) Für $a = 10$ erhält man die von der Schule bekannten und für die Zwecke des numerischen Rechnens vorzugsweise verwendeten „Briggschen" Logarithmen.

Für $x = 1$ ergibt sich speziell
$$f(1) = \log 1 = 0$$
und daher
$$\lim_{h \to 0} \frac{1}{h} \log(1 + h) = 1,$$
eine Beziehung, welche wir wegen $\frac{1}{h}\log(1 + h) = \log(1 + h)^{\frac{1}{h}}$ auch schreiben können
$$\lim_{h \to 0} \log(1 + h)^{\frac{1}{h}} = 1.$$

Setzen wir $\log(1 + h)^{\frac{1}{h}} = 1 + \delta$, so wird also für $h \to 0$ auch $\delta$ gegen Null streben. Nun wird $(1 + h)^{\frac{1}{h}} = e^{1+\delta} = e \cdot e^{\delta}$, und es ist $\lim_{\delta \to 0} e^{\delta} = 1$. Also ergibt sich die folgende wichtige Tatsache:

*Wenn die Größe h gegen Null strebt, so strebt der Ausdruck* $(1 + h)^{\frac{1}{h}}$ *gegen die Zahl e:*
$$\lim_{h \to 0} (1 + h)^{\frac{1}{h}} = e.$$

Ihr entnehmen wir, daß die jetzt mit $e$ bezeichnete Zahl gerade mit derjenigen Zahl $e$ übereinstimmt, die wir auf S. 32 als Beispiel für Grenzwertbildungen betrachtet hatten. Speziell können wir nämlich für die Zahl $h$ der Reihe nach die Werte $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$ einsetzen und gelangen so zu der Gleichung
$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n.$$

Für beliebiges reelles $x$ ergibt sich die entsprechende Limesgleichung
$$e^x = \lim_{h \to 0} (1 + xh)^{\frac{1}{h}}.$$
Zum Beweise ersetzen wir die Zahl $h$ in der obigen Gleichung für $e$ durch $x \cdot h$, das ja zugleich mit $h$ gegen 0 strebt. Man erhält für $h \to 0$ dann $(1 + xh)^{\frac{1}{xh}} \to e$ und somit
$$\left[(1 + xh)^{\frac{1}{xh}}\right]^x = (1 + xh)^{\frac{1}{h}} \to e^x. \text{ [1]}$$

Speziell erhalten wir wieder, wenn $h$ die obige Zahlenfolge $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$ durchläuft,
$$e^x = \lim_{n \to \infty} \left(1 + \frac{x}{n}\right)^n.$$

---

[1] Wenn eine Folge von Zahlen $a_1, a_2, \ldots$ gegen einen positiven Grenzwert $a$ konvergiert, so streben auch die Potenzen $a_1^\alpha, a_2^\alpha, \ldots$ gegen die Potenz $a^\alpha$ des Grenzwertes; denn die Funktion $x^\alpha$ ist im Punkte $x = a$ stetig.

Aus der Differentiationsformel für $a^x$

$$a^x \log a = \lim_{h \to 0} \frac{a^{x+h} - a^x}{h}$$

ergibt sich für $x = 0$

$$\log a = \lim_{h \to 0} \frac{a^h - 1}{h},$$

eine Formel, welche uns unmittelbar den Logarithmus von $a$ durch einen Grenzwert ausdrückt.

Ich schließe an diese Gleichung noch die Bemerkung an, daß durch sie die Auffassung unserer früher gewonnenen Beziehung

$$\int_a^b x^\alpha \, dx = \frac{1}{\alpha + 1}(b^{\alpha+1} - a^{\alpha+1}) \qquad (a > 0, \quad b > 0)$$

in befriedigender Weise ergänzt wird. Wir mußten hier stets den Fall $\alpha = -1$ ausschließen. Jetzt können wir aber verfolgen, was geschieht, wenn wir die auf beiden Seiten dieser Gleichung vorkommende Zahl $\alpha$ gegen den *Grenzwert* $-1$ streben lassen. Die linke Seite wird, wenn wir $a = 1$ setzen, auf Grund unserer Definition des Logarithmus gerade den Grenzwert

$$\int_1^b \frac{dx}{x} = \log b$$

haben[1]); die rechte Seite hat daher für $\alpha \to -1$ denselben Grenzwert. Und diese Tatsache steht im Einklang mit der Formel

$$\log b = \lim_{h \to 0} \frac{b^h - 1}{h}.$$

Wir brauchen nur $\alpha + 1 = h$ zu setzen.

Wir haben damit die Ausnahmestellung der Zahl $\alpha = -1$ in der so oft von uns betrachteten Integralformel beseitigt: Unsere obige Formel verliert zwar für $\alpha = -1$ ihren Sinn; sie behält aber für $\alpha \to -1$ als Grenzwertformel ihre Bedeutung bei.

### 7. Schlußbemerkungen.

Ich möchte nochmals kurz den Gedankengang dieses Paragraphen zusammenfassen: Wir haben zunächst den natürlichen Logarithmus $y = \log x$ für $x > 0$ durch ein Integral definiert und daraus sofort Differentiationsformel, Additionstheorem und Umkehrbarkeit gefolgert. Sodann haben wir die Umkehrfunktion $y = e^x$ untersucht — wobei die Zahl $e$ sich als diejenige Zahl ergab, deren Logarithmus 1 war —,

---

[1]) Wir haben den Grenzübergang $\alpha \to -1$ ohne weiteres unter dem Integralzeichen vollzogen (vgl. hierzu die Betrachtungen im zweiten Kapitel, § 7, Nr. 2).

ihre Differentiationsformel sowie Grenzwertdarstellungen für sie und den Logarithmus hergeleitet. Die Hineinziehung der Funktionen $y = x^\alpha = e^{\alpha \log x}$ und $y = a^x = e^{x \log a}$ ergab sich dabei von selbst.

In der hier gegebenen Darstellung im Gegensatz zu der „elementaren" entstehen bei den Stetigkeitsfragen keinerlei Schwierigkeiten, da von vornherein der Logarithmus als Integral und somit als eine stetige und differenzierbare Funktion seines Argumentes gekennzeichnet ist und die Stetigkeit der Umkehrfunktion sich von selbst versteht.

## § 7. Einige Anwendungen der Exponentialfunktion.

Wir wollen in diesem Paragraphen eine Reihe verschiedenartiger Beispiele für das Auftreten der Exponentialfunktion kennenlernen und so einen Einblick in die fundamentale Wichtigkeit dieser Funktion für die mannigfachsten Anwendungen gewinnen.

### 1. Charakterisierung der Exponentialfunktion durch eine Differentialgleichung.

Wir können die Exponentialfunktion durch einen einfachen Satz charakterisieren, dessen Benutzung uns viele Einzelbetrachtungen ersparen wird.

*Wenn eine Funktion $y = f(x)$ einer Gleichung der Form*

$$y' = \alpha y$$

*genügt, wo $\alpha \neq 0$ eine Konstante ist, so hat $y$ die Form*

$$y = f(x) = c e^{\alpha x},$$

*wo $c$ ebenfalls eine Konstante ist; und umgekehrt hat jede Funktion der eben hingeschriebenen Gestalt $c e^{\alpha x}$ die Eigenschaft, der obigen Gleichung zu genügen,* die man kurz als „*Differentialgleichung*" bezeichnet, da in ihr die Funktion und ihr Differentialquotient zueinander in Beziehung gesetzt werden.

Um uns den ausgesprochenen Satz klar zu machen, beachten wir zunächst, daß im einfachsten Falle, wo $\alpha = 1$ ist, unsere obige Gleichung in $y' = y$ übergeht. Wir wissen, daß $y = e^x$ dieser Gleichung genügt, und es ist klar, daß dasselbe für $y = c e^x$ gilt, wenn $c$ eine beliebige Konstante ist. Aber auch umgekehrt können wir leicht sehen, daß keine andere Funktion unsere Differentialgleichung befriedigt. Denn ist $y$ eine solche Funktion, so betrachten wir die Funktion $u = y e^{-x}$. Es muß dann sein:

$$u' = y' e^{-x} - y e^{-x} = e^{-x}(y' - y).$$

Die rechte Seite aber verschwindet auf Grund der vorausgesetzten Beziehung, und somit ist $u' = 0$, d. h. nach S. 91 $u$ gleich einer Konstanten $c$ und $y = c e^x$, wie behauptet wurde.

Genau so wie der Spezialfall $\alpha = 1$ läßt sich nun auch der Fall eines beliebigen nicht verschwindenden $\alpha$ behandeln. Führen wir $u = y \cdot e^{-\alpha x}$ als neu zu bestimmende Funktion ein, so ergibt sich $u' = y' e^{-\alpha x} - \alpha y e^{-\alpha x}$ und somit wegen der vorausgesetzten Differentialgleichung $u' = 0$, also $u = c$ und $y = c e^{\alpha x}$. Die Umkehrung ist klar.

Wir wollen nun die gewonnene Einsicht an einer Reihe von Beispielen anwenden und des näheren verständlich machen.

## 2. Stetige Verzinsung. Radioaktiver Zerfall.

Ein Kapital, dessen Zinsen in gewissen Zeitabständen immer wieder zum Kapital geschlagen werden, vermehrt sich in diesen Zeitpunkten sprungweise folgendermaßen: Ist $100\alpha$ der Zinsfuß in Prozenten, wird weiter das Kapital am Ende jedes Jahres um die aufgelaufenen Zinsen vermehrt, so ergibt sich nach $x$ Jahren aus dem Anfangskapital 1 die Summe

$$(1 + \alpha)^x.$$

Würde jedoch das Kapital nicht zum Schlusse jedes Jahres, sondern am Schlusse jedes $n$-ten Teiles eines Jahres um die aufgelaufenen Zinsen vermehrt, so würde das Kapital nach $x$ Jahren den Wert

$$\left(1 + \frac{\alpha}{n}\right)^{nx}$$

erreicht haben. Nehmen wir der Einfachheit halber $x = 1$, d. h. eine auf das Jahr gerechnete $100\alpha$-prozentige Verzinsung, so würde nach einem Jahre aus dem Kapital 1 bei der letzteren Art der Zinsberechnung das Kapital

$$\left(1 + \frac{\alpha}{n}\right)^{n}$$

geworden sein. Lassen wir nun $n$ über alle Grenzen wachsen, lassen wir also in immer kürzeren Abständen die Verzinsung erfolgen, so wird der Grenzfall bedeuten, daß die Verzinsung gewissermaßen stetig in jedem Zeitmoment wirksam wird, und wir sehen, daß der Betrag des Kapitals nach einem Jahre das $e^{\alpha}$-fache des Ausgangskapitals geworden ist. Ebenso wird sich bei dieser Art der Verzinsung nach Ablauf von $x$ Jahren ($x$ kann dabei eine beliebige ganze oder auch nicht ganze Zahl sein) als Kapital die Summe $e^{\alpha x}$ ergeben.

Dies Beispiel und ähnliche kann man im Rahmen der unter Nr. 1 gegebenen Überlegungen folgendermaßen verstehen: Man hat irgend eine durch die Zahl $y$ gegebene Menge vor sich, welche mit der Zeit sich vermehrt (oder auch vermindert). Die Geschwindigkeit, mit der diese Menge sich vermehrt oder vermindert, sei nun proportional der Gesamtmenge. Dann gilt für die Geschwindigkeit $y'$ dieser Vermehrung — als unabhängige Veränderliche $x$ werden wir dabei die Zeit betrachten — ein Gesetz der Form $y' = \alpha y$, wobei der Proportionalitätsfaktor $\alpha$ positiv oder negativ ist, je nachdem, ob es sich um eine Vermehrung oder Verminderung handelt. Die Größe $y$ selbst wird gemäß Nr. 1 dann durch eine Formel

$$y = c e^{\alpha x}$$

gegeben werden, wobei die Bedeutung der Konstanten $c$ sofort erkennbar wird, wenn wir den Zeitmoment $x = 0$ betrachten. Es wird dann $e^{\alpha x} = 1$, und es ergibt sich $c = y_0$ als die Menge zu Beginn der Betrachtung, so daß wir schreiben können:

$$y = y_0 e^{\alpha x}.$$

Ein charakteristisches Beispiel für diese Betrachtung bietet uns der radioaktive Zerfall. Die Geschwindigkeit, mit der sich die Gesamtmenge $y$ der radioaktiven Substanz vermindert, wird, wie von vornherein plausibel, der gesamten im Moment vorhandenen Menge proportional sein, da jeder Bestandteil der Substanz sich ebenso schnell verringern wird wie jeder andere. Wir werden also für die Menge $y$ der Substanz als Funktion der Zeit $x$ eine Beziehung der Form $y' = -ky$ ansetzen dürfen, wobei die Zahl $k$ positiv zu nehmen ist, da es sich um eine Substanzverminderung handelt. Für die Substanzmenge als Funktion der Zeit $x$ ergibt sich hieraus: $y = y_0 e^{-kx}$, wobei $y_0$ die Substanzmenge zum Beginn der Betrachtung ($x = 0$) ist.

Nach einer gewissen Zeit $\tau$ wird sich die radioaktive Substanz auf die Hälfte ihrer ursprünglichen Menge verringert haben. Diese sog. *Halbwertszeit* ergibt sich aus der Gleichung

$$\frac{y_0}{2} = y_0 e^{-k\tau},$$

woraus wir für $\tau$ sofort den Wert $\tau = \frac{\log 2}{k}$ erhalten.

### 3. Abkühlung oder Erwärmung eines Körpers in einem umgebenden Medium.

Ein anderes typisches Beispiel für das Auftreten der Exponentialfunktion bietet der Vorgang der Abkühlung eines Körpers, z. B. einer Metallplatte, welche in ein sehr großes Bad von gegebener Temperatur getaucht ist; bei der Betrachtung dieser Abkühlung machen wir die Voraussetzung, daß das umgebende Bad derart groß ist, daß seine Temperatur durch den Prozeß nicht beeinflußt wird. Wir setzen ferner voraus, daß der eingetauchte Körper jederzeit überall dieselbe Temperatur besitzt und daß die Geschwindigkeit, mit der die Temperatur sich ändert, proportional der Temperaturdifferenz zwischen Körper und umgebendem Bade ist (Newtonsches Erkaltungsprinzip).

Bezeichnen wir die Zeit mit $x$, die Temperaturdifferenz mit $y = y(x)$, so wird sich dieses Erkaltungsgesetz in der Gleichung

$$y' = -ky$$

ausdrücken, wobei wir unter $k$ eine positive, von dem Material abhängige Konstante verstehen. Es handelt sich nun darum, aus diesem Momentangesetz, welches die Tendenz des Erkaltungsvorganges für einen bestimmten Zeitmoment $x$ ausdrückt, ein „Integralgesetz" herzuleiten, welches erlaubt, von einem Anfangsmoment $x = 0$ auf einen beliebigen späteren Moment $x$ zu schließen. Dieses Integralgesetz wird uns nun sofort durch den Satz aus Nr. 1 geliefert, und zwar in der Form

$$y = c e^{-kx},$$

wobei $k$ die obige Materialkonstante ist. Es zeigt sich also, daß die Temperatur im Laufe der Zeit „exponentiell" sich senkt und der Außentemperatur zustrebt. Die Geschwindigkeit, mit der dies geschieht, wird durch die Materialkonstante $k$ charakterisiert. Die Bedeutung der Konstanten $c$ erhält man wieder, indem man den Zeitmoment $x = 0$ betrachtet, wobei sich $y_0 = c$ ergibt, so daß wir unser Erkaltungsgesetz schließlich in der Form

$$y = y_0 e^{-kx}$$

schreiben können.

Im übrigen sei hierzu bemerkt, daß dasselbe, was für die Abkühlung eines Körpers gilt, selbstverständlich auch für die Erwärmung gelten wird. Der einzige Unterschied ist dabei der, daß die anfängliche Temperaturdifferenz $y_0$ in diesem Falle nicht positiv, sondern negativ ist.

**4. Abhängigkeit des Luftdruckes von der Höhe über dem Erdboden.** Als weiteres Beispiel für das Auftreten der Exponentialfunktion leiten wir das Gesetz der Abhängigkeit des Luftdruckes von der Höhe ab, die *barometrische Höhenformel.* Wir benutzen dabei einmal die physikalische Tatsache, daß der Luftdruck gleich dem Gewicht der senkrecht über einer Fläche vom Inhalt 1 befindlichen Luftsäule der Atmosphäre ist; zweitens das Boylesche Gesetz, nach welchem der Luftdruck $p$ bei konstanter Temperatur proportional dem spezifischen Gewicht $\sigma$ der Luft ist — in einer Formel: $p = a\sigma$ mit konstantem $a$, das nur von einer speziellen physikalischen Eigenschaft der Luft abhängt und im übrigen, worauf es aber hier nicht ankommt, der absoluten Temperatur der Luft proportional ist. Unsere Aufgabe ist, $p = f(l)$ als Funktion der Höhe $l$ über dem Erdboden zu berechnen.

Bezeichnet man mit $p_0$ den Luftdruck am Erdboden, d. h. das ganze Gewicht der über einer Flächeneinheit lastenden Luftsäule, und mit $\sigma(\lambda)$ das spezifische Gewicht der Luft in der Höhe $\lambda$ über dem Erdboden, so wird das Gewicht der Säule bis zur Höhe $l$ durch das Integral $\int_0^l \sigma(\lambda)\,d\lambda$ gegeben werden. Der Druck $p$ in der Höhe $l$ wird also gleich

$$p = f(l) = p_0 - \int_0^l \sigma(\lambda)\,d\lambda$$

sein. Hieraus aber ergibt sich durch Differenzieren zwischen dem

Druck $p = f(l)$ und dem spezifischen Gewicht $\sigma(l)$ der Zusammenhang

$$\sigma(l) = -f'(l) = -p'.$$

Aus dieser Gleichung in Verbindung mit dem Boyleschen Gesetz eliminieren wir die Größe $\sigma$, so daß wir zu der die unbekannte Druckfunktion allein enthaltenden Gleichung

$$p' = -\frac{1}{a} p$$

gelangen. Aus Nr. 1 ergibt sich somit:

$$p = f(l) = c \cdot e^{-\frac{l}{a}}.$$

Bezeichnen wir wie oben den Druck an der Erdoberfläche, d. h. den Wert $f(0)$, mit $p_0$, so folgt sofort $c = p_0$ und daher

$$p = f(l) = p_0 e^{-\frac{l}{a}}.$$

Gehen wir zu Logarithmen über, so erhalten wir:

$$l = a \log \frac{p_0}{p}.$$

Diese beiden Formeln finden häufige Anwendung. Sie gestatten z. B., wenn man die Konstante $a$ kennt, aus der Messung des Luftdruckes die Höhe des Standortes zu bestimmen, bzw. aus der Messung der Luftdrucke zweier Orte ihre Höhendifferenz. Ebenso erlauben diese Formeln, wenn Luftdruck und Höhe $l$ bekannt sind, die Konstante $a$ zu bestimmen, welche in der Gastheorie eine große Rolle spielt.

### 5. Verlauf chemischer Reaktionen.

Wir betrachten noch ein Beispiel aus der Chemie, und zwar das der sog. *unimolekularen Reaktionen*. Nehmen wir an, daß ein Stoff in einem quantitativ sehr viel reichlicheren Lösungsmittel gelöst ist, etwa eine Menge von Rohrzucker in Wasser, so werden wir bei eintretenden chemischen Reaktionen das sog. Massenwirkungsgesetz der Chemie in diesem einfachen Falle so formulieren können: Die Reaktionsgeschwindigkeit ist proportional der noch vorhandenen Menge der sich umwandelnden Substanz. Denken wir uns, daß durch katalytische Wirkung der Rohrzucker sich in Invertzucker verwandelt, und bezeichnen wir die Menge des zur Zeit $x$ noch nicht umgewandelten Rohrzuckers mit $u(x)$, so ist die Reaktionsgeschwindigkeit $-\frac{du}{dx}$, und es gilt dann gemäß dem Massenwirkungsgesetz eine Gleichung der Form

$$\frac{du}{dx} = -ku,$$

wo $k$ eine Materialkonstante ist. Aus diesem Momentangesetz erhalten wir gemäß Nr. 1 sofort ein Integralgesetz, welches uns die Menge $u(x)$ des übrigbleibenden Rohrzuckers als Funktion der Zeit unmittelbar angibt:

$$u(x) = a e^{-kx},$$

eine Formel, welche uns deutlich vor Augen führt, in welcher Weise die chemische Reaktion asymptotisch ihrem Endzustand $u = 0$, der völligen Umwandlung, zustrebt. Die Konstante $a$ ist offenbar die zur Zeit $x = 0$ vorhandene Menge.

### 6. Ein- und Ausschalten eines elektrischen Stromes.

Als letztes Beispiel betrachten wir den Vorgang, der beim Einschalten (und entsprechend auch beim Ausschalten) eines elektrischen Gleichstromes sich abspielt. Ist $W$ der Widerstand des Stromkreises, $E$ die äußere Spannung, so wird die Stromstärke $J$ von ihrem Anfangswert 0 allmählich zu dem stationären Endwert $\frac{E}{W}$ anwachsen. Wir haben also $J$ als Funktion der Zeit $x$ zu betrachten. Bei diesem Einschaltvorgang spielt die Selbstinduktion eine Rolle; es gehört zu dem Stromkreis eine bestimmte Konstante $L$, der Selbstinduktions-Koeffizient, derart, daß bei jeder Änderung der Stromstärke zu der äußeren Spannung $E$ eine entgegengesetzt gerichtete Spannung von der Größe $L\frac{dJ}{dx}$ hinzutritt, deren Tendenz auf eine Verringerung der Stromstärke gerichtet ist. Zufolge des Ohmschen Gesetzes, nach welchem in jedem Moment das Produkt aus Stromstärke und Widerstand gleich der tatsächlich wirksamen Spannung ist, erhalten wir die Beziehung

$$J \cdot W = E - L\frac{dJ}{dx}.$$

Vergleichen wir diese mit unserem Satz aus Nr. 1, indem wir

$$f(x) = J(x) - \frac{E}{W}$$

setzen, so ergibt sich sofort $f'(x) = -\frac{W}{L}f(x)$, also $f(x) = f(0)\,e^{-\frac{W}{L}x}$ und somit wegen $J(0) = 0$ oder $f(0) = -\frac{E}{W}$ für die Stromstärke als Funktion der Zeit der Ausdruck

$$J = f(x) + \frac{E}{W} = \frac{E}{W}\left(1 - e^{-\frac{W}{L}x}\right).$$

Wir sehen an diesem Ausdruck, wie sich der Strom beim Einschalten asymptotisch seinem stationären Endwert $\frac{E}{W}$ annähert.

## § 8. Die Hyperbelfunktionen.

### 1. Analytische Definition.

Bei zahlreichen Anwendungen tritt die Exponentialfunktion nicht isoliert auf, sondern in Kombinationen von der Art

$$\frac{1}{2}(e^x + e^{-x}) \quad \text{oder} \quad \frac{1}{2}(e^x - e^{-x}).$$

Es ist zweckmäßig, solche und ähnliche Kombinationen als besondere Funktionen einzuführen; wir bezeichnen sie folgendermaßen:

$$\mathfrak{Sin}\, x = \frac{e^x - e^{-x}}{2}, \qquad \mathfrak{Cof}\, x = \frac{e^x + e^{-x}}{2},$$

$$\mathfrak{Tg}\, x = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \qquad \mathfrak{Ctg}\, x = \frac{e^x + e^{-x}}{e^x - e^{-x}},$$

und nennen sie den *hyperbolischen Sinus*, den *hyperbolischen Cosinus* bzw. *hyperbolischen Tangens* und *Cotangens*[1]). Die Funktionen $\mathfrak{Sin}\, x$, $\mathfrak{Cof}\, x$, $\mathfrak{Tg}\, x$ sind für alle Werte von $x$ definiert, während bei $\mathfrak{Ctg}\, x$ die Stelle $x = 0$ ausgeschlossen werden muß. Man bringt mit dieser Bezeichnung eine gewisse Analogie mit den trigonometrischen Funktionen zum Ausdruck; und gerade diese Analogie, die wir sogleich im einzelnen studieren werden, rechtfertigt die besondere Betrachtung unserer neuen Funktionen. In den Figuren 64, 65 und 66 ist der Verlauf der hyperbolischen Funktionen gezeichnet; zum Vergleich ist in Figur 64 punktiert der Verlauf der Kurven $y = \frac{1}{2} e^x$, $y = \frac{1}{2} e^{-x}$ angegeben, aus denen sich die gesuchten Kurven von $\mathfrak{Sin}\, x$ und $\mathfrak{Cof}\, x$ sofort konstruieren lassen.

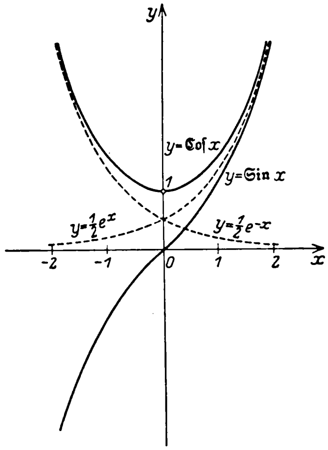


Fig. 64.

Man erkennt, daß $\mathfrak{Cof}\, x$ eine gerade Funktion ist, d. h. eine solche Funktion, die sich nicht ändert, wenn man $x$ in $-x$ verwandelt, während $\mathfrak{Sin}\, x$ eine ungerade Funktion ist, d. h. bei Ersetzung von $x$ durch $-x$ das Vorzeichen wechselt. (Vgl. S. 12 f.)

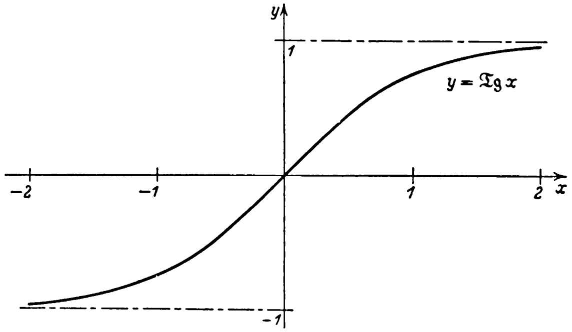


Fig. 65.

Die Funktion

$$\mathfrak{Cof}\, x = \frac{e^x + e^{-x}}{2}$$

ist ihrer Definition nach für alle Werte $x$ positiv. Sie hat ihren kleinsten Wert für $x = 0$; und zwar wird $\mathfrak{Cof}\, 0 = 1$.

Zwischen $\mathfrak{Cof}\, x$ und $\mathfrak{Sin}\, x$ besteht die grundlegende Beziehung

$$\mathfrak{Cof}^2 x - \mathfrak{Sin}^2 x = 1,$$

---

[1]) Man schreibt auch sinh $x$, cosh $x$, tgh $x$, ctgh $x$ für diese Funktionen.

wie man sofort aus der Definition dieser Funktionen erkennt. Bezeichnen wir die unabhängige Veränderliche nunmehr mit $t$ statt mit $x$ und setzen

$$x = \mathfrak{Cos}\, t, \quad y = \mathfrak{Sin}\, t,$$

so ergibt sich also

$$x^2 - y^2 = 1;$$

d. h. der Punkt mit den Koordinaten $x = \mathfrak{Cos}\, t$, $y = \mathfrak{Sin}\, t$ läuft auf der gleichseitigen Hyperbel $x^2 - y^2 = 1$, wenn $t$ die ganze Werteskala von $-\infty$ bis $+\infty$ durchläuft. Dabei ist zufolge unserer Definitionsgleichung $x \geqq 1$, und wir überzeugen uns leicht davon, daß $y$ zugleich mit $t$ durch die ganze Werteskala von $-\infty$ bis $+\infty$ geht; denn, wenn $t$ ins Unendliche wächst, so wächst $e^t$ ins Unendliche, während $e^{-t}$ gegen Null strebt. Daher können wir jetzt genauer sagen, daß durch die Gleichungen $x = \mathfrak{Cos}\, t$, $y = \mathfrak{Sin}\, t$ der eine, und zwar der rechts gelegene Ast unserer gleichseitigen Hyperbel geliefert wird, wenn $t$ von $-\infty$ bis $+\infty$ läuft.

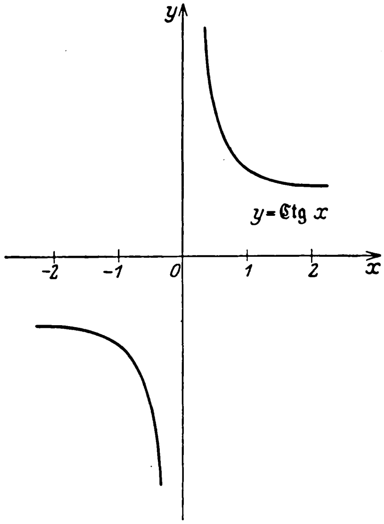



Fig. 66.

## 2. Additionstheoreme und Differentiationsformeln.

Aus der Definition unserer Funktionen folgen die nachstehenden als *Additionstheoreme* bezeichneten Formeln:

$$\mathfrak{Cos}\,(a+b) = \mathfrak{Cos}\, a\, \mathfrak{Cos}\, b + \mathfrak{Sin}\, a\, \mathfrak{Sin}\, b,$$
$$\mathfrak{Sin}\,(a+b) = \mathfrak{Sin}\, a\, \mathfrak{Cos}\, b + \mathfrak{Cos}\, a\, \mathfrak{Sin}\, b.$$

Die Beweise folgen sofort, wenn man schreibt

$$\mathfrak{Cos}\,(a+b) = \frac{e^a e^b + e^{-a} e^{-b}}{2}, \qquad \mathfrak{Sin}\,(a+b) = \frac{e^a e^b - e^{-a} e^{-b}}{2}$$

und hierin

$$e^a = \mathfrak{Cos}\, a + \mathfrak{Sin}\, a, \qquad e^{-a} = \mathfrak{Cos}\, a - \mathfrak{Sin}\, a,$$
$$e^b = \mathfrak{Cos}\, b + \mathfrak{Sin}\, b, \qquad e^{-b} = \mathfrak{Cos}\, b - \mathfrak{Sin}\, b$$

einsetzt.

Die Analogie unserer Formeln mit den entsprechenden trigonometrischen Formeln ist deutlich. Ein Unterschied der Additionstheoreme besteht nur in einem Vorzeichen der ersten Formel.

Eine entsprechende Analogie erhalten wir für die Differentiationsformeln. Es ergeben sich aus unseren Definitionen, wie man ohne Mühe auf Grund der Tatsache, daß $\frac{d e^x}{d x} = e^x$ ist, feststellt, die Differentiationsformeln

$$\frac{d}{dx}\mathfrak{Cof}\, x = \mathfrak{Sin}\, x, \qquad \frac{d}{dx}\mathfrak{Sin}\, x = \mathfrak{Cof}\, x,$$

$$\frac{d}{dx}\mathfrak{Tg}\, x = \frac{1}{\mathfrak{Cof}^2 x}, \qquad \frac{d}{dx}\mathfrak{Ctg}\, x = -\frac{1}{\mathfrak{Sin}^2 x}.$$

### 3. Die Umkehrfunktionen.

Zu den Hyperbelfunktionen $x = \mathfrak{Cof}\, t$, $y = \mathfrak{Sin}\, t$ gehören Umkehrfunktionen, die wir mit

$$t = \mathfrak{Ar}\,\mathfrak{Cof}\, x, \quad t = \mathfrak{Ar}\,\mathfrak{Sin}\, y$$

bezeichnen ($\mathfrak{Ar}$ spricht man „area“ aus). Da die Funktion $\mathfrak{Sin}\, t$ eine durchweg monoton mit $t$ wachsende Funktion ist, ist ihre Umkehrfunktion für alle $y$ eindeutig bestimmt; während, wie ein Blick auf die Kurven (vgl. Fig. 64, S. 149) lehrt, die Umkehrfunktion $t = \mathfrak{Ar}\,\mathfrak{Cof}\, x$ nicht eindeutig, sondern nur bis aufs Vorzeichen bestimmt ist, indem nämlich zu einem gegebenen Werte von $x$ zugleich mit dem Wert $t$ auch der Wert $-t$ gehört. Da stets $\mathfrak{Cof}\, t \geqq 1$ ist, ist $\mathfrak{Ar}\,\mathfrak{Cof}\, x$ nur für $x \geqq 1$ definiert.

Man kann diese Umkehrfunktionen sehr leicht mit Hilfe des Logarithmus ausdrücken, indem man in den Definitionsgleichungen

$$x = \frac{e^t + e^{-t}}{2}, \qquad y = \frac{e^t - e^{-t}}{2}$$

die Größe $e^t = u$ als Unbekannte auffaßt und die so entstehenden quadratischen Gleichungen für $u$ auflöst; es ergibt sich

$$u = x \pm \sqrt{x^2 - 1}; \qquad u = y + \sqrt{y^2 + 1}$$

und daher, wenn man zu den Logarithmen übergeht,

$$t = \log\left(x \pm \sqrt{x^2 - 1}\right) = \mathfrak{Ar}\,\mathfrak{Cof}\, x,$$
$$t = \log\left(y + \sqrt{y^2 + 1}\right) = \mathfrak{Ar}\,\mathfrak{Sin}\, y.$$

Die Variable $x$ ist bei $\mathfrak{Ar}\,\mathfrak{Cof}\, x$ auf das Intervall $x \geqq 1$ beschränkt, während $\mathfrak{Ar}\,\mathfrak{Sin}\, y$ für alle Werte von $y$ definiert ist.

In der Formel für $\mathfrak{Ar}\,\mathfrak{Sin}\, y$ ist die Quadratwurzel notwendig positiv zu nehmen (andernfalls würde die Klammer negativ werden und der Logarithmus keinen Sinn mehr haben); dagegen ist in der Formel für $\mathfrak{Ar}\,\mathfrak{Cof}\, x$ sowohl das positive als auch das negative Vorzeichen der Quadratwurzel zulässig. Wir haben also für die Funktion $\mathfrak{Ar}\,\mathfrak{Cof}\, x$ die beiden Werte $\log\left(x + \sqrt{x^2 - 1}\right)$ und $\log\left(x - \sqrt{x^2 - 1}\right)$, die den beiden Zweigen von $\mathfrak{Ar}\,\mathfrak{Cof}\, x$ entsprechen. Da

$$\left(x + \sqrt{x^2 - 1}\right)\left(x - \sqrt{x^2 - 1}\right) = 1$$

ist, so ist in der Tat die Summe dieser beiden Werte gleich Null in Übereinstimmung mit dem früher Gesagten.

Ganz analog kann man natürlich auch eine Umkehrfunktion zum hyperbolischen Tangens und Cotangens definieren und durch Logarithmen ausdrücken. Für diese Funktionen, die wir mit $\mathfrak{Ar}\,\mathfrak{Tg}\,x$ und $\mathfrak{Ar}\,\mathfrak{Ctg}\,x$ bezeichnen wollen, erhält man ohne Schwierigkeit, wenn durchweg die unabhängige Veränderliche mit $x$ bezeichnet wird, die Gleichungen:

$$\mathfrak{Ar}\,\mathfrak{Tg}\,x = \frac{1}{2}\log\frac{1+x}{1-x} \text{ im Intervall } -1 < x < 1,$$

$$\mathfrak{Ar}\,\mathfrak{Ctg}\,x = \frac{1}{2}\log\frac{x+1}{x-1} \text{ in den Intervallen } x < -1,\ x > 1.$$

Die Differentiation unserer Umkehrfunktionen mag der Leser selbst durchführen; man kann sich dabei entweder der Differentiationsregel für die Umkehrfunktion oder der direkten Darstellung der Umkehrfunktion durch den Logarithmus und der Kettenregel bedienen. Man erhält, wenn durchweg die unabhängige Veränderliche mit $x$ bezeichnet wird,

$$\frac{d}{dx}\mathfrak{Ar}\,\mathfrak{Cof}\,x = \pm\frac{1}{\sqrt{x^2-1}}, \qquad \frac{d}{dx}\mathfrak{Ar}\,\mathfrak{Sin}\,x = \frac{1}{\sqrt{x^2+1}},$$

$$\frac{d}{dx}\mathfrak{Ar}\,\mathfrak{Tg}\,x = \frac{1}{1-x^2}, \qquad \frac{d}{dx}\mathfrak{Ar}\,\mathfrak{Ctg}\,x = \frac{1}{1-x^2}.$$

Die beiden letzten Formeln stehen nicht miteinander im Widerspruch, da die erste nur im Intervall $-1 < x < 1$, die zweite nur für $x < -1$ und $1 < x$ gilt.

## 4. Weitere Analogien.

Bei der obigen Darstellung der gleichseitigen Hyperbel durch die Größe $t$ haben wir zunächst darauf verzichtet, für diese Größe eine geometrische Bedeutung aufzuzeigen; indem wir dies jetzt nachholen, werden wir unsere Einsicht in die Analogie zwischen trigonometrischen und hyperbolischen Funktionen noch weiter vervollständigen. Wenn wir einen Kreis mit der Gleichung $x^2 + y^2 = 1$ durch einen „Parameter" $t$ in der Form $x = \cos t$, $y = \sin t$ darstellen, so können wir die Größe $t$ als Winkel bzw. Bogen auf der Kreisperipherie deuten; wir können aber auch $t$ als den doppelten Flächeninhalt des zu dem betr. Winkel gehörenden Kreissektors auffassen, wobei dieser Flächeninhalt positiv oder negativ gerechnet wird, je nachdem der Winkel positiv oder negativ zu rechnen ist.

Nunmehr behaupten wir, daß ganz analog für die hyperbolischen Funktionen die Größe $t$ gleich dem doppelten Flächeninhalt des in Figur 67 schraffierten „Hyperbelsektors" ist. Der Beweis dieser Tatsache ergibt sich ohne Schwierigkeit, wenn wir die Hyperbel $x^2 - y^2 = 1$ durch die Koordinatentransformation $x - y = \sqrt{2}\,\xi$, $x + y = \sqrt{2}\,\eta$ oder

$$x = \frac{1}{\sqrt{2}}(\xi + \eta), \qquad y = \frac{1}{\sqrt{2}}(\eta - \xi)$$

auf ihre Asymptoten beziehen, wobei die Hyperbel die Gleichung $\xi\eta = \frac{1}{2}$ erhält. Es ergibt sich nun ohne weiteres, daß der fragliche Flächeninhalt dem Flächeninhalt der trapezförmigen Figur $ABQP$ gleich ist; denn die beiden rechtwinkligen Dreiecke $OPQ$ und $OAB$ sind wegen der Hyperbelgleichung flächengleich. Die beiden Punkte $A$ und $P$ entsprechen nun offenbar den Werten

$$\xi = \frac{1}{\sqrt{2}}, \quad \eta = \frac{1}{\sqrt{2}} \quad \text{bzw.} \quad \xi = \frac{x-y}{\sqrt{2}}, \quad \eta = \frac{x+y}{\sqrt{2}},$$

und es ergibt sich daher für den doppelten Flächeninhalt unserer Figur

$$2 \cdot \int\limits_{\frac{1}{\sqrt{2}}}^{\frac{1}{\sqrt{2}}(x+y)} \frac{1}{2\eta}\, d\eta = \log(x+y) = \log(x \pm \sqrt{x^2-1}).$$

Ein Vergleichen dieses Ausdruckes mit der Formel von S. 151 für die

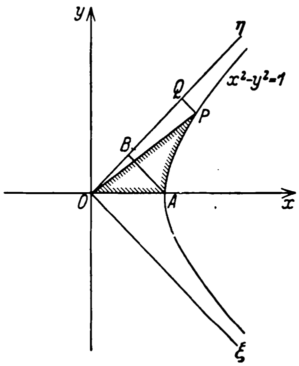



Fig. 67. Zur Parameterdarstellung der Hyperbel.

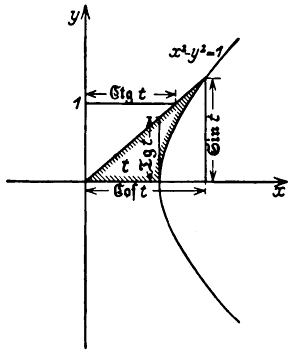



Fig. 68. Veranschaulichung der Hyperbelfunktionen.

Umkehrfunktion 𝔄𝔯 ℭ𝔬𝔣 $x$ zeigt uns, daß tatsächlich unsere Behauptung über die Größe $t$ zutrifft.

Hiermit ist übrigens auch die Bezeichnung Area = Flächeninhalt für die Umkehrfunktionen gerechtfertigt.

Endlich sei zum Schluß noch darauf hingewiesen, daß, wie in Figur 68 angedeutet ist, die hyperbolischen Funktionen sich an der gleichseitigen Hyperbel ganz ähnlich veranschaulichen lassen wie die trigonometrischen Funktionen am Kreise[1]).

---

[1]) Die Werte der Hyperbelfunktionen, deren Benutzung für viele numerische Rechnungen sehr bequem ist, sind vielfach in Tabellen zusammengestellt. Genannt seien die Tabellen von *Hayashi*, Fünfstellige Tafeln für die Kreis- und Hyperbelfunktionen, Berlin-Leipzig 1921.

## § 9. Die Größenordnung von Funktionen.

Die verschiedenen Funktionen, die uns in diesem Kapitel begegnet sind, unterscheiden sich voneinander sehr wesentlich hinsichtlich ihres Verhaltens für große Argumentwerte oder, wie man auch sagt, der „*Größenordnung*" ihres Anwachsens. Ich werde auf diese Dinge wegen ihrer großen Bedeutung hier noch kurz eingehen, obwohl sie unmittelbar nicht mit dem Begriff des Integrals oder des Differentialquotienten zusammenhängen.

### 1. Begriff der Größenordnung. Einfachste Fälle.

Wenn die Variable $x$ über alle Grenzen wächst, so werden mit ihr zugleich für $\alpha > 0$ auch die Funktionen $x^\alpha$, $\log x$, $e^x$, $e^{\alpha x}$ über alle Grenzen wachsen. Hinsichtlich der Art dieses Anwachsens können wir aber sofort wesentliche Unterschiede feststellen. Z. B. wird die Funktion $x^3$ von höherer Ordnung unendlich werden als $x^2$; wir meinen damit, daß der Quotient $\frac{x^3}{x^2}$ bei wachsendem $x$ selbst noch über alle Grenzen wächst. Entsprechend werden wir sagen, daß die Funktion $x^\alpha$ von stärkerer Ordnung unendlich wird als $x^\beta$, wenn $\alpha > \beta > 0$ ist, usw.

Ganz allgemein werden wir von zwei Funktionen $f(x)$ und $g(x)$, deren absolute Beträge mit $x$ über alle Grenzen wachsen, sagen: $f(x)$ *wird von höherer Ordnung unendlich als* $g(x)$, wenn bei wachsendem $x$ der Quotient $\left|\frac{f(x)}{g(x)}\right|$ über alle Grenzen wächst; wir werden sagen, daß $f(x)$ von geringerer Größenordnung als $g(x)$ unendlich wird, wenn der Quotient $\left|\frac{f(x)}{g(x)}\right|$ gegen 0 strebt; und wir werden sagen, daß beide Funktionen von derselben Größenordnung unendlich werden, wenn der Quotient $\left|\frac{f(x)}{g(x)}\right|$ bei wachsendem $x$ einen von 0 verschiedenen Grenzwert besitzt oder wenigstens zwischen zwei festen positiven Schranken bleibt. Es wird also z. B. die Funktion $a x^3 + b x^2 + c = f(x)$, wo $a \neq 0$ sei, von derselben Größenordnung sein wie die Funktion $x^3 = g(x)$; denn der Quotient $\left|\frac{f(x)}{g(x)}\right| = \left|\frac{a x^3 + b x^2 + c}{x^3}\right|$ hat den Grenzwert $|a|$. Dagegen wird die Funktion $x^3 + x + 1$ von höherer Größenordnung unendlich als die Funktion $x^2 + x + 1$.

Eine Summe von zwei Funktionen $f(x)$ und $\varphi(x)$, von denen $f(x)$ eine höhere Größenordnung als $\varphi(x)$ besitzt, hat dieselbe Größenordnung wie $f(x)$. Denn es ist $\left|\frac{f(x) + \varphi(x)}{f(x)}\right| = \left|1 + \frac{\varphi(x)}{f(x)}\right|$, und dieser Ausdruck strebt nach Voraussetzung bei wachsendem $x$ gegen 1.

Man könnte versucht sein, die Größenordnungen von Funktionen nach einer Skala zu messen, indem man der Größe $x$ die Größenordnung 1 und der Potenz $x^\alpha$ für positives $\alpha$ die Größenordnung $\alpha$ zuschreibt. Eine ganze rationale Funktion $n$-ten Grades hat dann offenbar die Größenordnung $n$; eine gebrochen rationale Funktion, deren Zähler einen um $h$ höheren Grad als der Nenner hat, würde die Größenordnung $h$ besitzen.

## 2. Die Größenordnung der Exponentialfunktion und des Logarithmus.

Es zeigt sich nun, daß ein Versuch, die Größenordnung beliebiger Funktionen durch die obige Skala festzulegen, scheitern muß. Es gibt nämlich Funktionen, welche stärker unendlich werden als jede noch so hohe Potenz $x^\alpha$ von $x$; ebenso gibt es Funktionen, welche schwächer unendlich werden als jede noch so kleine Potenz von $x$. Diese Funktionen würden sich also in unsere Skala gar nicht einreihen lassen.

Ohne auf eine genauere Theorie der Größenordnung hier einzugehen, will ich folgende Tatsache beweisen: *Wenn $a$ irgend eine Zahl größer als* 1 *ist, so strebt der Quotient $\frac{a^x}{x}$ bei wachsendem $x$ gegen Unendlich.*

Zum Beweise bilden wir die Funktion

$$\varphi(x) = \log\frac{a^x}{x} = x\log a - \log x;$$

offenbar genügt es, zu zeigen, daß sie über alle Grenzen wächst, wenn $x$ positiv unendlich wird. Hierzu betrachten wir die Ableitung

$$\varphi'(x) = \log a - \frac{1}{x}$$

und bemerken, daß diese für $x \geqq c = \frac{2}{\log a}$ nicht kleiner als die positive Zahl $\frac{1}{2}\log a$ ist. Hiernach ergibt sich für $x \geqq c$

$$\varphi(x) - \varphi(c) = \int_c^x \varphi'(t)\,dt \geqq \int_c^x \frac{1}{2}\log a\,dt = \frac{x-c}{2}\log a,$$

$$\varphi(x) \geqq \varphi(c) + \frac{x-c}{2}\log a,$$

und die rechte Seite strebt mit wachsendem $x$ gegen Unendlich.

Ich will für den so bewiesenen wichtigen Satz noch einen zweiten Beweis geben. Es ist

$$a^x = e^{x\log a} = \lim_{n\to\infty}\left(1 + \frac{x\log a}{n}\right)^n.$$

Andererseits gilt nach S. 23 für $n \geqq 2$, $x > 0$ die Abschätzung

$$\left(1 + \frac{x\log a}{n}\right)^n > \frac{n^2}{4}\cdot\frac{x^2(\log a)^2}{n^2}$$
$$= \frac{x^2(\log a)^2}{4},$$

also sicher auch
$$\frac{a^x}{x} \geqq \frac{x(\log a)^2}{4} \to \infty.$$

Aus der bewiesenen Tatsache folgt sogleich noch sehr viel mehr, nämlich: Für jeden positiven Exponenten $\alpha$ und jede Zahl $a > 1$ strebt der Quotient $\frac{a^x}{x^\alpha}$ bei wachsendem $x$ gegen Unendlich; d. h. *die Exponentialfunktion wird stärker unendlich als jede Potenz von $x$.* Um dies einzusehen, brauchen wir nur zu zeigen, daß die $\alpha$-te Wurzel aus dem Ausdruck $\frac{a^x}{x^\alpha}$, d. i.
$$\frac{a^{\frac{x}{\alpha}}}{x} = \frac{1}{\alpha} \cdot \frac{a^{\frac{x}{\alpha}}}{\frac{x}{\alpha}} = \frac{1}{\alpha} \cdot \frac{a^y}{y} \qquad \left(y = \frac{x}{\alpha}\right)$$
gegen Unendlich strebt. Das folgt aber unmittelbar aus dem vorangehenden Satz, wenn er auf $y = \frac{x}{\alpha}$ an Stelle von $x$ angewandt wird.

Ganz ähnlich können wir folgende Tatsache beweisen. Für jeden positiven Wert von $\alpha$ strebt die Größe $\frac{\log x}{x^\alpha}$ gegen Null, wenn $x$ gegen Unendlich strebt; d. h. *der Logarithmus wird schwächer unendlich als jede noch so niedrige positive Potenz.*

Der Beweis folgt sofort, wenn wir $\log x = y$ setzen, wodurch unser Quotient übergeht in $\frac{y}{e^{\alpha y}}$. Setzen wir $e^\alpha = a$, so wird $a$ eine Zahl, die größer ist als 1, und unser Quotient $\frac{y}{a^y}$ wird also bei wachsendem $y$ gegen Null streben. Da nun $y$ zugleich mit $x$ gegen Unendlich strebt, so ist damit unsere Behauptung bewiesen[1]).

Offenbar sind wir auf Grund unserer Resultate imstande, uns Funktionen von weit höherer Größenordnung als der der Exponentialfunktion und weit schwächerer als der des Logarithmus zu bilden. Z. B. wird die Funktion $e^{e^x}$ stärker anwachsen als die Exponentialfunktion und die Funktion $\log\log x$ schwächer als der Logarithmus, und wir können offenbar derartige Wiederholungsprozesse beliebig übereinander türmen und miteinander kombinieren.

### 3. Allgemeine Bemerkungen.

Unsere Überlegungen zeigen uns, daß es prinzipiell nicht möglich ist, jeder Funktion eine bestimmte Zahl als Größenordnung zuzuweisen,

---

[1]) Ein anderer sehr einfacher Beweis sei angedeutet: Es ist für $x > 1$ und $\varepsilon > 0$
$$\log x = \int_1^x \frac{d\xi}{\xi} < \int_1^x \xi^{\varepsilon - 1}\, d\xi = \frac{1}{\varepsilon}(x^\varepsilon - 1);$$
wählen wir $\varepsilon$ kleiner als $\alpha$ und dividieren wir die so erhaltene Ungleichung durch $x^\alpha$, so folgt für $x \to \infty$ sofort $\frac{\log x}{x^\alpha} \to 0$.

derart, daß einer Funktion mit höherer Größenordnung eine höhere Zahl zukommt. Wenn z. B. die Funktion $x$ die Größenordnung 1 hat und die Funktion $x^{1+\varepsilon}$ die Größenordnung $1+\varepsilon$, so müßte die Funktion $x \log x$ eine Größenordnung haben, die größer ist als 1 und kleiner als $1+\varepsilon$, wie klein auch immer wir die positive Zahl $\varepsilon$ wählen. Eine solche Zahl gibt es aber nicht. Aber auch abgesehen von dem eben erwähnten Umstande, ist es leicht zu sehen, daß Funktionen nicht eine klar definierte Größenordnung zu besitzen brauchen. Z. B. wird die Funktion $\frac{x^2(\sin x)^2+x}{x^2(\cos x)^2+x}$ bei wachsendem $x$ keinem bestimmten Grenzwert zustreben; vielmehr wird für $x = n\pi$ ($n$ ganzzahlig) der Funktionswert gleich $\frac{n\pi}{n^2\pi^2+n\pi} = \frac{1}{n\pi+1}$ werden, für $x = \left(n+\frac{1}{2}\right)\pi$ dagegen gleich $\left(n+\frac{1}{2}\right)\pi+1$. Obwohl Zähler und Nenner für sich unendlich werden, tritt also keiner der drei Fälle ein, daß bei wachsendem $x$ der Quotient zwischen festen positiven Schranken bleibt oder gegen Null oder gegen Unendlich strebt, so daß wir nach unserer Definition nicht sagen können, ob der Zähler oder der Nenner eine höhere Größenordnung besitzt oder ob sie beide von derselben Größenordnung sind.

**4. Die Größenordnung einer Funktion in der Umgebung eines beliebigen Punktes.**

Genau so, wie man das Verhalten von Funktionen bei unbegrenzt wachsendem $x$ untersuchen kann, wird man sich auch die Frage stellen, ob und wie man Funktionen, die an der Stelle $x=\xi$ unendlich werden, dort hinsichtlich ihres Anwachsens zu unterscheiden hat. Wir wollen wieder sagen: Die Funktion $f(x) = \frac{1}{|x-\xi|}$ wird an der Stelle $x=\xi$ von erster Ordnung unendlich und entsprechend die Funktion $\frac{1}{|x-\xi|^\alpha}$ von der Ordnung $\alpha$, sobald $\alpha$ positiv ist.

Man erkennt dann wiederum, daß die Funktion $e^{\frac{1}{|x-\xi|}}$ von höherer Ordnung, die Funktion $\log|x-\xi|$ von niedrigerer Ordnung unendlich wird als alle diese Potenzen, d. h. daß die Grenzwertbeziehungen

$$\lim_{x\to\xi}\left(|x-\xi|^\alpha \cdot e^{\frac{1}{|x-\xi|}}\right) = \infty \quad \text{und} \quad \lim_{x\to\xi}\left(|x-\xi|^\alpha \cdot \log|x-\xi|\right) = 0$$

gelten.

Um dies einzusehen, braucht man nur $\frac{1}{|x-\xi|} = y$ zu setzen und hat dann unsere Behauptungen auf die in der Nr. 2 bewiesenen Tatsachen zurückgeführt, da $|x-\xi|^\alpha \cdot e^{\frac{1}{|x-\xi|}} = \frac{e^y}{y^\alpha}$ und $|x-\xi|^\alpha \cdot \log|x-\xi| = -\frac{\log y}{y^\alpha}$ wird und einem gegen $\xi$ strebenden Werte von $x$ ein über alle

Grenzen wachsendes $y$ entspricht. — Die Methode, das Verhalten in einem endlichen Punkte auf das Verhalten im Unendlichen durch die „Substitution“ $\frac{1}{|x-\xi|} = y$ zurückzuführen, erweist sich übrigens auch sonst häufig als nützlich.

### 5. Größenordnung des Verschwindens einer Funktion.

Ganz ebenso, wie man das Unendlichwerden einer Funktion durch den Begriff der Größenordnung näher zu charakterisieren sucht, kann man auch das Null-Werden einer Funktion $f(x)$ kennzeichnen. Man sagt etwa: Die Größe $\frac{1}{x}$ verschwindet für $x \to \infty$ von der ersten Ordnung, die Größe $x^{-\alpha}$ bei positivem $\alpha$ von der Größenordnung $\alpha$. Dann zeigt sich wieder, daß *die Funktion* $\frac{1}{\log x}$ *von schwächerer Ordnung verschwindet als jede Potenz* $x^{-\alpha}$, d. h. daß für jedes positive $\alpha$ die Beziehung

$$\lim_{x \to \infty} (x^{-\alpha} \cdot \log x) = 0$$

gilt.

Entsprechend werden wir sagen, daß für $x = \xi$ die Größe $x - \xi$ von erster Ordnung, die Größe $|x - \xi|^\alpha$ von der Ordnung $\alpha$ verschwindet. Die nach dem Obigen leicht zu beweisenden Beziehungen

$$\lim_{x \to 0} (|x|^\alpha \cdot \log|x|) = 0, \qquad \lim_{x \to 0} \left(|x|^{-\alpha} \cdot e^{-\frac{1}{|x|}}\right) = 0$$

pflegt man dann folgendermaßen auszusprechen: Die Funktion $\frac{1}{\log|x|}$ verschwindet für $x \to 0$ von geringerer Ordnung als jede Potenz, *die Exponentialfunktion* $e^{-\frac{1}{|x|}}$ *verschwindet für* $x \to 0$ *von höherer Ordnung als jede Potenz.*

# Anhang zum dritten Kapitel.

## § 1. Betrachtung einiger spezieller Funktionen.

Wir haben uns gelegentlich an Beispielen klar gemacht, daß in dem allgemeinen Funktionsbegriff zahlreiche der naiven Anschauung fremdartig scheinende Möglichkeiten stecken. Diese Beispiele waren im allgemeinen nicht durch einheitliche analytische Ausdrücke gegeben. Ich möchte daher jetzt zeigen, daß man durch sehr einfache Ausdrücke mit Hilfe der elementaren Funktionen verschiedene typische Unstetigkeiten und anormale Erscheinungen darstellen kann. Ich beginne allerdings mit einem Beispiel, bei welchem keine Unstetigkeit auftritt.

### 1. Die Funktion $y = e^{-\frac{1}{x^2}}$.

Diese Funktion (vgl. Fig. 69), welche zunächst nur für alle von 0 verschiedenen Werte $x$ definiert ist, hat offenbar für $x \to 0$ selbst den Grenzwert 0. Denn durch die Transformation $\frac{1}{x^2} = \xi$ geht unsere Funktion in $y = e^{-\xi}$ über, und es ist $\lim\limits_{\xi \to \infty} e^{-\xi} = 0$. Um also die Funktion $y = e^{-\frac{1}{x^2}}$ zu einer auch für $x = 0$ stetigen Funktion zu ergänzen, setzen wir durch die Gleichung $y(0) = 0$ den Funktionswert an der Stelle $x = 0$ fest.

Der Differentialquotient unserer Funktion ergibt sich für $x \neq 0$ nach der Kettenregel als $y' = \frac{2}{x^3} e^{-\frac{1}{x^2}}$. Strebt $x$ gegen 0, so wird dieser Differentialquotient selbst den Grenzwert 0 haben, wie man ohne weiteres aus dem dritten Kapitel, § 9, entnehmen kann. An der Stelle $x = 0$ selbst ergibt sich der Differentialquotient $y'(0) = \lim\limits_{h \to 0} \frac{y(h) - y(0)}{h} = \lim\limits_{h \to 0} \frac{e^{-\frac{1}{h^2}}}{h}$ ebenfalls selbst als 0.

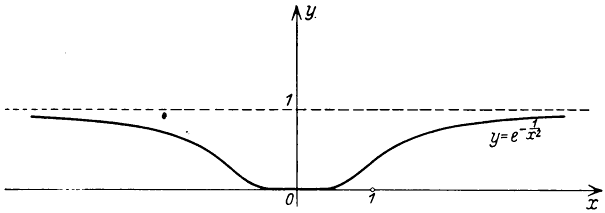



Fig. 69.

Bilden wir die weiteren Differentialquotienten, zunächst für $x \neq 0$, so erhalten wir offenbar stets das Produkt der Funktion $e^{-\frac{1}{x^2}}$ mit einer ganzen rationalen Funktion von $\frac{1}{x}$, und stets gibt der Grenzübergang $x \to 0$ den Wert 0. Auch alle höheren Differentialquotienten verschwinden also ebenso wie $y'$ an der Stelle $x = 0$.

Wir erkennen so, daß unsere Funktion eine für alle Werte von $x$ stetige und beliebig oft differenzierbare Funktion ist, welche an der Stelle $x = 0$ mit ihren sämtlichen Differentialquotienten verschwindet, ein Verhalten, dessen Merkwürdigkeit für uns später noch deutlich hervortreten wird (vgl. sechstes Kapitel, Anhang, § 1).

### 2. Die Funktion $y = e^{-\frac{1}{x}}$.

Diese Funktion hat, wie man sich leicht überlegen kann, für positive Werte von $x$ denselben Habitus wie die eben behandelte Funktion: wenn $x$ von positiven Werten her gegen 0 strebt, so nähert sich die Funktion dem Grenzwert 0, und dasselbe gilt für jeden Differentialquotienten. Setzt man für $x = 0$ als Funktionswert $y(0) = 0$ fest, so haben auch alle vorderen Differentialquotienten an der Stelle $x = 0$ den Wert 0. Ganz anders aber ist es, wenn $x$ sich von negativen Werten her der Null nähert; dann werden die Funktion und ihre sämtlichen Differentialquotienten unendlich werden, und hintere Differentialquotienten an der Stelle $x = 0$ existieren nicht. Die Funktion hat

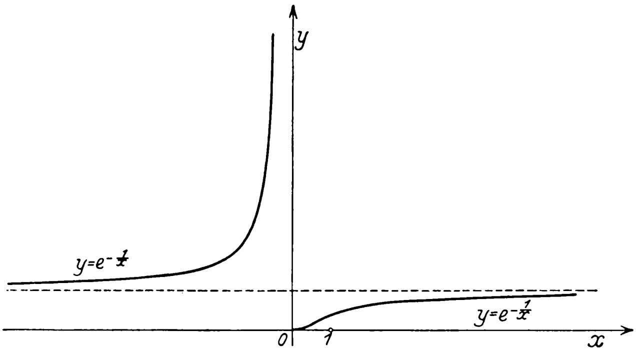



Fig. 70.

also an der Stelle $x = 0$ eine merkwürdige Art von Unstetigkeit; anders als die Unendlichkeitsstellen, die wir bei rationalen Funktionen im ersten Kapitel betrachtet haben (vgl. Fig. 70).

### 3. Die Funktion $y = \mathfrak{Tg}\,\frac{1}{x}$.

Wir haben schon in Kap. 1, § 5, Nr. 6 und § 8, Nr. 2 gesehen, daß wir Funktionen mit sprunghaften Unstetigkeiten vermittelst eines Grenzüberganges aus einfachen Funktionen erhalten können. Die im dritten Kapitel definierte Exponentialfunktion und das Prinzip der Zusammensetzung von Funktionen liefern uns ein weiteres Mittel, Funktionen mit derartigen Unstetigkeiten auch direkt ohne Benutzung neuer Grenzübergänge aus elementaren Funktionen aufzubauen. Ein Beispiel dafür ist die Funktion

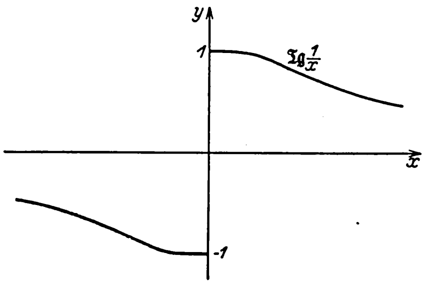



Fig. 71.

$$y = \mathfrak{Tg}\,\frac{1}{x} = \frac{e^{\frac{1}{x}} - e^{-\frac{1}{x}}}{e^{\frac{1}{x}} + e^{-\frac{1}{x}}}$$

und ihr Verhalten in der Umgebung der Stelle $x = 0$. An dieser Stelle ist die Funktion zunächst nicht definiert. Nähern wir uns der Stelle $x = 0$ von positiven $x$-Werten her, so erhalten wir offenbar den Grenzwert 1;

nähern wir uns dagegen der Stelle $x = 0$ von negativen $x$-Werten her, so erhalten wir den Grenzwert $-1$. Der Punkt $x = 0$ ist also für die Funktion eine Sprungstelle, bei deren Überschreiten der Funktionswert um die Zahl 2 springt (vgl. Fig. 71). Der Differentialquotient

$$y' = -\frac{1}{\mathfrak{Cof}^2 \frac{1}{x}} \frac{1}{x^2} = -\frac{1}{x^2} \frac{4}{\left(e^{\frac{1}{x}} + e^{-\frac{1}{x}}\right)^2}$$

nähert sich dagegen von beiden Seiten her dem Grenzwert 0, wie sich ebenfalls aus dem dritten Kapitel, § 9, leicht ergibt[1]).

## 4. Die Funktion $y = x\,\mathfrak{Tg}\frac{1}{x}$.

Bei der Funktion

$$y = x\,\mathfrak{Tg}\frac{1}{x} = x\,\frac{e^{\frac{1}{x}} - e^{-\frac{1}{x}}}{e^{\frac{1}{x}} + e^{-\frac{1}{x}}}$$

ist die oben betrachtete Unstetigkeit durch den Faktor $x$ beseitigt. Von beiden Seiten her hat diese Funktion für $x \to 0$ den Grenzwert 0, so daß wir wiederum zweckmäßig $y(0) = 0$ definieren wollen. Unsere Funktion ist dann zwar für $x = 0$ stetig, dagegen wird ihre erste Ableitung

$$y' = \mathfrak{Tg}\frac{1}{x} - \frac{1}{x}\frac{1}{\mathfrak{Cof}^2 \frac{1}{x}}$$

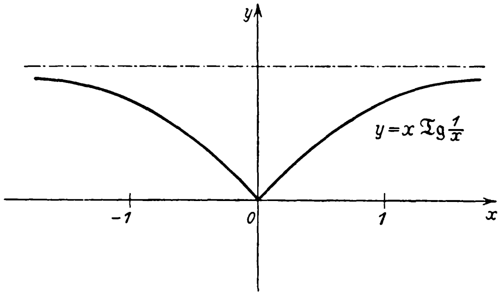



Fig. 72.

gerade die oben betrachtete Unstetigkeit haben; d. h. die Funktion stellt eine Kurve mit einer Ecke dar (vgl. Fig. 72): an der Stelle $x = 0$ selbst besitzt die Funktion keinen Differentialquotienten, aber einen vorderen mit dem Werte $+1$ und einen hinteren mit dem Werte $-1$.

## 5. Die Funktion $y = x \sin\frac{1}{x}$, $y(0) = 0$.

Von dieser Funktion haben wir schon erkannt, daß sie sich zwar nicht mehr aus einer endlichen Anzahl monotoner Stücke zusammensetzt, daß sie nicht mehr „abteilungsweise monoton" ist, wie man sich auch ausdrückt, daß sie aber nichtsdestoweniger stetig bleibt (S. 40).

[1]) Ein weiteres Beispiel für das Auftreten einer Sprungstelle gibt die Funktion $y = \operatorname{arc\,tg}\frac{1}{x}$ bei $x \to 0$.

Dagegen wird ihr erster Differentialquotient

$$y' = \sin\frac{1}{x} - \frac{1}{x}\cos\frac{1}{x} \qquad (x \neq 0)$$

bei Annäherung an die Stelle $x = 0$ eine Unstetigkeit haben, indem er dort fortwährend zwischen absolut genommen wachsenden positiven und negativen Grenzen hin und her schwankt. An der Stelle $x = 0$ selber ist der Differenzenquotient $\frac{y(h) - y(0)}{h} = \sin\frac{1}{h}$; da dieser für $h \to 0$ unendlich oft zwischen 1 und $-1$ hin und her pendelt, so besitzt die Funktion weder einen vorderen noch einen hinteren Differentialquotienten.

## § 2. Bemerkungen über die Differenzierbarkeit von Funktionen.

Wenn eine Funktion stetig ist und an jeder Stelle einen Differentialquotienten besitzt, so braucht dieser Differentialquotient noch keineswegs stetig zu sein. Als einfachstes Beispiel dafür betrachten wir die Funktion

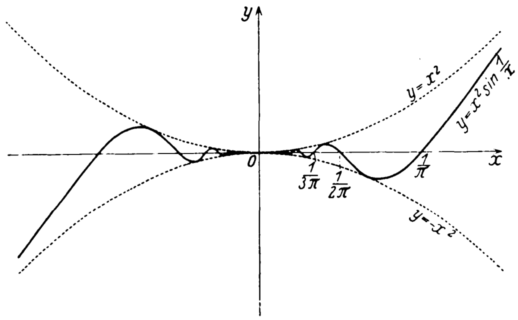



Fig. 73.

$$y = f(x) = x^2 \sin\frac{1}{x},$$

welche zunächst nur für $x \neq 0$ definiert ist und welcher wir durch besondere Festsetzung für $x = 0$ den Wert $f(0) = 0$ zuschreiben, so daß sie nunmehr eine überall definierte stetige Funktion darstellt. Für alle von Null verschiedenen Werte von $x$ wird der Differentialquotient durch den Ausdruck

$$f'(x) = -x^2\cos\frac{1}{x}\cdot\frac{1}{x^2} + 2x\sin\frac{1}{x} = -\cos\frac{1}{x} + 2x\sin\frac{1}{x}$$

gegeben. Strebt $x$ gegen 0, so existiert für $f'(x)$ kein bestimmter Grenzwert. Bilden wir dagegen den Ausdruck $\frac{f(h) - f(0)}{h} = \frac{h^2 \sin\frac{1}{h}}{h} = h\sin\frac{1}{h}$, so erkennen wir sofort, daß er bei abnehmendem $|h|$ gegen Null strebt. Es existiert also die Ableitung für $x = 0$, und zwar ist $f'(0) = 0$. Um dieses paradoxe Verhalten anschaulich zu erfassen, stellen wir die Kurve graphisch dar (vgl. Fig. 73). Sie pendelt zwischen den beiden Kurven $y = x^2$ und $y = -x^2$ hin und her, die sie abwechselnd berührt. Dabei wird zwar die Höhe der Wellenberge unserer Kurve im

Verhältnis zu ihrem Abstand vom Nullpunkt immer kleiner; aber sie werden trotzdem dabei nicht flacher, sondern ihre durch den Differentialquotienten $f'(x) = 2x\sin\frac{1}{x} - \cos\frac{1}{x}$ gemessene Steilheit wird in den Punkten $x = \frac{1}{2n\pi}$ $(n = 1, 2, \ldots)$, wo $\cos\frac{1}{x} = 1$ ist, gleich $-1$ und in den Punkten $x = \frac{1}{(2n-1)\pi}$ gleich $+1$.

Im Gegensatz zu der hier geschilderten Möglichkeit einer zwar überall existierenden, aber nicht überall stetigen Ableitung gilt offenbar der einfache Satz, welcher eine Reihe von früheren Beispielen und Überlegungen beleuchtet: Wenn wir wissen, daß die Ableitung $f'(x)$ überall in der Umgebung eines Punktes $x = a$ existiert und stetig ist, und wenn $\lim\limits_{x \to a} f'(x) = b$ gilt, dann existiert auch im Punkte $x = a$ die Ableitung $f'(a)$, und zwar ist $f'(a) = b$. Der Beweis folgt sofort aus dem Mittelwertsatz. Es ist nämlich $\frac{f(a+h) - f(a)}{h} = f'(\xi)$, wo $\xi$ ein Zwischenwert zwischen $a$ und $a + h$ ist. Strebt nun $h$ gegen 0, so strebt nach Voraussetzung auch $f'(\xi)$ gegen $b$, und daraus folgt sofort unsere Behauptung.

Ein Gegenstück hierzu ist der folgende genau entsprechend zu beweisende Satz: Wenn die Funktion $f(x)$ für $a \leqq x \leqq b$ stetig ist und für $a < x < b$ eine stetige Ableitung $f'(x)$ besitzt, die bei Annäherung des Punktes $x$ an den Punkt $a$ über alle Grenzen wächst, so wächst auch der vordere Differenzenquotient $\frac{f(a+h) - f(a)}{h}$ bei Abnahme von $h$ zu 0 über alle Grenzen, so daß kein endlicher vorderer Differentialquotient vorhanden ist. Die geometrische Deutung dieses Vorkommnisses ist ein im Endlichen gelegener Kurvenpunkt mit vertikaler Tangente.

## § 3. Verschiedene Einzelheiten.

### 1. Beweis des binomischen Satzes.

Auf Grund unserer Differentiationsregeln ergibt sich für den binomischen Satz ein einfacher Beweis, den ich als Beispiel einer für uns später wichtigen Betrachtungsweise, der „*Methode der unbestimmten Koeffizienten*“, hier anführen will. Wir suchen für beliebiges ganzzahliges positives $n$ eine Entwicklung der Größe $(1 + x)^n$ nach Potenzen von $x$. Wir erkennen sofort, daß die Funktion $(1 + x)^n$ eine ganze rationale Funktion $n$-ten Grades sein muß, d. h. die Gestalt $(1 + x)^n = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ hat, und es handelt sich lediglich um die Bestimmung der Koeffizienten $a_\nu$. Setzen wir $x = 0$, so erhalten wir sofort $a_0 = 1$. Differenzieren wir die linke und die rechte Seite der Gleichung nach $x$ einmal, zweimal, dreimal usw., so erhalten wir die Gleichungen

$$n(1+x)^{n-1} = a_1 + 2a_2x + \cdots + n\,a_n x^{n-1},$$
$$n(n-1)(1+x)^{n-2} = 2a_2 + 3\cdot 2a_3x + \cdots + n(n-1)\,a_n x^{n-2},$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots;$$

da diese Gleichungen für alle Werte von $x$ bestehen, so können wir in jeder von ihnen wiederum $x = 0$ setzen und erhalten so für die Koeffizienten $a_1, a_2, a_3, \ldots$ der Reihe nach die Ausdrücke

$$a_1 = n, \qquad a_2 = \frac{n(n-1)}{1\cdot 2}, \qquad a_3 = \frac{n(n-1)(n-2)}{1\cdot 2\cdot 3}, \quad \ldots,$$
$$a_k = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} = \binom{n}{k},$$

so daß wir schließlich den binomischen Satz in der Gestalt

$$(1+x)^n = 1 + nx + \binom{n}{2}x^2 + \cdots + \binom{n}{k}x^k + \cdots + x^n$$

gewinnen.

## 2. Fortgesetzte Differentiation.

Als Übungsaufgabe überlasse ich dem Leser, im Anschluß hieran zu beweisen, daß sich die mehrfache Differentiation eines Produktes nach der folgenden Formel (*Leibnizsche Regel*) vollzieht:

$$\frac{d^n}{dx^n}(fg) = \frac{d^n}{dx^n}f\cdot g + \binom{n}{1}\frac{d^{n-1}}{dx^{n-1}}f\cdot\frac{d}{dx}g + \binom{n}{2}\frac{d^{n-2}}{dx^{n-2}}f\cdot\frac{d^2}{dx^2}g + \cdots$$
$$+ \binom{n}{n-1}\frac{d}{dx}f\cdot\frac{d^{n-1}}{dx^{n-1}}g + f\cdot\frac{d^n}{dx^n}g$$

oder

$$(fg)^{(n)} = f^{(n)}g + \binom{n}{1}f^{(n-1)}g' + \binom{n}{2}f^{(n-2)}g'' + \cdots$$
$$+ \binom{n}{n-1}f'g^{(n-1)} + f\cdot g^{(n)}.$$

Keinem so übersichtlichen Bildungsgesetz genügt übrigens die fortgesetzte Differentiation einer zusammengesetzten Funktion $y = f(\varphi(x))$. Nach den Differentiationsregeln des vorigen Kapitels (Produktregel und Kettenregel) wird

$$\frac{dy}{dx} = \frac{df}{d\varphi}\cdot\frac{d\varphi}{dx} = f'\cdot\varphi',$$
$$\frac{d^2y}{dx^2} = f''\varphi'^2 + f'\varphi'',$$
$$\frac{d^3y}{dx^3} = f'''\varphi'^3 + 3f''\varphi'\varphi'' + f'\varphi''',$$
$$\dots\dots\dots\dots\dots$$

### 3. Weitere Beispiele für Anwendung der Kettenregel. Verallgemeinerter Mittelwertsatz.

Den Differentialquotienten der Funktion $x^x$ bildet man, indem man setzt $x^x = e^{x \log x}$, worauf man nach der Kettenregel erhält

$$\frac{d}{dx} x^x = x^x (\log x + 1).$$

Ganz ebenso vollzieht sich die Differentiation des allgemeineren Ausdruckes $f(x)^{f(x)} = e^{f(x) \log f(x)}$ nach der Kettenregel folgendermaßen:

$$\frac{d}{dx} f(x)^{f(x)} = f(x)^{f(x)} \cdot f'(x) (\log f(x) + 1).$$

Als weitere Anwendung der Kettenregel gebe ich einen Beweis für den schon in Kap. II, Anhang, § 2 genannten verallgemeinerten Mittelwertsatz der Differentialrechnung, wobei sich dieser Satz unter milderen Voraussetzungen ergibt. Es sei $G(x) = u$ eine im abgeschlossenen Intervall $a \leqq x \leqq b$ monotone stetige und im offenen Intervalle $a < x < b$ differenzierbare Funktion, deren Ableitung $G'(x)$ dort nirgends verschwindet, und es sei $F(x)$ eine in denselben Gebieten stetige bzw. differenzierbare Funktion. Führen wir in $F(x)$ durch die Umkehrfunktion $x = \Phi(u)$ von $G(x)$ anstatt $x$ die Größe $u$ als neue unabhängige Veränderliche ein, d. h. betrachten wir die zusammengesetzte Funktion $f(u) = F(\Phi(u))$, so wird nach der Kettenregel $f'(u) = F'(x)\Phi'(u) = \frac{F'(x)}{G'(x)}$. Der gewöhnliche Mittelwertsatz, angewandt auf die Funktion $f(u)$ und das Intervall zwischen $u_1 = G(a)$ und $u_2 = G(b)$, ergibt nun für einen Zwischenwert $\omega$

$$\frac{f(u_2) - f(u_1)}{u_2 - u_1} = f'(\omega)$$

oder

$$\frac{F(b) - F(a)}{G(b) - G(a)} = \frac{F'(\xi)}{G'(\xi)},$$

wo $\xi = \Phi(\omega)$ ein Zwischenwert zwischen $a$ und $b$ ist.

Viertes Kapitel.

# Weiterer Ausbau der Integralrechnung.

Wir sind im vorigen Kapitel durch Aufstellung der Differentiationsregeln zu einer weitgehenden Beherrschung der Aufgabe gelangt, gegebene Funktionen zu differenzieren. Aber gerade das umgekehrte Problem, das des Integrierens, geht fast überall an Wichtigkeit dem des Differenzierens voran. Demgemäß sind wir nunmehr genötigt, uns mit der Kunst des Integrierens gegebener Funktionen zu befassen.

Das mit Hilfe unserer Differentiationsregeln gewonnene Ergebnis können wir dahin zusammenfassen: *Jede mittels der elementaren Funktionen durch einen „geschlossenen Ausdruck“*[1]) *gebildete Funktion läßt sich differenzieren, und ihr Differentialquotient ist eine Funktion, die sich ebenfalls mit Hilfe der elementaren Funktionen geschlossen ausdrücken läßt.* Dagegen haben wir keine vollständig entsprechende Tatsache für das Integrationsproblem bei elementaren Funktionen kennengelernt. Wir wissen zwar, daß jede elementare Funktion, ja jede stetige Funktion, sich integrieren läßt, wir haben auch zahlreiche elementare Funktionen direkt oder durch Umkehrung von Differentiationsformeln mittels elementarer Funktionen integriert, aber wir sind weit davon entfernt, das folgende Problem allgemein lösen zu können: Gegeben ist eine Funktion $f(x)$, die irgendwie geschlossen mit Hilfe der elementaren Funktionen aufgebaut ist. Gesucht ist ein Ausdruck für ihr unbestimmtes Integral $F(x) = \int f(x)\,dx$, und zwar gesucht in dem Sinne, daß $F(x)$ wiederum durch elementare Funktionen geschlossen ausgedrückt wird.

Es besteht nun sogar die Tatsache, daß eine solche Aufgabe im allgemeinen nicht mehr lösbar ist; keineswegs führt das Integral jeder elementaren Funktion wiederum auf eine elementare Funktion oder ist, wie wir sagen wollen, *„elementar ausführbar“*. Trotzdem aber ist es außerordentlich wichtig, solche Integrationen, wo sie möglich sind, tatsächlich durchführen zu können und überhaupt eine gewisse technische Fertigkeit bei der Integration gegebener Funktionen zu gewinnen.

Hilfsmittel hierfür zu entwickeln, wird die Aufgabe des ersten Teiles dieses Kapitels bilden. Dabei möchte ich gerade den Anfänger ausdrücklich vor dem Versuche warnen, etwa rein gedächtnismäßig die Fülle der Formeln, welche man durch Anwendung dieser technischen Hilfsmittel erhält, sich einzuprägen. Das Bestreben darf lediglich darauf gerichtet sein, die Methoden der Integration innerlich verstehen und handhaben zu lernen.

Zweitens aber werden wir uns in diesem Kapitel — im wesentlichen unabhängig von der Frage der Technik des Integrierens — noch mit einigen mehr prinzipiellen Vertiefungen und Ergänzungen unserer Auffassung von Integration und Integral zu befassen haben.

## § 1. Zusammenstellung der elementaren Integrale.

Ich erinnere vorab nochmals daran, daß jeder der früher bewiesenen Differentiationsformeln eine gleichbedeutende Integrationsformel ent-

[1]) Wir verstehen darunter eine Funktion, welche sich durch mehrmalige Anwendung von Zusammensetzungsprozessen und rationalen Operationen aus den elementaren Funktionen bilden läßt.

Dabei mache ich darauf aufmerksam, daß die Unterscheidung zwischen „elementaren“ Funktionen und anderen etwas an sich recht Willkürliches ist.

spricht. Da diese elementaren Integrale immer wieder als Material der Integrierkunst auftreten, so stellen wir sie in einer Tabelle zusammen. In jeder Zeile dieser Tabelle steht rechts eine elementare Funktion, links ihr Differentialquotient. Lesen wir die Tabelle von links nach rechts, so erhalten wir zu der links stehenden Funktion rechts ein unbestimmtes Integral.

| | $F'(x) = f(x)$ | $F(x) = \int f(x)\,dx$ |
|---|---|---|
| 1 | $x^a \quad (a \neq -1)$ | $\dfrac{x^{a+1}}{a+1}$ |
| 2 | $\dfrac{1}{x}$ | $\log \lvert x \rvert$ |
| 3 | $e^x$ | $e^x$ |
| 4 | $a^x \quad (a \neq 1)$ | $\dfrac{a^x}{\log a}$ |
| 5 | $\sin x$ | $-\cos x$ |
| 6 | $\cos x$ | $\sin x$ |
| 7 | $\dfrac{1}{\sin^2 x}$ | $-\operatorname{ctg} x$ |
| 8 | $\dfrac{1}{\cos^2 x}$ | $\operatorname{tg} x$ |
| 9 | $\mathfrak{Sin}\, x$ | $\mathfrak{Cof}\, x$ |
| 10 | $\mathfrak{Cof}\, x$ | $\mathfrak{Sin}\, x$ |
| 11 | $\dfrac{1}{\mathfrak{Sin}^2 x}$ | $-\mathfrak{Ctg}\, x$ |
| 12 | $\dfrac{1}{\mathfrak{Cof}^2 x}$ | $\mathfrak{Tg}\, x$ |
| 13 | $\dfrac{1}{\sqrt{1-x^2}} \quad (\lvert x \rvert < 1)$ | $\begin{cases} \arcsin x \\ -\arccos x \end{cases}$ |
| 14 | $\dfrac{1}{1+x^2}$ | $\begin{cases} \operatorname{arc\,tg} x \\ -\operatorname{arc\,ctg} x \end{cases}$ |
| 15 | $\dfrac{1}{\sqrt{1+x^2}}$ | $\mathfrak{Ar}\,\mathfrak{Sin}\, x = \log\left(x + \sqrt{1+x^2}\right)$ |
| 16 | $\dfrac{1}{\pm\sqrt{x^2-1}} \quad (\lvert x \rvert > 1)$ | $\mathfrak{Ar}\,\mathfrak{Cof}\, x = \log\left(x \pm \sqrt{x^2-1}\right)$ |
| 17 | $\dfrac{1}{1-x^2} \begin{cases} \lvert x \rvert < 1 \\ \lvert x \rvert > 1 \end{cases}$ | $\mathfrak{Ar}\,\mathfrak{Tg}\, x = \dfrac{1}{2} \log \dfrac{1+x}{1-x}$<br>$\mathfrak{Ar}\,\mathfrak{Ctg}\, x = \dfrac{1}{2} \log \dfrac{x+1}{x-1}$ |

Ich erinnere ferner an die im zweiten Kapitel, § 4, bewiesenen fundamentalen Sätze der Differential- und Integralrechnung; insbesondere an die Tatsache, daß man aus dem unbestimmten Integral $F(x)$ das bestimmte Integral durch die Formel $\int_a^b f(x)\,dx = F(x)\Big|_a^b = F(b) - F(a)$ erhält. Endlich muß man für die Technik des Integrierens die im zweiten Kapitel in § 1 zusammengestellten elementaren Integrationsregeln bereit haben.

In den nächsten Paragraphen werden wir nun versuchen, die Ausführung von Integralen über gegebene Funktionen irgendwie auf die in dieser Tabelle zusammengestellten elementaren Integrale zurückzuführen. Abgesehen von Kunstgriffen, die der Anfänger doch nicht systematisch lernen kann, die vielmehr Sache längerer Erfahrung sein müssen, beruht eine solche Zurückführung immer wesentlich auf der Verwendung von zwei Hilfsmitteln. Jedes dieser beiden erlaubt uns, in mannigfacher Weise ein gegebenes Integral umzuformen, und das Ziel solcher Umformungen wird sein, mit einem Schritt oder mit einer Reihe von Schritten eine gegebene Integrationsaufgabe auf eine oder mehrere der oben angeschriebenen elementaren Integralformeln zu reduzieren.

## § 2. Die Substitutionsregel.

Das erste dieser Hilfsmittel für die Behandlung von Integrationsproblemen ist die Einführung einer neuen Veränderlichen (*Substitution* oder *Transformation*). Die entsprechende Integralformel ist nichts anderes als die Kettenregel der Differentialrechnung, in Integralform ausgedrückt.

### 1. Die Substitutionsformel.

Wir denken uns in einer Funktion $F(x)$ durch die Gleichung $x = \varphi(u)$ eine neue Veränderliche $u$ eingeführt, so daß $F(x)$ eine mittelbare Funktion von $u$ wird:

$$F(x) = F(\varphi(u)) = G(u).$$

Die Kettenregel der Differentialrechnung besagt nun

$$\frac{dG}{du} = \frac{dF}{dx}\varphi'(u).$$

Schreiben wir jetzt

$$F'(x) = f(x) \quad \text{und} \quad G'(u) = g(u)$$

oder damit gleichbedeutend

$$F(x) = \int f(x)\,dx \quad \text{und} \quad G(u) = \int g(u)\,du,$$

so geht einerseits die Kettenregel über in

$$g(u) = f(x)\,\varphi'(u),$$

anderseits ist nach Definition

$$G(u) = F(x), \quad \text{d. h.} \quad \int g(u)\,du = \int f(x)\,dx,$$

und wir erhalten die mit der Kettenregel äquivalente Integralformel

$$\int f(\varphi(u))\,\varphi'(u)\,du = \int f(x)\,dx \qquad (x = \varphi(u)).$$

*Diese Formel ist die Grundlage für die Einführung neuer Veränderlicher in ein Integral.* Sie besagt: Wenn wir ein unbestimmtes Integral über eine Funktion von $u$ zu bilden haben, welche in der besonderen Gestalt $f(\varphi(u))\,\varphi'(u)$ gegeben ist, so können wir statt dessen die Funktion $f(x)$ der Integrationsvariabeln $x$ unbestimmt nach $x$ integrieren und haben nach Ausführung dieser letzten Integration $x = \varphi(u)$ zu setzen und damit wieder $u$ einzuführen.

Wenden wir die Formel z. B. auf den Integranden $\frac{\varphi'(u)}{\varphi(u)}$ an, so erhalten wir sofort

$$\int \frac{\varphi'(u)}{\varphi(u)}\,du = \int \frac{dx}{x} = \log|x| = \log|\varphi(u)|$$

oder, indem wir jetzt wieder $x$ statt $u$ schreiben,

$$\int \frac{\varphi'(x)}{\varphi(x)}\,dx = \log|\varphi(x)|.$$

Setzen wir in diese wichtige Formel beispielsweise $\varphi(x) = \log x$ oder $\varphi(x) = \sin x$ oder $\varphi(x) = \cos x$ ein, so erhalten wir

$$\int \frac{dx}{x \log x} = \log|\log x|,$$

$$\int \operatorname{ctg} x\,dx = \log|\sin x|, \qquad \int \operatorname{tg} x\,dx = -\log|\cos x|.\text{[1]}$$

Ein weiteres Beispiel ist das Integral

$$\int \varphi(u)\,\varphi'(u)\,du = \int x\,dx = \frac{x^2}{2} = \frac{1}{2}[\varphi(u)]^2,$$

wo also $f(x) = x$ wird. Es ergibt sich so z. B. für $\varphi(u) = \log u$

$$\int \frac{\log u}{u}\,du = \frac{1}{2}(\log u)^2.$$

Endlich betrachten wir das Beispiel

$$\int \sin^n u \cos u\,du.$$

Hier ist $x = \sin u = \varphi(u)$, und es ergibt sich

$$\int \sin^n u \cos u\,du = \int x^n\,dx = \frac{x^{n+1}}{n+1} = \frac{\sin^{n+1} u}{n+1}.$$

In vielen Fällen wollen wir nun aber unsere obige allgemeine Formel in der umgekehrten Richtung anwenden, indem wir von der rechten Seite, dem Integrale $\int f(x)\,dx$, ausgehen. Jetzt handelt es sich also darum, ein vorgelegtes unbestimmtes Integral $F(x) = \int f(x)\,dx$ zu

[1]) Man bestätige diese und die folgenden Formeln, indem man zeigt, daß durch Differentiation des Ergebnisses der Integrand entsteht. Übrigens werden die Formeln natürlich nur behauptet, soweit die darin vorkommenden Ausdrücke Sinn haben.

berechnen oder zu vereinfachen, indem wir durch die Transformationsgleichung $x = \varphi(u)$ die neue Integrationsvariable $u$ einführen, demgemäß dann das unbestimmte Integral $G(u) = \int f(\varphi(u))\,\varphi'(u)\,du$ aufsuchen und schließlich in diesem die Variable $u$ durch $x$ ausdrücken. Um diesen letzten Schritt ausführen zu können, müssen wir eine Sicherheit haben, daß tatsächlich dem Werte $x$ ein bestimmter Wert $u$ zugehört, d. h. daß die Funktion $x = \varphi(u)$ sich umkehren läßt. Demgemäß setzen wir jetzt, indem wir die Variable $x$ als das Primäre ansehen, folgendes voraus: In dem betrachteten Intervalle sei $u = \psi(x)$ eine monotone differenzierbare Funktion mit dem Differentialquotienten $\psi'(x)$, der nirgends im Intervalle verschwindet. Die Umkehrfunktion — ihre eindeutige Bestimmtheit ist eine Folge der Voraussetzung — nennen wir $x = \varphi(u)$; ihre Ableitung ist dann gegeben durch $\varphi'(u) = \frac{1}{\psi'(x)}$. Als *Grundformel* für die Einführung einer neuen Veränderlichen $u$ in ein Integral erhalten wir daher

$$\int f(x)\,dx = \int f(\varphi(u))\,\varphi'(u)\,du \qquad (u = \psi(x)).$$

*Man erhält das unbestimmte Integral $\int f(x)\,dx$, indem man das unbestimmte Integral $\int f(\varphi(u))\,\varphi'(u)\,du$ bildet und am Schlusse anstatt $u$ wieder $x$ als Variable durch die Gleichung $u = \psi(x)$ einführt.* Es ist also nicht etwa so, daß man einfach in der zu integrierenden Funktion die alte Veränderliche $x$ durch die neue $u$ ausdrückt und dann nach dieser neuen Veränderlichen integriert; sondern man muß vor der Integration noch mit der Ableitung der ursprünglichen Veränderlichen $x$ nach der neuen Veränderlichen $u$ multiplizieren.

Als Integralformel für die bestimmte Integration zwischen zwei Grenzen haben wir entsprechend zu schreiben

$$\int_a^b f(x)\,dx = \int_{\psi(a)}^{\psi(b)} f(\varphi(u))\,\varphi'(u)\,du;$$

d. h. *wir haben in dem neuen Integral diejenigen Integrationsgrenzen zu wählen, welche durch die Transformation $x = \varphi(u)$, $u = \psi(x)$ den alten Integrationsgrenzen entsprechen.*

Meistens wird bei den Anwendungen der Integrand $f(x)$ von vornherein als zusammengesetzte Funktion erscheinen, etwa $f(x) = h(u)$, wo $u = \psi(x)$ ist. Dann ist es bequemer, unsere Integralformel ein wenig anders zu schreiben, indem wir den Ausdruck $f(\varphi(u))$ mit der Funktion $h(u)$ identifizieren. Machen wir für $u$ die Substitution $u = \psi(x)$, $x = \varphi(u)$, so lautet unsere Transformationsformel einfach

$$\int h(\psi(x))\,dx = \int h(u)\,\frac{dx}{du}\,du.$$

Als ein erstes Beispiel betrachten wir die Integration der Funktion $f(x) = \sin 2x$ mit $u = \psi(x) = 2x$ und $h(u) = \sin u$. Es ist dabei

$$\frac{du}{dx} = \psi'(x) = 2.$$

Führen wir $u = 2x$ in dem Integral als neue Veränderliche ein, so geht das Integral nicht über in $\int \sin u\, du$, sondern in

$$\frac{1}{2}\int \sin u\, du = -\frac{1}{2}\cos u = -\frac{1}{2}\cos 2x,$$

wie man im übrigen natürlich sofort durch Differenzieren der rechten Seite bestätigen kann.

Integrieren wir zwischen den Grenzen $x = 0$ und $x = \frac{\pi}{4}$, so sind die entsprechenden Grenzen $u = 0$ und $u = \frac{\pi}{2}$; wir erhalten

$$\int_0^{\frac{\pi}{4}} \sin 2x\, dx = \frac{1}{2}\int_0^{\frac{\pi}{2}} \sin u\, du = -\frac{1}{2}\cos u \Big|_0^{\frac{\pi}{2}} = \frac{1}{2}.$$

Ein anderes einfaches Beispiel ist das Integral $\int_1^4 \frac{dx}{\sqrt{x}}$. Hier nehmen wir $u = \psi(x) = \sqrt{x}$. Es wird also $x = \varphi(u) = u^2$. Da weiter $\varphi'(u) = 2u$ ist, so ergibt sich

$$\int_1^4 \frac{dx}{\sqrt{x}} = 2\int_1^2 \frac{u\, du}{u} = 2\int_1^2 du = 2.$$

## 2. Neuer Beweis der Substitutionsformel.

Wir können uns im übrigen unsere Integralformel auch noch auf eine andere, direkte Weise klar machen, indem wir auf die Formel für die bestimmte Integration abzielen und uns auf die Bedeutung des bestimmten Integrales als Grenzwert einer Summe stützen. Um (etwa im Falle $a < b$) das Integral

$$\int_a^b h(\psi(x))\, dx$$

zu berechnen, können wir in dem Intervall $a \leqq x \leqq b$ eine beliebige Einteilung zugrunde legen, die wir dann immer weiter verfeinern. Wir wählen diese Einteilung folgendermaßen. Dem Intervall $a \leqq x \leqq b$ der $x$-Achse entspricht, wenn $u = \psi(x)$ als monoton wachsend vorausgesetzt ist, in umkehrbar eindeutiger Weise ein Intervall $\alpha \leqq u \leqq \beta$ für die zugehörigen $u$-Werte $u = \psi(x)$, wobei $\alpha = \psi(a)$, $\beta = \psi(b)$ gesetzt ist. Dieses $u$-Intervall teilen wir in $n$ Teile der Länge $\Delta u$[1]) ein; dieser Einteilung entspricht eine nun im allgemeinen nicht mehr

[1]) Die Annahme der Gleichheit dieser Teilintervalle ist übrigens keineswegs wesentlich für den Beweis.

äquidistante Einteilung des $x$-Intervalles. Wir bezeichnen dessen Teilpunkte mit $a, x_1, \ldots, x_n = b$, die zugehörigen Intervallängen mit

$$\Delta x_1, \Delta x_2, \ldots, \Delta x_n.$$

Das betrachtete Integral ist nunmehr der Grenzwert[1]) der Summe

$$\sum_{\nu=1}^{n} h(\psi(\xi_\nu))\,\Delta x_\nu,$$

wo wir die Werte $\xi_\nu$ ganz beliebig im $\nu$-ten Teilintervall der $x$-Einteilung wählen können. Diese Summe schreiben wir jetzt $\sum\limits_{\nu=1}^{n} h(u_\nu)\frac{\Delta x_\nu}{\Delta u}\Delta u$, wo $u_\nu = \psi(\xi_\nu)$ gesetzt ist. Nun ist nach dem Mittelwertsatz der Differentialrechnung $\frac{\Delta x_\nu}{\Delta u} = \varphi'(\eta_\nu)$, wo $\eta_\nu$ ein passend zu wählender Zwischenwert der Variabeln $u$ im $\nu$-ten Teilintervall der $u$-Einteilung ist und $x = \varphi(u)$ wiederum die Umkehrfunktion von $u = \psi(x)$ bezeichnet. Verfügen wir nun über die Werte $\xi_\nu$ so, daß gerade $\xi_\nu$ und $\eta_\nu$ zusammengehören, d. h., daß $\xi_\nu = \varphi(\eta_\nu)$, $\eta_\nu = \psi(\xi_\nu)$ ist, so erhält unsere Summe die Gestalt

$$\sum_{\nu=1}^{n} h(\eta_\nu)\,\varphi'(\eta_\nu)\,\Delta u.$$

Machen wir hier den Grenzübergang, so erhalten wir unmittelbar als Grenzwert, d. h. als den Wert des betrachteten Integrals

$$\int_{\alpha}^{\beta} h(u)\frac{dx}{du}\,du$$

in Übereinstimmung mit der oben gegebenen Formel.

### 3. Beispiele. Integrationsformeln.

Mit Hilfe der Substitutionsregel können wir in vielen Fällen ein gegebenes Integral $\int f(x)\,dx$ auswerten, indem wir es durch eine geeignete Substitution $x = \varphi(u)$ auf eines der elementaren Integrale unserer Tabelle zurückführen. Ob es solche Substitutionen gibt und wie man sie findet, darüber lassen sich keine allgemeinen Aussagen machen; vielmehr ist hier ein Punkt, wo Übung und Geschicklichkeit gegenüber der systematischen Methode zu ihrem Rechte kommen.

Das Integral $\int \frac{dx}{\sqrt{a^2 - x^2}}$ berechnen wir z. B. vermittelst der Substitution[2]) $x = \varphi(u) = a u$, $u = \psi(x) = \frac{x}{a}$, $dx = a\,du$, durch welche wir

---

[1]) Dieser Grenzwert (für $\Delta u \to 0$) existiert und ist das Integral, weil wegen der gleichmäßigen Stetigkeit von $x = \varphi(u)$ mit $\Delta u \to 0$ auch das größte der $\Delta x$ gegen Null konvergiert.

[2]) Wir nehmen uns die Freiheit, der Kürze halber die Symbole $dx$ und $du$ getrennt zu schreiben, also $dx = \varphi'(u)\,du$ statt $\frac{dx}{du} = \varphi'(u)$ (vgl. S. 84 bis 85).

mit Rücksicht auf Nr. 13 der Tabelle erhalten:

$$\int \frac{dx}{\sqrt{a^2 - x^2}} = \int \frac{a\,du}{a\sqrt{1-u^2}} = \arcsin u = \arcsin \frac{x}{a} \quad \text{für} \quad |x| < |a|.$$

Ebenso ergibt sich durch dieselbe Substitution

$$\int \frac{dx}{a^2 + x^2} = \int \frac{a\,du}{a^2(1+u^2)} = \frac{1}{a} \operatorname{arc\,tg} u = \frac{1}{a} \operatorname{arc\,tg} \frac{x}{a},$$

$$\int \frac{dx}{\sqrt{a^2 + x^2}} = \mathfrak{Ar\,Sin} \frac{x}{a},$$

$$\int \frac{dx}{\sqrt{x^2 - a^2}} = \mathfrak{Ar\,Cof} \frac{x}{a} \quad \text{für} \quad |x| > |a|,$$

$$\int \frac{dx}{a^2 - x^2} = \begin{cases} \dfrac{1}{a} \mathfrak{Ar\,Tg} \dfrac{x}{a} & \text{für} \quad |x| < |a| \\[2ex] \dfrac{1}{a} \mathfrak{Ar\,Ctg} \dfrac{x}{a} & \text{für} \quad |x| > |a| \end{cases};$$

Formeln, die sehr häufig vorkommen und die im übrigen sich leicht durch Differentiation der rechten Seite bestätigen lassen.

Zum Schluß will ich noch einmal folgenden Umstand ausdrücklich hervorheben. Wir hatten bei unserer Substitution die Voraussetzung gemacht, daß in dem betrachteten Intervall überall $\psi'(x) \neq 0$ ist, woraus sich die eindeutige Umkehrbarkeit der Substitution durch die Gleichung $x = \varphi(u)$ ergibt. Wenn unsere Voraussetzung nicht erfüllt ist, kann man bei der Anwendung der Substitutionsformel leicht zu Fehlschlüssen verleitet werden. Will man solchen Schwierigkeiten entgehen, wenn im Integrationsintervall an einzelnen Stellen $\psi'(x) = 0$ ist, so muß man dieses Intervall in einzelne Teile einteilen, so daß nur an den Endpunkten der Teile $\psi'(x) = 0$ sein kann, und dann die Substitution für jedes Teilgebiet einzeln vornehmen.

## § 3. Weitere Beispiele zur Substitutionsmethode.

In den folgenden Zeilen stelle ich kurz eine Reihe von weiteren Beispielen zusammen, die der Leser als Übungsmaterial sorgfältig durchdenken mag.

Durch die Substitutionen $u = 1 \pm x^2$, $du = \pm 2x\,dx$ folgt

$$\int \frac{x\,dx}{\sqrt{1 \pm x^2}} = \pm \sqrt{1 \pm x^2},$$

$$\int \frac{x\,dx}{1 \pm x^2} = \pm \frac{1}{2} \log |1 \pm x^2|.$$

In diesen Formeln hat an allen drei Stellen entweder stets $+$ oder stets $-$ zu stehen.

Durch $u = ax + b$, $du = a\,dx$ $(a \neq 0)$ erhalten wir

$$\int \frac{dx}{ax + b} = \frac{1}{a} \log |ax + b|,$$

$$\int (ax+b)^\alpha dx = \frac{1}{a(\alpha+1)}(ax+b)^{\alpha+1} \qquad (\alpha \neq -1),$$

$$\int \sin(ax+b)\,dx = -\frac{1}{a}\cos(ax+b);$$

ebenso wird mittels $u = \cos x$, $du = -\sin x\,dx$

$$\int \operatorname{tg} x\,dx = -\log|\cos x|$$

und mittels $u = \sin x$, $du = \cos x\,dx$

$$\int \operatorname{ctg} x\,dx = \log|\sin x|$$

(vgl. S. 169). Durch die ganz analogen Substitutionen $u = \mathfrak{Cos}\, x$, $du = \mathfrak{Sin}\, x\,dx$ und $u = \mathfrak{Sin}\, x$, $du = \mathfrak{Cos}\, x\,dx$ erhalten wir die Formeln

$$\int \mathfrak{Tg}\, x\,dx = \log|\mathfrak{Cos}\, x|$$

und

$$\int \mathfrak{Ctg}\, x\,dx = \log|\mathfrak{Sin}\, x|.$$

Vermöge der Substitution $u = \frac{a}{b}\operatorname{tg} x$, $du = \frac{a}{b}\cdot\frac{dx}{\cos^2 x}$ gelangen wir zu den beiden Formeln

$$\int \frac{dx}{a^2\sin^2 x + b^2\cos^2 x} = \frac{1}{b^2}\int \frac{1}{\frac{a^2}{b^2}\operatorname{tg}^2 x + 1}\cdot\frac{dx}{\cos^2 x} = \frac{1}{ab}\operatorname{arc\,tg}\left(\frac{a}{b}\operatorname{tg} x\right)$$

und

$$\int \frac{dx}{a^2\sin^2 x - b^2\cos^2 x} = -\frac{1}{ab}\,\mathfrak{Ar\,Tg}\left(\frac{a}{b}\operatorname{tg} x\right) \text{ bzw. } = -\frac{1}{ab}\,\mathfrak{Ar\,Ctg}\left(\frac{a}{b}\operatorname{tg} x\right).$$

Das Integral

$$\int \frac{dx}{\sin x}$$

berechnen wir, indem wir $\sin x = 2\sin\frac{x}{2}\cos\frac{x}{2} = 2\operatorname{tg}\frac{x}{2}\cos^2\frac{x}{2}$ schreiben und $u = \operatorname{tg}\frac{x}{2}$, also $du = \frac{dx}{2\cos^2\frac{x}{2}}$ setzen; es wird dann

$$\int \frac{dx}{\sin x} = \int \frac{du}{u} = \log\left|\operatorname{tg}\frac{x}{2}\right|.$$

Ersetzen wir in dieser Formel $x$ durch $x + \frac{\pi}{2}$, so geht sie über in

$$\int \frac{dx}{\cos x} = \log\left|\operatorname{tg}\left(\frac{x}{2} + \frac{\pi}{4}\right)\right|.$$

Die Substitution $u = 2x$ liefert in Anbetracht der bekannten trigonometrischen Formeln $2\cos^2 x = 1 + \cos 2x$ und $2\sin^2 x = 1 - \cos 2x$ die oft gebrauchten Formeln

$$\int \cos^2 x\,dx = \frac{1}{2}(x + \sin x\cos x)$$

und

$$\int \sin^2 x\,dx = \frac{1}{2}(x - \sin x\cos x).$$

Durch die Substitution $x = \cos u$, also $u = \operatorname{arc} \cos x$, oder allgemeiner $x = a \cos u$ $(a \neq 0)$ werden

$$\int \sqrt{1 - x^2}\, dx \quad \text{bzw.} \quad \int \sqrt{a^2 - x^2}\, dx$$

auf diese Formeln zurückgeführt. Wir erhalten so

$$\int \sqrt{a^2 - x^2}\, dx = -\frac{a^2}{2} \operatorname{arc} \cos \frac{x}{a} + \frac{x}{2} \sqrt{a^2 - x^2}.$$

Ganz entsprechend erhalten wir auch durch die Substitution $x = a \operatorname{\mathfrak{Cof}} u$ die Formel

$$\int \sqrt{x^2 - a^2}\, dx = -\frac{a^2}{2} \operatorname{\mathfrak{Ar}} \operatorname{\mathfrak{Cof}} \frac{x}{a} + \frac{x}{2} \sqrt{x^2 - a^2}$$

und durch die Substitution $x = a \operatorname{\mathfrak{Sin}} u$

$$\int \sqrt{a^2 + x^2}\, dx = \frac{a^2}{2} \operatorname{\mathfrak{Ar}} \operatorname{\mathfrak{Sin}} \frac{x}{a} + \frac{x}{2} \sqrt{a^2 + x^2}.$$

Die Substitution $u = \frac{a}{x}$, $dx = -\frac{a}{u^2} du$ führt uns auf die Formeln:

$$\int \frac{dx}{x \sqrt{x^2 - a^2}} = -\frac{1}{a} \operatorname{arc} \sin \frac{a}{x},$$

$$\int \frac{dx}{x \sqrt{x^2 + a^2}} = -\frac{1}{a} \operatorname{\mathfrak{Ar}} \operatorname{\mathfrak{Sin}} \frac{a}{x},$$

$$\int \frac{dx}{x \sqrt{a^2 - x^2}} = -\frac{1}{a} \operatorname{\mathfrak{Ar}} \operatorname{\mathfrak{Cof}} \frac{a}{x}.$$

Wir betrachten schließlich noch die drei Integrale

$$\int \sin mx \sin nx\, dx, \qquad \int \sin mx \cos nx\, dx, \qquad \int \cos mx \cos nx\, dx,$$

wobei $m$ und $n$ positive ganze Zahlen sind. Nach bekannten trigonometrischen Formeln können wir diese Integrale je in zwei zerlegen, indem wir schreiben:

$$\sin mx \sin nx = \frac{1}{2} (\cos (m - n) x - \cos (m + n) x),$$

$$\sin mx \cos nx = \frac{1}{2} (\sin (m + n) x + \sin (m - n) x),$$

$$\cos mx \cos nx = \frac{1}{2} (\cos (m + n) x + \cos (m - n) x).$$

Führen wir nun die Substitutionen $u = (m + n) x$ bzw. $u = (m - n) x$ ein, so ergibt sich ohne weiteres das Formelsystem:

$$\int \sin mx \sin nx\, dx = \begin{cases} \frac{1}{2} \left( \frac{\sin (m-n) x}{m - n} - \frac{\sin (m+n) x}{m + n} \right), & \text{wenn } m \neq n, \\ \frac{1}{2} \left( x - \frac{\sin 2mx}{2m} \right), & \text{wenn } m = n, \end{cases}$$

$$\int \sin mx \cos nx\, dx = \begin{cases} -\frac{1}{2}\left(\frac{\cos(m+n)x}{m+n} + \frac{\cos(m-n)x}{m-n}\right), & \text{wenn } m \neq n, \\ -\frac{1}{2}\cdot\frac{\cos 2mx}{2m}, & \text{wenn } m = n, \end{cases}$$

$$\int \cos mx \cos nx\, dx = \begin{cases} \frac{1}{2}\left(\frac{\sin(m+n)x}{m+n} + \frac{\sin(m-n)x}{m-n}\right), & \text{wenn } m \neq n, \\ \frac{1}{2}\left(\frac{\sin 2mx}{2m} + x\right), & \text{wenn } m = n. \end{cases}$$

Integrieren wir insbesondere von $-\pi$ bis $+\pi$, so erhalten wir aus diesen Formeln die überaus wichtigen Beziehungen

$$\int_{-\pi}^{+\pi} \sin mx \sin nx\, dx = \begin{cases} 0, & \text{wenn } m \neq n, \\ \pi, & \text{wenn } m = n, \end{cases}$$

$$\int_{-\pi}^{+\pi} \sin mx \cos nx\, dx = 0,$$

$$\int_{-\pi}^{+\pi} \cos mx \cos nx\, dx = \begin{cases} 0, & \text{wenn } m \neq n, \\ \pi, & \text{wenn } m = n. \end{cases}$$

Es sind dies die „*Orthogonalitätsrelationen*" der trigonometrischen Funktionen, mit denen wir uns noch im neunten Kapitel zu beschäftigen haben werden.

## § 4. Die Produktintegration.

Das zweite technische Hilfsmittel für die Behandlung von Integrationsproblemen liefert uns die Formel für die Differentiation eines Produktes: $(fg)' = f'g + g'f$.

### 1. Allgemeines.

Schreiben wir diese Formel als Integralformel, so erhalten wir (vgl. drittes Kapitel, § 2, Nr. 1)

$$f(x)\,g(x) = \int g(x)\,f'(x)\,dx + \int f(x)\,g'(x)\,dx$$

oder

$$\int f(x)\,g'(x)\,dx = f(x)\,g(x) - \int g(x)\,f'(x)\,dx.$$

Diese Formel wird als die *Formel der Produktintegration* (auch *partielle Integration* oder *Teilintegration*) bezeichnet. Sie gestattet die Zurückführung eines Integrales auf ein anderes. Wenn man nämlich in einem Integrale $\int \omega(x)\,dx$ den Integranden in ein Produkt $\omega(x) = f(x)\,\varphi(x)$ zerlegt hat und wenn man von dem einen Faktor $\varphi(x)$ das unbestimmte Integral

$$g(x) = \int \varphi(x)\,dx$$

elementar ausführen kann, so ist $\varphi(x) = g'(x)$, und dann wird durch unsere Formel das Integral $\int \omega(x)\,dx = \int f(x)\,\varphi(x)\,dx = \int f(x)\,g'(x)\,dx$

auf das andere $\int g(x) f'(x)\,dx$ reduziert, das unter Umständen einfacher zu behandeln sein kann als das erste. Da man eine als Integrand auftretende gegebene Funktion $\omega(x)$ in mannigfacher Weise als ein Produkt $f(x)\,\varphi(x) = f(x)\,g'(x)$ auffassen kann, so bietet uns diese Formel eine weitreichende Handhabe zur Umformung von Integralen.

Als Formel der bestimmten Integration geschrieben, lautet die Formel der Produktintegration

$$\int_a^b f(x)\,g'(x)\,dx = f(x)\,g(x)\Big|_a^b - \int_a^b g(x)\,f'(x)\,dx$$
$$= f(b)\,g(b) - f(a)\,g(a) - \int_a^b g(x)\,f'(x)\,dx.$$

Denn wir brauchen nur in der unbestimmten Integralformel auf beiden Seiten für die auftretende Variable das eine Mal den Wert $x = b$, das andere Mal den Wert $x = a$ einzusetzen und die Differenz der entstehenden Ausdrücke zu bilden, um gemäß dem zweiten Kapitel, § 4, aus der Formel für die unbestimmte Integration eine solche für die bestimmte Integration zu erhalten.

Zur ersten Erläuterung diene folgendes Beispiel:

$$\int \log x\,dx = \int \log x \cdot 1 \cdot dx.$$

Wir deuten durch diese Schreibweise an, daß wir $f(x) = \log x$ und $g'(x) = 1$ setzen wollen, so daß wir also $f'(x) = \frac{1}{x}$ und etwa $g(x) = x$ haben. Die rechte Seite unserer Formel geht dann über in

$$\int \log x\,dx = x \log x - \int \frac{x}{x}\,dx = x \log x - x.$$

Dieser letztere Ausdruck ist also das Integral des Logarithmus, wie man nunmehr auch unmittelbar durch Differenzieren bestätigt.

### 2. Beispiele.

Folgende weiteren Beispiele mögen dem Leser die Handhabung der Methode näherbringen.

Setzen wir $f(x) = x$, $g'(x) = e^x$, so erhalten wir $f'(x) = 1$, $g(x) = e^x$,

$$\int x e^x\,dx = e^x (x - 1);$$

ganz ähnlich ergibt sich

$$\int x \sin x\,dx = -x \cos x + \sin x$$

und

$$\int x \cos x\,dx = x \sin x + \cos x.$$

Für $f(x) = \log x$, $g'(x) = x^a$ folgt die Relation

$$\int x^a \log x\,dx = \frac{x^{a+1}}{a+1}\left(\log x - \frac{1}{a+1}\right).$$

Hierbei muß $a \neq -1$ vorausgesetzt werden. Für $a = -1$ erhalten wir (vgl. S. 169)

$$\int \frac{1}{x} \log x \, dx = (\log x)^2 - \int \log x \cdot \frac{dx}{x},$$

also, wenn wir das Integral von der rechten Seite auf die linke bringen,

$$\int \frac{1}{x} \log x \, dx = \frac{1}{2} (\log x)^2.$$

Das Integral $\int \arcsin x \, dx$ berechnen wir, indem wir $f(x) = \arcsin x$, $g'(x) = 1$ setzen. Es wird dann

$$\int \arcsin x \, dx = x \arcsin x - \int \frac{x \, dx}{\sqrt{1-x^2}};$$

wir können nach § 3 die Integration rechts sofort ausführen und erhalten

$$\int \arcsin x \, dx = x \arcsin x + \sqrt{1-x^2}.$$

Auf die gleiche Art berechnet sich auch das Integral

$$\int \operatorname{arc\,tg} x \, dx = x \operatorname{arc\,tg} x - \frac{1}{2} \log (1 + x^2)$$

wie auch viele ähnlich gebauten.

Einen etwas anderen Charakter zeigen folgende Beispiele, bei denen man nach zweimaliger Anwendung der Produktintegration wieder auf das Ausgangsintegral zurückkommt und so zu einer Gleichung für dasselbe gelangt.

Durch zweimalige partielle Integration erhalten wir

$$\begin{aligned}\int e^{ax} \sin bx \, dx &= -\frac{1}{b} e^{ax} \cos bx + \frac{a}{b} \int e^{ax} \cos bx \, dx \\ &= -\frac{1}{b} e^{ax} \cos bx + \frac{a}{b^2} e^{ax} \sin bx - \frac{a^2}{b^2} \int e^{ax} \sin bx \, dx\end{aligned}$$

und daher, indem wir nun aus dieser Gleichung das Ausgangsintegral $\int e^{ax} \sin bx \, dx$ berechnen:

$$\int e^{ax} \sin bx \, dx = \frac{1}{a^2 + b^2} e^{ax} (a \sin bx - b \cos bx).$$

Auf die gleiche Art folgt auch

$$\int e^{ax} \cos bx \, dx = \frac{1}{a^2 + b^2} e^{ax} (a \cos bx + b \sin bx).$$

### 3. Rekursionsformeln.

In vielen Fällen hängt der Integrand außer von der unabhängigen Veränderlichen noch von einem ganzzahligen Index $n$ ab, und es gelingt durch Produktintegration, das Integral zwar nicht sofort auszuwerten, aber doch auf ein solches derselben Form, nur mit einem kleineren Wert des Index $n$ zurückzuführen, bis wir schließlich nach einer Anzahl von Schritten zu einem Integral gelangen, welches wir vermöge unserer Integrationstabelle beherrschen. Ein solches Verfahren nennt man *Rekursionsverfahren*. Ich erläutere es an einigen Beispielen: Durch wiederholte Anwendung der Produktintegration können wir die trigonometrischen Integrale

$$\int \cos^n x\,dx, \qquad \int \sin^n x\,dx, \qquad \int \sin^m x \cos^n x\,dx$$

berechnen, wenn $m$ und $n$ positive ganze Zahlen bedeuten. Es wird nämlich z. B.

$$\int \cos^n x\,dx = \cos^{n-1} x \sin x + (n-1) \int \cos^{n-2} x \sin^2 x\,dx;$$

die rechte Seite können wir auch in die Form

$$\cos^{n-1} x \sin x + (n-1) \int \cos^{n-2} x\,dx - (n-1) \int \cos^n x\,dx$$

setzen und erhalten daraus die Rekursionsformel

$$\int \cos^n x\,dx = \frac{1}{n} \cos^{n-1} x \sin x + \frac{n-1}{n} \int \cos^{n-2} x\,dx.$$

Diese Formel gestattet uns, den Exponenten des Integranden immer weiter zu vermindern, bis wir endlich zum Integral

$$\int \cos x\,dx = \sin x \quad \text{bzw.} \quad \int dx = x$$

gelangen, je nachdem $n$ ungerade oder gerade ist. Ebenso ergeben sich auch die analogen Rekursionsformeln

$$\int \sin^n x\,dx = -\frac{1}{n} \sin^{n-1} x \cos x + \frac{n-1}{n} \int \sin^{n-2} x\,dx$$

und

$$\int \sin^m x \cos^n x\,dx = \frac{\sin^{m+1} x \cos^{n-1} x}{m+n} + \frac{n-1}{m+n} \int \sin^m x \cos^{n-2} x\,dx.$$

Insbesondere gestatten diese Formeln, die Integrale

$$\int \sin^2 x\,dx = \frac{1}{2}(x - \sin x \cos x)$$

und

$$\int \cos^2 x\,dx = \frac{1}{2}(x + \sin x \cos x)$$

zu berechnen, was wir schon früher vermöge der Substitutionsmethode taten.

Es braucht wohl kaum erwähnt zu werden, daß sich die entsprechenden Integrale für die Hyperbelfunktionen ganz auf dieselbe Art berechnen lassen.

Weitere Rekursionsformeln liefern uns die folgenden Umformungen:

$$\int (\log x)^m\,dx = x (\log x)^m - m \int (\log x)^{m-1}\,dx,$$

$$\int x^m e^x\,dx = x^m e^x - m \int x^{m-1} e^x\,dx,$$

$$\int x^m \sin x\,dx = -x^m \cos x + m \int x^{m-1} \cos x\,dx,$$

$$\int x^m \cos x\,dx = x^m \sin x - m \int x^{m-1} \sin x\,dx,$$

$$\int x^a (\log x)^m\,dx = \frac{x^{a+1} (\log x)^m}{a+1} - \frac{m}{a+1} \int x^a (\log x)^{m-1}\,dx \quad (a \neq -1).$$

## 4. Die Wallissche Produktzerlegung von $\pi$.

Die Rekursionsformel für das Integral $\int \sin^n x\,dx$ führt in elementarer Weise zu einer höchst bemerkenswerten Darstellung der Zahl $\pi$ durch ein unendliches Produkt. Setzen wir für $n > 1$ in die Formel

$$\int \sin^n x\,dx = -\frac{1}{n}\sin^{n-1} x \cos x + \frac{n-1}{n}\int \sin^{n-2} x\,dx$$

die Grenzen 0 und $\frac{\pi}{2}$ ein, so ergibt sich

$$\int_0^{\frac{\pi}{2}} \sin^n x\,dx = \frac{n-1}{n}\int_0^{\frac{\pi}{2}} \sin^{n-2} x\,dx \qquad \text{für } n > 1.$$

Wenden wir nun auf das Integral der rechten Seite wieder die Rekursionsformel an und fahren so fort, so erhalten wir — wenn wir noch die Fälle $n = 2m$ und $n = 2m+1$ unterscheiden —:

$$\int_0^{\frac{\pi}{2}} \sin^{2m} x\,dx = \frac{2m-1}{2m}\cdot\frac{2m-3}{2m-2}\cdots\frac{1}{2}\cdot\int_0^{\frac{\pi}{2}} dx,$$

$$\int_0^{\frac{\pi}{2}} \sin^{2m+1} x\,dx = \frac{2m}{2m+1}\cdot\frac{2m-2}{2m-1}\cdots\frac{2}{3}\cdot\int_0^{\frac{\pi}{2}} \sin x\,dx,$$

somit

$$\int_0^{\frac{\pi}{2}} \sin^{2m} x\,dx = \frac{2m-1}{2m}\cdot\frac{2m-3}{2m-2}\cdots\frac{1}{2}\cdot\frac{\pi}{2},$$

$$\int_0^{\frac{\pi}{2}} \sin^{2m+1} x\,dx = \frac{2m}{2m+1}\cdot\frac{2m-2}{2m-1}\cdots\frac{2}{3}.$$

Durch Division ergibt sich hieraus:

$$\frac{\pi}{2} = \frac{2\cdot 2}{1\cdot 3}\cdot\frac{4\cdot 4}{3\cdot 5}\cdot\frac{6\cdot 6}{5\cdot 7}\cdots\frac{2m\cdot 2m}{(2m-1)\cdot(2m+1)}\;\frac{\int_0^{\frac{\pi}{2}} \sin^{2m} x\,dx}{\int_0^{\frac{\pi}{2}} \sin^{2m+1} x\,dx}.$$

Der Quotient der beiden Integrale auf der rechten Seite konvergiert nun mit wachsendem $m$ gegen 1, wie wir aus folgender Betrachtung erkennen. Im Intervall $0 < x < \frac{\pi}{2}$ ist $0 < \sin^{2m+1} x \leqq \sin^{2m} x \leqq \sin^{2m-1} x$;

folglich

$$0 < \int_0^{\frac{\pi}{2}} \sin^{2m+1} x\,dx \leqq \int_0^{\frac{\pi}{2}} \sin^{2m} x\,dx \leqq \int_0^{\frac{\pi}{2}} \sin^{2m-1} x\,dx.$$

Dividieren wir hier jedes Glied durch $\int_0^{\frac{\pi}{2}} \sin^{2m+1} x\,dx$ und beachten, daß nach der oben zuerst bewiesenen Formel

$$\frac{\int_0^{\frac{\pi}{2}} \sin^{2m-1} x\,dx}{\int_0^{\frac{\pi}{2}} \sin^{2m+1} x\,dx} = \frac{2m+1}{2m} = 1 + \frac{1}{2m}$$

ist, so finden wir

$$1 \leqq \frac{\int_0^{\frac{\pi}{2}} \sin^{2m} x\,dx}{\int_0^{\frac{\pi}{2}} \sin^{2m+1} x\,dx} \leqq 1 + \frac{1}{2m},$$

woraus die obige Behauptung folgt.

Infolgedessen besteht die Beziehung

$$\frac{\pi}{2} = \lim_{m\to\infty} \frac{2}{1}\,\frac{2}{3}\,\frac{4}{3}\,\frac{4}{5}\,\frac{6}{5}\,\frac{6}{7} \cdots \frac{2m}{2m-1}\,\frac{2m}{2m+1}.$$

Diese von *Wallis* herrührende Produktdarstellung stellt eine durch ihr übersichtliches Bildungsgesetz höchst merkwürdige Verknüpfung der Zahl $\pi$ mit den ganzen Zahlen dar. Wir können ihr noch verschiedene andere Gestalten geben. Beachten wir, daß $\lim\limits_{m\to\infty} \frac{2m}{2m+1} = 1$ ist, so können wir schreiben

$$\lim_{m\to\infty} \frac{2^2 \cdot 4^2 \cdots (2m-2)^2}{3^2 \cdot 5^2 \cdots (2m-1)^2}\, 2m = \frac{\pi}{2}$$

und, indem wir die Quadratwurzel ziehen und dann den Bruch mit $2 \cdot 4 \cdots (2m-2)$ erweitern,

$$\sqrt{\frac{\pi}{2}} = \lim_{m\to\infty} \frac{2 \cdot 4 \cdots (2m-2)}{3 \cdot 5 \cdots (2m-1)} \sqrt{2m} = \lim_{m\to\infty} \frac{2^2 \cdot 4^2 \cdots (2m-2)^2}{(2m-1)!} \sqrt{2m}$$
$$= \lim_{m\to\infty} \frac{2^2 \cdot 4^2 \cdots (2m)^2}{(2m)!} \frac{\sqrt{2m}}{2m}.$$

Hieraus erhalten wir endlich

$$\lim_{m\to\infty} \frac{(m!)^2\, 2^{2m}}{(2m)!\,\sqrt{m}} = \sqrt{\pi},$$

eine Gestalt der Wallisschen Formel, die wir später noch (vgl. Kap. VII, Anhang) verwenden werden.

## § 5. Integration der rationalen Funktionen.

Die wichtigste allgemeine Klasse von Funktionen, deren unbestimmte Integration sich mit Hilfe der elementaren Funktionen stets ausführen läßt, sind die rationalen Funktionen:

$$R(x) = \frac{f(x)}{g(x)},$$

wo

$$f(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_0,$$
$$g(x) = b_n x^n + b_{n-1} x^{n-1} + \cdots + b_0 \qquad (b_n \neq 0)$$

Polynome sind.

Ich erinnere daran, daß jede ganze rationale Funktion sich sofort integrieren läßt und dabei wieder zu einer ganzen rationalen Funktion führt.

Wir brauchen daher unsere Aufmerksamkeit lediglich auf die gebrochen rationalen Funktionen zu richten, bei welchen der Nenner $g(x)$ keine Konstante ist[1]). Um eine solche gebrochene Funktion zu integrieren, können wir weiter bemerken, daß sie sich als Summe der Funktionen $\frac{a_\nu x^\nu}{g(x)}$ darstellen läßt und daß wir daher nur Integrale über Funktionen der Form $\frac{x^\nu}{g(x)}$ zu betrachten brauchen.

### 1. Aufstellung der Grundtypen.

Wir werden die Integration zunächst noch nicht für die allgemeinste derartige rationale Funktion durchführen, sondern nur für solche, deren Nenner $g(x)$ einen besonders einfachen Typus zeigt; und zwar:

$$g(x) = x \quad \text{oder} \quad g(x) = 1 + x^2$$

oder allgemeiner

$$g(x) = x^n, \quad g(x) = (1 + x^2)^n$$

mit beliebigem ganzen positiven Exponenten $n$.

Auf diesen Fall läßt sich sofort die Integration in dem etwas allgemeineren Falle zurückführen, daß $g(x) = (\alpha x + \beta)^n$ eine Potenz eines linearen Ausdruckes $\alpha x + \beta$ $(\alpha \neq 0)$ oder $g(x) = (a x^2 + 2 b x + c)^n$ eine

---

[1]) Den Grad des Zählers $f(x)$ dürfen wir dabei stets als geringer als den Grad $n$ des Nenners $g(x)$ voraussetzen. Denn man kann bekanntlich sonst das Polynom $f(x)$ durch das Polynom $g(x)$ dividieren, so daß der Rest ein Polynom von geringerem als $n$-tem Grade ist: mit anderen Worten, man kann schreiben $f(x) = q(x)\,g(x) + r(x)$, wo auch $q(x)$ und $r(x)$ Polynome sind und $r(x)$ einen geringeren Grad als $n$ hat; es ist somit die Integration von $\frac{f(x)}{g(x)}$ auf die des Polynomes $q(x)$ und die des „echten" Bruches $\frac{r(x)}{g(x)}$ zurückgeführt.

Potenz eines definiten quadratischen Ausdruckes[1]) ist. Im ersten Falle führen wir $\xi = \alpha x + \beta$ als neue Veränderliche ein. Es wird dann $\frac{d\xi}{dx} = \alpha$, und $x = \frac{1}{\alpha}\xi - \frac{\beta}{\alpha}$ ist auch eine lineare Funktion von $\xi$. Somit geht jeder Zähler $f(x)$ in ein Polynom $\varphi(\xi)$ vom selben Grade über, und es wird

$$\int \frac{f(x)}{(\alpha x + \beta)^n}\, dx = \frac{1}{\alpha}\int \frac{\varphi(\xi)}{\xi^n}\, d\xi .$$

Im zweiten Falle schreiben wir

$$a x^2 + 2 b x + c = \frac{1}{a}(a x + b)^2 + \frac{d^2}{a} \qquad (d^2 = ac - b^2, \quad d > 0),$$

indem wir beachten, daß wegen der vorausgesetzten Definitheit unseres Ausdruckes $a \neq 0$ und $ac - b^2$ positiv sein muß. Durch Einführung der neuen Veränderlichen

$$\xi = \frac{a x + b}{d}$$

gelangen wir nun sofort zu einem Integral mit dem Nenner $\frac{d^2}{a}(1 + \xi^2)$ bzw. $\left[\frac{d^2}{a}(1 + \xi^2)\right]^n$.

Um also rationale Funktionen zu integrieren, deren Nenner eine Potenz eines linearen Ausdruckes oder eines definiten quadratischen Ausdruckes ist, genügt es, die folgenden Typen von Funktionen integrieren zu können:

$$\frac{1}{x^n}, \quad \frac{x^{2\nu}}{(x^2+1)^n}, \quad \frac{x^{2\nu+1}}{(x^2+1)^n}.$$

Wir werden sogar sehen, daß wir auch diese Typen gar nicht allgemein zu behandeln brauchen, daß wir nämlich alle Integrationen rationaler Funktionen auf die Integration dieser Funktionstypen für den Spezialfall $\nu = 0$ zurückführen können. Demgemäß wollen wir uns zunächst mit der Integration der drei Ausdrücke

$$\frac{1}{x^n}, \quad \frac{1}{(x^2+1)^n}, \quad \frac{x}{(x^2+1)^n}$$

befassen.

## 2. Integration der Grundtypen.

Die Funktion des ersten Typus $\frac{1}{x^n}$ ergibt für $n = 1$ integriert $\log|x|$, für $n > 1$ die Funktion $-\frac{1}{(n-1)\,x^{n-1}}$, also wieder eine rationale Funktion. Die Funktion des dritten Typus läßt sich sofort integrieren, wenn wir $x^2 + 1 = \xi$ als neue Veränderliche einführen, wodurch wir

[1]) Ein quadratischer Ausdruck $Q(x) = a x^2 + 2 b x + c$ heißt *definit*, wenn er für reelle Werte von $x$ nur Werte eines Vorzeichens annehmen kann, wenn also die Gleichung $Q(x) = 0$ keine reellen Wurzeln besitzt. Hierzu ist notwendig und hinreichend, daß $ac - b^2$ eine positive Zahl ist.

$2x\,dx = d\xi$, also

$$\int \frac{x}{(x^2+1)^n}\,dx = \frac{1}{2}\int \frac{d\xi}{\xi^n} = \begin{cases} \frac{1}{2}\log(x^2+1) & \text{für } n = 1\,, \\ -\frac{1}{2(n-1)(x^2+1)^{n-1}} & \text{für } n > 1 \end{cases}$$

erhalten. Um endlich das Integral

$$J_n = \int \frac{dx}{(x^2+1)^n}$$

für beliebiges $n > 1$ auszuwerten, bedienen wir uns eines Rekursionsverfahrens. Setzen wir nämlich

$$\frac{1}{(x^2+1)^n} = \frac{1}{(x^2+1)^{n-1}} - \frac{x^2}{(x^2+1)^n}\,,$$

also

$$\int \frac{dx}{(x^2+1)^n} = \int \frac{dx}{(x^2+1)^{n-1}} - \int \frac{x^2\,dx}{(x^2+1)^n}\,,$$

so können wir das zweite Integral auf der rechten Seite durch Produktintegration umformen, indem wir in die Formel der Produktintegration (S. 176) einsetzen:

$$f(x) = x, \quad g'(x) = \frac{x}{(x^2+1)^n}\,.$$

Es ist demnach (siehe oben):

$$g(x) = -\frac{1}{2}\,\frac{1}{(n-1)(x^2+1)^{n-1}}\,,$$

und infolgedessen erhalten wir

$$J_n = \int \frac{dx}{(x^2+1)^n} = \frac{x}{2(n-1)(x^2+1)^{n-1}} + \frac{2n-3}{2(n-1)}\int \frac{dx}{(x^2+1)^{n-1}}\,.$$

Es ist also das Integral $J_n$ auf das Integral $J_{n-1}$ zurückgeführt. Wenden wir (bei $n-1>1$) auf das letztere das Verfahren nochmals an und schreiten so fort, bis wir schließlich auf den Ausdruck $\int \frac{dx}{x^2+1} = \operatorname{arc\,tg} x$ stoßen, so erkennen wir, daß sich das Integral $J_n$ explizite durch rationale Funktionen und die Funktion $\operatorname{arc\,tg} x$ ausdrücken läßt [1]).

Beiläufig bemerke ich, daß wir zur Integration der Funktion $\frac{1}{(x^2+1)^n}$ auch ohne weiteres durch die Substitution $x = \operatorname{tg} t$ hätten gelangen können; wir würden dann erhalten $dx = \frac{dt}{\cos^2 t}$ und $\frac{1}{1+x^2} = \cos^2 t$, also

$$\int \frac{dx}{(x^2+1)^n} = \int \cos^{2n-2} t\,dt,$$

und dieses Integral haben wir schon in § 4, Nr. 3 auszuwerten gelernt.

---

[1]) In genau derselben Art können wir übrigens auch das Integral der Funktion $\frac{1}{(x^2-1)^n}$ berechnen, indem wir es durch ein entsprechendes Rekursionsverfahren auf das Integral $\int \frac{dx}{1-x^2} = \mathfrak{Ar}\,\mathfrak{Tg}\,x$ bzw. $= \mathfrak{Ar}\,\mathfrak{Ctg}\,x$ zurückführen.

### 3. Die Partialbruchzerlegung.

Die Integration der allgemeinsten rationalen Funktionen gelingt nun auf Grund der Tatsache, daß man jede solche Funktion als Summe von sog. *Partialbrüchen* darstellen kann, d. h. als Summe aus einer ganzen rationalen Funktion und endlich vielen rationalen Funktionen, deren jede entweder als Nenner nur eine Potenz eines linearen Ausdruckes und als Zähler eine Konstante oder als Nenner eine Potenz eines definiten quadratischen Ausdruckes und als Zähler eine lineare Funktion hat. Ist der Grad des Zählers $f(x)$ geringer als der Grad des Nenners $g(x)$, so tritt bei der Zerlegung keine ganze rationale Funktion auf. Alle diese Funktionen können wir nach dem Vorangehenden integrieren. Denn nach Nr. 1 lassen sich die Faktoren der Nenner auf die speziellen Formen $x^n$ und $(x^2+1)^n$ zurückführen, und durch nochmalige Partialbruchzerlegung stellt man sich hieraus die in Nr. 2 integrierten Grundtypen her.

Ich will darauf verzichten, den allgemeinen Beweis für die Möglichkeit einer solchen Partialbruchzerlegung vollständig durchzuführen. Vielmehr will ich mich darauf beschränken, dem Leser die Aussage dieses Satzes verständlich zu machen und an Beispielen zu zeigen, wie man tatsächlich im gegebenen Falle die Partialbruchzerlegung praktisch ausführt — es kommen in der Praxis dafür nur verhältnismäßig einfache Funktionen in Frage, da sonst die Rechnung allzu kompliziert wird.

Man kann, wie die elementare Algebra lehrt, jedes Polynom $g(x)$ in die Gestalt setzen

$$g(x) = a\,(x-\alpha_1)^{l_1}(x-\alpha_2)^{l_2}\cdots(x^2+2b_1x+c_1)^{r_1}(x^2+2b_2x+c_2)^{r_2}\cdots.$$

Dabei sind die Zahlen $\alpha_1, \alpha_2, \ldots$ die voneinander verschiedenen reellen Wurzeln der Gleichung $g(x)=0$, und die ganzen positiven Zahlen $l_1, l_2, \ldots$ geben ihre Vielfachheit an; die Faktoren $x^2+2b_\nu x+c_\nu$ bedeuten lauter verschiedene definite quadratische Ausdrücke mit konjugiert komplexen Wurzeln, und die positiven ganzen Zahlen $r_1, r_2, \ldots$ geben wiederum die Vielfachheit dieser Wurzeln an.

Stellen wir uns vor, daß der Nenner $g(x)$ entweder von vornherein in dieser zerlegten Form gegeben ist oder daß wir ihn durch Aufsuchen der reellen und imaginären Wurzeln in diese Form zerlegt haben. Setzen wir ferner voraus, daß der Zähler $f(x)$ einen niedrigeren als den Grad $n$ des Nenners besitzt (vgl. Anm. 1 auf S. 182), so lautet der Satz von der Partialbruchzerlegung: Zu jedem Faktor $(x-\alpha)^l$ — wo $\alpha$ gleich irgend einer der reellen Wurzeln $\alpha_\nu$ sein kann und $l$ die ihr entsprechende Zahl $l_\nu$ ist — läßt sich ein Ausdruck der Form

$$\frac{A_1}{x-\alpha} + \frac{A_2}{(x-\alpha)^2} + \cdots + \frac{A_l}{(x-\alpha)^l}$$

mit konstanten Koeffizienten $A_1, \ldots, A_l$ und zu jedem quadratischen Faktor $Q(x) = x^2+2bx+c$ unserer Produktzerlegung, dessen Vielfachheit $r$ ist, ein Ausdruck der Form

$$\frac{B_1 + C_1 x}{Q} + \frac{B_2 + C_2 x}{Q^2} + \cdots + \frac{B_r + C_r x}{Q^r}$$

bestimmen, so daß $\frac{f(x)}{g(x)}$ gleich der Summe aller dieser Ausdrücke wird. Mit anderen Worten, der Quotient $\frac{f(x)}{g(x)}$ läßt sich als Summe von „Stammbrüchen" darstellen, deren jeder einem der durch Nr. 1 und 2 erledigten Typen angehört[1]).

Im einzelnen Falle kann eine solche Partialbruchzerlegung sehr einfach aussehen. Ist z. B. $g(x) = x^2 - 1$ oder allgemeiner $g(x) = (x - \alpha)(x - \beta)$, d. h. ist $g(x)$ ein nicht definiter quadratischer Ausdruck mit zwei verschiedenen reellen Wurzeln $\alpha$ und $\beta$, so erhalten wir sofort

$$\frac{1}{x^2 - 1} = \frac{1}{2}\,\frac{1}{x - 1} - \frac{1}{2}\,\frac{1}{x + 1}$$

$$\text{bzw.}\quad \frac{1}{(x - \alpha)(x - \beta)} = \frac{1}{\alpha - \beta}\cdot\frac{1}{x - \alpha} - \frac{1}{\alpha - \beta}\cdot\frac{1}{x - \beta}$$

und daher

$$\int \frac{dx}{x^2 - 1} = \frac{1}{2}\log\left|\frac{x - 1}{x + 1}\right| \quad \text{bzw.} \quad \int \frac{dx}{(x - \alpha)(x - \beta)} = \frac{1}{\alpha - \beta}\log\left|\frac{x - \alpha}{x - \beta}\right|.$$

### 4. Beispiel. Chemische Reaktionen.

Für die Anwendung der letzten einfachen Partialbruchzerlegung liefern die sog. Bimolekülreaktionen ein einfaches Beispiel. Haben wir zwei Ausgangsstoffe mit den anfänglichen in Mol pro Volumeinheit gerechneten Konzentrationen $a$ und $b$, wobei wir $a < b$ voraussetzen, und bildet sich in der Volumeneinheit während der Zeit $t$ durch chemische Reaktionen eine Menge $x$ (in Mol) des Reaktionsproduktes, so wird nach dem Massenwirkungsgesetz (vgl. drittes Kapitel, § 7, Nr. 5) im einfachsten Fall — Zusammentreten je eines Moleküls von beiden

---

[1]) Der Gedanke für den Beweis der Möglichkeit unserer Zerlegung sei hier wenigstens kurz skizziert: Wenn $g(x) = (x - \alpha)^k\, h(x)$ und $h(\alpha) \neq 0$ ist, so verschwindet offenbar auf der rechten Seite der Gleichung

$$\frac{f(x)}{g(x)} - \frac{f(\alpha)}{h(\alpha)(x - \alpha)^k} = \frac{1}{h(\alpha)}\cdot\frac{f(x)\,h(\alpha) - f(\alpha)\,h(x)}{(x - \alpha)^k\, h(x)}$$

der Zähler für $x = \alpha$; er hat daher die Gestalt $h(\alpha)(x - \alpha)^m \cdot f_1(x)$, wo $f_1(x)$ wieder ein Polynom, $f_1(\alpha) \neq 0$ und die ganze Zahl $m \geqq 1$ ist. Mit $\frac{f(\alpha)}{h(\alpha)} = \beta$ folgt also

$$\frac{f(x)}{g(x)} - \frac{\beta}{(x - \alpha)^k} = \frac{f_1(x)}{(x - \alpha)^{k-m}\, h(x)}.$$

So weitergehend kann man den Grad der im Nenner auftretenden Potenz von $x - \alpha$ immer weiter erniedrigen, bis schließlich keine solche Potenz im Nenner mehr auftritt. Indem man auf den dann verbleibenden Bruch dasselbe Verfahren für eine andere Wurzel von $g(x)$ anwendet und dieses so oft wiederholt, als $g(x)$ verschiedene Faktoren enthält, gelangt man schließlich unter Berücksichtigung auch der komplexen Wurzeln zu der vollständigen Partialbruchzerlegung.

Reaktionskomponenten — die Größe $x$ als Funktion der Zeit sich mit der Reaktionsgeschwindigkeit $\frac{dx}{dt} = k(a-x)(b-x)$ vermehren; Aufgabe ist die Bestimmung der Funktion $x(t)$. Fassen wir umgekehrt die Zeit $t$ als Funktion von $x$ auf, so ergibt sich

$$\frac{dt}{dx} = \frac{1}{k(a-x)(b-x)} = \frac{1}{k(b-a)}\left(\frac{1}{a-x} - \frac{1}{b-x}\right);$$

also durch Integration

$$kt = \frac{1}{a-b}\log\frac{a-x}{b-x} + c, \quad \text{für } x < a < b,$$

wobei wir die Integrationskonstante $c$ aus der Forderung bestimmen, daß für $t = 0$ noch kein Reaktionsprodukt vorhanden ist, daß also $\frac{1}{a-b}\log\frac{a}{b} + c = 0$ wird. Wir erhalten schließlich $kt = \frac{1}{a-b}\log\frac{1-\frac{x}{a}}{1-\frac{x}{b}}$ oder durch Auflösung nach $x$ die gesuchte Funktion $x(t)$:

$$x = \frac{ab\,(1 - e^{(a-b)kt})}{b - a\,e^{(a-b)kt}}.$$

**5. Weitere Beispiele für Partialbruchzerlegung. (Methode der unbestimmten Koeffizienten.)**

Ist $g(x) = (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)$, wo $\alpha_i \neq \alpha_k$ für $i \neq k$, hat also $g(x)$ lauter einfache reelle Wurzeln, so lautet die Partialbruchzerlegung einfach

$$\frac{1}{g(x)} = \frac{a_1}{x-\alpha_1} + \frac{a_2}{x-\alpha_2} + \cdots + \frac{a_n}{x-\alpha_n}.$$

Für die Koeffizienten, z. B. $a_1$, erhalten wir eine explizite Darstellung sofort, indem wir diese Gleichung mit $x - \alpha_1$ multiplizieren, links und im ersten Gliede rechts den Linearfaktor $x - \alpha_1$ kürzen und danach $x = \alpha_1$ setzen. Es ergibt sich so

$$a_1 = \frac{1}{(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)\cdots(\alpha_1-\alpha_n)}\text{[1]}.$$

Als typisches Beispiel für Nenner $g(x)$ mit mehrfachen Wurzeln betrachten wir die Funktion $\frac{1}{x^2(x-1)}$, deren Nenner die Doppelwurzel $x = 0$ besitzt. Hier führt uns entsprechend Nr. 3 der Ansatz

$$\frac{1}{x^2(x-1)} = \frac{a}{x-1} + \frac{b}{x} + \frac{c}{x^2}$$

zum Ziel. Multiplizieren wir diese Gleichung mit $x^2(x-1)$, so erhalten wir zur Bestimmung der Koeffizienten $a, b, c$ die Gleichung

$$1 = (a+b)x^2 - (b-c)x - c,$$

[1] Der Leser mag sich überlegen, daß der Nenner rechts nichts anderes ist als $g'(\alpha_1)$, d. h. die Ableitung der Funktion $g(x)$, genommen an der Stelle $x = \alpha_1$.

die für alle Werte von $x$ erfüllt sein soll. Hierzu müssen alle Koeffizienten des Polynomes $(a+b)\,x^2-(b-c)\,x-c-1$ Null sein, d. h. es muß $a+b=b-c=c+1=0$ oder $c=-1$, $b=-1$, $a=1$ sein. Wir erhalten so die Zerlegung

$$\frac{1}{x^2(x-1)}=\frac{1}{x-1}-\frac{1}{x}-\frac{1}{x^2},$$

und es ist

$$\int\frac{dx}{x^2(x-1)}=\log|x-1|-\log|x|+\frac{1}{x}.$$

Die Funktion $\frac{1}{x(x^2+1)}$ (Beispiel für komplexe Wurzeln des Nenners) werden wir gemäß der Gleichung

$$\frac{1}{x(x^2+1)}=\frac{a}{x}+\frac{bx+c}{x^2+1}$$

zerlegen. Für die Koeffizienten gilt: $a+b=c=a-1=0$; also

$$\frac{1}{x(x^2+1)}=\frac{1}{x}-\frac{x}{x^2+1}$$

und hiernach

$$\int\frac{dx}{x(x^2+1)}=\log|x|-\frac{1}{2}\log(x^2+1).$$

Als drittes Beispiel betrachten wir die Funktion $\frac{1}{x^4+1}$, deren Integration noch Leibniz große Schwierigkeiten bereitete. Wir können den Nenner als Produkt zweier quadratischer Faktoren darstellen: $x^4+1=(x^2+1)^2-2x^2=(x^2+1+\sqrt{2}\,x)\,(x^2+1-\sqrt{2}\,x)$ und werden infolgedessen folgende Partialbruchzerlegung ansetzen:

$$\frac{1}{x^4+1}=\frac{ax+b}{x^2+\sqrt{2}\,x+1}+\frac{cx+d}{x^2-\sqrt{2}\,x+1}.$$

Zur Bestimmung der Konstanten $a, b, c, d$ erhalten wir die Gleichung

$$(a+c)\,x^3+(b+d-a\sqrt{2}+c\sqrt{2})\,x^2+(a+c-b\sqrt{2}+d\sqrt{2})\,x \\ +(b+d-1)=0,$$

die durch die Werte

$$a=\frac{1}{2\sqrt{2}},\quad b=\frac{1}{2},\quad c=-\frac{1}{2\sqrt{2}},\quad d=\frac{1}{2}$$

befriedigt wird. Es ist also

$$\frac{1}{x^4+1}=\frac{1}{2\sqrt{2}}\cdot\frac{x+\sqrt{2}}{x^2+\sqrt{2}\,x+1}-\frac{1}{2\sqrt{2}}\cdot\frac{x-\sqrt{2}}{x^2-\sqrt{2}\,x+1},$$

und wir erhalten unter Anwendung der in Nr. 1 angegebenen Methode

$$\int\frac{dx}{x^4+1}=\frac{1}{4\sqrt{2}}\log|x^2+\sqrt{2}\,x+1|-\frac{1}{4\sqrt{2}}\log|x^2-\sqrt{2}\,x+1| \\ +\frac{1}{2\sqrt{2}}\operatorname{arc\,tg}(\sqrt{2}\,x+1)+\frac{1}{2\sqrt{2}}\operatorname{arc\,tg}(\sqrt{2}\,x-1),$$

ein Ergebnis, das man hinterher leicht durch Differenzieren bestätigen kann.

# § 6. Integration einiger anderer Funktionenklassen.

## 1. Vorbemerkungen über die rationale Darstellung der trigonometrischen und Hyperbelfunktionen.

Auf die Integration der rationalen Funktionen läßt sich die Integration einiger anderer allgemeiner Funktionenklassen zurückführen. Wir verstehen diese Zurückführung am besten, wenn wir uns vorher einige elementare Tatsachen über die trigonometrischen und hyperbolischen Funktionen klar machen. Aus der elementaren Trigonometrie ergeben sich, wenn $t = \operatorname{tg}\frac{x}{2}$ gesetzt ist, die folgenden einfachen Formeln:

$$\sin x = \frac{2t}{1+t^2}, \quad \cos x = \frac{1-t^2}{1+t^2}\ ^{1)}.$$

D. h. $\sin x$ *und* $\cos x$ *lassen sich rational durch die Größe* $t = \operatorname{tg}\frac{x}{2}$ *ausdrücken.* Aus $t = \operatorname{tg}\frac{x}{2}$ folgt durch Differentiation

$$\frac{dt}{dx} = \frac{1}{2\cos^2\frac{x}{2}} = \frac{1+t^2}{2}, \quad \text{also} \quad \frac{dx}{dt} = \frac{2}{1+t^2};$$

also auch der Differentialquotient $\frac{dx}{dt}$ drückt sich rational durch $t$ aus.

Die geometrische Bedeutung und Veranschaulichung unserer Formeln gibt uns Figur 74. In ihr ist der Kreis $u^2 + v^2 = 1$ in einer $uv$-Ebene gezeichnet. Bedeutet $x$ den in der Figur gezeichneten Winkel $POT$, so wird $u = \cos x$, $v = \sin x$. Der Winkel des Dreiecks $OSP$ an der Spitze $S$ im Punkte $u = -1$, $v = 0$ ist nach einem elementaren geometrischen Satz $\frac{x}{2}$, und man entnimmt der Figur nunmehr sofort die geometrische Bedeutung der Größe $t$; es ist nämlich $t = \operatorname{tg}\frac{x}{2} = OR$. Durchläuft der Punkt $P$, von $S$ angefangen, im positiven Sinne einmal den Kreis, d. h. durchläuft $x$ das Intervall von $-\pi$ bis $+\pi$, so wird die Größe $t$ gerade einmal die ganze Werteskala von $-\infty$ bis $+\infty$ durchlaufen.

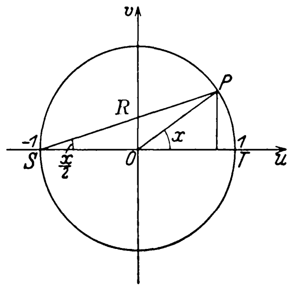


Fig. 74. Parameterdarstellung der trigonometrischen Funktionen.

Ganz entsprechend können wir auch die Hyperbelfunktionen $\mathfrak{Cof}\, x = \frac{1}{2}(e^x + e^{-x})$ und $\mathfrak{Sin}\, x = \frac{1}{2}(e^x - e^{-x})$ als rationale Funktionen einer dritten Größe darstellen. Das Nächstliegende ist, $e^x = \tau$

---

1) Es ist nämlich $\frac{1}{1+t^2} = \cos^2\frac{x}{2}$, $\frac{t^2}{1+t^2} = \sin^2\frac{x}{2}$, und hieraus ergeben sich wegen $\sin x = 2\cos^2\frac{x}{2}\operatorname{tg}\frac{x}{2}$ und $\cos x = \cos^2\frac{x}{2} - \sin^2\frac{x}{2}$ sofort die oben hingeschriebenen Formeln.

zu setzen, wodurch wir

$$\mathfrak{Cof}\,x = \frac{1}{2}\left(\tau + \frac{1}{\tau}\right), \quad \mathfrak{Sin}\,x = \frac{1}{2}\left(\tau - \frac{1}{\tau}\right),$$

also tatsächlich eine rationale Darstellung von $\mathfrak{Sin}\,x$ und $\mathfrak{Cof}\,x$ erhalten. Auch hier wird $\frac{dx}{d\tau} = \frac{1}{\tau}$ rational in $\tau$. Es ergibt sich jedoch eine größere Analogie zu den trigonometrischen Funktionen, wenn wir die Größe $t = \mathfrak{Tg}\,\frac{x}{2}$ einführen; wir gelangen dann zu den Formeln

$$\mathfrak{Sin}\,x = \frac{2t}{1 - t^2}, \quad \mathfrak{Cof}\,x = \frac{1 + t^2}{1 - t^2}.$$

Hier folgt wie S. 189 durch Differentiation von $t = \mathfrak{Tg}\,\frac{x}{2}$ die rationale Darstellung

$$\frac{dx}{dt} = \frac{2}{1 - t^2}$$

des Differentialquotienten $\frac{dx}{dt}$. Auch hier hat die Größe $t$ eine ganz ähnliche geometrische Bedeutung wie bei den trigonometrischen Funktionen, was man der Figur 75 leicht entnimmt.

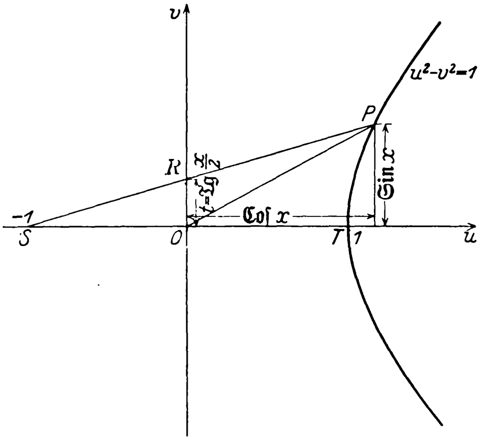



Fig. 75. Parameterdarstellung der Hyperbelfunktionen.

Während jedoch bei den trigonometrischen Funktionen $t$ das Intervall von $-\infty$ bis $+\infty$ durchlaufen muß, damit wir die Gesamtheit der Wertepaare von $\sin x$ und $\cos x$ bekommen, wird bei den hyperbolischen Funktionen die Größe $t$ auf das Intervall $-1 < t < 1$ beschränkt sein.

Nach diesen Vorbemerkungen gehen wir an unser Integrationsproblem.

## 2. Integration von $R(\cos x, \sin x)$.

Es bedeute $R(\cos x, \sin x)$ einen rationalen Ausdruck in den beiden Funktionen $\sin x$ und $\cos x$, d. h. einen Ausdruck, der aus diesen beiden Funktionen und Konstanten in rationaler Weise gebildet ist, wie z. B.

$$\frac{3 \sin^2 x + \cos x}{3 \cos^2 x + \sin x}.$$

Wenden wir die Substitution $t = \operatorname{tg} \frac{x}{2}$ an, so geht das Integral $\int R(\cos x, \sin x)\, dx$ in das Integral

$$\int R\left(\frac{1 - t^2}{1 + t^2}, \frac{2t}{1 + t^2}\right) \cdot \frac{2\,dt}{1 + t^2}$$

über, und unter dem Integralzeichen steht jetzt eine rationale Funktion von $t$. Die Integration unseres Ausdruckes ist daher im Prinzip geleistet, da wir das Integrationsproblem jetzt nach den Methoden des vorigen Paragraphen lösen können.

### 3. Integration von $R(\mathfrak{Cos}\, x, \mathfrak{Sin}\, x)$.

Ist ganz entsprechend $R(\mathfrak{Cos}\, x, \mathfrak{Sin}\, x)$ ein rationaler Ausdruck in den hyperbolischen Funktionen $\mathfrak{Cos}\, x$ und $\mathfrak{Sin}\, x$, so gelingt seine Integration durch die Substitution $t = \mathfrak{Tg}\frac{x}{2}$; es wird also unter Beachtung von

$$\frac{dx}{dt} = \frac{2}{1-t^2}$$

$$\int R(\mathfrak{Cos}\, x,\ \mathfrak{Sin}\, x)\, dx = \int R\left(\frac{1+t^2}{1-t^2},\ \frac{2t}{1-t^2}\right)\frac{2}{1-t^2}\, dt.$$

(Wir hätten im übrigen entsprechend der obigen Bemerkung auch einfach $\tau = e^x$ als neue unabhängige Veränderliche einführen und $\mathfrak{Cos}\, x$ und $\mathfrak{Sin}\, x$ durch $\tau$ ausdrücken können.) Wiederum ist die Integration auf die einer rationalen Funktion zurückgeführt.

### 4. Integration von $R(x, \sqrt{1-x^2})$.

Das Integral $\int R(x, \sqrt{1-x^2})\, dx$ führen wir durch die Substitution $x = \cos u$, $\sqrt{1-x^2} = \sin u$, $dx = -\sin u\, du$ auf den unter Nr. 2 behandelten Typus zurück; durch die Substitution $t = \operatorname{tg}\frac{u}{2}$ können wir von da aus sofort zu einem Integral über eine rationale Funktion gelangen. Wir können aber, wie ich beiläufig bemerke, diese Zurückführung statt mit zwei Schritten auch auf einmal vornehmen, indem wir sofort die Substitution

$$t = \sqrt{\frac{1-x}{1+x}};\quad x = \frac{1-t^2}{1+t^2};\quad \sqrt{1-x^2} = \frac{2t}{1+t^2};\quad \frac{dx}{dt} = \frac{-4t}{(1+t^2)^2}$$

ausführen, welche unser Integral in eines über eine rationale Funktion von $t$ verwandelt; d. h. wir führen sofort $t = \operatorname{tg}\frac{u}{2}$ als neue Veränderliche ein und erhalten damit ein Integral einer rationalen Funktion.

### 5. Integration von $R(x, \sqrt{x^2-1})$.

Das Integral $\int R(x, \sqrt{x^2-1})\, dx$ geht durch die Substitution $x = \mathfrak{Cos}\, u$ in den unter Nr. 3 behandelten Typus über. Hier könnten wir auch direkt durch Einführung von

$$t = \sqrt{\frac{x-1}{x+1}} = \mathfrak{Tg}\frac{u}{2}$$

zum Ziele kommen.

### 6. Integration von $R(x, \sqrt{x^2+1})$.

Das Integral $\int R(x, \sqrt{x^2+1})\,dx$ wird durch die Substitution $x = \mathfrak{Sin}\, u$ ebenfalls auf den Typus unter Nr. 3 zurückgeführt, ist somit elementar ausführbar. Anstatt die weitere Reduktion auf das Integral einer rationalen Funktion nun durch die Substitution $e^u = \tau$ oder $\mathfrak{Tg}\,\frac{u}{2} = t$ vorzunehmen, hätten wir übrigens auch mit einem Schlage durch die Substitution $\tau = x + \sqrt{x^2+1}$ oder auch $t = \frac{-1+\sqrt{x^2+1}}{x}$ das Integral einer rationalen Funktion erhalten können.

### 7. Integration von $R(x, \sqrt{ax^2+2bx+c})$.

Das Integral $\int R(x, \sqrt{ax^2+2bx+c})\,dx$ über einen Ausdruck, welcher sich rational aus $x$ und einer Quadratwurzel aus einem beliebigen Polynom zweiten Grades in $x$ zusammensetzt, läßt sich sofort auf einen der eben behandelten Typen zurückführen. Wir schreiben (vgl. zum folgenden S. 183)

$$ax^2 + 2bx + c = \frac{1}{a}(ax+b)^2 + \frac{ac-b^2}{a}$$

und führen im Falle $ac - b^2 > 0$ durch die Transformation $\xi = \frac{ax+b}{\sqrt{ac-b^2}}$ eine neue Integrationsvariable $\xi$ ein, wobei sich unsere Quadratwurzel in $\sqrt{\frac{ac-b^2}{a}}\sqrt{\xi^2+1}$ verwandelt. Das Integral wird also in der Variablen $\xi$ genau den Typus aus Nr. 6 zeigen. Die Konstante $a$ muß hierbei positiv sein, damit die Quadratwurzel überhaupt reelle Werte besitzt.

Ist $ac - b^2 = 0$, $a > 0$, so ist wegen $\sqrt{ax^2+2bx+c} = \sqrt{a}\left(x + \frac{b}{a}\right)$ der Integrand von vornherein rational in $x$.

Ist endlich $ac - b^2 < 0$, so setzen wir $\xi = \frac{ax+b}{\sqrt{b^2-ac}}$ und erhalten für unsere Quadratwurzel den Ausdruck $\sqrt{\frac{b^2-ac}{a}(\xi^2-1)}$. Ist $a$ positiv, so ist damit unser Integral auf den Typus unter Nr. 5 zurückgeführt; ist jedoch $a$ negativ, so schreiben wir unsere Wurzel in der Form $\sqrt{\frac{b^2-ac}{-a}}\sqrt{1-\xi^2}$ und erkennen, daß die Zurückführung auf den Typus Nr. 4 geleistet ist.

### 8. Weitere Beispiele für Zurückführung auf Integrale rationaler Funktionen.

Von weiteren Funktionstypen, deren Integration durch Zurückführung auf rationale Integranden gelingt, nenne ich kurz noch zwei: einmal rationale Ausdrücke in zwei verschiedenen Quadratwurzeln aus linearen Ausdrücken: $R(x, \sqrt{ax+b}, \sqrt{\alpha x+\beta})$, zweitens Ausdrücke der Form $R\left(x, \sqrt[n]{\frac{ax+b}{\alpha x+\beta}}\right)$, wobei $a, b, \alpha, \beta$ Konstante sind. Führen wir im ersten Falle $\xi = \sqrt{\alpha x+\beta}$ als neue unabhängige Veränderliche ein, so daß $\alpha x+\beta = \xi^2$, also $x = \frac{\xi^2-\beta}{\alpha}$, $\frac{dx}{d\xi} = \frac{2\xi}{\alpha}$ wird, so ergibt sich

$$\int R(x, \sqrt{ax+b}, \sqrt{\alpha x+\beta})\,dx = \int R\left(\frac{\xi^2-\beta}{\alpha}, \sqrt{\frac{1}{\alpha}(a\xi^2-(a\beta-b\alpha))}, \xi\right)\frac{2\xi}{\alpha}\,d\xi,$$

also der unter Nr. 7 behandelte Typus.

Führen wir im zweiten Falle als neue unabhängige Veränderliche die Größe $\xi = \sqrt[n]{\frac{ax+b}{\alpha x+\beta}}$ ein, so wird

$$\xi^n = \frac{ax+b}{\alpha x+\beta}, \quad x = \frac{-\beta\xi^n+b}{\alpha\xi^n-a}, \quad \frac{dx}{d\xi} = \frac{a\beta-b\alpha}{(\alpha\xi^n-a)^2}\cdot n\,\xi^{n-1},$$

und wir gelangen unmittelbar zu der Formel

$$\int R\left(x, \sqrt[n]{\frac{ax+b}{\alpha x+\beta}}\right)dx = \int R\left(\frac{-\beta\xi^n+b}{\alpha\xi^n-a}, \xi\right)\frac{a\beta-b\alpha}{(\alpha\xi^n-a)^2}\cdot n\,\xi^{n-1}\,d\xi,$$

also zu dem Integral einer rationalen Funktion.

### 9. Bemerkungen zu den Beispielen.

Die obigen Überlegungen haben vor allem ein prinzipielles theoretisches Interesse. Die wirkliche Durchführung bei verwickelten Ausdrücken würde häufig allzu kompliziert werden. Es ist daher zweckmäßig, unter Umständen die spezielle Natur der zu integrierenden Funktionen zu benutzen, um größere Einfachheit zu erzielen. Z. B. wird man für die Integration des Ausdrucks $\frac{1}{a^2\sin^2 x + b^2\cos^2 x}$ statt der unter Nr. 2 angegebenen Substitution besser $t = \operatorname{tg} x$ als neue Veränderliche einführen, da sich $\sin^2 x$ und $\cos^2 x$ schon rational durch $\operatorname{tg} x$ ausdrücken lassen und daher ein Zurückgehen auf $\operatorname{tg}\frac{x}{2}$ nicht mehr nötig ist. Dasselbe gilt für jeden Ausdruck, der sich aus $\sin^2 x$, $\cos^2 x$ und $\sin x \cos x$ rational aufbaut [1]). In vielen Fällen wird man übrigens für die Ausführung von Integralen nicht eine rationale Gestalt, sondern eine trigonometrische vorziehen, falls man von dieser aus durch ein übersichtliches Rekursionsverfahren zum Ziele kommen kann. Z. B. wird man das Integral $\int x^n(\sqrt{1-x^2})^m\,dx$ statt in eine rationale Form lieber durch die Substitution $x = \sin u$ in die Gestalt $\int \sin^n u \cos^{m+1} u\,du$ setzen, in der man es leicht durch ein Rekursionsverfahren nach § 4 behandeln kann (oder auch, indem man mittels der Additionstheoreme

[1]) Denn $\sin x \cos x = \operatorname{tg} x \cos^2 x$ läßt sich natürlich rational durch $\operatorname{tg} x$ ausdrücken.

die Potenzen der sin und cos auf die sin und cos der vielfachen Winkel zurückführt).

Zur Auswertung des Integrals

$$\int \frac{dx}{a\cos x + b\sin x} \qquad (a^2 + b^2 > 0)$$

werden wir, statt auf die allgemeine Theorie zurückzugehen, eine Zahl $A$ und einen Winkel $\vartheta$ so bestimmen, daß

$$a = A\sin\vartheta, \quad b = A\cos\vartheta$$

ist; d. h. wir setzen

$$A = \sqrt{a^2 + b^2}, \qquad \vartheta = \operatorname{arc\,tg}\frac{a}{b}.$$

Das Integral geht dann in

$$\frac{1}{A}\int \frac{dx}{\sin(x + \vartheta)}$$

über, als dessen Wert sich nach Einführung von $x + \vartheta$ als neuer Veränderlichen

$$\frac{1}{A}\operatorname{lcg}\left|\operatorname{tg}\frac{x+\vartheta}{2}\right|$$

ergibt (vgl. S. 174).

## § 7. Bemerkungen über Funktionen, die sich nicht mittels der elementaren Funktionen integrieren lassen.

### 1. Definition von Funktionen durch Integrale. Elliptische Integrale.

Mit den angegebenen Beispielen von Funktionentypen, deren Integration sich auf die von rationalen Funktionen zurückführen läßt, ist im wesentlichen der Bereich der elementar integrierbaren Funktionenklassen erschöpft. Die Bemühungen, z. B. allgemeine Integrale der folgenden Art:

$$\int \frac{dx}{\sqrt{a_0 + a_1 x + \cdots + a_n x^n}}, \quad \int \sqrt{a_0 + a_1 x + \cdots + a_n x^n}\, dx \quad \text{oder} \quad \int \frac{e^x}{x}\, dx$$

durch elementare Funktionen auszudrücken, sind stets gescheitert; im 19. Jahrhundert gelang sogar der Beweis für die prinzipielle Unmöglichkeit, diese Integrationen mittels der elementaren Funktionen auszuführen.

Wenn es also das Ziel der Integralrechnung wäre, Funktionen elementar zu integrieren, so wären wir rasch am Ende dieser Kunst angelangt. Aber ein solches Ziel hat tatsächlich keine innere Berechtigung; im Gegenteil, es haftet ihm etwas Künstliches an. Daß das Integral einer stetigen Funktion e x i s t i e r t und als Funktion der oberen Grenze wiederum eine neue stetige Funktion darstellt, ist eine Tatsache, die mit der Ausdrückbarkeit dieser letzten Funktion durch elementare Funktionen nichts zu tun hat. Was die elementaren Funktionen auszeichnet, ist im Grunde nur die Tatsache, daß sie in ihren Eigenschaften leicht übersehbar sind, daß ihre numerische Verwendung vielfach durch bequeme Tabellen erleichtert ist bzw. daß man sie, wie die rationalen Funktionen, auf einfache Weise mit beliebiger Genauigkeit berechnen kann.

Nichts hindert uns, wenn das Integral einer Funktion sich durch die uns geläufigen Funktionen nicht ausdrücken läßt, dieses Integral als eine neue, „höhere" Funktion in die Analysis einzuführen, d. h. im Grunde nur, ihm einen Namen zu geben. Ob die Einführung einer solchen neuen Funktion zweckmäßig oder unzweckmäßig ist, wird davon abhängen, was für Eigenschaften sie besitzt, wie häufig man auf sie geführt wird und wie leicht man sie theoretisch oder numerisch beherrschen kann. In diesem Sinne bildet also der Integrationsprozeß ein Prinzip zur Erzeugung neuer Funktionen.

Im Grunde genommen kennen wir dieses Prinzip schon von den elementaren Funktionen her. So sahen wir uns genötigt (drittes Kapitel), das zunächst noch nicht bekannte Integral von $\frac{1}{x}$ als neue Funktion einzuführen, die wir als Logarithmus bezeichnet haben und deren Eigenschaften wir dann leicht feststellen konnten. Ganz ähnlich hätten wir auch, lediglich auf die rationale Funktion, den Integrationsprozeß und den Umkehrungsprozeß gestützt, die trigonometrischen Funktionen einführen können; man braucht dazu nur etwa eine der beiden Gleichungen

$$\operatorname{arc\,tg} x = \int_0^x \frac{dt}{1+t^2} \quad \text{oder} \quad \arcsin x = \int_0^x \frac{dt}{\sqrt{1-t^2}}$$

als Definition der Funktion $\operatorname{arc\,tg} x$ bzw. $\arcsin x$ an die Spitze zu stellen, um dann aus ihr durch Umkehrung die trigonometrischen Funktionen zu gewinnen; ein Verfahren, durch welches man die Definition dieser Funktionen von der Geometrie loslöst, sich aber natürlich die Verpflichtung auferlegt, ihre Eigenschaften nunmehr auch unabhängig von der Geometrie aus der Definition durch Integrale zu entwickeln[1]).

Das erste und wichtigste Beispiel, welches über den Bereich der elementaren Funktionen hinausführt, geben uns die *elliptischen Integrale*. Es sind dies Integrale, bei denen der Integrand in rationaler Weise eine Quadratwurzel aus einem Ausdruck dritten oder vierten Grades enthält. Als besonders wichtig hat sich unter diesen Integralen die Funktion

$$u(s) = \int_0^s \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}$$

[1]) Auf die Ausführung dieses Gedankens will ich hier nicht eingehen. Das Wesentliche ist, daß man die Additionstheoreme der Umkehrfunktionen, d. h. des Sinus und des Tangens, beweist.

erwiesen. Ihre Umkehrfunktion $s(u)$ spielt ebenfalls eine große Rolle. Für $k = 0$ erhalten wir speziell $u(s) = \arcsin s$ bzw. $s(u) = \sin u$. Die Funktionen $s(u)$ hat man in ihren Eigenschaften genau so gut erforscht und in Tabellen festgelegt wie die elementaren Funktionen. Dies führt uns jedoch aus dem Rahmen dieser Betrachtung hinaus in das Gebiet der sog. elliptischen Funktionen, welches ein Kernstück der Funktionentheorie bildet.

Hier möchte ich nur erwähnen, daß das Wort elliptische Integrale von der Tatsache herrührt, daß solche Integrale bei dem Problem der Längenbestimmung eines Ellipsenbogens auftreten. (Vgl. fünftes Kapitel, S. 229.)

Ich möchte ferner bemerken, daß Integrale von scheinbar ganz anderem Aussehen sich durch einfache Substitutionen als elliptische Integrale erweisen. Beispielsweise geht das Integral

$$\int \frac{dx}{\sqrt{\cos\alpha - \cos x}}$$

durch die Substitution $u = \cos\frac{x}{2}$ in das Integral

$$-k\sqrt{2}\int \frac{du}{\sqrt{(1-u^2)(1-k^2u^2)}}, \quad k = \frac{1}{\cos\frac{\alpha}{2}},$$

über; das Integral

$$\int \frac{dx}{\sqrt{\cos 2x}}$$

durch die Substitution $u = \sin x$ in

$$\int \frac{du}{\sqrt{(1-u^2)(1-2u^2)}};$$

schließlich verwandelt sich das Integral

$$\int \frac{dx}{\sqrt{1-k^2\sin^2 x}}$$

durch $u = \sin x$ in

$$\int \frac{du}{\sqrt{(1-u^2)(1-k^2u^2)}}.$$

## 2. Grundsätzliches über Differentiation und Integration.

Noch eine allgemeine Bemerkung über das Verhältnis der Differentiation zur Integration sei eingeschaltet. Die Differentiation ist, da sie aus dem Bereich des Gegebenen nicht herausführt, der elementarere Prozeß gegenüber der Integration. Andererseits aber müssen wir uns vor Augen halten, daß die Differenzierbarkeit irgend einer stetigen Funktion keineswegs selbstverständlich ist, sondern eine sehr einschneidende Voraussetzung darstellt. Wir haben ja gesehen, daß es stetige Funktionen gibt, die an einzelnen Stellen nicht mehr differenzierbar sind, und ich erwähne ohne Beweis die Tatsache, daß man seit *Weierstraß* zahlreiche Beispiele für stetige Funktionen konstruieren kann, welche sogar nirgends einen Differentialquotienten besitzen[1]). (Es steckt also in der mathematischen Definition der Stetigkeit viel weniger, als die naive Anschauung zunächst vermuten läßt.) Im Gegensatz dazu ist zwar die Integration im allgemeinen nicht mehr elementar ausführbar, dafür aber ist man unter allen Umständen der Existenz des Integrales einer stetigen Funktion sicher.

Alles in allem erkennen wir, daß Differentiation und Integration nicht schlechthin als mehr oder weniger elementar einander gegenüberstehen, sondern daß in der einen Hinsicht der eine, in der anderen der andere Prozeß als der elementarere bezeichnet zu werden verdient.

Was den Integralbegriff betrifft, so werden wir sogleich im nächsten Paragraphen erkennen, daß er nicht einmal an die Voraussetzung der Stetigkeit der zu integrierenden Funktion geknüpft ist, sondern sich auf weite Klassen von Funktionen mit Unstetigkeiten ausdehnen läßt.

## § 8. Erweiterung des Integralbegriffes. Uneigentliche Integrale.

### 1. Funktionen mit Sprungstellen.

Zunächst sehen wir sofort, daß der Erweiterung des Integralbegriffs keinerlei Schwierigkeit entgegensteht, wenn die zu integrierende Funktion $f(x)$ an einer oder mehreren Stellen des Intervalles sprunghaft unstetig ist. Dann brauchen wir als Integral der Funktion nur die Summe der Integrale über die einzelnen Teilintervalle zu verstehen, in denen die Funktion stetig bleibt[2]). Auch anschaulich behält das Integral als Flächeninhalt seine Bedeutung bei (vgl. Fig. 76).

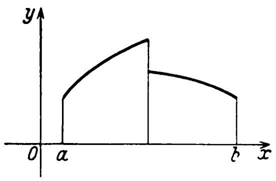


Fig. 76. Integral über unstetige Funktion.

### 2. Funktionen mit Unendlichkeitsstellen.

Anders liegt jedoch die Sache, wenn das Integrationsintervall im Innern oder an einem Endpunkt eine Unendlichkeitsstelle der Funktion besitzt. Um den Integralbegriff für diesen Fall noch formulieren zu können, müssen wir einen weiteren Grenzübergang heranziehen.

---

[1]) Vgl. *F. Klein*, Elementarmathematik vom höheren Standpunkt aus, III. S. 39ff. Berlin: Julius Springer 1928.

[2]) Eigentlich müßten wir beachten, daß wir früher bei der Integraldefinition die Intervalle als abgeschlossen betrachteten und die Funktion als stetig in dem abgeschlossenen Intervalle voraussetzten. Es entsteht aber hieraus jetzt keine Schwierigkeit für uns, da wir für jedes abgeschlossene Teilintervall die Funktion $f(x)$ zu einer stetigen ergänzen können, indem wir einfach die Grenzwerte der Funktion bei Annäherung an die Endpunkte vom Inneren des Intervalles als Funktionswerte in diesen Endpunkten hinzunehmen.

Wir erläutern die möglichen Vorkommnisse zunächst an einigen Beispielen, bevor wir die allgemeinen Begriffsbildungen formulieren, und zwar betrachten wir das Integral

$$\int \frac{dx}{x^\alpha},$$

wo $\alpha$ eine positive Zahl ist. Der Integrand $\frac{1}{x^\alpha}$ wird für $x \to 0$ unendlich, und wir können das Integral daher nicht von der unteren Grenze 0 an erstrecken. Dagegen können wir untersuchen, was herauskommt, wenn wir es von der positiven Grenze $\varepsilon$ etwa bis zur Grenze 1 erstrecken und zum Schluß $\varepsilon$ gegen 0 streben lassen. Nach den elementaren Integrationsregeln erhalten wir, außer im Falle $\alpha = 1$,

$$\int_\varepsilon^1 \frac{dx}{x^\alpha} = \frac{1}{1-\alpha}\left(1 - \varepsilon^{1-\alpha}\right).$$

Wir erkennen nun sofort, daß folgende Möglichkeiten bestehen. Erstens: Es ist $\alpha > 1$. Dann strebt die rechte Seite mit abnehmendem $\varepsilon$ gegen $\infty$. Zweitens: $\alpha < 1$. Dann strebt die rechte Seite gegen den Grenzwert $\frac{1}{1-\alpha}$. Im zweiten Falle werden wir also diesen Grenzwert einfach als das Integral zwischen den Grenzen 0 und 1 betrachten. Im ersten Falle werden wir sagen, daß dieses Integral nicht existiert. Im dritten Falle $\alpha = 1$ wird das Integral gleich $-\log \varepsilon$ sein und daher ebenfalls mit abnehmendem $\varepsilon$ keinem endlichen Grenzwert zustreben, sondern unendlich werden, d. h. nicht existieren.

Ein anderes Beispiel dafür, daß man eine Integration bis in eine Unendlichkeitsstelle einer Funktion hineinerstrecken kann, gibt der Integrand $\frac{1}{\sqrt{1-x^2}}$. Es wird

$$\int_0^{1-\varepsilon} \frac{dx}{\sqrt{1-x^2}} = \arcsin(1-\varepsilon).$$

Läßt man $\varepsilon$ gegen 0 streben, so wird die rechte Seite gegen einen bestimmten Grenzwert, nämlich $\frac{\pi}{2}$, konvergieren; man wird daher diesen Wert als $\int_0^1 \frac{dx}{\sqrt{1-x^2}}$ bezeichnen, obwohl der Integrand an der Stelle $x = 1$ unendlich wird.

Um aus diesen Beispielen einen allgemeinen Begriff zu abstrahieren, bedenken wir zunächst, daß es ganz gleichgültig ist, ob die Unstetigkeit einer zu integrierenden Funktion am oberen oder unteren Ende des Integrationsintervalles liegt; denn wir können obere und untere Grenze eines Integrales bei gleichzeitiger Änderung des Vorzeichens vertauschen. Nunmehr sagen wir: *Wenn in einem Intervall $a \leqq x \leqq b$*

*die Funktion $f(x)$ höchstens mit Ausnahme des Endpunktes $b$ stetig ist, so definieren wir als $\int_a^b f(x)\,dx$ den Grenzwert*

$$\lim_{\varepsilon \to 0} \int_a^{b-\varepsilon} f(x)\,dx$$

*— wobei die Stelle $b - \varepsilon$ vom Innern des Intervalles her gegen den Endpunkt $b$ strebt —, vorausgesetzt, daß dieser Grenzwert existiert.*

Wir sagen in diesem Falle, das Integral $\int_a^b f(x)\,dx$ ist ein *konvergentes uneigentliches Integral.* Wenn aber unser Grenzwert nicht existiert, sagen wir, daß das Integral $\int_a^b f(x)\,dx$ nicht existiert oder nicht konvergiert oder daß es *divergiert.*

Eine analoge Definition gilt nach dem Gesagten, wenn nicht die obere, sondern die untere Grenze des Integrals der Ausnahmepunkt ist.

Auch ein uneigentliches Integral kann man durch einen Flächeninhalt deuten. Es hat zwar zunächst keinen Sinn, von dem Flächeninhalt eines sich ins Unendliche erstreckenden Gebietes zu reden; aber man kann doch versuchen, einen solchen Flächeninhalt zu definieren, indem man einen Grenzübergang von beschränkten Gebieten mit endlichem Flächeninhalt vornimmt. Beispielsweise besagt das obige, auf die Funktionen $\frac{1}{x^\alpha}$ bezügliche Resultat, daß der Flächeninhalt, der von der $x$-Achse, der Geraden $x = 1$, der Geraden $x = \varepsilon$ und der Kurve $y = \frac{1}{x^\alpha}$ begrenzt ist, für $\varepsilon \to 0$ einem endlichen Grenzwert zustrebt, sobald $\alpha < 1$ ist, daß er aber unendlich wird, sobald $\alpha \geqq 1$ ist. Man drückt diese Tatsache einfach so aus: Der Flächeninhalt zwischen $x$-Achse, $y$-Achse, unserer Kurve und der Geraden $x = 1$ ist endlich bzw. unendlich.

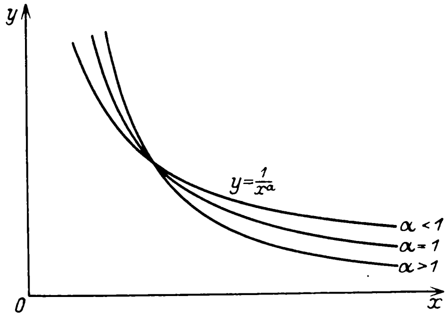


Fig. 77. Zur Konvergenz und Divergenz uneigentlicher Integrale.

Die Anschauung kann uns natürlich über Endlichkeit oder Unendlichkeit des Flächeninhaltes eines sich bis ins Unendliche erstreckenden Flächenstückes nichts Präzises aussagen. Man kann nur sagen, daß ein Flächenstück, das ins Unendliche reicht, um so eher doch noch einen endlichen Flächeninhalt besitzen wird, je schmaler es sich zusammenzieht. In diesem Sinne veranschaulicht uns die Figur 77 die Tatsache, daß für $\alpha < 1$ die Flächeninhalte unter unseren Kurven endlich bleiben, während sie für $\alpha \geqq 1$ unendlich werden.

Um zu erkennen, ob eine Funktion $f(x)$, welche für $x = b$ eine Unendlichkeitsstelle besitzt, sich in die Stelle $x = b$ hineinintegrieren läßt, wird man nicht stets eine besondere Untersuchung anstellen wollen, sondern sich häufig des folgenden *Kriteriums* bedienen können:

Es sei im Intervall $a \leqq x < b$ die Funktion $f(x)$ positiv[1]), und es gelte $\lim\limits_{x \to b} f(x) = \infty$. Dann *konvergiert das Integral* $\int\limits_a^b f(x)\,dx$, falls es eine unterhalb 1 liegende positive Zahl $\mu$ und eine feste, von $x$ unabhängige Schranke $M$ gibt, derart, daß in dem Intervall $a \leqq x < b$ $f(x) \leqq \frac{M}{(b-x)^\mu}$ bleibt; mit anderen Worten, *wenn die Funktion* $f(x)$ *an der Stelle* $x = b$ *von geringerer Größenordnung als von der ersten unendlich wird. Das Integral divergiert dagegen,* wenn es eine Zahl $\nu \geqq 1$ und eine positive Schranke $N$ gibt, derart, daß $f(x) \geqq \frac{N}{(b-x)^\nu}$ wird; mit anderen Worten, *wenn die Funktion an der Stelle* $x = b$ *von mindestens erster Ordnung unendlich wird.*

Der Beweis folgt fast unmittelbar aus dem Vergleich mit dem einfachsten schon oben diskutierten Fall. Um etwa den ersten Teil des Satzes zu beweisen, beachten wir, daß für $0 < \varepsilon < b - a$

$$0 \leqq \int\limits_a^{b-\varepsilon} f(x)\,dx \leqq \int\limits_a^{b-\varepsilon} \frac{M}{(b-x)^\mu}\,dx$$

ist und daß das Integral rechts für $\varepsilon \to 0$ — es entsteht aus dem Integral zu Anfang dieser Nummer durch einfache Bezeichnungsänderung — einen Grenzwert hat, also beschränkt bleibt, daß ferner die Werte $\int\limits_a^{b-\varepsilon} f(x)\,dx$ für $\varepsilon \to 0$ monoton zunehmen. Diese Werte besitzen also einen Grenzwert, d. h., das Integral $\int\limits_a^b f(x)\,dx$ konvergiert.

Den genau parallel laufenden Beweis für den zweiten Teil des Satzes kann ich dem Leser selbst überlassen.

Ebenso erkennt man sofort, daß genau die entsprechenden Sätze auch gelten, wenn die *untere* Grenze des Integrales eine Unendlichkeitsstelle ist. Liegt eine Unendlichkeitsstelle im Innern des betrachteten Intervalles, so braucht man dieses nur durch diese Stelle in zwei Teilintervalle zu zerlegen und auf jedes der beiden die obigen Betrachtungen anzuwenden.

Als weiteres Beispiel betrachten wir das elliptische Integral

$$\int\limits_0^1 \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} \qquad (k^2 < 1).$$

---

[1]) Im achten Kapitel, Anhang, werden wir übrigens sehen, daß eine solche Vorzeichenbeschränkung leicht beseitigt werden kann.

Man sieht sofort, daß für $x = 1$ der Integrand nur von der Größenordnung $\frac{1}{2}$ unendlich wird, woraus die Existenz des uneigentlichen Integrales folgt.

### 3. Unendliches Integrationsintervall.

Eine andere, ebenso wichtige Erweiterung des Integralbegriffes besteht darin, daß wir eine Integrationsgrenze ins Unendliche verlegen. Wir führen, um diese Erweiterung des Begriffes zu präzisieren, folgende Bezeichnungen ein: Wenn das Integral

$$\int_a^A f(x)\,dx$$

bei festem $a$ einem bestimmten Grenzwert zustrebt, sobald $A$ positiv über alle Grenzen wächst, bezeichnen wir diesen Grenzwert mit

$$\int_a^\infty f(x)\,dx$$

und nennen ihn das bis ins Unendliche erstreckte Integral der Funktion $f(x)$. Selbstverständlich braucht ein solches Integral nicht immer zu existieren oder, wie man sagt, zu *konvergieren*. Einfache Beispiele für die möglichen Vorkommnisse liefern uns wieder die Funktionen $f(x) = \frac{1}{x^\alpha}$:

$$\int_1^A \frac{dx}{x^\alpha} = \frac{1}{1-\alpha}(A^{1-\alpha} - 1);$$

wir erkennen hieraus, wenn wir wieder den Fall $\alpha = 1$ ausschließen, daß im Falle $\alpha > 1$ das ins Unendliche erstreckte Integral existiert, und zwar, daß genau

$$\int_1^\infty \frac{dx}{x^\alpha} = \frac{1}{\alpha - 1}$$

wird; daß dagegen im Falle $\alpha < 1$ das Integral nicht mehr existiert. Für den Fall $\alpha = 1$ existiert das Integral selbstverständlich auch nicht, da $\log x$ mit $x$ gegen Unendlich strebt. Wir sehen also, daß die Funktionen $\frac{1}{x^\alpha}$ hinsichtlich der ins Unendliche erstreckten Integrale sich anders verhalten als bei Integration in den Nullpunkt hinein. Auch diese Tatsache wird durch einen Blick auf die Figur 77 plausibel gemacht. Denn wir sehen, daß, je größer $\alpha$ ist, desto enger sich die Kurven für große Werte von $x$ an die $x$-Achse anschmiegen, so daß die Konvergenz des betreffenden Flächeninhaltes für größere Werte von $A$ verständlich wird.

Für die Existenz eines ins Unendliche erstreckten Integrales ist oft das folgende Kriterium nützlich, bei dem wir wiederum voraussetzen, daß für hinreichend große Werte von $x$, etwa für $x \geqq a$, der Integrand $f(x)$ von einem Zeichen bleibt — wir dürfen ihn ohne Beschränkung der Allgemeinheit positiv wählen[1] —. Dann gilt: *Das Integral* $\int_a^\infty f(x)\,dx$ *konvergiert, wenn die Funktion* $f(x)$ *im Unendlichen von höherer als erster Ordnung verschwindet,* d. h. wenn es eine Zahl $\nu > 1$ gibt, derart, daß für alle noch so großen Werte von $x$ die Beziehung $0 < f(x) \leqq \frac{M}{x^\nu}$ besteht, wobei $M$ eine von $x$ unabhängige feste Schranke bedeutet. *Ebenso divergiert das Integral, wenn die Funktion* $f(x)$ *positiv bleibt und im Unendlichen von nicht höherer als erster Ordnung verschwindet,* d. h. wenn es eine feste Schranke $N > 0$ gibt, derart, daß $x f(x) \geqq N$ bleibt. Der Beweis dieser Kriterien, der dem obigen vollständig parallel läuft, kann dem Leser überlassen bleiben.

Das einfachste Beispiel ist das Integral $\int_a^\infty \frac{1}{x^2}\,dx$ $(a>0)$. Der Integrand verschwindet im Unendlichen von zweiter Ordnung. Tatsächlich sehen wir sofort, daß das Integral konvergiert, denn es ist $\int_a^A \frac{1}{x^2}\,dx = \frac{1}{a} - \frac{1}{A}$, und so erhalten wir

$$\int_a^\infty \frac{1}{x^2}\,dx = \frac{1}{a}\,.$$

Ebenso nahe liegt das Beispiel

$$\int_0^\infty \frac{1}{1+x^2}\,dx = \lim_{A\to\infty} (\operatorname{arc\,tg} A - \operatorname{arc\,tg} 0) = \frac{\pi}{2}\,.$$

Ein weiteres, für die Analysis besonders wichtiges Beispiel bieten die sog. $\Gamma$-Integrale

$$\Gamma(n) = \int_0^\infty e^{-x} x^{n-1}\,dx \qquad (n \geqq 1)\,.$$

Auch bei ihnen ist das Konvergenzkriterium erfüllt; denn z. B. für $\nu = 2$ ist $\lim_{x\to\infty} x^\nu \cdot e^{-x} x^{n-1} = 0$, da ja die Exponentialfunktion $e^{-x}$ von höherer Ordnung Null wird als jede Potenz $\frac{1}{x^m}$ $(m > 0)$. Diese $\Gamma$-Integrale, welche wir als Funktionen $\Gamma(n)$ der (nicht notwendig ganzen) Zahl $n$ auffassen können, erfüllen eine bemerkenswerte Beziehung, zu welcher wir durch Produktintegration folgendermaßen gelangen:

$$\int e^{-x} x^{n-1}\,dx = -\,e^{-x} x^{n-1} + (n-1) \int e^{-x} x^{n-2}\,dx\,.$$

Nehmen wir die Formel zwischen den Grenzen 0 und $A$ und lassen

[1]) Die Aufhebung dieser Vorzeichenbeschränkung ergibt sich von selbst im Anhang zum achten Kapitel.

dann $A$ über alle Grenzen wachsen, so erhalten wir sofort

$$\Gamma(n) = (n-1)\int_0^\infty e^{-x} x^{n-2}\,dx = (n-1)\,\Gamma(n-1)$$

und aus dieser Rekursionsformel, falls $\mu$ eine ganze Zahl und $0 < \mu < n$ ist,

$$\Gamma(n) = (n-1)\cdot(n-2)\cdots(n-\mu)\int_0^\infty e^{-x} x^{n-\mu-1}\,dx.$$

Ist $n$ eine ganze positive Zahl, so ergibt sich

$$\Gamma(n) = (n-1)\,(n-2)\cdots 3\cdot 2\cdot 1\int_0^\infty e^{-x}\,dx,$$

und da

$$\int_0^\infty e^{-x}\,dx = 1$$

ist, folgt schließlich

$$\Gamma(n) = (n-1)\,(n-2)\cdots 2\cdot 1 = (n-1)!\,.$$

Diese Darstellung der Fakultäten durch Integrale spielt in sehr vielen Anwendungen eine große Rolle.

Auch die Integrale

$$\int_0^\infty e^{-x^2}\,dx, \qquad \int_0^\infty x^n e^{-x^2}\,dx$$

konvergieren, wovon man sich sofort nach unseren Kriterien überzeugt.

Ein für viele Anwendungen wichtiges konvergentes Integral, dessen Konvergenz wir nicht direkt nach dem obigen Kriterium nachweisen können, ist das „*Integral von Dirichlet*"

$$J = \int_0^\infty \frac{\sin x}{x}\,dx.$$

Seine Konvergenz beruht auf dem periodischen Vorzeichenwechsel des Integranden, wobei die von benachbarten Integrationsintervallen der Länge $\pi$ herrührenden Bestandteile sich gegenseitig nahezu zerstören. Um diesen Umstand auszunutzen, schreiben wir den Ausdruck

$$D_{AB} = \int_A^B \frac{\sin x}{x}\,dx$$

— dessen Verschwinden im Limes für unbegrenzt wachsendes $A$ und $B$ gleichbedeutend mit der zu beweisenden Konvergenz ist — in der Form

$$D_{AB} = \int_A^{A+\pi} \frac{\sin x}{x}\,dx - \int_B^{B+\pi} \frac{\sin x}{x}\,dx + \int_{A+\pi}^{B+\pi} \frac{\sin t}{t}\,dt,$$

führen in dem letzten der drei Integrale rechts $x = t - \pi$ als neue Integrationsvariable ein, wobei $\sin t = -\sin x$ wird, und erhalten

$$D_{AB} = \int_A^{A+\pi} \frac{\sin x}{x}\,dx - \int_B^{B+\pi} \frac{\sin x}{x}\,dx - \int_A^B \frac{\sin x}{x+\pi}\,dx.$$

Addition zu dem ursprünglichen Ausdruck für $D_{AB}$ ergibt

$$2D_{AB} = \int_A^{A+\pi} \frac{\sin x}{x}\,dx - \int_B^{B+\pi} \frac{\sin x}{x}\,dx + \pi \int_A^B \frac{\sin x}{x(x+\pi)}\,dx.$$

Hieraus folgt sofort, wenn wir etwa $B > A$ annehmen,

$$|2D_{AB}| < \frac{2\pi}{A} + \pi \int_A^B \frac{dx}{x^2}.$$

Da die rechte Seite mit wachsendem $A$ und $B$ gegen Null strebt (siehe die Betrachtungen auf S. 202), so ist damit die Konvergenz des Integrales $J$ bewiesen. Wir werden später in Kap. VIII, Anhang, § 3 einen weiteren Beweis dieser Tatsache kennenlernen und in Kap. IX, § 5 darüber hinaus feststellen, daß $J$ den Wert $\frac{\pi}{2}$ besitzt.

Selbstverständlich behalten alle Regeln über Substitution neuer Veränderlicher usw. auch bei konvergenten uneigentlichen Integralen ihre Gültigkeit. Um beispielsweise das Integral $\int_0^\infty x e^{-x^2}\,dx$ zu berechnen, führen wir $u = x^2$ als neue Veränderliche ein und erhalten

$$\int_0^\infty x e^{-x^2}\,dx = \frac{1}{2}\int_0^\infty e^{-u}\,du = \lim_{A\to\infty} \frac{1}{2}(1 - e^{-A}) = \frac{1}{2}.$$

Ein anderes Beispiel für die Anwendung der Transformationstheorie zur Untersuchung uneigentlicher Integrale liefern die in der Theorie der Lichtbeugung auftretenden „*Fresnelschen Integrale*“

$$F_1 = \int_0^\infty \sin(x^2)\,dx, \qquad F_2 = \int_0^\infty \cos(x^2)\,dx.$$

Die Substitution $x^2 = u$ liefert

$$F_1 = \frac{1}{2}\int_0^\infty \frac{\sin u}{\sqrt{u}}\,du, \qquad F_2 = \frac{1}{2}\int_0^\infty \frac{\cos u}{\sqrt{u}}\,du.$$

Nun wird nach Ausführung einer Produktintegration

$$\int_A^B \frac{\sin u}{\sqrt{u}}\,du = \frac{\cos A}{\sqrt{A}} - \frac{\cos B}{\sqrt{B}} - \frac{1}{2}\int_A^B \frac{\cos u}{u^{\frac{3}{2}}}\,du.$$

Hier strebt mit wachsendem $A$ und $B$ der erste Bestandteil rechts gegen Null, ebenso der zweite Bestandteil nach dem Kriterium von S. 202. Es ist somit die Konvergenz unseres Integrales $F_1$ bewiesen.

In genau derselben Weise wird der Konvergenzbeweis für $F_2$ geführt.

Vielfach wird durch eine Substitution ein uneigentliches Integral in ein eigentliches übergehen. Z. B. liefert die Transformation $x = \sin u$

$$\int_0^1 \frac{dx}{\sqrt{1-x^2}} = \int_0^{\frac{\pi}{2}} du = \frac{\pi}{2}.$$

Andererseits gehen ebenso Integrale über stetige Funktionen $f(x)$ in uneigentliche Integrale über, wenn in dem Endpunkt des Integrationsintervalles bei einer Transformation $u = \varphi(x)$ die Ableitung $\varphi'(x)$ verschwindet, also $\frac{dx}{du}$ unendlich wird.

Fünftes Kapitel.

# Anwendungen.

Wir wollen, nachdem wir eine gewisse Bewegungsfreiheit gewonnen haben, in diesem Kapitel die Anwendbarkeit des Gelernten nach verschiedenartigen Richtungen hin in Geometrie und Physik dartun.

## § 1. Darstellung von Kurven.

### 1. Die Parameterdarstellung.

Bei der Darstellung einer Kurve durch eine Funktion $y = f(x)$ müssen wir uns, wie wir im ersten Kapitel sahen, jeweils auf einen eindeutigen Zweig beschränken. Es ist daher vielfach, insbesondere, wenn es sich um geschlossene Kurven handelt, bequemer, andere analytische Darstellungen heranzuziehen. Die allgemeinste und zugleich handlichste Darstellungsform gibt uns die *Parameterdarstellung* von Kurven. Man betrachtet nicht die eine rechtwinklige Koordinate als Funktion der anderen, sondern man faßt beide Koordinaten $x$ und $y$ als Funktionen einer dritten unabhängigen Veränderlichen $t$, einer sog. Hilfsvariabeln oder eines Parameters, auf; dabei durchläuft dann der Punkt mit den Koordinaten $x$ und $y$ die Kurve, wenn $t$ ein bestimmtes Intervall durchläuft. Solche Parameterdarstellungen sind uns schon begegnet. Z. B. erhalten wir für den Kreis $x^2 + y^2 = a^2$ eine Parameterdarstellung in der Form $x = a\cos t$, $y = a\sin t$. Hier hat $t$ in der bekannten Weise die geometrische Bedeutung eines zum Kreise gehörigen Zentriwinkels. Ebenso ergibt sich für die Ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

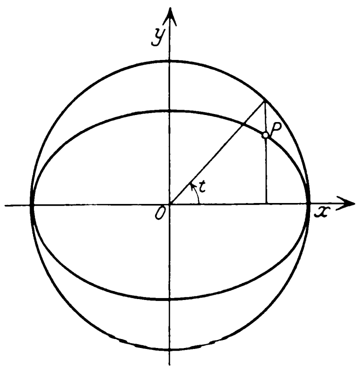


Fig. 78.

die Parameterdarstellung $x = a\cos t$, $y = b\sin t$, wobei $t$ die sog. exzentrische Anomalie bedeutet, nämlich den Zentriwinkel, der zu dem senkrecht über bzw. unter dem Ellipsenpunkt $P$ liegenden Punkte des umbeschriebenen Kreises gehört. (Fig. 78.) In diesen beiden Fällen beschreibt der Punkt mit den Koordinaten $x, y$ den ganzen Kreis bzw. die ganze Ellipse, wenn der Parameter $t$ das Intervall von 0 bis $2\pi$ durchläuft.

Allgemein können wir nun eine Kurve darzustellen versuchen, indem wir setzen

$$x = \varphi(t) = x(t), \qquad y = \psi(t) = y(t),$$

d. h. zwei Funktionen eines Parameters $t$ betrachten — die kürzere symbolische Bezeichnungsweise $x(t)$ und $y(t)$ werden wir fortan benutzen, wo sie zu keinem Mißverständnis Anlaß geben kann —. Und zwar müssen zu einer gegebenen Kurve diese beiden Funktionen $\varphi(t)$ und $\psi(t)$ so hinzubestimmt werden, daß man durch die Gesamtheit der Paare von Funktionswerten $x(t)$ und $y(t)$ eines gegebenen Intervalles für $t$ gerade die Kurvenpunkte und nur diese erhält. Ist eine Kurve zunächst in der Gestalt $y = f(x)$ gegeben, so kann man zu einer solchen Parameterdarstellung gelangen, indem man zuerst $x = \varphi(t)$ setzt, wo $\varphi(t)$ eine beliebige stetige monotone Funktion ist, welche in einem bestimmten Intervalle alle in Frage kommenden Werte von $x$ gerade einmal annimmt; es wird dann $y = f(\varphi(t)) = \psi(t)$, d. h. die zweite Funktion $\psi(t)$ bestimmt sich durch Zusammensetzung von $f$ und $\varphi$. Wir sehen hieraus, daß wir wegen der Willkür bei der Wahl der Funktion $\varphi$ noch eine große Freiheit in der Parameterdarstellung einer gegebenen Kurve haben; insbesondere können wir $t = x$ selbst wählen und daher die ursprüngliche Darstellung $y = f(x)$ als Parameterdarstellung mit dem Parameter $t = x$ auffassen.

Der Vorteil der Parameterdarstellung ist nun der, daß man die verbleibende Willkür zu einer Vereinfachung ausnutzen kann. Z. B. werden wir die Kurve $y = \sqrt[3]{x^2}$ darstellen, indem wir setzen: $x = t^3$, $y = t^2$, also $\varphi(t) = t^3$, $\psi(t) = t^2$. Es wird dann der Punkt mit den Koordinaten $x$, $y$ die ganze Kurve (Neilsche Parabel) durchlaufen, wenn $t$ von $-\infty$ bis $\infty$ variiert.

Ist umgekehrt eine Kurve von vornherein in der Parameterdarstellung $x = \varphi(t)$, $y = \psi(t)$ gegeben und will man die Kurvengleichung in rechtwinkeligen Koordinaten erhalten, so braucht man nur aus den beiden Gleichungen $x = \varphi(t)$, $y = \psi(t)$ den Parameter $t$ zu eliminieren. Bei der obigen Parameterdarstellung für Kreis und Ellipse gelingt dies ohne weiteres durch Quadrieren und Berücksichtigung der Gleichung $\sin^2 t + \cos^2 t = 1$. (Weiteres Beispiel siehe unten.) Allgemein hätte man $t$ aus der Gleichung $x = \varphi(t)$ durch die Umkehrfunktion $t = \Phi(x)$ auszudrücken und in $y = \psi(t)$ einzusetzen, um die Darstellung $y = \psi(\Phi(x)) = f(x)$ zu erhalten[1]. Allerdings muß man

---

[1]) Es kann dabei aber vorkommen, daß die so erhaltene Gleichung $y = f(x)$ mehr Punkte darstellt als die ursprüngliche Parameterdarstellung. So liefert z. B. $x = a \sin t$, $y = b \sin t$ nur das endliche zwischen den Punkten $x = -a$, $y = -b$ und $x = a$, $y = b$ gelegene Stück der Geraden $y = \frac{b}{a} x$, während die letztere Gleichung die ganze Gerade darstellt.

bei dieser Elimination sich im allgemeinen auf ein Stück der Kurve, nämlich einen eindeutig über der $x$-Achse liegenden Zweig, beschränken. Mit der Parameterdarstellung ist ein wachsenden Parameterwerten entsprechender *Durchlaufungssinn* der Kurve verbunden; wir werden diesen Durchlaufungssinn häufig als *positiven Durchlaufungssinn* bezeichnen.

In sehr vielen Fällen läßt sich dem Parameter $t$ eine unmittelbare physikalische Bedeutung geben, nämlich die Bedeutung der Zeit. Jede Bewegung eines Punktes in der Ebene wird mathematisch ihren Ausdruck darin finden, daß die Koordinaten $x$ und $y$ als Funktionen der Zeit erscheinen. Diese beiden Funktionen geben also in Parameterdarstellung die *Bewegung auf einer Bahnkurve*, z. B. bei den Zykloiden, die entstehen, wenn ein Kreis auf einer Geraden oder auf einem anderen Kreise abrollt. Jeder Punkt der rollenden Kreisscheibe beschreibt dabei eine Zykloide. Wir beschränken uns hier auf den einfachsten Fall, daß ein Kreis vom Radius $a$ auf der $x$-Achse rollt und ein Punkt der Kreisperipherie betrachtet wird. Dieser Punkt beschreibt dann eine „*gewöhnliche*“ *Zykloide*.

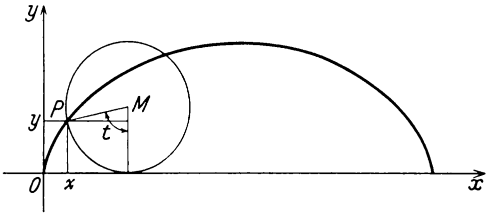


Fig. 79. Zykloide.

Wählen wir den Anfangspunkt des Koordinatensystemes und den der Zeitrechnung so, daß der Kurvenpunkt für $t = 0$ gerade in den Ursprung fällt, so ergibt sich (vgl. Fig. 79) für die Zykloide die Parameterdarstellung

$$x = a\,(t - \sin t)\,, \qquad y = a\,(1 - \cos t)\,;$$

dabei bedeutet $t$ den Winkel, um den sich der rollende Kreis aus seiner Anfangslage heraus gedreht hat und der bei gleichförmiger Rollbewegung der Zeit proportional ist.

Man kann nach Elimination des Parameters $t$ die Gleichung der Zykloide auch in rechtwinkligen Koordinaten schreiben, allerdings unter Preisgabe der Übersichtlichkeit. Man erhält

$$\cos t = \frac{a-y}{a}\,, \qquad t = \arccos\frac{a-y}{a}\,, \qquad \sin t = \pm\sqrt{1 - \frac{(a-y)^2}{a^2}}\,,$$

also

$$x = a \arccos\frac{a-y}{a} \mp \sqrt{(2a-y)\,y}\,,$$

d. h. $x$ als Funktion von $y$.

Für die Parameterdarstellung einer geometrisch gegebenen Kurve bleibt uns, wie schon auf S. 206 gesagt wurde, in der Wahl des Parameters noch eine große Freiheit. Man könnte z. B. statt der Zeit $t$ auch die Größe $\tau = t^2$ als Parameter wählen oder schließlich einen ganz beliebigen Parameter $\tau$, welcher mit dem ursprünglich gegebenen

Parameter $t$ durch eine beliebige Gleichung der Form $\tau = \omega(t)$ verknüpft ist, wobei wir voraussetzen, daß diese Gleichung durch eine Gleichung $t = \varkappa(\tau)$ in einem gewissen Intervall sich eindeutig umkehren läßt. Wenn dabei wachsenden Werten von $t$ wachsende Werte von $\tau$ entsprechen, so bleibt der positive Durchlaufungssinn erhalten; andernfalls wird er umgekehrt.

Natürlich kann man nicht nur die rechtwinkligen Koordinaten einer Kurve in Parameterdarstellung geben, sondern ebenso gut auch z. B. die *Polarkoordinaten* $r$ und $\vartheta$, welche in der bekannten Weise mit den rechtwinkligen Koordinaten durch die Gleichungen $x = r\cos\vartheta$, $y = r\sin\vartheta$ oder $r = \sqrt{x^2+y^2}$, $\vartheta = \operatorname{arc\,tg}\frac{y}{x}$ verknüpft sind, also $r = r(t)$, $\vartheta = \vartheta(t)$. Beispielsweise erhalten wir eine Gerade durch die Parameterdarstellung (siehe Fig. 80)

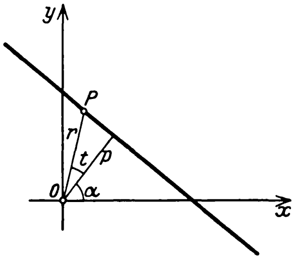


Fig. 80.

$$r = \frac{p}{\cos t}, \qquad \vartheta = \alpha + t$$

($p$ und $\alpha$ sind Konstanten), aus welcher durch Elimination des Parameters $t$ sich sofort

$$r = \frac{p}{\cos(\vartheta - \alpha)}$$

als Gleichung der Geraden in Polarkoordinaten ergibt.

## 2. Die zu einer Kurve gehörigen Differentialquotienten bei Parameterdarstellung.

Ist eine Kurve einmal durch eine Gleichung $y = f(x)$, andererseits in Parameterdarstellung durch $x = x(t)$, $y = y(t)$ gegeben, so muß $y(t) = f(x(t))$ sein. Nach der Kettenregel der Differentialrechnung ist dann

$$\frac{dy}{dt} = \frac{dy}{dx}\cdot\frac{dx}{dt}$$

oder

$$y' = \frac{dy}{dx} = \frac{\dot y}{\dot x},$$

wobei wir zur Abkürzung für die Differentiation nach dem Parameter $t$ an Stelle des Zeichens $'$ einen über die Größe gesetzten Punkt verwendet haben (im Anschluß an Newton).

Beispielsweise ergibt sich für die Zykloide

$$\dot x = a(1-\cos t) = 2a\sin^2\frac{t}{2}, \qquad \dot y = a\sin t = 2a\sin\frac{t}{2}\cos\frac{t}{2}.$$

Diese Formeln lassen erkennen, daß die Zykloide in den Punkten $t = 0, \pm 2\pi, \pm 4\pi, \ldots$, in denen sie die $x$-Achse trifft, eine Spitze mit senkrechter Tangente besitzt; denn es wird bei Annäherung an diese Stellen der Differentialquotient $y' = \frac{\dot y}{\dot x} = \operatorname{ctg}\frac{t}{2}$ unendlich werden; im Punkte selbst ist $y = 0$ und in der Umgebung sonst überall $y > 0$.

Die Gleichung der Tangente an unsere Kurve wird, wenn auf der Tangente die laufenden Koordinaten mit $\xi$ und $\eta$ bezeichnet werden, durch

$$(\xi - x)\,\dot{y} - (\eta - y)\,\dot{x} = 0$$

gegeben. Ebenso ist die Gleichung für die Normale der Kurve, d. h. die in einem Kurvenpunkt auf der Tangente senkrecht stehende Gerade,

$$(\xi - x)\,\dot{x} + (\eta - y)\,\dot{y} = 0\,.$$

Die *„Richtungskosinus" der Tangente*, d.h. die Cosinus der Winkel $\alpha$, $\beta$, welche die Tangente mit der $x$-Achse bzw. der $y$-Achse bildet, werden, wie man elementar bestätigt, durch die Ausdrücke

$$\cos\alpha = \frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}, \qquad \cos\beta = \frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}}$$

gegeben, während die entsprechenden *Richtungskosinus der Normalen* durch

$$\cos\alpha' = \frac{-\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}}, \qquad \cos\beta' = \frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}$$

geliefert werden (Fig. 81).

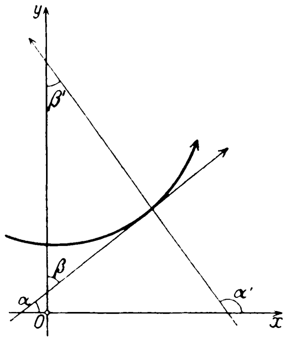


Fig. 81. Richtungskosinus für Tangente und Normale.

Aus einer bekannten Formel der Trigonometrie bzw. der analytischen Geometrie ergibt sich nun für den Winkel $\delta$ zwischen zwei Kurven mit den Darstellungen $x = \varphi(t)$, $y = \psi(t)$ bzw. $x_1 = \varphi_1(t)$, $y_1 = \psi_1(t)$ (d. h. für den Winkel zwischen ihren Tangenten oder Normalen) der Ausdruck

$$\cos\delta = \frac{\dot{x}\dot{x}_1 + \dot{y}\dot{y}_1}{\sqrt{\dot{x}^2 + \dot{y}^2}\sqrt{\dot{x}_1^2 + \dot{y}_1^2}}\,.$$

Die Unbestimmtheit des Vorzeichens der hier überall auftretenden Quadratwurzel deutet an, daß unsere Winkel nicht völlig bestimmt sind, da man auf der Tangente bzw. Normalen noch einen beliebigen Richtungssinn als „positiv" auszeichnen kann. Wenn wir, wie üblich, die Quadratwurzel positiv nehmen, so bedeutet dies, daß wir als *positive Tangentenrichtung* die nach wachsenden Parameterwerten weisende bezeichnen und als *positive Normalenrichtung* die hieraus durch positive Drehung [1]) um $\frac{\pi}{2}$ hervorgehende Drehung.

Die zweite Ableitung $y'' = \frac{d^2 y}{dx^2}$ erhalten wir durch Benutzung der Kettenregel und der Quotientenregel folgendermaßen:

$$y'' = \frac{dy'}{dx} = \frac{dy'}{dt}\cdot\frac{dt}{dx} = \frac{d}{dt}\left(\frac{\dot{y}}{\dot{x}}\right)\cdot\frac{1}{\dot{x}} = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{\dot{x}^2}\cdot\frac{1}{\dot{x}},$$

also

$$y'' = \frac{d^2y}{dx^2} = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{\dot{x}^3}\,.$$

[1]) D. h. eine Drehung im selben Sinne wie die kürzeste Drehung, welche die positive $x$-Achse in die positive $y$-Achse überführt.

**3. Übergang zu neuen Koordinatensystemen bei Parameterdarstellung.**

Drehen wir das Koordinatensystem um den Winkel $\alpha$ in positivem Sinne, so bestehen zwischen den neuen rechtwinkeligen Koordinaten $\xi$, $\eta$ und den alten $x$, $y$ die Beziehungen

$$x = \xi \cos\alpha - \eta \sin\alpha\,, \qquad \xi = x\cos\alpha + y\sin\alpha\,,$$
$$y = \xi \sin\alpha + \eta \cos\alpha\,, \qquad \eta = -x\sin\alpha + y\cos\alpha\,.$$

Es sind also zugleich mit $x$ und $y$ auch die neuen Koordinaten $\xi$ und $\eta$ als Funktionen des Parameters $t$ bekannt. Durch Differentiation erhalten wir unmittelbar

$$\dot{x} = \dot{\xi} \cos\alpha - \dot{\eta} \sin\alpha\,, \qquad \dot{\xi} = \dot{x}\cos\alpha + \dot{y}\sin\alpha\,,$$
$$\dot{y} = \dot{\xi} \sin\alpha + \dot{\eta} \cos\alpha\,, \qquad \dot{\eta} = -\dot{x}\sin\alpha + \dot{y}\cos\alpha \quad \text{usw.}$$

Ist die Kurve durch Polarkoordinaten gegeben und sind sowohl die Polarkoordinaten als auch die rechtwinkligen Koordinaten als Funktionen eines Parameters $t$ dargestellt, so folgen aus $x = r\cos\vartheta$, $y = r\sin\vartheta$ durch Differentiation nach $t$ die Beziehungen

$$(*) \qquad \begin{aligned} \dot{x} &= \dot{r}\cos\vartheta - r\sin\vartheta\cdot\dot{\vartheta}\,,\\ \dot{y} &= \dot{r}\sin\vartheta + r\cos\vartheta\cdot\dot{\vartheta}\,, \end{aligned}$$

die bei dem Übergang von rechtwinkligen Koordinaten zu Polarkoordinaten vielfach Anwendung finden. Als ein Beispiel hierfür betrachten wir die Gleichung einer Kurve in Polarkoordinaten $r = f(\vartheta)$, die etwa aus einer Parameterdarstellung $r = r(t)$, $\vartheta = \vartheta(t)$ durch Elimination des Parameters $t$ entstanden ist. Dann wird der Winkel $\mu$ zwischen dem Radiusvektor nach einem Kurvenpunkte und der dort an die Kurve gelegten Tangente durch

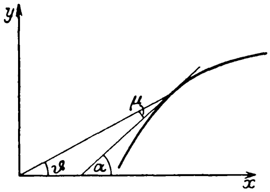


Fig. 82.

$$\operatorname{tg}\mu = \frac{f(\vartheta)}{f'(\vartheta)}$$

gegeben. Wir überzeugen uns von dieser Tatsache z. B. folgendermaßen. Denken wir uns die Kurve durch eine Gleichung $y = F(x)$ gegeben und als Parameter speziell $t = \vartheta$ verwandt, so daß $\dot{\vartheta} = 1$, $\dot{r} = f'(\vartheta)$ wird, so ist

$$y' = \operatorname{tg}\alpha = \frac{\dot{y}}{\dot{x}} = \frac{\dot{r}\operatorname{tg}\vartheta + r}{\dot{r} - r\operatorname{tg}\vartheta}$$

(vgl. Fig. 82 und (*)). Ferner ist $\mu = \alpha - \vartheta$, also

$$\operatorname{tg}\mu = \frac{y' - \operatorname{tg}\vartheta}{1 + y'\operatorname{tg}\vartheta} = \frac{r + r\operatorname{tg}^2\vartheta}{\dot{r} + \dot{r}\operatorname{tg}^2\vartheta} = \frac{r}{\dot{r}}\,,$$

eine Formel, die sich leicht auch geometrischen Überlegungen entnehmen läßt.

### 4. Allgemeine Bemerkungen.

Bei der Diskussion gegebener Kurven betrachtet man einmal Eigenschaften, die nichts über die Gestalt der Kurve selbst, sondern lediglich etwas über ihre Lage zum Koordinatensystem aussagen: z. B. das Auftreten einer horizontalen Tangente, ausgedrückt durch die Gleichung $\dot{y} = 0$, oder das Auftreten einer vertikalen Tangente, ausgedrückt durch $\dot{x} = 0$. Bei einer Drehung des Koordinatensystemes werden solche Eigenschaften nicht bestehen bleiben.

Dagegen wird ein Wendepunkt der Kurve auch nach einer Drehung ein Wendepunkt bleiben. Die Bedingung für einen Wendepunkt lautet nämlich in Parameterdarstellung mit Rücksicht auf unsere obige Umrechnungsformel

$$\dot{x}\ddot{y} - \ddot{x}\dot{y} = 0.$$

Ersetzen wir links die Ausdrücke $\dot{x}, \dot{y}, \ddot{x}, \ddot{y}$ durch ihre Werte in den neuen Koordinaten $\xi, \eta$, so finden wir sehr leicht

$$\dot{x}\ddot{y} - \ddot{x}\dot{y} = \dot{\xi}\ddot{\eta} - \ddot{\xi}\dot{\eta},$$

und so ergibt sich, daß aus $\dot{x}\ddot{y} - \ddot{x}\dot{y} = 0$ auch $\dot{\xi}\ddot{\eta} - \ddot{\xi}\dot{\eta} = 0$ folgt, daß also unsere Gleichung eine vom Koordinatensystem unabhängige Eigenschaft des Punktes der Kurve ausdrückt.

Wir werden noch öfter sehen, daß eigentliche geometrische Eigenschaften ihren Ausdruck in Formeln finden, deren Gestalt sich bei einer Koordinatendrehung nicht ändert.

## § 2. Anwendung auf die Theorie der ebenen Kurven.

Wir werden bei Kurven zwei verschiedene Arten von geometrischen Eigenschaften oder Größen betrachten, solche, die nur von dem *Verhalten der Kurve im Kleinen*, d. h. in der unmittelbaren Umgebung eines Punktes, abhängen und die sich analytisch mit Hilfe der Differentialquotienten in diesem Punkte ausdrücken lassen, und solche, die mit dem *Gesamtverlauf der Kurve* oder eines Kurvenstückes zusammenhängen und ihre analytische Formulierung mit Hilfe des Integralbegriffes finden. Wir beschäftigen uns zunächst mit Eigenschaften des zweiten Typus.

### 1. Der Flächeninhalt in rechtwinkligen Koordinaten.

Der Flächeninhalt war unser Ausgangspunkt für die Integraldefinition; aber dem Zusammenhang zwischen bestimmtem Integral und Flächeninhalt haftet noch etwas Unbefriedigendes an. Das, worauf es uns in der Geometrie ankommt, ist der Flächeninhalt, der von beliebig gegebenen geschlossenen Kurven eingegrenzt wird; dagegen besteht bei dem Integral $\int_{x_0}^{x_1} y\,dx$ die Begrenzung des Flächeninhaltes nur zu einem Teil aus der jeweils vorgegebenen Kurve $y = f(x)$, zum anderen Teile aber aus Linien, welche von der Willkür des Koordinatensystemes abhängen. Will man den Flächeninhalt einer geschlossenen Kurve, wie Kreis oder Ellipse, durch Integrale dieser Art bestimmen, so muß man etwa die Fläche in mehrere Teile zerlegen, deren jeder von einem eindeutig über der $x$-Achse liegenden Ast der Kurve und der $x$-Achse sowie den zugehörigen Ordinaten begrenzt ist.

Für die Diskussion dieses allgemeinen Falles ist es zweckmäßig, einige Bemerkungen über die Bestimmung des Vorzeichens der betrachteten Flächeninhalte vorauszuschicken. Wir haben im zweiten Kapitel gesehen, daß durch bestimmte Integrale nicht Flächeninhalte an und für sich gegeben werden, sondern daß den dort betrachteten Flächenstücken durch den Integralausdruck ein bestimmtes Vorzeichen zugeschrieben wird. Wir können nun diese Vorzeichenbestimmung des Flächeninhaltes — und zwar sogleich für ein beliebiges von einer geschlossenen Kurve begrenztes Flächenstück — mit dem rein geometrischen Begriff des Umlaufungssinnes durch folgende Festsetzungen in Verbindung bringen: Wir sagen, daß *ein Flächenstück positiv umlaufen wird*, wenn man seine Berandung so durchläuft, daß dabei das Innere des Flächenstückes zur Linken bleibt[1]; den entgegengesetzten Umlaufssinn nennen wir negativ. Bei einem mit einem Umlaufungssinn versehenen Flächenstück, einem sog. *orientierten Flächenstück*, soll dann der Flächeninhalt negativ bzw. positiv gerechnet werden, je nachdem, ob der Umlaufungssinn positiv oder negativ ist (vgl. Fig. 83). Daß man dem Umlaufungssinn gerade das entgegengesetzte Vorzeichen wie dem Flächeninhalt zuschreibt, ist eine an sich willkürliche, aber, wie sich zeigt, durchaus zweckmäßige Festsetzung.

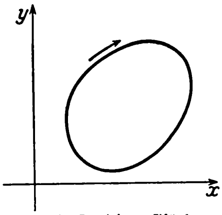


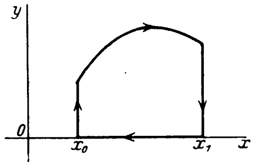
Fig. 83. Positiver Flächeninhalt.

Nunmehr betrachten wir speziell den Linienzug, der aus dem vom Punkte $x = b = x_1$ bis zum Punkte $x = a = x_0$ durchlaufenen Stück der $x$-Achse, der anschließenden Ordinate $x = a = x_0$ bis zur Kurve $y = f(x)$, dem oben betrachteten Kurvenstück und schließlich dem Stück der Ordinate $x = b = x_1$ von der Kurve bis zur $x$-Achse besteht (vgl. Fig. 84 und 85). Dieser Linienzug grenzt ein oder mehrere Flächenstücke ein, die positiv oder negativ umlaufen sein können. Ihre Flächeninhalte, mit dem nach obiger Regel bestimmten Vorzeichen versehen und addiert, werden dann gerade durch das Integral $F_{01} = \int_a^b f(x)\,dx$ ge-

[1] Will man die Worte „links" und „rechts" bei einer solchen Erklärung vermeiden, so sagt man: Das Dreieck, dessen Eckpunkte der Reihe nach der Nullpunkt, der Punkt $x = 1$, $y = 0$, endlich der Punkt $x = 0$, $y = 1$ sind, wird *positiv* umlaufen, wenn die angegebene Reihenfolge der Ecken innegehalten wird. Jede im selben Sinne umlaufene Fläche heißt in diesem Koordinatensystem positiv, jede im entgegengesetzten Sinne umlaufene Fläche negativ umlaufen.

geben, wie ein Vergleich mit den Festsetzungen vom zweiten Kapitel, § 1, zeigt.

Nach diesen Vorbemerkungen ist es nun in einfachster Weise möglich, den zu Anfang genannten Schwierigkeiten zu entgehen, indem man nämlich die Parameterdarstellung $x = x(t)$, $y = y(t)$ für unsere Kurve heranzieht. Führt man nach der Substitutionsregel formal $t$



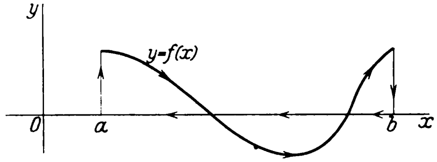




Fig. 84. Vorzeichen des Flächeninhaltes. Fig. 85.

als neue unabhängige Veränderliche in dem obigen Integral ein, so erhält man die Darstellung

$$F_{01} = \int_{t_0}^{t_1} y(t)\,\dot{x}(t)\,dt,$$

wo $t_0$ und $t_1$ die den Abszissen $x_0 = a$ und $x_1 = b$ entsprechenden Parameterwerte sind. Dabei ist zunächst vorausgesetzt, daß dem betreffenden Zweig der Kurve $y = f(x)$ in eindeutiger Weise ein Intervall $t_0 \leqq t \leqq t_1$ zugeordnet ist, daß etwa $\dot{x}(t)$ in diesem Intervall nicht verschwindet. Unser Ausdruck stellt dann in der eben geschilderten Weise das zum Kurvenbogen gehörige Flächenstück dar, d. h. den Flächeninhalt, der begrenzt wird von diesem Kurvenbogen, den beiden Geraden $x = x_0$ und $x = x_1$ sowie der $x$-Achse. Es ist nun das Eigentümliche unseres Ausdruckes für den Flächeninhalt bei Parameterdarstellung, daß er auch für geschlossene Kurven gültig bleibt.

Eine geschlossene Kurve ist in Parameterdarstellung durch zwei Funktionen $x = x(t)$ und $y = y(t)$ gegeben, welche den Bedingungen $x(t_0) = x(t_1)$ und $y(t_0) = y(t_1)$ genügen müssen, wenn bei Durchlaufung des Parameterintervalles $t_0 \leqq t \leqq t_1$ die ganze Kurve einmal umlaufen wird. Wenn die Kurve nirgends Ecken hat, so dürfen wir die Ableitungen $\dot{x}(t)$ und $\dot{y}(t)$ als durchweg vorhanden und stetig voraussetzen, während etwa vorhandenen Ecken Unstetigkeitspunkte dieser Ableitungen entsprechen. Durchläuft der Parameter $t$ sein Intervall $t_0 \leqq t \leqq t_1$, so entspricht dem ein bestimmter Umlaufungssinn der Kurve, den wir, wie gesagt, positiv rechnen wollen, wenn er im entgegengesetzten Sinne des Uhrzeigers geht, andernfalls negativ. — In Figur 83 ist der negative Umlaufungssinn durch einen Pfeil angedeutet.

Sehen wir zunächst von Ecken ab und setzen die Kurve als konvex voraus, so daß sie von einer geraden Linie in höchstens zwei Punkten geschnitten wird. Diejenigen Stellen, an welchen die Kurve vertikale Tangenten oder, wie man auch sagt, „*Stützgeraden*" besitzt, an denen also $\dot{x}(t) = 0$ ist, seien $P_1$ und $P_2$. Wir können dann ihren Flächeninhalt wie in Figur 86 auffassen als die Summe des unterhalb des oberen Bogens $P_1 P_2$ liegenden Flächeninhaltes $F_{12}$ und des unterhalb des unteren Bogens $P_2 P_1$ liegenden Flächeninhaltes $F_{21}$. Dabei nehmen wir den Umlaufssinn, wie in der Figur, negativ an; $F_{\alpha\beta}$ bedeutet wie oben den mit einem Vorzeichen versehenen zu dem betreffenden

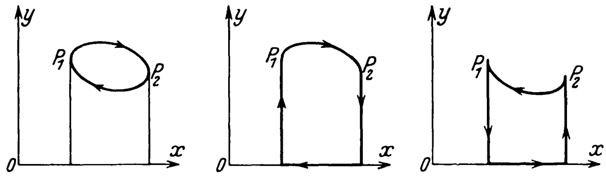



Fig. 86. Flächeninhalt geschlossener Kurven.

Kurvenbogen gehörigen Flächeninhalt, wobei hier $F_{12}$ positiv, $F_{21}$ negativ sein wird. Wir nehmen dabei an, daß der Punkt $x(t)$, $y(t)$ auf dem oberen Teil der Kurve von $P_1$ bis $P_2$ läuft, wenn $t$ von $t_0$ bis $\tau$ geht, und auf dem unteren Teil von $P_2$ bis $P_1$ bei $\tau \leqq t \leqq t_1$. Wir erhalten dann sofort

$$F_{12} = \int_{t_0}^{\tau} y(t)\,\dot{x}(t)\,dt$$

und

$$F_{21} = \int_{\tau}^{t_1} y(t)\,\dot{x}(t)\,dt;$$

es ergibt sich daher für den gesamten Flächeninhalt der konvexen Kurve

$$F = \int_{t_0}^{t_1} y(t)\,\dot{x}(t)\,dt.$$

Dieser Ausdruck stellt uns nun stets bis aufs Vorzeichen den Flächeninhalt dar. Ändert man den Durchlaufungssinn der Kurve, so bedeutet dies, daß man in dem Integral den Parameter $t$ nicht von $t_0$ bis $t_1$, sondern umgekehrt von $t_1$ nach $t_0$ laufen läßt; das Integral wechselt dann sein Vorzeichen, und wir erkennen: *Der durch unsere Formel dargestellte Flächeninhalt besitzt ein positives oder negatives Vorzeichen, je nachdem der Umlaufungssinn der Kurve negativ oder positiv ist*[1]).

[1]) Wir haben in der Figur angenommen, daß längs der Kurve überall $y > 0$ ist. Tatsächlich liegt hierin keine Beschränkung der Allgemeinheit. Denn verschieben wir die Kurve parallel der $y$-Achse um die Strecke $a$, ersetzen mit anderen Worten $y$ durch $y + a$, so ändert sich der Flächeninhalt nicht; zugleich bleibt der Wert des Integrals erhalten, da bei einer solchen Verschiebung an Stelle des obigen Integrales das Integral $\int_{t_0}^{t_1} (y + a)\,\dot{x}\,dt$ tritt und da wegen der Bedingung für die Geschlossenheit der Kurve $\int_{t_0}^{t_1} a\,\dot{x}\,dt = a(x(t_1) - x(t_0)) = 0$ ist.

Wir erweitern nun unser Ergebnis mit Hilfe zweier Bemerkungen. Einmal können wir ohne weiteres auch Ecken der Kurve zulassen, ohne daß die Gültigkeit unserer Formel aufhört. Es wird dann lediglich die Ableitung $\dot{x}(t)$ oder $\dot{y}(t)$ an den Ecken sprunghafte Unstetigkeiten aufweisen; unser Integral behält dann nach dem vierten Kapitel, § 8, seinen Sinn. Zweitens gilt unsere Formel auch noch dann, wenn die Kurve nicht mehr die einfache konvexe Gestalt aus Figur 86 hat, sondern eine allgemeinere geschlossene Form zeigt, wie z. B. Figur 87. Wir zerlegen dann einfach die Kurve durch die Punkte $P_0, P_1, P_2, \ldots,$ in welchen vertikale Stützgeraden auftreten, in eindeutig über der $x$-Achse liegende Äste und erhalten dann sofort wie in der Figur den umgrenzten Flächeninhalt $F$ in der Gestalt $F = F_{01} + F_{12} + F_{23} + F_{30}$. Drücken wir jeden dieser Teilflächeninhalte in Parameterdarstellung aus und fügen diese Ausdrücke zu

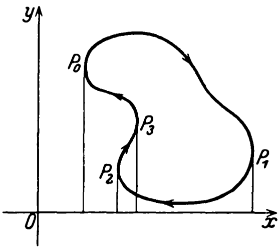



Fig. 87.

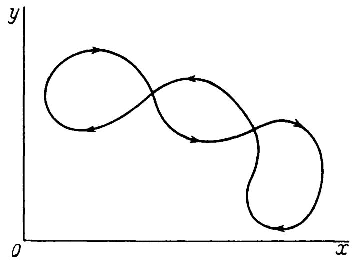



Fig. 88.

einem einzigen Integrale zusammen, so erhalten wir genau wie oben als Flächeninhalt der geschlossenen Kurve den Ausdruck

$$\int_{t_0}^{t_1} y\,\dot{x}\,dt,$$

dessen entgegengesetzt genommenes Vorzeichen uns gleichzeitig den Umlaufungssinn angibt.

Das Integral auf der rechten Seite unserer Formel hat sogar dann noch einen Sinn, *wenn die geschlossene Kurve sich überschlägt* (vgl. Fig. 88). *Es stellt dann die Summe der negativ umlaufenen vermindert um die Summe der positiv umlaufenen von der Kurve begrenzten Flächenstücke dar.* Unsere Darstellung des Flächeninhaltes gilt für jede geschlossene Kurve, wenn $x(t)$, $y(t)$ stetige Funktionen sind, deren Ableitungen bis auf höchstens endlich viele Sprungstellen stetig bleiben und für welche es nur endlich viele vertikale Stützgeraden gibt.

Man kann unsere Formel für den Flächeninhalt in eine elegantere symmetrische Gestalt bringen, wenn man das Integral zunächst durch Produktintegration umformt:

$$\int_{t_0}^{t_1} y\,\dot{x}\,dt = -\int_{t_0}^{t_1} x\,\dot{y}\,dt + x\,y\Big|_{t_0}^{t_1},$$

woraus wegen

$$x(t_0) = x(t_1), \quad y(t_0) = y(t_1)$$

sich sofort

$$F = \int_{t_0}^{t_1} y\dot{x}\,dt = -\int_{t_0}^{t_1} x\dot{y}\,dt$$

ergibt [1]). Bilden wir das arithmetische Mittel beider Ausdrücke, so erhalten wir die *symmetrische Darstellung*

$$F = \frac{1}{2}\int_{t_0}^{t_1} (y\dot{x} - x\dot{y})\,dt. \text{ [2])}$$

---

[1]) Statt den zweiten Ausdruck für den Flächeninhalt durch Produktintegration zu finden, hätten wir auch von der Bemerkung ausgehen können, daß für die Definition des Inhaltes die $x$-Achse und die $y$-Achse gleichberechtigt sind, abgesehen davon, daß der Drehungssinn, der die $x$-Achse auf dem kürzesten Wege in die $y$-Achse überführt, demjenigen entgegengesetzt ist, durch den umgekehrt die $y$-Achse auf dem kürzesten Wege in die $x$-Achse übergeführt wird.

[2]) An diese Formel knüpfe ich beiläufig eine Bemerkung von grundsätzlichem Interesse an: Man bestätigt, daß unsere Flächeninhaltsdefinition durch ein Integral tatsächlich von der speziellen Wahl des rechtwinkeligen Koordinatensystemes nicht abhängt, wie es bei jeder echten geometrischen Größe der Fall sein muß. Denken wir uns nämlich das Koordinatensystem irgendwie um den Winkel $\alpha$ gedreht, indem wir statt $x$ und $y$ durch die Gleichungen $x = \xi\cos\alpha - \eta\sin\alpha$, $y = \xi\sin\alpha + \eta\cos\alpha$ neue Veränderliche $\xi$ und $\eta$ einführen, die dann ihrerseits wieder Funktionen des Parameters $t$ werden; berücksichtigen wir dann, daß $\dot{x} = \dot{\xi}\cos\alpha - \dot{\eta}\sin\alpha$ und $\dot{y} = \dot{\xi}\sin\alpha + \dot{\eta}\cos\alpha$ ist, so finden wir durch eine kurze Rechnung $y\dot{x} - x\dot{y} = \eta\dot{\xi} - \xi\dot{\eta}$, so daß

$$F = \frac{1}{2}\int_{t_0}^{t_1} (y\dot{x} - x\dot{y})\,dt = \frac{1}{2}\int_{t_0}^{t_1} (\eta\dot{\xi} - \xi\dot{\eta})\,dt$$

wird, eine Gleichung, welche die behauptete Unabhängigkeit der Flächeninhaltsdefinition vom Koordinatensystem ausspricht, indem sie zeigt, daß der Flächeninhalt sich im $\xi\eta$-System durch dieselbe Formel wie im ursprünglichen $xy$-System ausdrückt. Ebenso wird der Flächeninhalt nach unserer Definition von der Wahl des Parameters $t$ unabhängig; d. h. unsere Formel behält ihre Gestalt, wenn wir statt $t$ durch die Gleichung $\tau = \tau(t)$ einen neuen Parameter $\tau$ einführen. Es wird nämlich

$$\frac{dx}{dt} = \frac{dx}{d\tau}\cdot\frac{d\tau}{dt}, \quad \frac{dy}{dt} = \frac{dy}{d\tau}\cdot\frac{d\tau}{dt},$$

also

$$\int_{t_0}^{t_1}\left(y\frac{dx}{dt} - x\frac{dy}{dt}\right)dt = \int_{t_0}^{t_1}\left(y\frac{dx}{d\tau} - x\frac{dy}{d\tau}\right)\frac{d\tau}{dt}\,dt = \int_{\tau_0}^{\tau_1}\left(y\frac{dx}{d\tau} - x\frac{dy}{d\tau}\right)d\tau,$$

wobei $\tau_0$ und $\tau_1$ die den Parameterwerten $t_0$ und $t_1$ entsprechenden Anfangs- und Endwerte des neuen Parameters sind.

Wir waren bisher bei der Definition des Flächeninhaltes vom Integralbegriff ausgegangen und haben nun gezeigt, daß diese analytische Flächeninhaltsdefinition

Als Beispiel für die Anwendung unserer Flächeninhaltsformel betrachten wir die Ellipse $y = \frac{b}{a}\sqrt{a^2 - x^2}$. Um ihren Inhalt in rechtwinkligen Koordinaten auszudrücken, können wir die obere und die untere Hälfte der Ellipse getrennt betrachten und ihn demgemäß durch das Integral

$$2\frac{b}{a}\int_{-a}^{a}\sqrt{a^2 - x^2}\,dx$$

darstellen. In der Parameterdarstellung $x = a\cos t$, $y = b\sin t$ jedoch erhalten wir unmittelbar für den Flächeninhalt, absolut genommen, den Ausdruck

$$ab\int_{0}^{2\pi}\sin^2 t\,dt,$$

der sich zufolge dem vierten Kapitel, § 4, integrieren läßt und den Wert $ab\pi$ besitzt.

## 2. Flächeninhalt in Polarkoordinaten.

Für viele Zwecke ist die Darstellung des Flächeninhaltes in Polarkoordinaten von Wichtigkeit. Es sei also $r = f(\vartheta)$ die Gleichung einer Kurve in Polarkoordinaten. $F(\vartheta)$ sei der Inhalt der Fläche, die von der $x$-Achse, dem Strahl durch den Nullpunkt, welcher mit ihr den Winkel $\vartheta$ bildet, und dem zwischenliegenden Stück der Kurve begrenzt wird. Dann ist

$$F'(\vartheta) = \frac{1}{2}r^2.$$

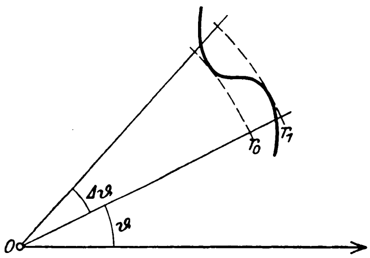


Fig. 89. „Flächenelement" bei Polarkoordinaten.

Denn betrachten wir außer dem zu $\vartheta$ gehörigen Radiusvektor den zum Winkel $\vartheta + \Delta\vartheta$ gehörigen und ist $r_0$ der kleinste, $r_1$ der größte Radiusvektor in diesem Winkelintervall (siehe Fig. 89), so wird der zwischen dem Radiusvektor $\vartheta$ und dem Radiusvektor $\vartheta + \Delta\vartheta$ gelegene Sektor einen Flächeninhalt $\Delta F$ besitzen, welcher zwischen den Grenzen $\frac{1}{2}r_0^2\Delta\vartheta$ und $\frac{1}{2}r_1^2\Delta\vartheta$ liegt; es ist also $\frac{1}{2}r_0^2 \leqq \frac{\Delta F}{\Delta\vartheta} \leqq \frac{1}{2}r_1^2$, und der Grenzüber-

tatsächlich geometrischen Charakter trägt, indem sie ihren Ausdruck unabhängig vom Koordinatensystem findet. Man kann aber leicht auch unmittelbar geometrisch den Flächeninhalt der oben betrachteten Kurven folgendermaßen definieren: Der Flächeninhalt ist der obere Häufungswert der Flächeninhalte aller im Innern der Kurve liegenden Polygone. Den an sich einfachen Beweis für die Äquivalenz beider Definitionen übergehe ich hier.

gang $\Delta\vartheta \to 0$ liefert uns sofort die obige Relation. Aus ihr folgt nach dem Fundamentalsatz der Integralrechnung für den Inhalt des Sektors zwischen den Polarwinkeln $\alpha$ und $\beta$ der Ausdruck

$$\frac{1}{2}\int_{\alpha}^{\beta} r^2\,d\vartheta.$$

Als Beispiel betrachten wir etwa die Fläche, die von einer Lemniskatenschleife begrenzt ist. Dabei läuft der Winkel $\vartheta$ von $-\frac{\pi}{4}$ bis $+\frac{\pi}{4}$, und wir erhalten für diesen Flächeninhalt gemäß der Lemniskatengleichung $r^2 = 2a^2\cos 2\vartheta$ (vgl. S. 55) den Ausdruck

$$a^2\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \cos 2\vartheta\,d\vartheta,$$

der sich durch Einführung der neuen Veränderlichen $u = 2\vartheta$ sofort ausrechnen läßt und den Wert $a^2$ hat.

### 3. Länge einer Kurve.

Der zweite wichtige, mit einer Kurve zusammenhängende geometrische Begriff, der auf eine Integration führt, ist die *Bogenlänge*.

Wir machen uns zunächst geometrisch klar, wie wir die Länge einer beliebigen Kurve zu definieren haben. Der elementare Prozeß des Messens einer Länge besteht darin, daß man die zu messenden Längen mit geradlinigen Maßstäben vergleicht; das einfachste Verfahren wird dann das sein, daß man einen solchen Maßstab auf der zu messenden Linie immer wieder abträgt und zählt, wie oft diese Abtragung hintereinander möglich ist; daß man sodann je nach Bedarf diesen Messungsprozeß verfeinert, indem man zum Gebrauch immer kürzerer Maßstäbe übergeht. Dieser elementaren anschaulichen Vorstellung entsprechend werden wir bei der Definition der Länge einer Kurve folgendermaßen vorgehen: Wir schreiben der Kurve ein geradliniges Polygon ein und messen dessen Länge. Diese wird davon abhängen, in welcher Art dieses Polygon gewählt wird, z. B. wie groß die Anzahl seiner Ecken genommen wird. Lassen wir die Anzahl der Seiten eines solchen, einem Kurvenbogen einbeschriebenen Polygones über alle Grenzen wachsen, während gleichzeitig die Länge der längsten Polygonseite gegen Null strebt, so werden wir den Grenzwert der Gesamtlängen der Polygone als die Länge des betreffenden Kurvenbogens bezeichnen. Diese Definition der Länge setzt voraus, daß einmal der beschriebene Grenzwert existiert und daß er zweitens unabhängig von der speziellen Wahl der Polygonfolge ist, mit der man die Kurve immer feiner annähert. Nur wenn diese Voraussetzung der sogenannten „*Rektifizierbarkeit*" erfüllt ist, werden wir von der Länge einer Kurve sprechen können. Wir werden bald sehen, daß unter sehr weiten Voraussetzungen die Rektifizierbarkeit sich leicht beweisen läßt.

Um die Länge durch einen analytischen Ausdruck, und zwar durch ein Integral darzustellen, denken wir uns die Kurve zunächst durch eine Funktion $y = f(x)$ mit stetiger Ableitung $y'$ repräsentiert und das Intervall $a \leqq x \leqq b$ der $x$-Achse, welches dem betrachteten Kurvenbogen entspricht, durch die Punkte $a = x_1, \ldots, x_n = b$ in $n - 1$ Teile von der Länge $\Delta x_1, \ldots, \Delta x_{n-1}$ eingeteilt. Über diesen Teilpunkten der $x$-Achse mögen die Ecken des einbeschriebenen Polygones liegen. Die Gesamtlänge des einbeschriebenen Polygones wird dann gemäß dem pythagoräischen Lehrsatz (vgl. Fig. 90) durch den Ausdruck

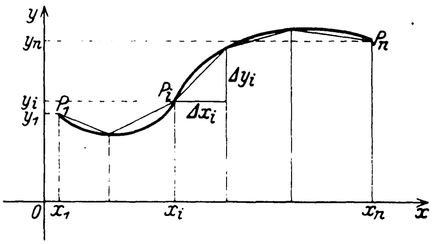


Fig. 90. Zur Rektifizierung von Kurven.

$$\sum_{\nu=1}^{n-1} \sqrt{\Delta x_\nu^2 + \Delta y_\nu^2} = \sum_{\nu=1}^{n-1} \sqrt{1 + \left(\frac{\Delta y_\nu}{\Delta x_\nu}\right)^2}\, \Delta x_\nu$$

gegeben. Der Differenzenquotient $\frac{\Delta y_\nu}{\Delta x_\nu}$ ist aber nach dem Mittelwertsatz der Differentialrechnung gerade gleich $f'(\xi_\nu)$, wo $\xi_\nu$ ein Zwischenwert im Intervall $\Delta x_\nu$ ist. Läßt man nun $n$ über alle Grenzen wachsen und dabei das längste Intervall $\Delta x_\nu$ gegen Null streben, so wird gemäß der Definition des Integrales unser Ausdruck gegen den Grenzwert

$$\int_a^b \sqrt{1 + y'^2}\, dx$$

streben, der somit die Länge unserer Kurve zwischen den Punkten mit den Abszissen $a$ und $b$ darstellt.

Da unser Grenzübergang von der Summe zum Integral stets zum selben Resultat führt, unabhängig von der Art der Einteilung des Intervalles, so zeigen uns unsere Betrachtungen die Gültigkeit des folgenden Satzes: *Jede Kurve $y = f(x)$ mit stetigem Differentialquotienten ist rektifizierbar, und ihre Länge zwischen den Punkten $x = a$ und $x = b \geqq a$ ist gegeben durch:*

$$s(a, b) = \int_a^b \sqrt{1 + y'^2}\, dx.$$

Für den Differentialquotienten der Bogenlänge nach der unabhängigen Veränderlichen $x$ entnimmt man, wenn man diese Bogenlänge, von irgend einem Anfangspunkt aus bis zur Abszisse $x$ gezählt, mit $s$ bezeichnet, aus unserer obigen Gleichung sofort die Beziehung

$$\frac{ds}{dx} = \sqrt{1 + y'^2}.$$

Unserem Ausdruck für die Bogenlänge haftet noch die spezielle und künstliche Voraussetzung an, daß der betrachtete Kurvenbogen als eindeutiger Zweig über der $x$-Achse liegt. Von dieser Einschränkung befreit uns die Parameterdarstellung. Ist unsere Kurve in Parameterdarstellung $x = x(t)$, $y = y(t)$ gegeben, so finden wir durch Einführung des Parameters $t$ in unsern obigen Ausdruck sofort die Parameterdarstellung der Bogenlänge

$$s(\alpha, \beta) = \int_\alpha^\beta \sqrt{\dot{x}^2 + \dot{y}^2}\, dt,$$

wobei $\alpha$ und $\beta$ die Werte von $t$ sind, die zu unseren Kurvenpunkten $x = a$ und $x = b$ gehören.

Diese Parameterdarstellung der Länge besitzt gegenüber der früheren den großen Vorzug, daß sie wiederum nicht mehr auf eindeutige Zweige der Kurven — dargestellt durch Gleichungen der Form $y = f(x)$ — beschränkt ist, sondern für beliebige Kurvenbögen, auch für geschlossene Kurven gilt, vorausgesetzt, daß längs dieser Kurvenbögen $\dot{x}(t)$ und $\dot{y}(t)$ stetig sind.

Wir erkennen dies am einfachsten, indem wir auf die obige Ausgangsformel für die Länge des einem Kurvenbogen einbeschriebenen Polygones zurückgehen. Es seien längs dieses Bogens $x(t)$ und $y(t)$ mit stetigen Differentialquotienten $\dot{x}(t)$ und $\dot{y}(t)$ versehen. Es mögen den Eckpunkten des eingeschriebenen Polygones die Parameterwerte $t_1, t_2, \ldots, t_n$ mit den Differenzen $\Delta t_i$ entsprechen, und bei dem Grenzübergang $n \to \infty$ möge die jeweils größte dieser Differenzen gegen Null streben. Schreibt man nun die Länge des Polygones in der Form

$$\sum \sqrt{\Delta x_\nu^2 + \Delta y_\nu^2} = \sum \sqrt{\left(\frac{\Delta x_\nu}{\Delta t_\nu}\right)^2 + \left(\frac{\Delta y_\nu}{\Delta t_\nu}\right)^2}\, \Delta t_\nu,$$

so erkennt man sofort, daß diese Summe gegen das Integral $\int \sqrt{\dot{x}^2 + \dot{y}^2}\, dt$ strebt. Man braucht sich hierzu nur der allgemeinen Integralbildung (vgl. S. 107) zu erinnern. Ist eine Kurve aus mehreren derartigen Bögen zusammengesetzt, welche in Ecken oder Spitzen aneinander grenzen dürfen, so erhält man als Ausdruck für die Bogenlänge die Summe der entsprechenden Integrale. Als Resultat fassen wir zusammen: *Wenn längs eines Kurvenbogens $\alpha \leqq t \leqq \beta$ die Funktionen $x(t)$, $y(t)$ stetig sind und auch die Ableitungen $\dot{x}(t)$, $\dot{y}(t)$ bis auf endlich viele Sprungstellen stetig bleiben, dann besitzt der Bogen eine durch*

$$\int_\alpha^\beta \sqrt{\dot{x}^2 + \dot{y}^2}\, dt$$

*gegebene Länge,* wobei dieses Integral im Sinne des vierten Kapitels nötigenfalls als uneigentliches Integral zu verstehen ist[1]). Auf Grund dieser Formel, in welcher $\alpha < \beta$ sein muß, hat es Sinn, einem im Sinne fallender Parameterwerte $t$ durchlaufenen Bogen eine negative Länge zuzuschreiben, die sich alsdann ebenfalls durch das obige Integral ausdrückt. Das Vorzeichen der Bogenlänge hängt also von der Wahl des Kurvenparameters ab.

Ich gebe noch den Ausdruck für die *Bogenlänge* an, falls *Polarkoordinaten* zugrunde liegen. Wir brauchen dann nur in dem zuletzt gefundenen Ausdruck für $\dot{x}$ und $\dot{y}$ die Werte aus der Formel (*) von S. 210 einzutragen und erhalten

$$\dot{x}^2 + \dot{y}^2 = \dot{r}^2 + r^2 \dot{\vartheta}^2$$

und somit

$$s(\alpha, \beta) = \int_\alpha^\beta \sqrt{\dot{r}^2 + r^2 \dot{\vartheta}^2}\, dt.$$

Gehen wir hier von der Parameterdarstellung zu der Darstellung $r = f(\vartheta)$ über, indem wir $t = \vartheta$ selbst als Parameter einführen, so ergibt sich sofort für die Bogenlänge wegen $\dot{\vartheta} = 1$

$$s(\vartheta_0, \vartheta_1) = \int_{\vartheta_0}^{\vartheta_1} \sqrt{\dot{r}^2 + r^2}\, d\vartheta.$$

Ein einfaches Beispiel zur expliziten Berechnung der Bogenlänge gibt uns die Parabel $y = \frac{1}{2} x^2$; für ihre Bogenlänge erhalten wir sofort das Integral $\int_a^b \sqrt{1 + x^2}\, dx$, welches durch die Substitution $x = \operatorname{Sin} u$ übergeht in

$$\int_{\operatorname{Ar Sin} a}^{\operatorname{Ar Sin} b} \operatorname{Cof}^2 u\, du = \frac{1}{2} \int_{\operatorname{Ar Sin} a}^{\operatorname{Ar Sin} b} (1 + \operatorname{Cof} 2u)\, du = \frac{1}{2} (u + \operatorname{Sin} u \operatorname{Cof} u) \Big|_{\operatorname{Ar Sin} a}^{\operatorname{Ar Sin} b},$$

so daß für die Länge des Parabelbogens zwischen den Abszissen $a$ und $b$ der Ausdruck

$$s(a, b) = \frac{1}{2} \left\{ \operatorname{Ar Sin} b + b \sqrt{1 + b^2} - \operatorname{Ar Sin} a - a \sqrt{1 + a^2} \right\}$$

entsteht.

---

[1]) Daß der Ausdruck für die Bogenlänge von der Wahl des Parameters unabhängig ist, bestätigen wir sofort, wenn wir durch $\tau = \tau(t)$ einen neuen Parameter $\tau$ einführen. Es wird dann, falls $\frac{d\tau}{dt} \neq 0$,

$$\int \sqrt{\dot{x}^2 + \dot{y}^2}\, dt = \int \sqrt{\left(\frac{dx}{d\tau}\right)^2 \left(\frac{d\tau}{dt}\right)^2 + \left(\frac{dy}{d\tau}\right)^2 \left(\frac{d\tau}{dt}\right)^2}\, dt = \int \sqrt{\left(\frac{dx}{d\tau}\right)^2 + \left(\frac{dy}{d\tau}\right)^2}\, d\tau.$$

Für die Kettenlinie $y = \mathfrak{Cof}\, x$ ergibt sich $s(a, b) = \int\limits_a^b \sqrt{1 + \mathfrak{Sin}^2 x}\, dx$ $= \int\limits_a^b \mathfrak{Cof}\, x\, dx$ oder $s(a, b) = \mathfrak{Sin}\, b - \mathfrak{Sin}\, a$.

Schließlich sei bemerkt, daß es in vielen Fällen bequem ist, die von einem festen Anfangspunkt $P_0$ der Kurve aus gerechnete *Bogenlänge s als Parameter* zu wählen, d. h. $x = x(s)$ und $y = y(s)$ zu betrachten. Den Kurvenpunkten auf den beiden Seiten von $P_0$ werden dabei Werte von $s$ mit verschiedenen Vorzeichen entsprechen. Es ist dann

$$\dot{x}^2 + \dot{y}^2 = \left(\frac{ds}{dt}\right)^2 = 1,$$

woraus durch Differentiation

$$\dot{x}\ddot{x} + \dot{y}\ddot{y} = 0$$

folgt; zwei Relationen, die vielfach Anwendung finden.

## 4. Die Krümmung einer Kurve.

Während der Flächeninhalt und die Bogenlänge von dem Gesamtverlauf der Kurve abhängen, möchte ich hier die Besprechung eines Begriffes einschalten, der sich auf das Verhalten der Kurve nur in der Umgebung eines Punktes bezieht, nämlich der *Krümmung*.

Denken wir uns die Kurve gleichförmig in positivem Sinne durchlaufen, derart, daß in gleichen Zeiten gleiche Bogenlängen zurückgelegt werden, so wird sich die Richtung der Kurve mit einer bestimmten Geschwindigkeit ändern. Diese Geschwindigkeit betrachten wir als Maß für die Krümmung der Kurve an der betreffenden Stelle. Ist also $\alpha$ der Winkel zwischen der positiven Tangente und der positiven $x$-Achse und fassen wir $\alpha$ als Funktion der Bogenlänge $s$ auf, so werden wir die Krümmung $k$ an der Stelle, zu der die Bogenlänge $s$ gehört, durch die Gleichung $k = \frac{d\alpha}{ds}$ definieren. Nun ist $\alpha = \operatorname{arc\,tg} y'$, und daher gilt nach der Kettenregel $\frac{d\alpha}{ds} = \frac{d\alpha}{dx} : \frac{ds}{dx} = \frac{y''}{1 + y'^2} \cdot \frac{1}{\sqrt{1 + y'^2}}$ (wobei positives Zeichen der Quadratwurzel bedeutet, daß wachsendem $x$ wachsendes $s$ entspricht); mithin wird die Krümmung durch den Ausdruck

$$k = \frac{y''}{(1 + y'^2)^{\frac{3}{2}}}$$

geliefert.

Mit Hilfe der in § 1, Nr. 2 gegebenen Umrechnungsformeln für $y'$ und $y''$ auf Parameterdarstellung ergibt sich nunmehr in Parameterdarstellung für die Krümmung der einfache Ausdruck

$$k = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}},$$

der natürlich auch sofort aus der Gleichung

$$\alpha = \operatorname{arc\,tg} \frac{\dot{y}}{\dot{x}} = \operatorname{arc\,ctg} \frac{\dot{x}}{\dot{y}}$$

erhalten werden kann. Im Gegensatz zu dem obigen Ausdruck, der an die Kurvengleichung $y = f(x)$ geknüpft ist und daher eine spezielle Voraussetzung über die Lage des betrachteten Kurvenbogens zur $x$-Achse macht, gilt die Parameterdarstellung der Krümmung wieder für beliebige Kurvenbögen, längs deren $\dot{x}$, $\dot{y}$, $\ddot{x}$, $\ddot{y}$ stetige Funktionen von $t$ sind, insbesondere auch für solche Stellen, wo $\dot{x} = 0$ ist, wo also $\frac{dy}{dx}$ unendlich wird.

Führen wir die Bogenlänge $s$ als Parameter ein, so wird wegen $\dot{x}^2 + \dot{y}^2 = 1$ und $\dot{x}\ddot{x} + \dot{y}\ddot{y} = 0$ einfach

$$k = \dot{x}\ddot{y} - \dot{y}\ddot{x} = \ddot{y}\left(\dot{x} + \dot{y}\frac{\dot{y}}{\dot{x}}\right) = \frac{\ddot{y}}{\dot{x}} = -\frac{\ddot{x}}{\dot{y}}.$$

Wir erhalten also besonders einfache Ausdrücke für die Krümmung.

Das *Vorzeichen der Krümmung* ändert sich, wenn wir den Durchlaufungssinn umkehren, d. h. analytisch: $s$ durch $\sigma = -s$ ersetzen. Man muß also stets Kurven mit bestimmtem Durchlaufungssinn betrachten, um die Krümmung nicht nur nach ihren Absolutwerten, sondern auch nach ihrem Vorzeichen eindeutig festzulegen.

Als Beispiel betrachten wir die Krümmung eines positiv umlaufenen Kreises mit dem Radius $a$. Gehen wir von der Parameterdarstellung $x = a \cos t$, $y = a \sin t$ aus, so ergibt sich sofort

$$k = \frac{1}{a}.$$

*Die Krümmung eines positiv umlaufenen Kreises ist also gleich seinem reziproken Radius.* Dies Ergebnis bestätigt uns, daß unsere Krümmungsdefinition wirklich zweckmäßig war; denn bei einem Kreise werden wir ganz naturgemäß als Maß für die Krümmung den reziproken Radius ansehen.

Setzt man $\varrho = \frac{1}{k}$, so nennt man allgemein $|\varrho| = \frac{1}{|k|}$ den *Krümmungsradius* der Kurve an der betreffenden Stelle. Denjenigen die Kurve in einem gegebenen Punkte berührenden und dort in derselben Richtung wie die Kurve durchlaufenen Kreis, welcher dieselbe Krümmung $k$ wie die Kurve besitzt und dessen Mittelpunkt auf der Seite der positiven oder negativen Normalenrichtung liegt, je nachdem $k$ positiv oder negativ ist, nennt man den zum Kurvenpunkte gehörigen *Krümmungskreis.* Denken wir uns seine Gleichung in der Form $y = g(x)$ geschrieben, so muß für den betreffenden Kurvenpunkt nicht nur $g(x) = f(x)$, $g'(x) = f'(x)$ sein, was die Berührung zwischen Kreis und Kurve ausdrückt, sondern es muß auch wegen

$$\frac{f''(x)}{(\sqrt{1 + f'(x)^2})^3} = k = \frac{g''(x)}{(\sqrt{1 + g'(x)^2})^3}$$

die Gleichung

$$f''(x) = g''(x)$$

bestehen.

Den Mittelpunkt des Krümmungskreises nennt man den zum Kurvenpunkte gehörigen *Krümmungsmittelpunkt.* Seine Koordinaten $\xi$ und $\eta$ drücken sich in Parameterdarstellung durch

$$\xi = x - \frac{\varrho \dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}}, \qquad \eta = y + \frac{\varrho \dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}$$

aus; wir brauchen hierzu nur die oben gegebenen Formeln für die Richtungskosinus der Normalen — auf welcher der Krümmungsmittelpunkt ja im Abstande $\frac{1}{|k|} = |\varrho|$ von der Tangente liegt — anzuwenden. Die obigen Formeln stellen uns den Krümmungsmittelpunkt durch den Parameter $t$ ausgedrückt dar. Durchwandert $t$ sein Intervall, so beschreibt der Krümmungsmittelpunkt eine Kurve, die sogenannte *Evolute* der gegebenen Kurve, und unsere Formeln geben uns, da wir zugleich mit $x$ und $y$ auch $\dot{x}$, $\dot{y}$ und $\varrho$ als bekannte Funktionen von $t$ anzusehen haben, diese Evolute in Parameterdarstellung.

Hinsichtlich einzelner Beispiele verweise ich auf den nächsten Paragraphen und den Anhang.

### 5. Schwerpunkt und statisches Moment einer Kurve.

Wir kommen nunmehr zu einigen Anwendungen, die uns in das Gebiet der Mechanik hineinführen. Wir betrachten ein System von $n$ in einer Ebene liegenden Massenpunkten. Es seien $m_1, m_2, \ldots, m_n$ die Massen dieser einzelnen Punkte; es seien weiter $y_1, y_2, \ldots, y_n$ die $y$-Koordinaten der einzelnen Punkte. Man bezeichnet dann

$$T = \sum_{\nu=1}^{n} m_\nu y_\nu = m_1 y_1 + m_2 y_2 + \cdots + m_n y_n$$

als das *statische Moment unseres Punktsystems in bezug auf die x-Achse.* Der Ausdruck $\eta = \frac{T}{M}$ gibt uns die *Höhe des Schwerpunktes* unseres Punktsystems über der $x$-Achse an, wobei

$$M = m_1 + m_2 + \cdots + m_n$$

die Gesamtmasse des Punktsystems darstellt. Entsprechend definiert man das statische Moment in bezug auf die $y$-Achse und die $x$-Koordinate des Schwerpunktes.

Wir wollen uns nunmehr überlegen, was man unter dem statischen Moment einer gleichmäßig mit Masse belegten Kurve $y = f(x)$ zu verstehen hat und wie man dementsprechend die Koordinaten $\xi$ und $\eta$ des Schwerpunktes einer solchen Kurve bestimmt. Dabei machen wir der Kürze halber die Voraussetzung, daß die Massendichte auf der Kurve konstant, etwa gleich $\mu$ sei.

Wir führen unsere Aufgabe zurück auf die Betrachtung eines Systems von endlich vielen Punkten und auf einen Grenzübergang. Zu diesem Zwecke denken wir uns auf der Kurve die Bogenlänge $s$ als Parameter eingeführt und den betrachteten Kurvenbogen durch $n-1$ Teilpunkte in Teile der Längen $\Delta s_1, \Delta s_2, \ldots, \Delta s_n$ zerlegt. Die Massen $\mu \Delta s_i$ dieser Teilstücke stellen wir uns jeweils in einem beliebigen Punkt des Intervalles mit der Ordinate $y_i$ konzentriert vor.

Gemäß der Definition des Momentes erhalten wir also für das Moment dieses Punktsystemes in bezug auf die $x$-Achse den Wert

$$T = \mu \sum y_i \Delta s_i .$$

Strebt nun das größte $\Delta s_i$ gegen 0, so liefert uns der Grenzübergang für das Moment der Kurve in bezug auf die $x$-Achse den Ausdruck

$$T = \mu \int_{s_0}^{s_1} y\, ds = \mu \int_{x_0}^{x_1} y \sqrt{1 + y'^2}\, dx .$$

Da die Gesamtmasse der Kurve gleich ihrer Länge mal $\mu$:

$$\mu \int_{s_0}^{s_1} ds = \mu (s_1 - s_0)$$

zu setzen ist, so ergeben sich für die Koordinaten des Schwerpunktes die Ausdrücke

$$\eta = \frac{\int_{s_0}^{s_1} y\, ds}{s_1 - s_0}, \qquad \xi = \frac{\int_{s_0}^{s_1} x\, ds}{s_1 - s_0} .$$

### 6. Flächeninhalt und Volumen einer Rotationsfläche.

Läßt man die Kurve $y = f(x)$, wobei wir $f(x) \geqq 0$ voraussetzen, um die $x$-Achse rotieren, so beschreibt sie eine sog. Rotationsfläche. Der Flächeninhalt dieser Oberfläche, soweit ihre $x$-Koordinaten zwischen den Grenzen $x_0$ und $x_1 > x_0$ liegen, läßt sich unmittelbar durch eine der obigen ganz analoge Betrachtung ableiten. Ersetzt man nämlich zunächst die Kurve durch ein einbeschriebenes Polygon, so wird an Stelle der krummen Fläche ein Gebilde entstehen, das aus einer Anzahl schmaler Kreiskegelstümpfe zusammengesetzt ist. Der Flächeninhalt des Mantels eines solchen Kegelstumpfes ist aber bekanntlich gleich der Länge der Seitenlinie, multipliziert mit dem Umfang des mittleren Kreisquerschnittes. Addiert man alle diese Ausdrücke und führt dann den Grenzübergang von dem Polygon zur Kurve aus, so erhält man als Flächeninhalt den Ausdruck

$$O = 2\pi \int_{x_0}^{x_1} y \sqrt{1 + y'^2}\, dx = 2\pi \int_{s_0}^{s_1} y\, ds .$$

Man kann dieses Resultat in Worten dahin aussprechen, daß der Flächeninhalt der Rotationsfläche gleich der Länge der Kurve multipliziert mit dem Wege des Schwerpunkts der Kurve ist (*Guldinsche Regel*).

Ganz ebenso ergibt sich für das von unserer Rotationsfläche eingeschlossene Volumen, welches durch die Ebenen $x = x_0$ und $x = x_1 > x_0$

abgeschlossen wird, der Ausdruck

$$V = \pi \int_{x_0}^{x_1} y^2\, dx.$$

Die Herleitung beruht auf der aus der Anschauung entnommenen Festsetzung, daß man das fragliche Volumen als Grenzwert der Summe der Volumina der eben betrachteten Kegelstümpfe auffassen muß. Hiernach darf ich dem Leser den Beweis der Formel selbst überlassen.

## 7. Trägheitsmoment.

Bei allen Drehbewegungen spielen in der Mechanik gewisse Größen eine wichtige Rolle, welche man Trägheitsmomente nennt. Auch diese Ausdrücke sollen hier nur kurz erwähnt werden.

Wir denken uns, daß ein Massenpunkt $m$ mit dem Abstand $y$ von der als Rotationsachse betrachteten $x$-Achse gleichförmig um diese Achse rotiert, und zwar mit der *Winkelgeschwindigkeit* $\omega$ (dies bedeutet, daß in der Zeiteinheit eine Drehung um den Winkel $\omega$ erfolgt). Die *lebendige Kraft* oder *kinetische Energie* des Punktes, ausgedrückt als das Produkt aus halber Masse und Quadrat der Geschwindigkeit, wird offenbar

$$\frac{m}{2}(y\,\omega)^2.$$

Wir nennen den Faktor von $\frac{1}{2}\omega^2$, d. h. die Größe $m y^2$, das *Trägheitsmoment des Punktes um die $x$-Achse.*

Ebenso bezeichnen wir, wenn $n$ Massenpunkte mit den Massen $m_1, m_2, \ldots, m_n$ und den Ordinaten $y_1, y_2, \ldots, y_n$ gegeben sind, den Ausdruck

$$T = \sum_i m_i y_i^2$$

als das Trägheitsmoment des Massensystems um die $x$-Achse. Das Trägheitsmoment ist eine Größe, welche dem Massensystem als solchem ohne Bezug auf seinen Bewegungszustand zukommt. Seine Bedeutung besteht darin, daß wir durch Multiplikation mit dem halben Quadrat der Winkelgeschwindigkeit aus dem Trägheitsmoment die kinetische Energie des Massensystems erhalten, sobald es in Rotation um die betreffende Achse versetzt wird, ohne daß dabei die gegenseitigen Abstände geändert werden. Das Trägheitsmoment um eine Achse spielt bei Drehbewegungen um sie eine ähnliche Rolle wie die gesamte Masse des Massensystems bei geradliniger Bewegung.

Um nun das Trägheitsmoment für eine beliebige mit einer Masse von der Dichte 1 belegte Kurve $y = f(x)$ zwischen den Abszissen $x_0$ und $x_1\,(> x_0)$ zu definieren, müssen wir ganz entsprechend verfahren wie bei der Definition des statischen Momentes; es ergibt sich in genau derselben Art für das Trägheitsmoment um die $x$-Achse die Definition

$$T = T_x = \int_{s_0}^{s_1} y^2\, ds = \int_{x_0}^{x_1} y^2 \sqrt{1 + y'^2}\, dx.$$

Entsprechend erhalten wir als Trägheitsmoment um die $y$-Achse den Ausdruck

$$T_y = \int_{s_0}^{s_1} x^2\,ds = \int_{x_0}^{x_1} x^2 \sqrt{1+y'^2}\,dx.$$

## § 3. Beispiele.

Die Theorie der ebenen Kurven mit ihren mannigfaltigen speziellen Gestalten und Eigenschaften bietet in Fülle Beispiele für alle unsere Begriffsbildungen. Ich kann mich aber hier nicht in allzu viele Einzelheiten verlieren und muß mich auf einige wenige charakteristische Anwendungen beschränken.

### 1. Die gemeine Zykloide.

Aus den Gleichungen (vgl. § 1, Nr. 1) $x = a(t - \sin t)$, $y = a(1 - \cos t)$ erhalten wir ohne weiteres $\dot{x} = a(1 - \cos t)$, $\dot{y} = a \sin t$ und daraus die Bogenlänge

$$s = \int_0^\alpha \sqrt{\dot{x}^2 + \dot{y}^2}\,dt = \int_0^\alpha \sqrt{2a^2(1-\cos t)}\,dt.$$

Nun ist aber $1 - \cos t = 2\sin^2\frac{t}{2}$; damit ergibt sich für $0 \leqq \alpha \leqq 2\pi$

$$s = 2a\int_0^\alpha \sin\frac{t}{2}\,dt = -4a\cos\frac{t}{2}\Big|_0^\alpha = 4a\left(1-\cos\frac{\alpha}{2}\right) = 8a\sin^2\frac{\alpha}{4}.$$

Betrachten wir insbesondere die Bogenlänge zwischen zwei aufeinanderfolgenden Spitzen, so ist $\alpha = 2\pi$ einzusetzen, da diese Bogenlänge der Zykloide einem vollen Umlaufe des rollenden Kreises entspricht. Wir erhalten somit $8a$; d. h. der Umfang der Zykloide zwischen zwei aufeinanderfolgenden Spitzen ist gleich dem vierfachen Durchmesser des rollenden Kreises.

Auf dieselbe Art berechnen wir auch den Flächeninhalt, der von einem Zykloidenbogen und der $x$-Achse eingeschlossen wird:

$$J = \int_0^{2\pi} y\dot{x}\,dt = a^2\int_0^{2\pi}(1-\cos t)^2\,dt = a^2\int_0^{2\pi}(1 - 2\cos t + \cos^2 t)\,dt$$

$$= a^2\left(t - 2\sin t + \frac{t}{2} + \frac{\sin 2t}{4}\right)\Big|_0^{2\pi} = 3a^2\pi.$$

Der berechnete Flächeninhalt ist somit gleich dem dreifachen Inhalt des rollenden Kreises.

Für die reziproke Krümmung $\varrho = \frac{1}{k}$ findet man

$$\varrho = \frac{(\dot{x}^2+\dot{y}^2)^{\frac{3}{2}}}{\dot{x}\ddot{y} - \dot{y}\ddot{x}} = -2a\sqrt{2(1-\cos t)} = -4a\,\sin\frac{t}{2};$$

an den Stellen $t = 0$, $t = \pm 2\pi$, ... wird dieser Ausdruck gleich Null. Es sind dies gerade die Spitzen, in denen die Zykloide unter rechtem Winkel auf die $x$-Achse trifft.

Die Oberfläche des Rotationskörpers, welcher entsteht, wenn sich der Zykloidenbogen um die $x$-Achse dreht, ergibt sich gemäß unserer Formel (S. 225) als

$$O = 2\pi \int_0^{8a} y\,ds = 2\pi \int_0^{2\pi} a(1-\cos t)\cdot 2a \sin\frac{t}{2}\,dt$$

$$= 8a^2\pi \int_0^{2\pi} \sin^3\frac{t}{2}\,dt = 16a^2\pi \int_0^{\pi} \sin^3 u\,du = 16a^2\pi \int_0^{\pi} (1-\cos^2 u)\sin u\,du.$$

Das letzte Integral berechnet sich durch die Substitution $\cos u = v$, und man findet

$$O = 16a^2\pi \left(-\cos u + \frac{1}{3}\cos^3 u\right)\Bigg|_0^{\pi} = \frac{64a^2\pi}{3}.$$

Zur Übung mag der Leser selbst noch die Höhe $\eta$ des Schwerpunktes der Zykloide über der $x$-Achse und das Trägheitsmoment $T_x$ berechnen. Das Resultat ist:

$$\eta = \frac{4}{3}a = \frac{O}{2\pi s} \quad \text{und} \quad T_x = \frac{256}{15}a^3.$$

## 2. Kettenlinie.

Die Bogenlänge der Kettenlinie haben wir schon als Beispiel im vorigen Paragraphen berechnet und haben den Wert

$$s = \int_a^b \operatorname{Cof} x\,dx = \operatorname{Sin} b - \operatorname{Sin} a$$

gefunden.

Für den Flächeninhalt der Rotationsfläche, die durch Drehung der Kettenlinie um die $x$-Achse entsteht, des sog. Katenoids, finden wir

$$O = 2\pi \int_a^b \operatorname{Cof}^2 x\,dx = 2\pi \int_a^b \frac{1+\operatorname{Cof} 2x}{2}\,dx = \pi\left(b - a + \frac{1}{2}\operatorname{Sin} 2b - \frac{1}{2}\operatorname{Sin} 2a\right).$$

Daraus folgt wieder die Höhe des Schwerpunktes des Bogens von $a$ bis $b$:

$$\eta = \frac{O}{2\pi s} = \frac{b - a + \frac{1}{2}\operatorname{Sin} 2b - \frac{1}{2}\operatorname{Sin} 2a}{2(\operatorname{Sin} b - \operatorname{Sin} a)}.$$

Endlich ergibt sich für die Krümmung

$$k = \frac{y''}{\left(1+y'^2\right)^{\frac{3}{2}}} = \frac{\operatorname{Cof} x}{\operatorname{Cof}^3 x} = \frac{1}{\operatorname{Cof}^2 x}.$$

### 3. Ellipse und Lemniskate.

Die Bogenlängen dieser beiden Kurven lassen sich nicht mehr auf elementare Funktionen zurückführen, sondern gehören schon unter die im vierten Kapitel, § 7, erwähnten „elliptischen Integrale".

Wir erhalten nämlich für die Ellipse $y = \frac{b}{a}\sqrt{a^2 - x^2}$

$$s = \frac{1}{a}\int\sqrt{\frac{a^4 - (a^2 - b^2)\,x^2}{a^2 - x^2}}\,dx = a\int\frac{1 - \varkappa^2\,\xi^2}{\sqrt{(1 - \xi^2)\,(1 - \varkappa^2\,\xi^2)}}\,d\xi,$$

wenn wir $\frac{x}{a} = \xi$, $1 - \frac{b^2}{a^2} = \varkappa^2$ setzen, ein Integral, welches sich vermöge der Substitution $\frac{x}{a} = \sin\varphi$ noch auf die Form

$$s = \int\sqrt{a^2 - (a^2 - b^2)\sin^2\varphi}\,d\varphi = a\int\sqrt{1 - \varkappa^2\sin^2\varphi}\,d\varphi$$

bringen läßt. Für den halben Ellipsenbogen hat man hierin $x$ von $-a$ bis $+a$ laufen zu lassen, was dem Intervall $-1 \leqq \xi \leqq +1$ bzw. dem Intervall $-\frac{\pi}{2} \leqq \varphi \leqq +\frac{\pi}{2}$ entspricht.

Für die Lemniskate, deren Gleichung in Polarkoordinaten $r^2 = 2a^2\cos 2t$ war, erhalten wir analog

$$s = \int\sqrt{r^2 + \dot r^2}\,dt = \int\sqrt{2a^2\cos 2t + 2a^2\frac{\sin^2 2t}{\cos 2t}}\,dt$$

$$= a\sqrt{2}\int\frac{dt}{\sqrt{\cos 2t}} = a\sqrt{2}\int\frac{dt}{\sqrt{1 - 2\sin^2 t}}.$$

Führen wir in dem letzten Integral $u = \operatorname{tg} t$ als unabhängige Veränderliche ein, so wird

$$\sin^2 t = \frac{u^2}{1 + u^2}, \quad dt = \frac{du}{1 + u^2},$$

und es ergibt sich

$$a\sqrt{2}\int\frac{du}{\sqrt{1 - u^4}}.$$

Für eine volle Lemniskatenschleife läuft $u$ von $-1$ bis $1$, und die Bogenlänge wird daher gleich

$$a\sqrt{2}\int\limits_{-1}^{+1}\frac{du}{\sqrt{1 - u^4}},$$

einem speziellen elliptischen Integral, welches in den Untersuchungen von Gauß eine große Rolle gespielt hat.

## § 4. Die einfachsten Probleme der Mechanik.

Neben der Geometrie ist es vor allen Dingen die Mechanik gewesen, welcher die Integral- und Differentialrechnung ihre erste Entwicklung verdankt. Die Mechanik beruht auf einigen Grundprinzipien, die von

Newton aufgestellt wurden, zu deren Formulierung der Begriff des Differentialquotienten und zu deren Verwertung die Theorie der Integration notwendig ist. Ohne die Grundprinzipien näher zu analysieren, will ich an einfachen Beispielen die Anwendung der Integral- und Differentialrechnung in der Mechanik erläutern.

## 1. Grundvoraussetzungen aus der Mechanik.

Wir beschränken uns auf die Betrachtung eines Massenpunktes, d. h. eines Punktes, in dem wir uns eine Masse von der Größe $m$ konzentriert vorstellen. Wir wollen weiter annehmen, daß die Bewegung nur auf einer fest vorgegebenen Kurve vor sich gehen kann, auf welcher wir die Lage der Massenpunkte durch die von einem festen Anfangspunkt gezählte Bogenlänge $s$ charakterisieren; speziell kann es sich auch um eine gerade Linie handeln, auf welcher wir dann statt $s$ als Ortskoordinate die Abszisse $x$ einführen. Die Bewegung ist durch die Angabe der Koordinate $s = \varphi(t)$ als Funktion der Zeit $t$ charakterisiert. Unter der *Geschwindigkeit* der Bewegung haben wir den Differentialquotienten $\varphi'(t)$ oder, wie wir auch schreiben wollen,

$$\frac{ds}{dt} = \varphi'(t) = \dot{s}$$

zu verstehen. Als *Beschleunigung* bezeichnen wir den zweiten Differentialquotienten

$$\frac{d^2 s}{dt^2} = \varphi''(t) = \ddot{s}.$$

Die Mechanik geht von der Vorstellung aus, daß die Bewegung der Massenpunkte durch *Kräfte* zu erklären bzw. zu beschreiben ist, denen eine bestimmte Richtung und Größe zukommt. Das Newtonsche Grundgesetz der Mechanik sprechen wir dann für den Fall der Bewegung auf unserer gegebenen Kurve so aus: *Die Masse $m$ multipliziert mit der Beschleunigung $\ddot{s}$ ist gleich der auf den Massenpunkt in der Richtung der Kurve wirkenden Kraft*, die wir mit $S$ bezeichnen, in Formeln

$$m\ddot{s} = S.$$

Die Richtung der Kraft ist hiernach immer dieselbe wie die der Beschleunigung: sie liegt in Richtung wachsender $s$-Werte, wenn die Geschwindigkeit in dieser Richtung wächst, andernfalls ist sie der $s$-Richtung entgegengesetzt.

In diesem Newtonschen Gesetz liegt zunächst nur eine Definition des Kraftbegriffes. Die linke Seite unserer Gleichung ist eine durch Beobachtung der Bewegung feststellbare Größe, durch welche wir die Kraft messen. Aber diese Gleichung hat eine viel weitergehende Bedeutung. Es zeigt sich nämlich in Wirklichkeit, daß wir in sehr vielen Fällen ohne Kenntnis der betreffenden Bewegung die wirkenden Kräfte aus anderen physikalischen Voraussetzungen heraus von vornherein bestimmen können; dann wird das obige Newtonsche Grundgesetz nicht mehr Kraftdefinition sein, sondern eine Beziehung darstellen, aus der wir wichtige Schlüsse auf die Bewegung ziehen können.

Das wichtigste Beispiel einer von vornherein bekannten Kraft gibt uns die Schwere. Aus direkten Messungen wissen wir, daß die Schwerkraft, welche auf eine Masse $m$ in der Vertikalrichtung nach unten wirkt, gleich $mg$ ist, wobei die Konstante $g$, die sog. Erdbeschleunigung, ungefähr gleich 981 ist, wenn wir die Zeit in Sekunden, die Länge in Zentimetern messen. Bewegt sich eine Masse auf einer gegebenen Kurve, so entnehmen wir der Erfahrung, daß die Schwerkraft in Richtung dieser Kurve gleich $mg \cos \alpha$ ist, wenn $\alpha$ den Winkel der Vertikalen mit der Tangente in dem betreffenden Punkte der Kurve bedeutet (vgl. Fig. 91).

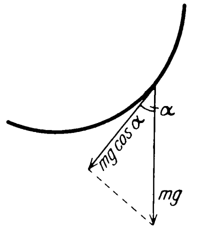


Fig. 91. Schwerkraft in Richtung einer Kurve.

Das Grundproblem der Mechanik ist für den Fall einer Bewegung auf unserer gegebenen Kurve folgendes: Wir kennen irgendwoher die auf den Massenpunkt wirkende Kraft, z. B. die Schwerkraft, und sollen die Lage des Massenpunktes, d. h. seine Koordinate $s$ oder $x$ als Funktion der Zeit bestimmen.

Beschränken wir uns auf den einfachsten Fall, daß wir diese Kraft $S = mf(s)$[1]) von vornherein als Funktion der Bogenlänge kennen — daß sie also von der Zeit unabhängig ist —, so wollen wir zeigen, daß und wie wir aus der Gleichung $\ddot{s} = \frac{1}{m} S = f(s)$ den Verlauf der Bewegung auf der Kurve entnehmen können.

Wir haben es hier mit einer *Differentialgleichung* zu tun, d. h. einer Gleichung, aus der man eine unbekannte Funktion — hier $s(t)$ — zu bestimmen hat und in welcher außer dieser unbekannten Funktion auch noch Differentialquotienten von ihr vorkommen (vgl. drittes Kapitel, § 7).

## 2. Freier Fall. Reibung.

Beim freien Fall eines Massenpunktes auf der vertikalen $x$-Achse liefert uns das Newtonsche Gesetz die Differentialgleichung

$$\ddot{x} = g .$$

Aus ihr folgt $\dot{x}(t) = gt + v_0$, wo $v_0$ eine Integrationskonstante ist. Ihre Bedeutung ergibt sich, wenn wir $t = 0$ setzen. Es wird $\dot{x}(0) = v_0$; d. h. $v_0$ ist die Geschwindigkeit des Massenpunktes zum Beginn der

[1]) Die Herausnahme des Faktors $m$ aus dem Ausdruck für die gegebene Kraft erfolgt lediglich der formalen Einfachheit wegen.

Zeitzählung, die *Anfangsgeschwindigkeit.* Durch nochmalige Integration erhalten wir

$$x(t) = \frac{1}{2} g t^2 + v_0 t + x_0,$$

wo auch $x_0$ eine Integrationskonstante ist, deren Bedeutung sich wiederum ergibt, wenn wir $t = 0$ setzen: $x_0$ ist die *Anfangskoordinate,* d. h. die Koordinate des Punktes zu Beginn der Bewegung.

Umgekehrt können wir bei der Bewegung die Anfangslage $x_0$ und die Anfangsgeschwindigkeit $v_0$ willkürlich wählen und erhalten dann stets in der Gleichung $x = \frac{1}{2} g t^2 + v_0 t + x_0$ die vollständige Darstellung des Bewegungsvorganges.

Wollen wir den Einfluß der auf den fallenden Massenpunkt wirkenden *Reibung* berücksichtigen, so müssen wir die Reibung als eine Kraft betrachten, welche entgegengesetzt der Schwerkraft wirkt und über die wir bestimmte physikalische Hypothesen zu machen haben. Wir wollen zwei verschiedenartige physikalische Annahmen verfolgen: a) Die Reibungskraft sei proportional der Geschwindigkeit, sie werde gegeben durch einen Ausdruck der Form $-r\dot{x}$, wo $r$ eine positive Konstante ist. b) Die Reibungskraft sei proportional dem Quadrat der Geschwindigkeit; sie hat dann die Form $-r\dot{x}^2$. — Als Bewegungsgleichungen gemäß dem Newtonschen Grundgesetz erhalten wir dann

$$\text{a)}\quad m\ddot{x} = mg - r\dot{x}, \qquad \text{b)}\quad m\ddot{x} = mg - r\dot{x}^2.$$

Betrachten wir zunächst einmal $\dot{x} = u(t)$ als die gesuchte Funktion, so ergibt sich $\ddot{x}(t) = \dot{u}(t)$, also

$$\text{a)}\quad m\dot{u} = mg - ru, \qquad \text{b)}\quad m\dot{u} = mg - ru^2.$$

Aus diesen Gleichungen bestimmen wir jetzt nicht $u$ als Funktion von $t$, sondern umgekehrt $t$ als Funktion von $u$, indem wir von der Schreibweise

$$\text{a)}\quad \frac{dt}{du} = \frac{1}{g - \frac{r}{m}u}, \qquad \text{b)}\quad \frac{dt}{du} = \frac{1}{g - \frac{r}{m}u^2}$$

unserer Differentialgleichungen ausgehen. Wir können dann mit Hilfe der Integrationsmethode des vorigen Kapitels ohne weiteres die Integration ausführen und erhalten

$$\text{a)}\quad t(u) = -\frac{m}{r}\log\left(1 - \frac{r}{mg}u\right) + t_0,$$

$$\text{b)}\quad t(u) = -\frac{1}{2}\varkappa \log\frac{\varkappa g - u}{\varkappa g + u} + t_0,$$

wobei wir $\sqrt{\frac{m}{rg}} = \varkappa$ gesetzt haben und wo $t_0$ eine Integrationskonstante ist. Durch Auflösung unserer Gleichungen nach $u$ ergibt sich jetzt

$$\text{a)}\quad u(t) = -\frac{mg}{r}\left(e^{-\frac{r}{m}(t-t_0)} - 1\right),$$

$$\text{b)}\quad u(t) = -g\varkappa\,\frac{e^{-\frac{2(t-t_0)}{\varkappa}} - 1}{e^{-\frac{2(t-t_0)}{\varkappa}} + 1}.$$

Wir erkennen aus diesen Gleichungen schon eine wichtige Eigenschaft unserer Bewegungen: Die Geschwindigkeit wächst mit der Zeit nicht über alle Grenzen, sondern sie nähert sich bestimmten, übrigens von der Masse $m$ abhängigen Grenzwerten. Es wird nämlich

$$\text{a)}\quad \lim_{t\to\infty} u(t) = \frac{mg}{r}, \qquad \text{b)}\quad \lim_{t\to\infty} u(t) = \sqrt{\frac{mg}{r}}.$$

Durch nochmalige Integration unserer gewonnenen Ausdrücke erhalten wir, was sich ebenfalls mit Hilfe der Methoden des vorigen Kapitels leicht durchführen und durch Differentiation bestätigen läßt.

$$\text{a)}\quad x(t) = \frac{m^2}{r^2}\,g\,e^{-\frac{r}{m}(t-t_0)} + \frac{mg}{r}\,t + c,$$

$$\text{b)}\quad x(t) = \frac{m}{r}\log \mathfrak{Cof}\sqrt{\frac{rg}{m}}\,(t-t_0) + c,$$

wo $c$ eine neue Integrationskonstante ist. Die beiden Integrationskonstanten $t_0$ und $c$ bestimmen sich im übrigen sofort, wenn wir Anfangslage $x(0) = x_0$ und Anfangsgeschwindigkeit $\dot{x}(0) = u(0) = v_0$ unseres fallenden Massenpunktes kennen.

### 3. Die einfachste elastische Schwingung.

Als zweites Beispiel betrachten wir die Bewegung eines in Richtung der $x$-Achse beweglichen und durch eine elastische Kraft an den Nullpunkt gebundenen Massenpunktes. Von der elastischen Kraft nehmen wir an, daß sie stets auf den Nullpunkt zu gerichtet ist und daß ihre Größe proportional dem Abstand vom Nullpunkt ist. Mit anderen Worten: Wir setzen die Kraft gleich $-kx$, wobei die Größe des Koeffizienten $k$ die Festigkeit der elastischen Bindung mißt. Die Kraft ist, da $k$ als positiv angenommen wird, negativ, wenn $x$ positiv ist, und positiv, wenn $x$ negativ ist. Die Newtonsche Gleichung besagt also jetzt

$$m\ddot{x} = -kx.$$

Wir können nicht erwarten, daß diese Differentialgleichung den Vorgang vollständig bestimmt; vielmehr ist plausibel, daß wir in einem bestimmten Moment, etwa zur Zeit $t = 0$, die Anfangskoordinate $x(0) = x_0$ und die Anfangsgeschwindigkeit $\dot{x}(0) = v_0$ willkürlich vorschreiben dürfen, d. h., physikalisch ausgedrückt, den Massenpunkt zum Beginn von einer beliebigen Anfangslage mit einer beliebigen Geschwindigkeit ablaufen lassen dürfen und daß dann erst der Vorgang gemäß der Bewegungsgleichung festgelegt ist. Mathematisch findet dies seinen Ausdruck darin, daß die allgemeinste Lösung unserer Differentialgleichung zwei zunächst unbestimmte Integrationskonstanten enthält, die erst durch die beiden Anfangsbedingungen zu bestimmen sind, eine Tatsache, die wir sogleich beweisen werden.

Wir können sehr leicht direkt eine solche Lösung angeben. Setzen wir $\omega = \sqrt{\frac{k}{m}}$, so wird nämlich unsere Differentialgleichung, wie man durch Differenzieren sofort bestätigt, durch alle Funktionen

$$x(t) = c_1 \cos \omega t + c_2 \sin \omega t$$

befriedigt, bei welchen $c_1$ und $c_2$ beliebig gewählte Konstanten bedeuten. Wir werden in Nr. 4 sehen, daß es andere Lösungen unserer Differentialgleichung nicht gibt und daß daher jede solche Bewegung unter dem Einfluß einer elastischen Kraft durch den obigen Ausdruck dargestellt wird, den wir übrigens auch leicht in die Form

$$x(t) = a \sin \omega (t - \delta) = -a \sin \omega \delta \cos \omega t + a \cos \omega \delta \sin \omega t$$

bringen können, wenn wir $-a \sin \omega \delta = c_1$, $a \cos \omega \delta = c_2$ setzen und somit statt der Konstanten $c_1$ und $c_2$ die neuen Konstanten $a = \sqrt{c_1^2 + c_2^2}$ und $\delta = -\frac{1}{\omega} \operatorname{arc\,tg} \frac{c_1}{c_2}$ einführen. Bewegungen dieser Art nennt man *reine Sinus-* oder *Kosinusschwingungen.* Sie stellen periodische Bewegungen dar; jeder Zustand (d.h. Lage $x(t)$ und Geschwindigkeit $\dot{x}(t)$) kehrt nämlich nach der Zeit $T = \frac{2\pi}{\omega}$, der „*Schwingungsdauer*“, wieder, weil die Funktionen $\sin \omega t$ und $\cos \omega t$ die Periode $T$ besitzen. Die Größe $a$ heißt die *Amplitude* oder der *maximale Ausschlag* der Schwingung; die Zahl $\omega$ heißt die *Frequenz* der Schwingung, sie mißt die Geschwindigkeit, mit welcher die periodische Bewegung sich wiederholt. Im übrigen komme ich im zehnten Kapitel auf die Theorie der Schwingungen noch zurück.

### 4. Die allgemeine Bewegung auf einer vorgegebenen Kurve.

Endlich erörtere ich noch das oben aufgestellte Problem in seiner allgemeinsten Form, nämlich die Frage der Bewegung auf einer Kurve bei beliebig vorgegebener Kraft $m f(s)$.

Es handelt sich hier einfach um die Bestimmung der Funktion $s(t)$ als Funktion von $t$ aus der Differentialgleichung

$$\ddot{s} = f(s),$$

wo $f(s)$ eine gegebene Funktion ist. Man kann diese Differentialgleichung zur Bestimmung von $s$ durch folgenden Kunstgriff vollständig auflösen.

Wir betrachten zunächst irgend eine primitive Funktion $F(s)$ von $f(s)$, so daß also $F'(s) = f(s)$ wird, und multiplizieren nunmehr die Differentialgleichung $\ddot{s} = f(s) = F'(s)$ beiderseits mit $\dot{s}$. Dann können wir die linke Seite in der Form schreiben $\frac{d}{dt}\left(\frac{1}{2}\dot{s}^2\right)$, wie man ohne weiteres durch Differentiation des Ausdruckes $\dot{s}^2$ erkennt; die rechte Seite $F'(s)\dot{s}$ aber ist gerade der Differentialquotient von $F(s)$ nach der Zeit $t$, wenn man in $F(s)$ die Größe $s$ als Funktion von $t$ auffaßt. Es ergibt sich also sofort

$$\frac{d}{dt}\left(\frac{1}{2}\dot{s}^2\right) = \frac{d}{dt}F(s)$$

oder durch Integration

$$\frac{1}{2}\dot{s}^2 = F(s) + c,$$

wo $c$ eine noch zu bestimmende Konstante bedeutet.

Schreiben wir diese Gleichung in der Form $\frac{ds}{dt} = \sqrt{2(F(s) + c)}$, so erkennen wir, daß wir zwar hieraus nicht unmittelbar $s$ als Funktion von $t$ durch Integration gewinnen können, daß die Lösung der Aufgabe aber gelingt, wenn wir uns zunächst mit der Auffindung der Umkehrfunktion $t(s)$ begnügen, d. h. der Zeit $t = t(s)$, die der Massenpunkt braucht, um zu einer bestimmten Stelle $s$ zu gelangen. Für diese ergibt sich die Gleichung

$$\frac{dt}{ds} = \frac{1}{\sqrt{2(F(s) + c)}},$$

d. h. der Differentialquotient der Funktion $t(s)$ ist uns bekannt, und es wird

$$t = \int \frac{ds}{\sqrt{2(F(s) + c)}} + c_1,$$

wo $c_1$ eine neue Integrationskonstante bedeutet. Sobald wir also dieses letzte Integral ausführen, haben wir das Problem gelöst, indem wir zwar nicht die Ortskoordinate $s$ als Funktion der Zeit, sondern umgekehrt die Zeit als Funktion der Ortskoordinate $s$ bestimmt haben. Die Tatsache, daß die beiden Integrationskonstanten $c$ und $c_1$ noch verfügbar sind, ermöglicht es, die allgemeine Lösung speziellen Anfangsbedingungen anzupassen.

In unserem obigen Beispiel der elastischen Bewegung haben wir $x$ mit $s$ zu identifizieren; es wird $f(s) = -\omega^2 s$ und entsprechend etwa $F(s) = -\frac{1}{2}\omega^2 s^2$. Wir erhalten also

$$\frac{dt}{ds} = \frac{1}{\sqrt{2c - \omega^2 s^2}}$$

und weiter

$$t = \int \frac{ds}{\sqrt{2c - \omega^2 s^2}} + c_1.$$

Dieses Integral aber können wir leicht auswerten, indem wir $\frac{\omega s}{\sqrt{2c}}$ als neue Veränderliche einführen, und es ergibt sich

$$t = \frac{1}{\omega} \arcsin \frac{\omega s}{\sqrt{2c}} + c_1$$

oder, indem wir wieder zu den Umkehrfunktionen übergehen,

$$s = \frac{\sqrt{2c}}{\omega} \sin \omega (t - c_1).$$

Wir sehen, daß wir gerade die schon oben direkt angegebene Lösung erhalten.

Wir erkennen aber auch an diesem Beispiel wieder, was die Integrationskonstanten bedeuten und wie sie zu bestimmen sind. Verlangen wir z. B., daß zur Zeit $t = 0$ der Massenpunkt sich gerade an der Stelle $s = 0$ befindet und daß seine Geschwindigkeit $\dot{s}(0)$ in diesem Augenblick den Wert 1 hat, so erhalten wir die beiden Gleichungen $0 = \frac{\sqrt{2c}}{\omega} \sin \omega c_1$, $1 = \sqrt{2c} \cos \omega c_1$, woraus sich für die Konstanten die bestimmten Werte $c_1 = 0$, $c = \frac{1}{2}$ ergeben. Genau ebenso lassen sich die Integrationskonstanten $c$ und $c_1$ bestimmen, wenn wir im Zeitmoment $t = 0$ die Anfangslage $s_0$ und die Anfangsgeschwindigkeit $\dot{s}_0$ ganz beliebig vorschreiben.

## § 5. Weitere Anwendungen: Fall eines Massenpunktes auf einer Kurve.

### 1. Allgemeines.

In besonders einfacher Weise läßt sich nach der zuletzt geschilderten Methode die Bewegung eines Massenpunktes behandeln, welcher unter dem Einfluß der Schwerkraft reibungslos auf einer ebenen Kurve gleitet. Ich will diese Bewegung zunächst allgemein und dann an dem Beispiel des gewöhnlichen Pendels und des Zykloidenpendels diskutieren. Wir legen das Koordinatensystem $x$, $y$ so, daß die $y$-Achse senkrecht nach oben, d. h. entgegengesetzt der Richtung der Schwerkraft weist, und denken uns die Kurve durch einen Parameter $\vartheta$ in der Parameterdarstellung $x = \varphi(\vartheta) = x(\vartheta)$, $y = \psi(\vartheta) = y(\vartheta)$ gegeben.

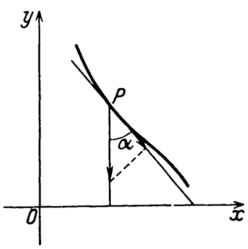



Fig. 92.

Ein Stück der Kurve, auf dem wir die Bewegung gerade betrachten wollen, ist in Figur 92 angedeutet. In jedem Punkt der Kurve wirkt auf die Masse $m$ unseres Massenpunktes die Schwerkraft vertikal nach unten, d. h. entgegengesetzt der $y$-Richtung, mit der Stärke $mg$. Verstehen wir unter $\alpha$ den Winkel zwischen dieser

Vertikalrichtung und der Tangente an die Kurve, so ist nach der in § 4, Nr. 1 formulierten Voraussetzung die Schwerkraft längs der Kurvenrichtung

$$m g \cos \alpha = - m g \frac{y'}{\sqrt{x'^2 + y'^2}},$$

wo

$$x' = \frac{d\varphi}{d\vartheta} = \varphi'(\vartheta), \qquad y' = \frac{d\psi}{d\vartheta} = \psi'(\vartheta)$$

ist — man beachte, daß wir hier durch den Strich nicht die Ableitung nach $x$, sondern die Ableitung nach $\vartheta$ bezeichnet haben —. Führen wir insbesondere die Bogenlänge $s$ als Parameter an Stelle von $\vartheta$ ein, so finden wir für die in Richtung der Kurve wirkende Kraft den Ausdruck $-mg\frac{dy}{ds}$, und somit ergibt sich für die Funktion $s(t)$ auf Grund des Newtonschen Gesetzes sofort die Differentialgleichung

$$\ddot{s} = -g \frac{dy}{ds}.$$

Auf der rechten Seite steht eine bekannte Funktion von $s$, da wir die Kurve kennen, also die Größen $x$ und $y$ als gegebene Funktionen von $s$ anzusehen haben.

Wenn wir wieder wie im vorigen Paragraphen diese Differentialgleichung mit $\dot{s}$ multiplizieren, so erhalten wir links den Differentialquotienten von $\frac{1}{2}\dot{s}^2$ nach $t$, rechts den Differentialquotienten von $-gy$ nach $t$ — wir haben uns dabei in der Funktion $y(s)$ die Größe $s$ durch $t$ ausgedrückt zu denken — und gewinnen nun durch Integration

$$\frac{1}{2}\dot{s}^2 = -gy + c,$$

wobei $c$ eine Integrationskonstante ist. Um die Bedeutung der Integrationskonstanten von vornherein zu fixieren, nehmen wir an, daß unser Massenpunkt zur Zeit $t = 0$ sich an einer Stelle der Kurve mit dem Parameterwert $\vartheta = \vartheta_0$ und den Koordinaten $x_0 = \varphi(\vartheta_0)$, $y_0 = \psi(\vartheta_0)$ befindet und daß in diesem Moment seine Geschwindigkeit Null ist, d. h. daß $\dot{s}(0) = 0$ wird. Dann ergibt sich sofort, wenn wir $t = 0$ oben einsetzen, $-gy_0 + c = 0$, also

$$\frac{1}{2}\dot{s}^2 = -g(y - y_0).$$

Ebenso wie wir bei der Bewegung $s$ als Funktion von $t$ auffassen können, dürfen wir auch die Umkehrfunktion $t(s)$ betrachten und erhalten für diese sofort

$$\frac{dt}{ds} = \pm \frac{1}{\sqrt{2g(y_0 - y)}},$$

eine Gleichung, welche gleichbedeutend mit

$$t = c_1 \pm \int \frac{ds}{\sqrt{2g(y_0 - y)}}$$

ist, wo $c_1$ eine neue Integrationskonstante bedeutet. Rechts steht nun unter dem Integralzeichen ein Ausdruck, den wir als Funktion des Parameters $\vartheta$ kennen, da uns die Kurve bekannt ist. Wir erhalten durch Einführung von $\vartheta$ als unabhängiger Veränderlichen

$$t = c_1 \pm \int \frac{ds}{d\vartheta} \frac{d\vartheta}{\sqrt{2g(y_0 - y)}} = c_1 \pm \int \sqrt{\frac{x'^2 + y'^2}{2g(y_0 - y)}}\, d\vartheta,$$

und hier sind uns die Funktionen $x' = \varphi'(\vartheta)$, $y' = \psi'(\vartheta)$, $y = \psi(\vartheta)$ bekannt. Um die Integrationskonstante $c_1$ zu bestimmen, beachten wir, daß für $t = 0$ der Parameterwert gleich $\vartheta_0$ sein soll. Wir erhalten daher sofort unsere Lösung in der Form

$$t = \pm \int_{\vartheta_0}^{\vartheta} \sqrt{\frac{x'^2 + y'^2}{2g(y_0 - y)}}\, d\vartheta.$$

Diese Gleichung stellt uns mit Hilfe eines Integrationsprozesses die Zeit $t$ dar, welche der Massenpunkt bei seiner Bewegung vom Parameterwerte $\vartheta_0$ bis zum Parameterwerte $\vartheta$ braucht. Die Umkehrfunktion $\vartheta(t)$ der so definierten Funktion $t(\vartheta)$ erlaubt, den Bewegungsvorgang vollständig zu beschreiben; denn wir können für jeden Zeitmoment $t$ den Punkt $x = \varphi(\vartheta(t))$, $y = \psi(\vartheta(t))$ der Kurve bestimmen, welchen unser beweglicher Massenpunkt gerade passiert.

## 2. Diskussion der Bewegung.

Aus unseren Gleichungen läßt sich, ohne daß wir die Integration mit Hilfe elementarer Funktionen auszuführen brauchen, die allgemeine Art der Bewegung durch eine einfache anschauliche Diskussion entnehmen. Wir setzen voraus, daß unsere Kurve den in Figur 93 gekennzeichneten Typus besitzt, d. h. aus einem nach unten konvexen Bogen besteht. Wenn zu Beginn der Bewegung der Massenpunkt an der links oben in der Figur gezeichneten Stelle $A$ mit den Koordinaten $x = x_0, y = y_0$, entsprechend $\vartheta = \vartheta_0$, losgelassen wird, nimmt seine Geschwindigkeit zu — die Beschleunigung $\ddot{s}$ ist nach unserer Bewegungsgleichung positiv. Der Massenpunkt wird bis zum tiefsten Punkt mit immer wachsender Geschwindigkeit herunterfallen. Nach dem Passieren dieses tiefsten Punktes aber wird die Beschleunigung negativ, da nunmehr die rechte Seite $-g\frac{dy}{ds}$ der Bewegungsgleichung negativ ist. Somit muß die Geschwindigkeit wieder abnehmen. Wir erkennen sofort aus der Gleichung $\dot{s}^2 = -2g(y - y_0)$, daß die Geschwindigkeit erst dann 0 wird, wenn der Punkt in die Höhe der Ausgangslage an der Stelle $B$

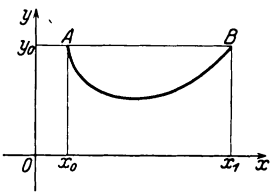


Fig. 93.

zurückgekehrt ist. Da die Beschleunigung noch immer negativ ist, so muß der Massenpunkt an dieser Stelle umkehren und bis zur Stelle $A$ zurückpendeln, und dieses Spiel muß sich — die Reibung haben wir ja dabei nicht berücksichtigt — fortwährend wiederholen. Die Zeit, welche der Punkt bei dieser Pendelbewegung braucht, um von $B$ nach $A$ zurückzugelangen, muß offenbar dieselbe wie für die Bewegung von $A$ nach $B$ sein. Nennen wir die Zeit für einen solchen Hin- und Rückweg $T$, so wird sich die Bewegung offenbar periodisch mit der Periode $T$ wiederholen. Bezeichnen wir mit $\vartheta_0$ und $\vartheta_1$ die Parameterwerte, welche zu den Punkten $A$ und $B$ gehören, so ist die halbe Schwingungsdauer unserer Pendelbewegung durch den Ausdruck

$$\frac{T}{2} = \frac{1}{\sqrt{2g}} \int_{\vartheta_0}^{\vartheta_1} \sqrt{\frac{x'^2 + y'^2}{y_0 - y}}\, d\vartheta = \frac{1}{\sqrt{2g}} \int_{\vartheta_0}^{\vartheta_1} \sqrt{\frac{\varphi'^2(\vartheta) + \psi'^2(\vartheta)}{\psi(\vartheta_0) - \psi(\vartheta)}}\, d\vartheta$$

dargestellt. Ist $\vartheta_2$ der Parameterwert, welcher dem tiefsten Punkt der Kurve entspricht, so wird die Fallzeit von $A$ nach diesem tiefsten Punkte durch den Ausdruck

$$\sqrt{\frac{1}{2g}} \left| \int_{\vartheta_0}^{\vartheta_2} \sqrt{\frac{x'^2 + y'^2}{y_0 - y}}\, d\vartheta \right|$$

geliefert.

### 3. Das gewöhnliche Pendel.

Das einfachste Beispiel einer solchen Gleichung liefert uns das sogenannte mathematische Pendel. Hier ist die betrachtete Kurve ein Kreis vom Radius $l$:

$$x = l \sin\vartheta, \quad y = -l\cos\vartheta.$$

Dabei haben wir den Winkel $\vartheta$ in positivem Sinne vom tiefsten Punkte aus gemessen. Aus unserem obigen allgemeinen Ausdruck erhalten wir sofort

$$T = \sqrt{\frac{2l}{g}} \int_{-\alpha}^{\alpha} \frac{d\vartheta}{\sqrt{\cos\vartheta - \cos\alpha}} = \sqrt{\frac{l}{g}} \int_{-\alpha}^{\alpha} \frac{d\vartheta}{\sqrt{\sin^2\frac{\alpha}{2} - \sin^2\frac{\vartheta}{2}}}.$$

Es bedeutet dabei $\alpha$ $(0 < \alpha < \pi)$ den Ausschlagswinkel des Pendels, d. h. diejenige Winkelstellung, von der aus wir etwa zur Zeit $t = 0$ den Massenpunkt mit der Geschwindigkeit 0 losgelassen denken. Durch die Substitution

$$u = \frac{\sin\frac{\vartheta}{2}}{\sin\frac{\alpha}{2}}, \quad \frac{du}{d\vartheta} = \frac{\cos\frac{\vartheta}{2}}{2\sin\frac{\alpha}{2}}$$

geht unser Ausdruck für die Schwingungsdauer des Pendels über in

$$T = 2\sqrt{\frac{l}{g}} \int_{-1}^{1} \frac{du}{\sqrt{(1-u^2)\left(1-u^2 \sin^2 \frac{\alpha}{2}\right)}}.$$

Wir haben also die Schwingungsdauer des Pendels durch ein elliptisches Integral ausgedrückt.

Nehmen wir an, daß der Ausschlagswinkel $\alpha$ klein ist und daß wir daher den zweiten Faktor unter dem Wurzelzeichen mit hinreichender Genauigkeit durch 1 ersetzen dürfen, so erhalten wir als Annäherung für die Schwingungsdauer den Ausdruck

$$2\sqrt{\frac{l}{g}} \int_{-1}^{1} \frac{du}{\sqrt{1-u^2}}.$$

Wir können rechts das Integral nach Formel **13** unserer Integrationstabelle (S. **167**) auswerten und bekommen als angenäherten Wert für $T$ den Ausdruck $2\pi\sqrt{\frac{l}{g}}$.

## 4. Das Zykloidenpendel.

Die Tatsache, daß die Schwingungsdauer des gewöhnlichen Pendels nicht streng unabhängig von dem Ausschlagswinkel ist, hat *Christian Huygens* im Zusammenhang mit seinen Bemühungen um die Konstruktion von Präzisionsuhren dazu veranlaßt, nach einer Kurve zu suchen, bei welcher die Schwingungsdauer streng unabhängig davon wird, an welcher Stelle der Kurve der schwingende Massenpunkt seine Bewegung beginnt. Als solche Kurve erkannte Huygens die Zykloide.

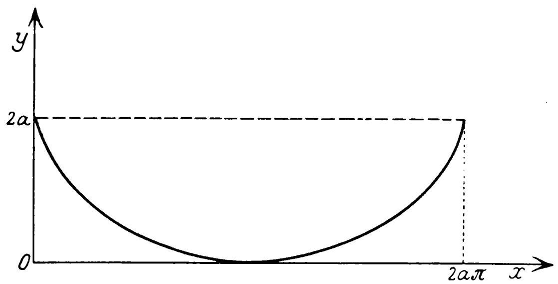



Fig. 94. Bahn des Zykloidenpendels.

Damit ein Massenpunkt überhaupt auf einer Zykloide schwingen kann, müssen die Spitzen der Zykloide entgegengesetzt der Richtung der Schwerkraft weisen; d. h. wir haben die Zykloide, wie wir sie bis jetzt betrachteten, etwa um die $x$-Achse umzuklappen (vgl. Fig. 94). Wir schreiben daher jetzt die Gleichung der Zykloide in der Form

$$x = a(\vartheta - \sin\vartheta),$$
$$y = a(1 + \cos\vartheta),$$

was noch einer Verschiebung um $2a$ in der positiven $y$-Richtung entspricht. Die Zeit, welche der Massenpunkt braucht, um von einem Punkt von der Höhe

$$y_0 = a(1 + \cos\alpha) \qquad (0 < \alpha < \pi)$$

bis zum tiefsten Punkt zu gelangen, ist gemäß der in Nr. 2 entwickelten Formel

$$\frac{T}{4} = \sqrt{\frac{1}{2g}} \int_\alpha^\pi \sqrt{\frac{x'^2 + y'^2}{y_0 - y}}\, d\vartheta = \sqrt{\frac{a}{g}} \int_\alpha^\pi \sqrt{\frac{1 - \cos\vartheta}{\cos\alpha - \cos\vartheta}}\, d\vartheta.$$

Nun benutzen wir die Gleichung

$$\cos\alpha - \cos\vartheta = 2\left(\cos^2\frac{\alpha}{2} - \cos^2\frac{\vartheta}{2}\right);$$

es wird

$$\frac{T}{4} = \sqrt{\frac{a}{g}} \int_\alpha^\pi \frac{\sin\frac{\vartheta}{2}}{\sqrt{\cos^2\frac{\alpha}{2} - \cos^2\frac{\vartheta}{2}}}\, d\vartheta.$$

Wir berechnen zuerst das unbestimmte Integral, und zwar durch die Substitution

$$\cos\frac{\vartheta}{2} = u\cos\frac{\alpha}{2}, \qquad \sin\frac{\vartheta}{2}\, d\vartheta = -2\cos\frac{\alpha}{2}\, du.$$

Es wird

$$\int \frac{\sin\frac{\vartheta}{2}}{\sqrt{\cos^2\frac{\alpha}{2} - \cos^2\frac{\vartheta}{2}}}\, d\vartheta = -2\int \frac{du}{\sqrt{1 - u^2}} = -2 \arcsin u,$$

und daher erhalten wir

$$T = -8\sqrt{\frac{a}{g}} \arcsin \frac{\cos\frac{\vartheta}{2}}{\cos\frac{\alpha}{2}} \Bigg|_\alpha^\pi = 4\pi\sqrt{\frac{a}{g}}.$$

Es ist also die Schwingungsdauer $T$ in der Tat unabhängig von dem Ausschlagswinkel $\alpha$.

## § 6. Arbeit.

### 1. Allgemeines.

Auf die Betrachtung der letzten Paragraphen und auf viele andere Fragen der Mechanik und Physik wird ein neues Licht geworfen durch den Begriff der *Arbeit*.

Denken wir uns wieder den Massenpunkt auf einer Kurve unter dem Einfluß einer in dieser Kurve wirkenden Kraft bewegt und die Punkte der Kurve durch die von irgend einer Anfangsstelle gemessene Bogenlänge $s$ charakterisiert. Dann wird die Kraft im allgemeinen selbst von $s$ abhängig sein. Wir nehmen an, daß sie eine stetige Funktion $f(s)$ der Bogenlänge sei. Diese Funktion möge positive Werte besitzen, sobald die Richtung der Kraft nach der Richtung wachsender $s$-Werte weist, negative Werte, wenn die Kraftrichtung der Richtung wachsender $s$ entgegengesetzt ist.

Wenn die Größe der wirkenden Kraft längs des Weges konstant ist, so versteht man unter der von der Kraft geleisteten Arbeit das Produkt aus der Kraft und dem von dem Massenpunkt zurückgelegten Wege $s_1 - s_0$, wobei $s_1$ den Endpunkt, $s_0$ den Anfangspunkt der Bewegung bezeichnet. Ist die Kraft nicht konstant, so hat man die geleistete Arbeit durch einen Grenzübergang zu definieren. Man zerlegt das Intervall von $s_0$ bis $s_1$ in $n$ gleich oder verschieden lange Teile und stellt sich vor, daß in jedem Teil die Kraft einen beinahe konstanten Wert besitzt, etwa gleich der tatsächlichen Größe der Kraft in irgend einem willkürlich im betreffenden Teilintervall gewählten Punkt $\sigma_\nu$. Wir können dann für eine solche sprunghaft von Intervall zu Intervall sich verändernde Kraft den Ausdruck für die Arbeit in der Form

$$\sum_{\nu=1}^{n} f(\sigma_\nu)\,\Delta s_\nu$$

hinschreiben, wobei wir unter $\sigma_1, \sigma_2, \ldots, \sigma_n$ die benutzten Zwischenwerte der Kraft im ersten, zweiten usw. Intervall verstehen und mit $\Delta s_1, \Delta s_2, \ldots$ die Längen der zugehörigen Teilintervalle bezeichnen. Machen wir nunmehr den Grenzübergang, indem wir $n$ über alle Grenzen wachsen und dabei die Länge des jeweils längsten Teilintervalles gegen Null streben lassen, so wird unsere Summe auf Grund der Integraldefinition einfach gegen das Integral

$$A = \int_{s_0}^{s_1} f(s)\,ds$$

streben, und dieses Integral werden wir naturgemäß als die bei der Bewegung geleistete Arbeit bezeichnen.

Sind die Kraftrichtung und die Bewegungsrichtung dieselbe, so ist die geleistete Arbeit positiv; man sagt dann auch kurz, daß von der Kraft (positive) Arbeit geleistet wird. Sind dagegen Kraftrichtung und Bewegungsrichtung entgegengesetzt, so wird die Arbeit negativ; man spricht dann von gewonnener Arbeit[1]).

[1]) Man muß beachten, daß es bei dieser Ausdrucksweise noch ganz darauf ankommt, welche Kraft man ins Auge faßt. Z. B. wird beim Heben eines Gewichtes die von der Schwerkraft geleistete Arbeit negativ. Vom Standpunkte des Hebenden

Betrachten wir die Ortskoordinate $s$ und die Kraft $f(s) = p$ beide als Funktionen der Zeit $t$ als Parameter, so wird in einer Ebene mit den rechtwinkligen Koordinaten $s$ und $p$ der Punkt mit diesen Koordinaten $s(t)$, $p(t)$ als Funktion der Zeit eine Kurve beschreiben; diese Kurve nennen wir das *Arbeitsdiagramm des Vorganges.* Wenn es sich um einen periodischen Vorgang handelt, wie bei allen Maschinen, so wird der mit der Zeit sich bewegende Punkt $s(t)$, $p(t)$ nach der Zeit $T$ (einer Periode) immer wieder an seine Stelle zurückkehren, d. h. das Arbeitsdiagramm eine geschlossene Kurve sein. Die Kurve kann in diesem Falle einfach aus einem und demselben vorwärts und rückwärts durchlaufenen Bogen bestehen, nämlich dann, wenn die Kraft bei dem Vorgang als eindeutige Funktion von $s$ gegeben ist, wie z. B. bei elastischen Schwingungen; sie kann aber auch eine eigentliche geschlossene Kurve darstellen, wie z. B. bei jeder Maschine, bei welcher der Druck auf einen Kolben beim Hingang nicht derselbe wie beim Rückgang ist. Die geleistete Arbeit wird dann einfach durch den Flächeninhalt des Arbeitsdiagrammes gegeben, nämlich durch das Integral

$$\int_{t_0}^{t_0+T} p(t) \frac{ds}{dt} dt,$$

wo das Zeitintervall von $t_0$ bis $t_0 + T$ gerade eine Periode der Bewegung darstellt. Wird ein Flächeninhalt positiv umlaufen, so entspricht ihm eine negative, wird er negativ umlaufen, so entspricht ihm eine positive Arbeit. Besteht die Kurve aus mehreren Schleifen, die teils positiv, teils negativ umlaufen werden, so ist die gesamte geleistete Arbeit durch die Summe der mit dem richtigen Vorzeichen zu nehmenden einzelnen Schleifeninhalte dargestellt.

### 2. Erstes Beispiel. Massenanziehung.

Ein Massenpunkt wirke nach dem Newtonschen Anziehungsgesetz auf einen zweiten Massenpunkt; wir betrachten dann als erstes Beispiel die Arbeit, welche von dieser Anziehungskraft geleistet wird, wenn sich der zweite Massenpunkt im Anziehungsbereich des ersten auf der Verbindungslinie bewegt. Das Newtonsche Anziehungsgesetz besagt, daß die anziehenden Kräfte umgekehrt proportional dem Quadrat der Entfernung sind. Denken wir uns den anziehenden Punkt im Nullpunkte ruhend, den angezogenen Punkt in der Entfernung $r$ vom Nullpunkt, so wird die anziehende Kraft gegeben durch einen Ausdruck

$$f(r) = -\mu \frac{1}{r^2},$$

aus würde man die Arbeit als positiv geleistet und nicht als gewonnen bezeichnen. Der Mensch, welcher hebt, leistet nämlich Arbeit mit einer der Schwere entgegengesetzten Kraft.

wo $\mu$ eine positive Konstante ist. Die Arbeit, welche von dieser Kraft geleistet wird, wenn der Punkt unter ihrem Einfluß aus der Entfernung $r$ in die Entfernung $r_1 < r$ gelangt, ist daher positiv und gleich dem Integral

$$-\mu \int_r^{r_1} \frac{ds}{s^2} = \mu \left( \frac{1}{r_1} - \frac{1}{r} \right).$$

Wird der Punkt durch eine Gegenkraft vom Nullpunkt fort bewegt, gelangt er also aus der Entfernung $r$ in eine Entfernung $r_1 > r$, so wird die von der Anziehungskraft geleistete Arbeit natürlich wieder durch dieses (nunmehr negative) Integral gegeben. Die von der Gegenkraft geleistete Arbeit hat den entgegengesetzt gleichen, also positiven Wert $\mu\left(\frac{1}{r} - \frac{1}{r_1}\right)$. Denken wir uns die Endlage immer weiter und weiter entfernt, so erhalten wir als Grenzwert für die Arbeit, welche gegen die Anziehung geleistet werden muß, um den beweglichen Punkt aus dem Anziehungsbereich des festen vollständig ins „Unendliche" zu entfernen, den Ausdruck $\frac{\mu}{r}$. Man nennt diesen wichtigen Ausdruck das *Potential* der beiden Punkte aufeinander. Das Potential ist also hier definiert als Arbeit, die zur vollständigen Trennung zweier einander anziehender Punkte gebraucht wird; beispielsweise die Arbeit, die geleistet werden muß, um ein Elektron aus dem Verbande eines Atoms vollständig herauszureißen (Ionisierungsspannung, bezogen auf das Elektron).

### 3. Zweites Beispiel. Spannen einer Feder.

Als weiteres Beispiel betrachten wir die Arbeit, welche beim Spannen einer Feder geleistet wird. Wir machen (vgl. auch § 4) die in der Elastizitätstheorie übliche Annahme, daß die Kraft, die zur Federspannung notwendig ist, proportional der Vergrößerung $x$ der Federlänge ist, d. h. $p = kx$, wo $k$ eine Konstante bedeutet. Die Arbeit, die wir leisten müssen, um die Feder aus der Ruhelage $x = 0$ in eine Endlage $x = x_1$ zu bringen, wird also durch das Integral

$$\int_0^{x_1} k\,x\,dx = \frac{k\,x_1^2}{2}$$

gegeben.

### 4. Drittes Beispiel. Aufladen eines Kondensators.

Ähnlich steht es mit dem Begriff der Arbeit in anderen Gebieten der Physik. Betrachten wir z. B. das Aufladen eines Kondensators. Bezeichnen wir mit $Q$ die Elektrizitätsmenge auf dem Kondensator, mit $C$ seine Kapazität, mit $V$ seine Spannung, so besteht die Relation $Q = CV$. Hätte die Spannung des Kondensators bei der Aufladung einen konstanten Wert $V$, so würde die „elektrische Arbeit", welche nötig ist, um eine Elektrizitätsmenge $Q$ auf den ungeladenen Kondensator zu bringen, durch das Produkt $QV$ gemessen werden. Da sich jedoch bei der Aufladung die Spannung selbst vermehrt, so erhalten wir durch einen dem bei der obigen Definition der Arbeit in Nr. 1 durchgeführten ganz analogen Grenzübergang als Arbeit beim Aufladen eines Kondensators den Ausdruck

$$\int_0^{Q_1} V\,dQ = \frac{1}{C}\int_0^{Q_1} Q\,dQ = \frac{Q_1^2}{2C} = \frac{Q_1 V_1}{2},$$

wo $Q_1$ die gesamte auf den Kondensator gebrachte Elektrizitätsmenge, $V_1$ die Spannung am Schluß des Ladevorganges bedeutet.

# Anhang zum fünften Kapitel.

## Eigenschaften der Evolute.

Unsere obige Parameterdarstellung

$$\xi = x - \varrho\frac{\dot y}{\sqrt{\dot x^2 + \dot y^2}}, \qquad \eta = y + \varrho\frac{\dot x}{\sqrt{\dot x^2 + \dot y^2}}$$

für die Evolute einer gegebenen Kurve $x = x(t)$, $y = y(t)$ (vgl. S. 224) erlaubt uns, einige interessante geometrische Beziehungen zwischen ihr und der gegebenen Kurve abzulesen. Der Bequemlichkeit halber bedienen wir uns dabei wieder der Bogenlänge $s$ als Parameter, so daß

$$\dot x^2 + \dot y^2 = 1 \quad \text{und} \quad \dot x\ddot x + \dot y\ddot y = 0,$$

$$\frac{1}{\varrho} = k = \frac{\ddot y}{\dot x} = -\frac{\ddot x}{\dot y} \quad \text{oder} \quad \varrho\ddot y = \dot x \quad \text{und} \quad \varrho\ddot x = -\dot y$$

wird. Es ist dann also

$$\xi = x - \varrho\dot y, \qquad \eta = y + \varrho\dot x;$$

durch Differentiation ergibt sich

$$\dot\xi = \dot x - \varrho\ddot y - \dot\varrho\dot y = -\dot\varrho\dot y, \qquad \dot\eta = \dot y + \varrho\ddot x + \dot\varrho\dot x = \dot\varrho\dot x$$

und somit

$$\dot\xi\dot x + \dot\eta\dot y = 0.$$

Da die Richtungskosinus der Kurvennormale durch $-\dot y$ und $\dot x$ gegeben werden, so folgt hieraus: *Die Kurvennormale berührt die Evolute im Krümmungsmittelpunkt;* oder: Die Tangenten an die Evolute sind die Normalen der gegebenen Kurve; oder: ***Die Evolute ist die von den Normalen „eingehüllte" Kurve*** (vgl. Fig. 95).

Nennen wir weiter die von einem beliebigen Anfangspunkt der Evolute aus gezählte Bogenlänge auf ihr $\sigma$, so wird

$$\left(\frac{d\sigma}{ds}\right)^2 = \dot\sigma^2 = \dot\xi^2 + \dot\eta^2.$$

Aus den obigen Formeln ergibt sich sofort wegen $\dot x^2 + \dot y^2 = 1$

$$\dot\sigma^2 = \dot\varrho^2,$$

also bei geeigneter Zählung der Bogenlänge $\sigma$ — jedenfalls solange $\dot{\sigma} \neq 0$ bleibt —

$$\dot{\sigma} = \dot{\varrho}$$

oder integriert

$$\sigma_1 - \sigma_0 = \varrho_1 - \varrho_0 .$$

D. h. *die Bogenlänge der Evolute zwischen zwei Punkten ist gleich der Differenz der zugehörigen Krümmungsradien, solange für den betreffenden Bogen $\dot{\varrho}$ von Null verschieden bleibt.*

Die letzte Bedingung ist nicht überflüssig. Wenn nämlich $\dot{\varrho}$ das Vorzeichen wechselt, so würde die Formel $\dot{\sigma} = \dot{\varrho}$ beim Weiterwandern über den betreffenden Punkt der Evolute hinaus bedeuten, daß die Bogenlänge $\sigma$ ein Maximum oder Minimum hat, d. h. daß man $\sigma$ nicht einfach weiterzählt, sondern auch bei dieser Zählung das Zeichen umkehren muß. Will man das vermeiden, so hat man beim Überschreiten dieser Punkte in der Formel das Zeichen zu wechseln, d. h. jetzt $\dot{\sigma} = -\dot{\varrho}$ zu setzen.

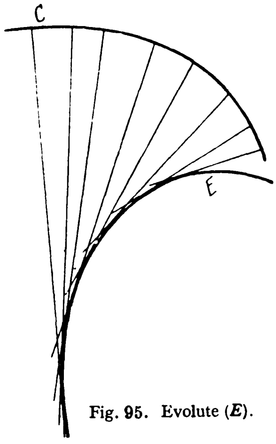


Fig. 95. Evolute ($E$).

Im übrigen erwähne ich noch, daß die Krümmungsmittelpunkte, welche zu einem Maximum oder Minimum des Krümmungsradius gehören, *Spitzen für die Evolute* sind. (Den Beweis übergehe ich hier.)

Man kann dem eben gefundenen geometrischen Zusammenhang noch einen andern Ausdruck geben: Denkt man sich einen biegsamen, nicht dehnbaren Faden um einen Bogen der Evolute gelegt und so gespannt, daß ein Teil desselben sich von der Kurve ablöst und tangential zu ihr steht, so wird der Endpunkt $Q$, falls er zuerst auf der Ausgangskurve $C$ liegt, auch beim Abwickeln des Fadens einen Bogen dieser Ausgangskurve beschreiben. Diese nennen wir deshalb eine zur Evolute $E$ gehörige Evolvente (evolvere = abwickeln). Man kann diesen Zusammenhang umkehren, von einer beliebigen Kurve $E$ ausgehen und zu ihr durch den beschriebenen Abwickelungsprozeß eine Evolvente $C$ konstruieren. Dann zeigt sich, daß umgekehrt $E$ die Evolute von $C$ wird.

Zum Beweis denken wir uns die jetzt als gegeben betrachtete Kurve $E$ in der Form $\xi = \xi(\sigma)$, $\eta = \eta(\sigma)$ dargestellt, wobei die laufenden rechtwinkeligen Koordinaten mit $\xi$ und $\eta$ bezeichnet sind und als Parameter die Bogenlänge $\sigma$ angenommen werden möge. Die Aufwickelung möge erfolgen wie in Figur 96 angedeutet; d.h. der Endpunkt $Q$ des im veränderlichen, zur Bogenlänge $\sigma$ gehörigen Kurvenpunkte $P$ mit den Koordinaten $\xi$, $\eta$ berührenden Fadens möge bei voller Aufwickelung gerade in den festen, zur Bogenlänge $a \geqq \sigma$ gehörigen Kurvenpunkt $A$ fallen.

Dann ist die Strecke $PQ$ gleich $a - \sigma$, und die Richtungskosinus der Tangente $PQ$ sind $\dot{\xi}$ und $\dot{\eta}$, wenn nunmehr der Punkt die Ableitung nach $\sigma$ bedeutet. Somit ergeben sich für die Koordinaten $x$ und $y$ des Punktes $Q$ die Ausdrücke

$$x = \xi + (a - \sigma)\dot{\xi}, \qquad y = \eta + (a - \sigma)\dot{\eta},$$

welche die Parameterdarstellung der vom Punkte $Q$ beschriebenen Evolvente mit $\sigma$ als Parameter geben. Durch Differentiation nach $\sigma$ folgt sofort

$$\dot{x} = \dot{\xi} - \dot{\xi} + (a - \sigma)\ddot{\xi} = (a - \sigma)\ddot{\xi},$$
$$\dot{y} = \dot{\eta} - \dot{\eta} + (a - \sigma)\ddot{\eta} = (a - \sigma)\ddot{\eta}.$$

Somit ergibt sich wegen $\dot{\xi}\ddot{\xi} + \dot{\eta}\ddot{\eta} = 0$

$$\dot{x}\dot{\xi} + \dot{y}\dot{\eta} = 0,$$

und diese Gleichung besagt, daß die Gerade $PQ$ auf der Evolvente $C$ normal steht. Wir können somit sagen: Die Normalen der Kurve $C$ berühren die Kurve $E$. Durch diese Beziehung aber ist $E$ als die Evolute zu $C$ charakterisiert. Also: *Jede Kurve ist die Evolute jeder ihrer Evolventen.*

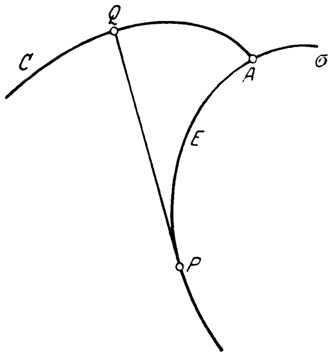


Fig. 96. Evolvente (C).

Als Beispiel für unsere allgemeinen Erörterungen betrachten wir die Evolute der Zykloide $x = t - \sin t$, $y = 1 - \cos t$. Es ist (siehe S. 222ff.)

$$\xi = x - \dot{y}\,\frac{\dot{x}^2 + \dot{y}^2}{\dot{x}\ddot{y} - \dot{y}\ddot{x}}, \qquad \eta = y + \dot{x}\,\frac{\dot{x}^2 + \dot{y}^2}{\dot{x}\ddot{y} - \dot{y}\ddot{x}};$$

also erhalten wir die Evolute durch den Parameter $t$ dargestellt in der Form $\xi = t + \sin t$, $\eta = -1 + \cos t$. Setzt man $t = \tau + \pi$, so

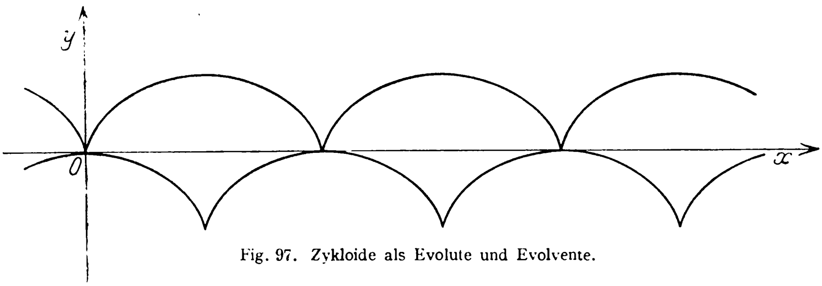


Fig. 97. Zykloide als Evolute und Evolvente.

wird $\xi - \pi = \tau - \sin\tau$, $\eta + 2 = 1 - \cos\tau$, und diese Gleichungen zeigen, daß die Evolute wieder eine der ursprünglichen kongruente Zykloide ist, die durch Verschiebung, wie in Figur 97 angedeutet, aus der gegebenen entsteht.

Als weiteres Beispiel stelle ich die Gleichung der Kreisevolvente auf. Wir gehen demgemäß von dem Kreise $\xi = \cos t$, $\eta = \sin t$ aus und wickeln die Tangente auf, wie in Figur 97a angedeutet. Dann ergibt sich die Kreisevolvente in der Gestalt

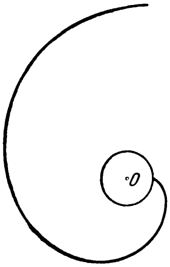


Fig. 97a. Kreisevolvente.

$$x = \cos t + t \sin t, \qquad y = \sin t - t \cos t$$

durch den Parameter $t$ dargestellt.

Endlich bestimmen wir noch die Evolute der Ellipse $x = a \cos t$, $y = b \sin t$. Es ergibt sich sofort

$$\xi = x - \dot{y}\frac{\dot{x}^2 + \dot{y}^2}{\dot{x}\ddot{y} - \dot{y}\ddot{x}} = \frac{a^2 - b^2}{a}\cos^3 t$$

und

$$\eta = y + \dot{x}\frac{\dot{x}^2 + \dot{y}^2}{\dot{x}\ddot{y} - \dot{y}\ddot{x}} = -\frac{a^2 - b^2}{b}\sin^3 t.$$

Eliminiert man $t$ in der üblichen Weise, so erhält man aus der gefundenen Parameterdarstellung der Evolute deren Gleichung in rechtwinkeligen Koordinaten

$$(a\xi)^{\frac{2}{3}} + (b\eta)^{\frac{2}{3}} = (a^2 - b^2)^{\frac{2}{3}}.$$

Diese Kurve ist eine sogenannte *Asteroide*. Ihre Gestalt ist durch Figur 98 angegeben. Daß tatsächlich die Krümmungsmittelpunkte für die Scheitel der Ellipse Spitzen der Asteroide sind, erkennt man übrigens leicht aus deren Parameterdarstellung.

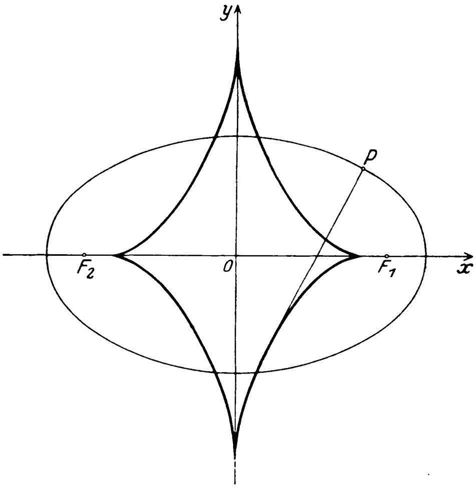


Fig. 98. Evolute der Ellipse.

Sechstes Kapitel.

# Die Taylorsche Formel und die Annäherung von Funktionen durch ganze rationale.

Die rationalen Funktionen stellen sich in vieler Hinsicht als die prinzipiell einfachsten Funktionen der Analysis dar; sie entstehen, indem man eine endliche Anzahl von Malen die rationalen Rechenoperationen auf die Variable $x$ anwendet, während letzten Endes die

Bildung jeder anderen Funktion mehr oder weniger versteckt die Ausführung eines Grenzüberganges von rationalen Funktionen aus verlangt. Es ist daher eine Frage von großer theoretischer wie praktischer Bedeutung, ob und wie genau man eine gegebene Funktion $f(x)$ durch rationale und insbesondere durch ganze rationale Funktionen angenähert darstellen oder, wie man auch sagt, *approximieren* kann.

## § 1. Der Logarithmus und der Arcustangens.

### 1. Der Logarithmus.

Zunächst betrachten wir einige spezielle Fälle, bei denen uns die Integration der geometrischen Reihe fast unmittelbar die gewünschte Annäherung liefert. Ich erinnere vorab noch einmal an die folgende Tatsache. Es ist für $q \neq 1$ und ganzes $n > 0$

$$\frac{1}{1-q} = 1 + q + q^2 + \cdots + q^{n-1} + r_n,$$

wo

$$r_n = \frac{q^n}{1-q}$$

gesetzt ist. Im Falle $|q| < 1$ wird mit wachsendem $n$ der Rest $r_n$ gegen 0 streben, und wir erhalten dann, wie im ersten Kapitel, S. 26, auseinandergesetzt wurde, die *unendliche geometrische Reihe*

$$1 + q + q^2 + \cdots \text{ mit der Summe } \frac{1}{1-q}.$$

Nunmehr gehen wir von der Formel

$$\log(1+x) = \int_0^x \frac{dt}{1+t}$$

aus und entwickeln den Integranden gemäß der obigen Formel, wo $q = -t$ zu setzen ist. Dann ergibt sich sofort durch Integration

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + - \cdots + (-1)^{n-1}\frac{x^n}{n} + R_n,$$

wobei

$$R_n = \int_0^x r_n\,dt = (-1)^n \int_0^x \frac{t^n\,dt}{1+t}$$

gesetzt ist. Wir haben damit die Funktion $\log(1+x)$ bei beliebigem positivem ganzzahligem $n$ durch eine ganze rationale Funktion $n$-ten Grades, nämlich die Funktion $x - \frac{x^2}{2} + \frac{x^3}{3} - + \cdots + (-1)^{n-1}\frac{x^n}{n}$, approximiert; die Größe $R_n$, der *Rest*, gibt uns an, einen wie großen *Fehler* wir bei dieser Approximation begehen.

Um die Genauigkeit dieser Approximation abzuschätzen, brauchen wir nur eine Abschätzung für den Rest $R_n$; und diese Abschätzung wird uns ohne weiteres durch die Integralabschätzungen aus Kap. II,

§ 7, Nr. 1 gegeben. Nehmen wir zunächst an, es sei $x \geqq 0$, dann ist der Integrand in dem ganzen Integrationsintervall nirgends negativ und überall $\leqq t^n$. Es wird also

$$|R_n| \leqq \int_0^x t^n\,dt = \frac{x^{n+1}}{n+1},$$

und wir sehen hieraus, daß für alle Werte von $x$, welche in dem Intervall $0 \leqq x \leqq 1$ liegen, dieser Rest so klein wird, wie wir wollen, wenn wir nur $n$ hinreichend groß wählen (vgl. erstes Kapitel, § 5, Nr. 5). Liegt die Größe $x$ dagegen in dem Intervall $-1 < x \leqq 0$, so wird der Integrand das Vorzeichen nicht wechseln und wird absolut genommen $\leqq \frac{|t|^n}{1+x}$ sein, und wir erhalten für den Rest sofort die Abschätzung

$$|R_n| \leqq \frac{1}{1+x}\int_0^{|x|} t^n\,dt = \frac{|x|^{n+1}}{(1+x)(n+1)}.$$

Wir sehen also, daß auch hier der Rest bei hinreichend großem $n$ beliebig klein wird. Unsere Abschätzung versagt aber naturgemäß, wenn wir $x = -1$ setzen.

Zusammenfassend können wir sagen: Es ist

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - + \cdots + (-1)^{n-1}\frac{x^n}{n} + R_n,$$

wobei der Rest $R_n$ mit wachsendem $n$ gegen 0 strebt, sobald $x$ in dem Gebiet $-1 < x \leqq 1$ liegt[1]). Wir können für den Rest aus unseren obigen Ungleichungen sogar eine und dieselbe Abschätzung für alle Werte $x$ eines Intervalles $-1 + h \leqq x \leqq 1$ geben, wobei $h$ eine Zahl bedeutet, die der Ungleichung $0 < h \leqq 1$ genügt. Dann wird nämlich

$$|R_n| \leqq \frac{1}{h}\,\frac{1}{n+1},$$

eine Formel, die uns zeigt, daß in dem ganzen Intervall die Funktion $\log(1+x)$ durch unser Polynom $n$-ten Grades mindestens mit der Genauigkeit $\frac{1}{h}\,\frac{1}{n+1}$ approximiert wird. Ich überlasse es dem Leser, sich davon zu überzeugen, daß für alle Werte von $x$, für welche $|x| > 1$ ist, der Rest nicht nur nicht gegen 0 streben kann, sondern sogar absolut genommen mit wachsendem $n$ über alle Grenzen wachsen muß, so daß für solche Werte von $x$ durch unsere Polynome keine angenäherte Darstellung des Logarithmus geliefert wird.

Die Tatsache, daß in dem obigen Intervall $R_n$ gegen 0 strebt, drückt man auch dadurch aus, daß man sagt, in diesem Intervall haben wir

[1]) Man beachte, daß dieses Gebiet nach der linken Seite offen, nach der rechten abgeschlossen ist.

*für den Logarithmus die unendliche Reihe*

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + - \cdots$$

hergeleitet. Setzen wir in dieser Reihe speziell $x = 1$, so erhalten wir die merkwürdige Formel

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + - \cdots,$$

eine der Beziehungen, deren Entdeckung auf die Gemüter der ersten Pioniere der Differential- und Integralrechnung einen tiefen Eindruck gemacht hat.

Die oben gegebene Approximation für den Logarithmus führt zu einer anderen, für viele Zwecke, insbesondere auch für numerische Rechnungen nützlichen Formel, wenn man im Falle $-1 < x < 1$ noch den Ausdruck

$$\log(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \cdots - R_n^*$$

bildet und durch Subtraktion für gerade $n$ zu der Gleichung

$$\frac{1}{2}\log\frac{1+x}{1-x} = x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots + \frac{x^{n-1}}{n-1} + \overline{R}_n$$

übergeht. Hierbei ist der Rest durch den Ausdruck

$$\overline{R}_n = \frac{1}{2}(R_n + R_n^*) = \frac{1}{2}\int_0^x t^n\left\{\frac{1}{1+t} + \frac{1}{1-t}\right\}dt = \int_0^x \frac{t^n}{1-t^2}\,dt$$

gegeben. Der Rest $\overline{R}_n$ strebt wegen der Beziehung

$$|\overline{R}_n| \leqq \frac{|x^{n+1}|}{n+1}\,\frac{1}{1-x^2}$$

bei wachsendem $n$ gegen Null, was wir wiederum zum Ausdruck bringen, indem wir die für $|x| < 1$ gültige Entwicklung

$$\frac{1}{2}\log\frac{1+x}{1-x} = \mathfrak{Ar}\,\mathfrak{Tg}\,x = x + \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7} + \cdots$$

in eine unendliche Reihe hinschreiben.

## 2. Der Arcustangens.

Ganz ähnlich läßt sich die Funktion Arcustangens behandeln, indem wir von der für jedes ganze positive $n$ gültigen Formel

$$\frac{1}{1+t^2} = 1 - t^2 + t^4 - + \cdots + (-1)^{n-1} t^{2n-2} + r_n$$

mit

$$r_n = (-1)^n \frac{t^{2n}}{1+t^2}$$

ausgehen. Durch Integration erhalten wir

$$\operatorname{arc\,tg} x = x - \frac{x^3}{3} + \frac{x^5}{5} - + \cdots + (-1)^{n-1} \frac{x^{2n-1}}{2n-1} + R_n,$$

$$R_n = (-1)^n \int_0^x \frac{t^{2n}}{1+t^2}\, dt,$$

und wir erkennen sofort, daß im Intervall $-1 \leqq x \leqq 1$ der Rest $R_n$ mit wachsendem $n$ gegen $0$ strebt; denn für diesen Rest gilt zufolge des Mittelwertsatzes der Integralrechnung

$$|R_n| \leqq \int_0^{|x|} t^{2n}\, dt = \frac{|x|^{2n+1}}{2n+1}.$$

Wir sehen aus der Restdarstellung auch leicht, daß für $|x| > 1$ der absolute Betrag des Restes bei wachsendem $n$ über alle Grenzen wachsen wird. Für $|x| \leqq 1$ haben wir somit die unendliche Reihe

$$\operatorname{arc\,tg} x = x - \frac{x^3}{3} + \frac{x^5}{5} - + \cdots$$

abgeleitet, aus der wir für den speziellen Wert $x = 1$, wegen $\operatorname{arc\,tg} 1 = \frac{\pi}{4}$, die Reihe

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - + \cdots$$

erhalten. Ein ebenso merkwürdiges Resultat wie das oben für $\log 2$ gefundene.

## § 2. Die allgemeine Taylorsche Formel.

Eine angenäherte Darstellung durch rationale Funktionen, wie in den oben betrachteten speziellen Fällen, gelingt nun auch in dem Fall einer beliebigen Funktion $f(x)$, von der wir nur voraussetzen, daß sie für alle betrachteten Werte der unabhängigen Veränderlichen eines vorgegebenen abgeschlossenen Intervalles stetige Ableitungen bis mindestens zur $(n+1)$-ten Ordnung besitzt. In den meisten vorkommenden Fällen wird von vornherein die Existenz und Stetigkeit aller Ableitungen der Funktion feststehen, so daß wir für $n$ jede beliebige Zahl wählen können.

Die Annäherungsformel, die ich sogleich ableiten werde, ist in den ersten Zeiten der Differential- und Integralrechnung von *Taylor*, einem Schüler Newtons, entdeckt worden und trägt den Namen Taylorsche Formel[1]).

[1]) Häufig bezeichnet man einen Spezialfall von ihr ohne jeden sachlichen oder historischen Grund als Formel von *MacLaurin*, ein Brauch, dem wir uns nicht anschließen.

**1. Die Taylorsche Formel für ganze rationale Funktionen.**

Wir betrachten zunächst, um uns zu orientieren, den Fall, daß $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ selbst eine ganze rationale Funktion $n$-ten Grades ist. Dann können wir die Koeffizienten dieser Funktion in einfacher Weise durch die Ableitungen von $f(x)$ an der Stelle $x = 0$ ausdrücken. Differenzieren wir nämlich unsere Gleichung einmal, zweimal usw. nach $x$ und setzen dann $x = 0$ ein, so ergibt sich sofort für die Koeffizienten die Darstellung

$$a_0 = f(0), \quad a_1 = f'(0), \quad a_2 = \frac{1}{2!} f''(0), \quad \ldots, \quad a_n = \frac{1}{n!} f^{(n)}(0).$$

Jede ganze rationale Funktion $n$-ten Grades $f(x)$ läßt sich also in der Form

$$f(x) = f(0) + x f'(0) + \frac{x^2}{2!} f''(0) + \frac{x^3}{3!} f'''(0) + \cdots + \frac{x^n}{n!} f^{(n)}(0)$$

schreiben, eine Formel, welche lediglich besagt, daß und wie die Koeffizienten $a_\nu$ sich durch die Ableitungen am Nullpunkt ausdrücken lassen.

Wir können diese „*Taylorsche Darstellung*“ *der ganzen rationalen Funktionen* noch ein wenig verallgemeinern, wenn wir $x$ durch $\xi = x + h$ ersetzen und nunmehr die Funktion $f(\xi) = f(x + h) = g(h)$ als Funktion der Größe $h$ auffassen, indem wir uns für den Moment $x$ als feste Zahl vorstellen und $h$ als die unabhängige Veränderliche betrachten. Es folgt dann

$$g'(h) = f'(\xi), \quad \ldots, \quad g^{(n)}(h) = f^{(n)}(\xi),$$

also, wenn wir $h = 0$ setzen,

$$g'(0) = f'(x), \quad \ldots, \quad g^{(n)}(0) = f^{(n)}(x).$$

Wenden wir unsere Taylorsche Darstellungsformel auf die Funktion $f(x + h) = g(h)$ an, die ja ebenfalls eine ganze rationale Funktion $n$-ten Grades in $h$ ist, so erhalten wir sofort die Taylorsche Darstellung

$$f(\xi) = f(x+h) = f(x) + h f'(x) + \frac{h^2}{2!} f''(x) + \frac{h^3}{3!} f'''(x) + \cdots + \frac{h^n}{n!} f^{(n)}(x).$$

**2. Die Taylorsche Formel für eine beliebige Funktion.**

Die gewonnene Formel gibt uns den Fingerzeig, auch für eine beliebige, nicht notwendig ganz rationale Funktion $f(x)$ eine ähnliche Darstellung aufzusuchen, die aber nunmehr, im Gegensatz zu der Darstellung ganzer rationaler Funktionen, nur noch zu einem angenäherten Ausdruck der Funktion durch eine ganze rationale führen kann.

Wir wollen die Funktionswerte von $f$ an der Stelle $x$ und an der Stelle $\xi = x + h$ miteinander vergleichen, wo also $h = \xi - x$ gesetzt ist. Jetzt wird, wenn $n$ irgendeine natürliche Zahl ist, der Ausdruck

$$f(x) + (\xi - x) f'(x) + \cdots + \frac{(\xi - x)^n}{n!} f^{(n)}(x)$$

im allgemeinen nicht mehr eine exakte Darstellung des Funktionswertes $f(\xi)$ sein. Wir werden also setzen müssen

$$f(\xi) = f(x) + (\xi - x) f'(x) + \frac{(\xi - x)^2}{2!} f''(x) + \cdots + \frac{(\xi - x)^n}{n!} f^{(n)}(x) + R_n,$$

wobei der Ausdruck $R_n$ den *Rest* bei der Ersetzung von $f(\xi)$ durch den Ausdruck $f(x) + f'(x)\,(\xi - x) + \cdots$ bedeutet. Diese Gleichung ist zunächst nichts anderes als eine bloß formale Definition des Ausdruckes $R_n$. Ihre Bedeutung besteht aber darin, daß wir für diesen Rest $R_n$ sehr leicht eine übersichtliche und handliche Darstellung finden können. Zu diesem Zwecke denken wir uns die Größe $\xi$ fest und die Größe $x$ als unabhängige Veränderliche. Der Rest wird dann eine Funktion $R_n(x)$. Diese Funktion verschwindet, wie sofort aus unserer Definitionsgleichung hervorgeht, für $x = \xi$:

$$R_n(\xi) = 0.$$

Weiter erhalten wir durch Differentiation

$$R_n'(x) = -\frac{(\xi - x)^n}{n!} f^{(n+1)}(x).$$

Denn differenzieren wir unsere Definitionsgleichung nach $x$, so entsteht links Null, weil $f(\xi)$ nicht mehr von $x$ abhängt und also als Funktion von $x$ angesehen konstant ist; rechts differenzieren wir jeden Summanden außer $R_n(x)$ nach der Produktregel und erkennen, daß dabei alle Ausdrücke sich gegenseitig zerstören bis auf den oben mit dem Minuszeichen hingeschriebenen letzten.

Nunmehr besagt der Fundamentalsatz der Integralrechnung:

$$R_n(x) = R_n(x) - R_n(\xi) = \int_\xi^x R_n'(t)\,dt = -\int_x^\xi R_n'(t)\,dt;$$

also erhalten wir für den Rest die Darstellung

$$R_n(x) = \int_x^{x+h} \frac{(x + h - t)^n}{n!} f^{(n+1)}(t)\,dt,$$

welche, wenn wir eine neue Integrationsveränderliche $\tau$ durch die Gleichung $\tau = t - x$ einführen, in

$$R_n = \frac{1}{n!} \int_0^h (h - \tau)^n f^{(n+1)}(x + \tau)\,d\tau$$

übergeht.

Ich fasse unser Ergebnis noch einmal zusammen: *Wenn die Funktion $f(x)$ in dem betrachteten Intervalle stetige Ableitungen bis zur $(n+1)$-ten Ordnung besitzt, so gilt*

$$f(x + h) = f(x) + h f'(x) + \frac{h^2}{2!} f''(x) + \frac{h^3}{3!} f'''(x) + \cdots + \frac{h^n}{n!} f^{(n)}(x) + R_n$$

*oder damit gleichbedeutend für* $h = \xi - x$

$$f(\xi) = f(x) + (\xi - x) f'(x) + \frac{(\xi - x)^2}{2!} f''(x) + \cdots + \frac{(\xi - x)^n}{n!} f^{(n)}(x) + R_n,$$

*wobei der Rest* $R_n$ *durch die Formel*

$$R_n = \frac{1}{n!} \int_0^h (h - \tau)^n f^{(n+1)}(x + \tau)\, d\tau$$

*dargestellt wird.*

Setzen wir speziell $x = 0$ und schreiben dann statt $h$ wieder $x$, so erhalten wir die Formel

$$f(x) = f(0) + \frac{x}{1!} f'(0) + \frac{x^2}{2!} f''(0) + \cdots + \frac{x^n}{n!} f^{(n)}(0) + R_n$$

mit dem Rest

$$R_n = \frac{1}{n!} \int_0^x (x - \tau)^n f^{(n+1)}(\tau)\, d\tau.$$

Unsere Formeln heißen die *Taylorschen Formeln.* Sie geben einen Ausdruck für die Funktion $f(x + h)$ bzw. für die Funktion $f(x)$ durch ein Polynom $n$-ten Grades in $h$ bzw. in $x$, das sogenannte *$n$-te Näherungs-* oder *Approximationspolynom,* und ein Restglied. Das Näherungspolynom ist dadurch charakterisiert, daß es selbst sowie seine $n$ ersten Ableitungen für $h = 0$ bzw. $x = 0$ mit der Funktion bzw. ihren ersten $n$ Ableitungen übereinstimmt. Zum Unterschiede von der Taylorschen Darstellung ganzer rationaler Funktionen[1]) sind hier das Restglied und seine Darstellung wesentlich. Die Bedeutung der Formeln liegt darin, daß dieses Restglied, wenn es auch eine kompliziertere Gestalt als die übrigen Glieder der Formel hat, doch ein handliches Instrument zu einer Abschätzung der Genauigkeit bietet, mit welcher die Summe der ersten $n + 1$ Glieder:

$$f(0) + \frac{x}{1!} f'(0) + \frac{x^2}{2!} f''(0) + \cdots + \frac{x^n}{n!} f^{(n)}(0),$$

die Funktion $f(x)$ darstellt.

### 3. Abschätzung des Restgliedes.

Ob die ersten $n + 1$ Glieder der Taylorschen Formel tatsächlich eine hinreichend gute Annäherung an die Funktion geben, wird natürlich davon abhängen, ob das Restglied hinreichend klein bleibt, und die Aufmerksamkeit wird sich also auf die Abschätzung dieses Restgliedes zu konzentrieren haben. Für eine solche Abschätzung bietet sich als naturgemäßes Hilfsmittel der Mittelwertsatz der Integralrechnung (zweites Kapitel, § 7).

---

[1]) Bei deren Darstellung tritt nämlich überhaupt kein Restglied auf.

Wenden wir diesen Mittelwertsatz in der Form

$$\int_0^h p(\tau)\,\varphi(\tau)\,d\tau = \varphi(\vartheta h)\int_0^h p(\tau)\,d\tau$$

an — es ist hierbei $p(\tau)$ als eine in dem Integrationsintervall stetige nicht negative, $\varphi(\tau)$ lediglich als eine dort stetige Funktion vorausgesetzt, und $\vartheta$ bedeutet einen Wert aus dem Intervall $0 < \vartheta < 1$ —, einmal, indem wir $p(\tau) = (h-\tau)^n$, und zweitens, indem wir $p(\tau) = 1$ setzen, so erhalten wir für das Restglied das eine Mal den Ausdruck

$$R_n = \frac{h^{n+1}}{(n+1)!} f^{(n+1)}(x+\vartheta h)$$

und das andere Mal die für uns weniger wichtige und hier nur der Vollständigkeit wegen erwähnte Darstellung

$$R_n = \frac{h^{n+1}}{n!}(1-\vartheta)^n f^{(n+1)}(x+\vartheta h).$$

Es bedeutet in diesen Formeln $\vartheta$ einen gewissen nicht näher festzustellenden Wert in dem Intervall $0 < \vartheta < 1$; dieser Wert ist natürlich im allgemeinen in den beiden Formen des Restgliedes verschieden und zudem von $n$, $x$ und $h$ abhängig. Die erste Form des Restgliedes rührt von *Lagrange*, die zweite von *Cauchy* her; sie werden entsprechend benannt [1]).

Das Hauptinteresse wird sich einmal darauf richten, ob mit wachsendem $n$ der Rest $R_n$ gegen 0 strebt; wenn dies der Fall ist, so wird die Funktion $f(x+h)$ mit um so größerer Genauigkeit durch die entsprechende ganze rationale Funktion von $h$ dargestellt werden, je größer $n$ ist. Wir sagen dann, daß wir für die Funktion eine *Entwicklung in eine unendliche Taylorsche Reihe*

$$f(x+h) = f(x) + \frac{h}{1!}f'(x) + \frac{h^2}{2!}f''(x) + \frac{h^3}{3!}f'''(x) + \cdots$$

---

[1]) Man kann übrigens diese — wie auch noch andere — Ausdrücke für den Rest auch ohne weiteres aus dem Mittelwertsatz der Differentialrechnung bzw. aus dem verallgemeinerten Mittelwertsatz (S. 165) erhalten. Man hat diesen auf die Funktion $R_n(x) = R_n(x) - R_n(\xi)$ bzw. auf die Funktionen $R_n(x)$ und $(x-\xi)^{n+1}$ anzuwenden, wobei $\xi$ als fester Wert anzunehmen und die Formel

$$R_n'(x) = -\frac{(\xi - x)^n}{n!} f^{(n+1)}(x)$$

zu benutzen ist. Bei dieser Ableitung der Restformeln tritt der Charakter der Taylorschen Darstellung als Verallgemeinerung des Mittelwertsatzes mehr hervor; man hat überdies den für manche theoretischen Zwecke wesentlichen Vorteil, daß man nur die Existenz, nicht aber die Stetigkeit der $(n+1)$-ten Ableitung vorauszusetzen braucht. Dagegen verzichtet man hierbei auf die exakte Darstellung des Restes durch eine Integralformel.

bzw. speziell, wenn wir wieder erst $x = 0$ setzen und dann $x$ statt $h$ schreiben,

$$f(x) = f(0) + \frac{x}{1!} f'(0) + \frac{x^2}{2!} f''(0) + \frac{x^3}{3!} f'''(0) + \cdots$$

erhalten haben. Wir werden im nächsten Paragraphen Beispiele hierfür kennen lernen.

Zunächst will ich jedoch den zweiten wesentlichen Gesichtspunkt bezeichnen, welcher sich aus der Betrachtung der Taylorschen Reihe ergibt. Denken wir uns in der ersten Formel die Größe $h$ kleiner und kleiner werdend und gegen 0 strebend, so werden in der Ausdrucksweise des dritten Kapitels, § 9, die einzelnen Glieder der Reihe von verschiedenen Größenordnungen klein werden; wir nennen dementsprechend den Ausdruck $f(x)$ das Glied nullter Ordnung in der Taylorschen Formel, den Ausdruck $h f'(x)$ das Glied erster Ordnung, den Ausdruck $\frac{h^2}{2!} f''(x)$ das Glied zweiter Ordnung usw. Wir sehen aus der Form unseres Restgliedes, *daß wir bei einer Entwicklung bis zu Gliedern $n$-ter Ordnung einen Fehler begehen, der von der $(n+1)$-ten Ordnung mit $h$ klein wird.* Auf dieser Tatsache beruhen viele wichtige Anwendungen. Sie zeigt uns, daß man durch das Approximationspolynom eine um so bessere Darstellung der Funktion $f(x+h)$ erhält, je näher die Stelle $x+h$ an der Stelle $x$ liegt, und daß man diese Annäherung in der unmittelbaren Umgebung der Stelle $x$ gegebenenfalls durch Vergrößerung von $n$ verbessern kann.

## § 3. Anwendungen. Entwicklung der elementaren Funktionen.

Wir benutzen die allgemeinen Resultate des vorigen Paragraphen dazu, um die elementaren Funktionen durch ganze rationale Funktionen zu approximieren bzw. in Taylorsche Reihen zu entwickeln. Dabei will ich mich allerdings auf diejenigen Funktionen beschränken, bei denen sich die Koeffizienten der Reihenentwicklungen nach einfachen Bildungsgesetzen ergeben. Auf die Reihenentwicklungen einiger anderer Funktionen werde ich erst im achten Kapitel eingehen.

### 1. Die Exponentialfunktion.

Das einfachste Beispiel bietet die Exponentialfunktion $f(x) = e^x$. Hier sind alle Ableitungen mit $f(x)$ identisch, besitzen also für $x = 0$ den Wert 1, und wir erhalten daher unter Benutzung der Lagrangeschen Form des Restgliedes für die Exponentialfunktion gemäß § 2 sofort die Formel

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!} e^{\vartheta x}.$$

Lassen wir $n$ über alle Grenzen wachsen, so wird das Restglied gegen 0 streben, was auch immer der fest gewählte Wert $x$ sei. Denn zunächst ist $|e^{\vartheta x}| \leqq e^{|x|}$. Sodann wählen wir eine feste ganze Zahl $m$ größer als $2\,|x|$. Dann wird für $n \geqq m$

$$\frac{|x|}{n} < \frac{1}{2}; \quad \left|\frac{x^{n+1}}{(n+1)!}\right| = \frac{|x^m|}{m!} \cdot \frac{|x|}{m+1} \cdots \frac{|x|}{n+1} \leqq \frac{|x^m|}{m!}\,\frac{1}{2^{n+1-m}} \leqq \frac{|2x|^m}{m!} \cdot \frac{1}{2^n},$$

also

$$|R_n| \leqq \frac{|2x|^m}{m!} \cdot e^{|x|} \cdot \frac{1}{2^n},$$

und da die beiden ersten rechts stehenden Faktoren von $n$ unabhängige feste Zahlen sind, während die Zahl $\frac{1}{2^n}$ bei wachsendem $n$ gegen 0 strebt, so folgt unmittelbar unsere Behauptung. Wenn wir die Zahl $x$ nicht als fest ansehen, sondern frei im Intervall $-a \leqq x \leqq a$ variieren lassen, wo $a$ eine bestimmt gewählte positive Zahl ist, so folgt aus unserer Betrachtung, sobald wir dann nur $m > 2a$ wählen,

$$|R_n| < \frac{(2a)^m}{m!} e^a \cdot \frac{1}{2^n}.$$

Wir haben damit für den Rest eine von $x$ unabhängige, nur von $a$ abhängige Schranke angegeben, welche für $n \to \infty$ gegen Null strebt. Wir können also für die Funktion $e^x$ sofort die Entwicklung in eine unendliche Reihe

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{\nu=0}^{\infty} \frac{x^\nu}{\nu!}$$

hinschreiben[1]) und bemerken, daß diese Reihenentwicklung für alle Werte von $x$ gilt. Damit ist aufs neue bestätigt, daß die im ersten Kapitel betrachtete Größe $e$ (vgl. S. 32) mit der Basis der natürlichen Logarithmen übereinstimmt (vgl. drittes Kapitel, § 6). Für numerische Zwecke allerdings müssen wir uns der abbrechenden Taylorschen Formel mit dem Restglied bedienen; z. B. liefert sie für $x = 1$

$$e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + \frac{e^\vartheta}{(n+1)!}.$$

Wollen wir $e$ mit einem Fehler von höchstens $^1|_{10\,000}$ berechnen, so brauchen wir z. B. nur $n$ so groß zu wählen, daß das Restglied sicher kleiner als $^1/_{10000}$ wird; und da dieses Restglied sicher kleiner als $\frac{3}{(n+1)!}$ ist[2]), so genügt es, $n = 7$ zu wählen, weil $8! > 30000$ ist. Wir finden so den

---

[1]) Das Summenzeichen $\sum\limits_{\nu=0}^{\infty}$ ist auch hier und im folgenden eine Abkürzung der Vorschrift: Man setze in dem unter dem Summenzeichen stehenden Ausdruck für $\nu$ alle Werte 0, 1, 2, . . . ein und addiere dann.

[2]) Wir haben hierbei von der Tatsache Gebrauch gemacht, daß $e < 3$ ist. Diese folgt sofort (vgl. auch S. 32) aus unserer Reihe für $e$; denn es ist sicher $\frac{1}{n!} \leqq \frac{1}{2^{n-1}}$ und daher $e < 1 + 1 + \frac{1}{2} + \frac{1}{4} + \cdots = 1 + \frac{1}{1 - \frac{1}{2}} = 3.$

angenäherten Wert

$$e = 2{,}71822$$

mit einem Fehler, der sicher kleiner als 0,0001 ist. Die Abrundungsfehler haben wir dabei nicht berücksichtigt.

## 2. $\sin x$, $\cos x$, $\mathfrak{Sin}\, x$, $\mathfrak{Cof}\, x$.

Für die Funktionen $\sin x$, $\cos x$, $\mathfrak{Sin}\, x$ und $\mathfrak{Cof}\, x$ erhalten wir die folgenden Formeln[1]:

$$\begin{array}{lllll}
f(x) & = \quad \sin x & \cos x & \mathfrak{Sin}\, x & \mathfrak{Cof}\, x, \\
f'(x) & = \quad \cos x & -\sin x & \mathfrak{Cof}\, x & \mathfrak{Sin}\, x, \\
f''(x) & = -\sin x & -\cos x & \mathfrak{Sin}\, x & \mathfrak{Cof}\, x, \\
f'''(x) & = -\cos x & \sin x & \mathfrak{Cof}\, x & \mathfrak{Sin}\, x, \\
f^{(IV)}(x) & = \quad \sin x & \cos x & \mathfrak{Sin}\, x & \mathfrak{Cof}\, x, \\
\dots & \dots & \dots & \dots & \dots
\end{array}$$

Bei den Approximationspolynomen für $\sin x$ und $\mathfrak{Sin}\, x$ werden also die Koeffizienten der geraden Potenzen von $x$ verschwinden, bei den Approximationspolynomen für $\cos x$ und $\mathfrak{Cof}\, x$ die der ungeraden, so daß im ersten Falle das $(2n+1)$-te und das $(2n+2)$-te Polynom identisch sind, im zweiten Falle das $2n$-te und das $(2n+1)$-te. Denken wir uns die jeweils höchsten dieser Polynome benutzt, so erhalten wir sofort mit dem Restglied von Lagrange

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + (-1)^{n+1} \frac{x^{2n+3}}{(2n+3)!} \cos(\vartheta x),$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - + \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + (-1)^{n+1} \frac{x^{2n+2}}{(2n+2)!} \cos(\vartheta x),$$

$$\mathfrak{Sin}\, x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{x^{2n+1}}{(2n+1)!} + \frac{x^{2n+3}}{(2n+3)!} \mathfrak{Cof}(\vartheta x),$$

$$\mathfrak{Cof}\, x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{x^{2n}}{(2n)!} + \frac{x^{2n+2}}{(2n+2)!} \mathfrak{Cof}(\vartheta x),$$

wo natürlich $\vartheta$ im allgemeinen in jeder der vier Formeln eine andere Zahl zwischen 0 und 1 bedeutet, die außerdem noch von $n$ und $x$ abhängt. Auch in diesen Formeln kann man für jeden Wert von $x$ die Annäherung beliebig genau machen, da der Rest mit wachsendem $n$ gegen 0 strebt. Wir erhalten so die vier Reihenentwicklungen

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - + \cdots = \sum_{\nu=0}^{\infty} (-1)^\nu \frac{x^{2\nu+1}}{(2\nu+1)!},$$

[1] Ist $f(x) = \sin x$ oder $f(x) = \cos x$, so läßt sich übrigens stets die $n$-te Ableitung durch den Ausdruck

$$f^{(n)}(x) = f\left(x + n\frac{\pi}{2}\right)$$

einheitlich darstellen.

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - + \cdots = \sum_{\nu=0}^{\infty} (-1)^\nu \frac{x^{2\nu}}{(2\nu)!},$$

$$\mathfrak{Sin}\, x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots = \sum_{\nu=0}^{\infty} \frac{x^{2\nu+1}}{(2\nu+1)!},$$

$$\mathfrak{Cof}\, x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots = \sum_{\nu=0}^{\infty} \frac{x^{2\nu}}{(2\nu)!},$$

von denen übrigens die beiden letzten sich formal auch aus der Reihenentwicklung von $e^x$ gemäß den Definitionsformeln der Hyperbelfunktionen ergeben.

### 3. Die binomische Reihe.

Während ich darauf verzichten kann, die Funktionen $\log(1+x)$ und $\operatorname{arc\,tg} x$, die wir schon im § 1 direkt behandelt haben, nochmals mit Hilfe der Taylorschen Formel zu entwickeln, muß ich noch auf die Verallgemeinerung des binomischen Satzes für beliebige Exponenten eingehen, die eine der folgenreichsten mathematischen Entdeckungen von Newton war und eines der wichtigsten Beispiele für die Taylorsche Reihenentwicklung überhaupt darstellt. Es handelt sich darum, die Funktion

$$f(x) = (1+x)^\alpha$$

bei beliebigem positiven oder negativen, rationalen oder irrationalen $\alpha$ für $x > -1$ nach der Taylorschen Formel zu entwickeln. Daß wir nicht die Funktion $x^\alpha$, sondern gerade die obige Funktion gewählt haben, hat natürlich seinen Grund in der Tatsache, daß für $x^\alpha$ an der Stelle $x = 0$ nicht mehr sämtliche Ableitungen stetig sind, abgesehen von dem trivialen Falle eines nicht negativen ganzen $\alpha$. Wir berechnen zunächst die Ableitungen von $f(x)$ und erhalten

$$f'(x) = \alpha(1+x)^{\alpha-1}, \quad f''(x) = \alpha(\alpha-1)(1+x)^{\alpha-2}, \quad \ldots,$$
$$f^{(\nu)}(x) = \alpha(\alpha-1)\cdots(\alpha-\nu+1)(1+x)^{\alpha-\nu}.$$

Speziell für $x = 0$ ergibt sich

$$f'(0) = \alpha, \quad f''(0) = \alpha(\alpha-1), \quad \ldots, \quad f^{(\nu)}(0) = \alpha(\alpha-1)\cdots(\alpha-\nu+1).$$

Es ergibt sich also die Taylorsche Formel

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \cdots + \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!} x^n + R_n,$$

wo uns nun noch die Aufgabe der Diskussion des Restes bleibt. Diese Aufgabe ist zwar nicht schwierig, jedoch nicht ganz so einfach wie in den schon behandelten Fällen. Ich will daher auf die Durchführung der Restabschätzung hier verzichten, da ich im übernächsten Kapitel den allgemeinen binomischen Satz auf eine etwas andere und einfachere Art vollständig ableiten werde. Das Ergebnis, welches ich schon hier nenne, ist, daß jedenfalls für $|x| < 1$ das Restglied gegen 0 strebt und also der Ausdruck $(1 + x)^\alpha$ in die unendliche ***binomische Reihe***

$$(1 + x)^\alpha = 1 + \frac{\alpha}{1!} x + \frac{\alpha(\alpha-1)}{2!} x^2 + \cdots = \sum_{\nu=0}^{\infty} \binom{\alpha}{\nu} x^\nu$$

entwickelt werden kann, wobei wir zur Abkürzung die allgemeinen Binomialkoeffizienten $\binom{\alpha}{\nu} = \frac{\alpha(\alpha-1)\cdots(\alpha-\nu+1)}{\nu!}$ (für $\nu > 0$), $\binom{\alpha}{0} = 1$ eingeführt haben.

## § 4. Geometrische Anwendungen.

Die Taylorsche Formel erlaubt uns, das Verhalten einer Funktion $f(x)$ in der Umgebung einer Stelle $x = a$ bzw. das Verhalten einer gegebenen Kurve in der Umgebung eines Punktes genauer zu untersuchen, da sie den Zuwachs der Funktion bei Übergang zu einer Nachbarstelle $x = a + h$ in eine Summe von Größen erster, zweiter, dritter usw. Ordnung auflöst.

### 1. Berührung von Kurven.

Wir machen hiervon Gebrauch, um den Begriff der *Berührung* zweier Kurven zu analysieren.

Wenn zwei Kurven $y = f(x)$ und $y = g(x)$ an einer Stelle, etwa an der Stelle $x = a$, sich nicht nur treffen, sondern noch eine gemeinsame Tangente besitzen, so sagen wir, daß sie einander an dieser Stelle berühren. Die Taylorschen Entwicklungen der Funktionen $f(a + h)$ und $g(a + h)$ stimmen dann in den Gliedern 0-ter und 1-ter Ordnung in $h$ überein. Wenn an der Stelle $x = a$ auch noch die zweiten Ableitungen von $f(x)$ und $g(x)$ miteinander übereinstimmen, so sprechen wir von einer *Berührung zweiter Ordnung*. Die Taylorschen Entwicklungen stimmen alsdann auch noch in den Gliedern zweiter Ordnung überein, und die Differenz $D(x) = f(x) - g(x)$ wird sich, wenn wir die Stetigkeit der Ableitungen bis mindestens zur dritten Ordnung voraussetzen, in der Form

$$D(a + h) = f(a + h) - g(a + h) = \frac{h^3}{3!} D'''(a + \vartheta h) = \frac{h^3}{3!} F(h)$$

darstellen lassen, wobei der Ausdruck $F(h)$ für $h \to 0$ gegen $f'''(a) - g'''(a)$ strebt. Die Differenz $D(a + h)$ wird also mit $h$ von mindestens dritter Ordnung Null.

So können wir weitergehen und den allgemeinen Fall betrachten, daß die Taylorschen Formeln für $f(x)$ und $g(x)$ in den Gliedern bis zur $n$-ten Ordnung übereinstimmen, d. h. daß

$$f(a) = g(a), \quad f'(a) = g'(a), \quad f''(a) = g''(a), \quad \ldots, \quad f^{(n)}(a) = g^{(n)}(a)$$

ist, wobei wir weiter voraussetzen, daß noch die $(n+1)$-ten Ableitungen stetig sein sollen. Wir sagen unter diesen Voraussetzungen, daß die Kurven an dieser Stelle eine *Berührung n-ter Ordnung* besitzen. Die Differenz der beiden Funktionen wird dann die Form haben

$$f(a+h) - g(a+h) = \frac{h^{n+1}}{(n+1)!} F(h),$$

wo $F(h) = D^{(n+1)}(a + \vartheta h)$ wegen $0 < \vartheta < 1$ für $h \to 0$ gegen $f^{(n+1)}(a) - g^{(n+1)}(a)$ strebt. Man erkennt aus diesen Formeln, daß die Differenz $f(x) - g(x)$ im Berührungspunkt von mindestens $(n+1)$-ter Ordnung Null wird.

Die Taylorschen Polynome sind geometrisch einfach dadurch definiert, daß sie diejenigen Parabeln $n$-ter Ordnung darstellen, welche an der betreffenden Stelle mit der zur gegebenen Funktion gehörigen Kurve eine Berührung möglichst hoher Ordnung haben. Man nennt sie darum gelegentlich *Schmiegungsparabeln* oder *oskulierende Parabeln*. Figur 99 gibt uns für das Beispiel $y = e^x$ an der Stelle $x = 0$ die ersten Schmiegungsparabeln.

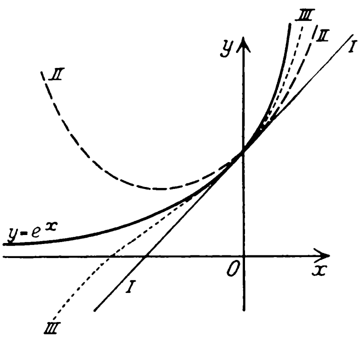


Fig. 99. Schmiegungsparabeln von $e^x$.

Wenn zwei Kurven $y = f(x)$ und $y = g(x)$ sich von $n$-ter Ordnung berühren, so ist durch die Definition nicht ausgeschlossen, daß die Berührung sogar von noch höherer Ordnung ist, d. h. daß auch noch $f^{(n+1)}(a) = g^{(n+1)}(a)$ wird. Ist dies jedoch nicht der Fall, ist also $f^{(n+1)}(a) \neq g^{(n+1)}(a)$, so sprechen wir von einer Berührung genau $n$-ter Ordnung oder sagen, die Ordnung der Berührung[1]) ist genau $n$.

Wir können aus unseren Formeln wie aus unseren Figuren eine bemerkenswerte, gerade von Anfängern häufig nicht beachtete Tatsache ablesen. Wenn die Berührung zweier Kurven von genau gerader Ordnung ist, d. h. wenn eine gerade Anzahl $n$ von Ableitungen der beiden Funktionen an der betreffenden Stelle miteinander übereinstimmt, die $(n+1)$-ten Ableitungen aber nicht mehr, so wird die Differenz $f(a+h) - g(a+h)$ der beiden Funktionen gemäß der obigen Formel für hinreichend kleine positive $h$ und für absolut genommen hinreichend kleine negative $h$ verschiedene Vorzeichen besitzen. Es werden sich die beiden einander berührenden Kurven bei der Berührung durchschneiden. Dieser Fall tritt z. B. ein bei einer Berührung zweiter

[1]) Daß die Ordnung der Berührung zweier Kurven eine echte geometrische Beziehung ist, die durch eine Bewegung des Koordinatensystems nicht beeinflußt wird, ist eine Tatsache, die man leicht den Formeln für die Bewegung des Koordinatensystemes entnehmen kann.

Ordnung, wenn tatsächlich die Ableitungen dritter Ordnung nicht mehr übereinstimmen. Betrachten wir aber den Fall einer Berührung ungerader Ordnung, z. B. den Fall einer gewöhnlichen Berührung erster Ordnung, so wird die Differenz $f(a+h)-g(a+h)$ für absolut genommen hinreichend kleine positive und negative $h$ dasselbe Vorzeichen haben; die beiden einander berührenden Kurven werden also in der Umgebung der Berührungsstelle einander nicht durchschneiden. Die Berührung einer Kurve mit ihrer Tangente ist dafür das einfachste Beispiel; nur an Punkten, wo wir eine Berührung zweiter Ordnung haben, muß die Tangente die Kurve durchsetzen, es sei denn, daß für die Kurve an der betreffenden Stelle auch noch die Ableitung dritter Ordnung verschwindet (oder allgemeiner alle Ableitungen bis zu einer ungeraden Ordnung einschließlich), wie z. B. für die Kurve $y = x^4$ an der Stelle $x = 0$.

### 2. Der Krümmungskreis als Oskulationskreis.

In dieser Auffassung gewinnt der Begriff der Krümmung einer Kurve $y = f(x)$ eine neue anschauliche Bedeutung. Betrachten wir einen bestimmten Kurvenpunkt mit den Koordinaten $x = a$ und $y = b$, so gibt es durch diesen Punkt unendlich viele Kreise, welche die Kurve dort berühren. Die Mittelpunkte dieser Kreise liegen auf der Kurvennormale, und zu jedem Punkt dieser Normalen gehört genau ein solcher berührender Kreis. Man darf erwarten, daß wir durch geeignete Wahl des Kreismittelpunktes eine Berührung *zweiter Ordnung* zwischen Kreis und Kurve erzielen können.

Nun wissen wir in der Tat aus dem fünften Kapitel, daß für den Krümmungskreis im Punkte $x = a$, dessen Gleichung $y = g(x)$ sei, nicht nur $g(a) = f(a)$ und $g'(a) = f'(a)$, sondern außerdem noch $g''(a) = f''(a)$ wird. Der *Krümmungskreis* ist also zugleich der *Schmiegungskreis* oder *oskulierende Kreis* für den betreffenden Kurvenpunkt, d. h. der Kreis, welcher dort mit der Kurve eine Berührung von zweiter Ordnung besitzt. Im Grenzfall eines Wendepunktes oder allgemein einer Stelle, wo die Krümmung Null, der Krümmungsradius unendlich wird, artet der Krümmungskreis in die Tangente aus. Der Krümmungskreis wird im allgemeinen, d. h. wenn die Berührung nicht zufällig an der betreffenden Stelle von höherer als zweiter Ordnung ist, *die Kurve* nicht nur berühren, sondern an der Berührungsstelle noch *durchschneiden*.

### 3. Zur Theorie der Maxima und Minima.

Wie wir früher im dritten Kapitel gesehen haben, liefert eine Stelle $x = a$, für welche $f'(a) = 0$ ist, ein Maximum bzw. ein Minimum für die Funktion $f(x)$, wenn $f''(a)$ negativ bzw. positiv ist. Die letzten Bedingungen sind also hinreichend für das Auftreten eines Maximums oder Minimums. Sie sind aber keineswegs notwendig; denn im Falle

$f''(a) = 0$ bleibt noch jede der drei Möglichkeiten offen, daß ein Maximum oder ein Minimum oder keine der beiden Erscheinungen vorliegt. Beispiele für die drei Erscheinungen bieten die Funktionen $y = -x^4$, $y = x^4$ und $y = x^3$ an der Stelle $x = 0$. Die Taylorsche Formel gibt uns sofort die Möglichkeit einer allgemeinen Formulierung hinreichender Bedingungen. Wir brauchen nur die Funktion $f(a+h)$ nach Potenzen von $h$ zu entwickeln; dann kommt es darauf an, ob das erste nicht verschwindende Glied eine gerade oder eine ungerade Potenz von $h$ enthält. Im ersten Falle haben wir ein Maximum oder Minimum, je nachdem der Koeffizient von $h$ negativ oder positiv ist; im zweiten Falle haben wir eine horizontale Wendetangente und weder ein Maximum noch ein Minimum. Der Leser möge sich diesen Zusammenhang an Hand des Restgliedes selber genauer durchdenken[1]).

# Anhang zum sechsten Kapitel.

## § 1. Beispiel einer Funktion, die sich nicht in eine Taylorsche Reihe entwickeln läßt.

Die Möglichkeit der Taylorschen Darstellung mit Restglied $(n+1)$-ter Ordnung beruhte wesentlich auf der Differenzierbarkeit der Funktion an der betreffenden Stelle. Deshalb ist z. B. die Funktion $\log x$ nicht durch eine Taylorsche Formel nach Potenzen von $x$ darstellbar, ebenso nicht die Funktion $\sqrt[3]{x}$, deren Ableitungen für $x = 0$ unendlich werden.

Damit eine Funktion in eine unendliche Taylorsche Reihe entwickelbar ist, müssen jedenfalls an der betreffenden Stelle alle Ableitungen existieren; aber diese notwendige Bedingung ist keineswegs hinreichend. Eine Funktion, deren sämtliche Ableitungen in einem Intervall vorhanden und stetig sind, braucht sich trotzdem keineswegs in eine Taylorsche Reihe entwickeln zu lassen; d. h. das Restglied $R_n$ der Taylorschen Formel braucht nicht mit wachsendem $n$ gegen 0 zu streben, nicht einmal, wenn das Gebiet, in welchem wir entwickeln wollen, hinreichend klein gewählt wird. Das einfachste Beispiel für diese Erscheinung bietet uns die Funktion $y = f(x) = e^{-\frac{1}{x^2}}$ für $x \neq 0$, $f(0) = 0$, die wir schon im Anhang zum dritten Kapitel betrachtet haben. Die Funktion ist mit ihren sämtlichen Ableitungen in jedem Intervalle stetig, auch für $x = 0$, und wir haben gesehen, daß in diesem Punkte alle Ableitungen ver-

[1]) Im übrigen ist die früher (S. 129) gegebene notwendige und hinreichende Bedingung allgemeiner und für Anwendungen bequemer: Notwendig und hinreichend für das Vorliegen eines Maximums oder Minimums ist, daß beim Durchgang durch die betreffende Stelle die erste Ableitung $f'(x)$ ihr Vorzeichen wechselt.

schwinden, d. h. daß $f^{(n)}(0) = 0$ ist. In der Taylorschen Formel verschwinden also alle Koeffizienten des Approximationspolynomes, wie wir auch $n$ wählen; mit anderen Worten: Das Restglied ist und bleibt gleich der Funktion selbst, strebt also, außer für $x = 0$, nicht gegen 0, da die Funktion für jedes $x \neq 0$ positiv ist.

## § 2. Beweis der Irrationalität von $e$.

Aus der Formel $e = 2 + \frac{1}{2!} + \cdots + \frac{1}{n!} + \frac{1}{(n+1)!}e^{\vartheta}$ ergibt sich sofort die Tatsache, daß die Zahl $e$ irrational ist. Wäre das Gegenteil richtig, nämlich $e = \frac{p}{q}$, wo $p$ und $q$ ganze Zahlen bedeuten, so könnten wir gewiß $n$ größer als $q$ wählen. Dann muß $n!e = n!\frac{p}{q}$ eine ganze Zahl sein. Andererseits ist $n!e = 2n! + \frac{n!}{2!} + \cdots + \frac{n!}{n!} + \frac{1}{n+1}e^{\vartheta}$, und da $e^{\vartheta} < e < 3$ ist, wird $0 < \frac{e^{\vartheta}}{n+1} < 1$. Es müßte also die ganze Zahl $n!e$ gleich der ganzen Zahl $2n! + \frac{n!}{2!} + \cdots + 1$ plus einem nicht verschwindenden echten Bruche sein, was nicht möglich ist.

## § 3. Nullstellen, Unendlichkeitsstellen von Funktionen und sogenannte unbestimmte Ausdrücke.

Die Taylorsche Entwicklung einer Funktion in der Umgebung einer Stelle $x = a$ gibt uns Veranlassung, das Verhalten der Funktion in der Umgebung dieser Stelle durch folgende Definition zu kennzeichnen. Wir sagen: eine Funktion $f(x)$ hat für $x = a$ eine *genau n-fache Nullstelle*, oder: sie verschwindet dort *genau von der Ordnung n*, wenn zugleich $f(a) = 0, f'(a) = 0, f''(a) = 0, \ldots, f^{(n-1)}(a) = 0$ und $f^{(n)}(a) \neq 0$ ist. Dabei setzen wir voraus, daß die Funktion in der Umgebung dieser Stelle stetige Ableitungen bis mindestens zur $n$-ten Ordnung besitzt. Mit unserer Definition wollen wir andeuten, daß die Taylorsche Entwicklung der Funktion in der Umgebung dieser Stelle sich in die Form setzen läßt

$$f(a+h) = \frac{h^n}{n!}F(h),$$

wobei der Faktor $F(h)$ für $h \to 0$ gegen einen von Null verschiedenen Grenzwert, nämlich den Wert $f^{(n)}(a)$, strebt.

Ist eine Funktion $\varphi(x)$ in allen Punkten der Umgebung einer Stelle $x = a$ definiert, nicht aber notwendig für $x = a$ selbst, und zwar durch einen Ausdruck der Form

$$\varphi(x) = \frac{f(x)}{g(x)},$$

wo an der Stelle $x = a$ der Zähler nicht verschwindet, während der Nenner eine $\nu$-fache Nullstelle besitzt, so sagen wir: die Funktion

$\varphi(x)$ wird an der Stelle $x = a$ *von der $\nu$-ten Ordnung unendlich.* Besitzt an der Stelle $x = a$ auch der Zähler eine $\mu$-fache Nullstelle und ist $\mu > \nu$, so schreiben wir der Funktion dort eine $(\mu - \nu)$-fache Nullstelle zu, während wir ihr im Falle $\mu < \nu$ eine $(\nu - \mu)$-fache Unendlichkeitsstelle zusprechen.

Alle diese Definitionen stehen im übrigen in Übereinstimmung mit den Festsetzungen, die wir schon früher im dritten Kapitel, § 9, hinsichtlich des Verhaltens von Funktionen getroffen haben. Um diesen Zusammenhang zu präzisieren, stellen wir uns Zähler und Nenner gemäß der Taylorschen Formel mit dem Restglied von Lagrange entwickelt dar; die Funktion erhält dann die Gestalt

$$\varphi(a+h) = \frac{f(a+h)}{g(a+h)} = \frac{\nu!}{\mu!}\,\frac{h^{\mu} f^{(\mu)}(a+\vartheta h)}{h^{\nu} g^{(\nu)}(a+\vartheta_1 h)},$$

wobei $\vartheta$ und $\vartheta_1$ zwei Zahlen zwischen 0 und 1 sind und die Faktoren von $\frac{h^{\mu}}{\mu!}$ bzw. $\frac{h^{\nu}}{\nu!}$ in der Grenze $h \to 0$ nicht verschwinden, indem sie in die von Null verschiedenen Ausdrücke $f^{(\mu)}(a)$ und $g^{(\nu)}(a)$ übergehen. Es wird dann für $\mu > \nu$

$$\lim_{h\to 0} \varphi(a+h) = \lim_{h\to 0} \frac{\nu!}{\mu!} h^{\mu-\nu} \frac{f^{(\mu)}(a)}{g^{(\nu)}(a)} = 0.$$

Der Ausdruck $\varphi(x)$ wird also von der Ordnung $\mu - \nu$ Null. Für $\nu > \mu$ erkennen wir, daß der Ausdruck $\varphi(a+h)$ für $h \to 0$ von der Ordnung $\nu - \mu$ unendlich wird. Im Falle $\mu = \nu$ erhalten wir die Gleichung

$$\lim_{h\to 0} \varphi(a+h) = \frac{f^{(\mu)}(a)}{g^{(\mu)}(a)}.$$

Den Inhalt der letzten Gleichungen können wir in folgender Form aussprechen: Wenn Zähler und Nenner eines Bruches $\varphi(x) = \frac{f(x)}{g(x)}$ für $x = a$ verschwinden, so bestimmt man den Grenzwert des Bruches für $x \to a$ einfach, indem man Zähler und Nenner gleich oft so lange differenziert, bis mindestens einer der Differentialquotienten von Null verschieden ist. Tritt dies gleichzeitig für Zähler und Nenner ein, so ist der gesuchte Grenzwert gleich dem Quotienten dieser beiden Ableitungen. Verschwindet als erster der Differentialquotient des Nenners nicht, so strebt der Bruch gegen 0; verschwindet als erster der Differentialquotient des Zählers nicht, so wächst der absolute Betrag des Bruches über alle Grenzen.

Wir haben damit eine Regel zur Festlegung der sog. unbestimmten Ausdrücke $\frac{0}{0}$ vor uns; ein Gegenstand, der in manchen Darstellungen der Differential- und Integralrechnung mit übertriebener Breite behandelt wird. In Wirklichkeit handelt es sich um nichts als die sehr einfache Bestimmung eines *Grenzwertes* eines Quotienten, bei dem Zähler und Nenner für sich gegen 0 streben. Die in der Literatur übliche Bezeichnung „unbestimmte Ausdrücke" ist eine irreführende unpräzise Ausdrucksweise.

Wir können übrigens unserer Betrachtung noch eine etwas andere Wendung geben, indem wir uns anstatt auf die Taylorsche Formel auf den verallgemeinerten Mittelwertsatz stützen (vgl. S. 108)[1]. Nach diesem gilt, wenn $g'(x) \neq 0$ ist, allgemein

$$\frac{f(a+h)-f(a)}{g(a+h)-g(a)} = \frac{f'(a+\vartheta h)}{g'(a+\vartheta h)}$$

mit demselben $\vartheta$ in Zähler und Nenner, und also speziell für $f(a) = g(a) = 0$

$$\frac{f(a+h)}{g(a+h)} = \frac{f'(a+\vartheta h)}{g'(a+\vartheta h)}.$$

Dabei ist $\vartheta$ ein Wert des Intervalles $0 < \vartheta < 1$, und es wird somit, wenn wir $k = \vartheta h$ setzen,

$$\lim_{h\to 0} \frac{f(a+h)}{g(a+h)} = \lim_{k\to 0} \frac{f'(a+k)}{g'(a+k)},$$

vorausgesetzt, daß der Grenzwert rechts existiert. Ist auch $f'(a) = g'(a) = 0$, so kann man in derselben Weise weiter schließen, bis man zu einem ersten Index $\mu$ kommt, für den nicht $f^{(\mu)}(a) = g^{(\mu)}(a) = 0$ ist. Es gilt dann stets

$$\lim_{h\to 0} \frac{f(a+h)}{g(a+h)} = \lim_{l\to 0} \frac{f^{(\mu)}(a+l)}{g^{(\mu)}(a+l)},$$

wobei wir auch den Fall einbeziehen, daß auf beiden Seiten der „Grenzwert Unendlich" steht.

Als Beispiele betrachten wir

$$\frac{\sin x}{x}, \quad \frac{1-\cos x}{x}, \quad \frac{e^{2x}-1}{\log(1+x)}, \quad \frac{x^2 \operatorname{tg} x}{\sqrt{1-x^2}-1}$$

für $x \to 0$. Es wird

$$\lim_{x\to 0} \frac{\sin x}{x} = \frac{\cos 0}{1} = 1, \qquad \lim_{x\to 0} \frac{1-\cos x}{x} = \frac{\sin 0}{1} = 0,$$

$$\lim_{x\to 0} \frac{e^{2x}-1}{\log(1+x)} = \lim_{x\to 0} \frac{2e^{2x}}{\frac{1}{1+x}} = 2,$$

$$\lim_{x\to 0} \frac{x^2 \operatorname{tg} x}{\sqrt{1-x^2}-1} = \lim_{x\to 0} \frac{2x \operatorname{tg} x + \frac{x^2}{\cos^2 x}}{-\frac{x}{\sqrt{1-x^2}}} = -\lim_{x\to 0}\left(2 \operatorname{tg} x + \frac{x}{\cos^2 x}\right)\sqrt{1-x^2} = 0.$$

[1] Diese Wendung unserer Regel hat den Vorteil, daß bei ihr von der Existenz der Ableitung im Punkte $x = a$ selbst kein Gebrauch gemacht wird; ferner umfaßt man so auch den Fall mit, daß $\varphi(x)$ nur für $x \geqq a$ definiert ist und man also nur von einer Seite her den Grenzübergang $x \to a$ oder $h \to 0$ zu machen hat.

Ich bemerke ferner, daß auch andere sog. unbestimmte Formen sich genau auf unseren Fall zurückführen lassen. Z. B. erscheint der Grenzwert von $\frac{1}{\sin x} - \frac{1}{x}$ für $x \to 0$ als Grenzwert der Differenz zweier Ausdrücke, die beide unendlich werden, als eine „unbestimmte" Form $\infty - \infty$. Durch die Umformung

$$\frac{1}{\sin x} - \frac{1}{x} = \frac{x - \sin x}{x \sin x}$$

gelangen wir sofort zu einem Ausdruck, dessen Grenzwert für $x \to 0$ wir durch unsere Regel als

$$\lim_{x \to 0} \frac{1 - \cos x}{x \cos x + \sin x} = \lim_{x \to 0} \frac{\sin x}{2 \cos x - x \sin x} = 0$$

bestimmen.

## § 4. Das Problem der Interpolation und sein Zusammenhang mit der Taylorschen Formel.

Die Taylorsche Formel läßt sich als Grenzfall einer allgemeinen Interpolationsformel auffassen, welche uns nicht nur einen genaueren Einblick in das Wesen der Taylorschen Formel geben wird, sondern auch an und für sich eine große theoretische und praktische Bedeutung besitzt. Ich möchte auf diese Dinge hier um so lieber kurz eingehen, als sie in den üblichen Darstellungen gewöhnlich vernachlässigt werden.

### 1. Problemstellung und Vorbemerkungen.

Wir gehen von folgender Aufgabe aus: Ein Polynom, d. h. eine ganze rationale Funktion $\varphi(x)$, von $n$-tem Grade soll so bestimmt werden, daß es an $n+1$ verschiedenen gegebenen Stellen $x_0, x_1, \ldots, x_n$ bzw. die $n+1$ gegebenen Werte $f_0, f_1, \ldots, f_n$ annimmt, so daß also

$$\varphi(x_0) = f_0, \quad \varphi(x_1) = f_1, \ldots, \quad \varphi(x_n) = f_n$$

wird. Dabei ist es bequem, sich die gegebenen Werte $f_i$ als diejenigen Werte vorzustellen, die von einer gegebenen Funktion $f(x)$ in den Punkten $x = x_i$ angenommen werden, d. h. $f_i = f(x_i)$ zu setzen. Wir werden dann das Polynom $\varphi(x)$ oder $\varphi_n(x)$ das *Interpolationspolynom n-ten Grades* der Funktion $f(x)$ für die Stellen $x_0, x_1, \ldots, x_n$ nennen.

Ich bemerke zuerst, daß es höchstens ein einziges solches Polynom $n$-ten Grades geben kann. Hätten wir nämlich in $\varphi(x)$ und $\psi(x)$ zwei verschiedene solche Polynome vor uns, so wäre bei passendem ganzen $m$ mit $0 \leqq m \leqq n$ ihre Differenz

$$D(x) = \varphi(x) - \psi(x) = c_0 x^m + c_1 x^{m-1} + \cdots + c_m \quad (c_0 \neq 0)$$

ein in den Punkten $x_1, \ldots, x_m$ verschwindendes Polynom $m$-ten Grades, also nach einem bekannten Satze der Algebra

$$D(x) = c_0 (x - x_1)(x - x_2) \cdots (x - x_n).$$

Da aber auch $D(x_0) = 0$ ist, also

$$c_0(x_0 - x_1)(x_0 - x_2)\cdots(x_0 - x_n) = 0$$

gilt, so folgt (da die Werte $x_0, x_1, \ldots, x_n$ nach Voraussetzung sämtlich verschieden sind), daß die Konstante $c_0$ verschwindet, entgegen der Annahme. Damit ist die Eindeutigkeit der interpolierenden Funktion gesichert.

Wir können im übrigen die eindeutige Bestimmtheit des Interpolationspolynomes noch nach einer anderen, unmittelbar an den Begriff des Differentialquotienten anschließenden Methode beweisen, welche auf dem allgemeinen *Satz von Rolle* beruht: *Wenn eine in einem Intervall mit stetigen Ableitungen bis zur n-ten Ordnung versehene Funktion $F(x)$ an mindestens $n+1$ verschiedenen Stellen $x_0, x_1, \ldots, x_n$ des Intervalles verschwindet, so gibt es im Innern des Intervalles sicherlich eine Stelle $\xi$, für welche $F^{(n)}(\xi) = 0$ ist.* Der Beweis dieses Hilfssatzes ergibt sich einfach folgendermaßen. Denken wir uns die Zahlen $x_0, x_1, \ldots, x_n$ nach wachsender Größe geordnet. Dann muß nach dem Mittelwertsatz der Differentialrechnung die erste Ableitung $F'(x)$ im Innern eines jeden der $n$ Teilintervalle $(x_i \ldots x_{i+1})$ mindestens einmal verschwinden. Dieselbe Betrachtung für die Funktion $F'(x)$ und die Intervalle zwischen ihren Nullstellen ergibt das Vorhandensein von $n-1$ Stellen, für welche die zweite Ableitung $F''(x)$ verschwindet; indem man so weiter schließt, gelangt man unmittelbar zu dem behaupteten Ergebnis.

Wenden wir diesen Satz auf die Differenz $F(x) = D(x) = \varphi(x) - \psi(x) = d_0 x^n + d_1 x^{n-1} + \cdots + d_n$ an, welche ja nach Voraussetzung an $n+1$ Stellen verschwindet, so folgt das Verschwinden der $n$-ten Ableitung $D^{(n)}(\xi)$ an einer Stelle $\xi$ des Intervalles. Diese $n$-te Ableitung aber ist gerade $n!\,d_0$. Es ist also $d_0 = 0$; d. h. die Differenz ist ein Polynom von höchstens $(n-1)$-tem Grade, das an unseren $n+1$ Stellen verschwindet. In derselben Weise erkennen wir, indem wir auf dieses Polynom den Rolleschen Satz anwenden, daß $d_1 = 0$ ist, und so fortfahrend, daß auch alle anderen Koeffizienten des Polynomes $D(x)$ verschwinden, was die behauptete Eindeutigkeit ausdrückt.

## 2. Konstruktion der Lösung. Die Steigungen einer Funktion. Die Newtonsche Interpolationsformel.

Wir gehen nun an die Aufgabe, ein Polynom $n$-ten Grades $\varphi(x)$ zu bilden, für welches die Gleichungen $\varphi(x_0) = f_0, \ldots, \varphi(x_n) = f_n$ bestehen. Um dieses Polynom schrittweise aufzubauen, gehen wir von der Konstanten $f_0$ aus, einem Polynom „0-ten Grades“ $\varphi_0(x)$, das überall, also auch für $x = x_0$ den Wert $f_0 = A_0$ annimmt. Zu ihm addieren wir ein Polynom ersten Grades, welches für $x = x_0$ verschwindet, also die

Form $A_1(x - x_0)$ besitzt, und bestimmen $A_1$ so, daß die Summe auch noch für $x = x_1$ den richtigen Wert $f_1$ erhält. Das entstehende Polynom ersten Grades nennen wir $\varphi_1(x)$. Weiter addieren wir zu $\varphi_1(x)$ ein Polynom zweiten Grades, welches für $x = x_0$ und $x = x_1$ verschwindet, also die Form $A_2(x - x_0)(x - x_1)$ besitzt, dessen Hinzufügung an diesen beiden Stellen also nichts mehr ändert, bei dem wir jedoch den Faktor $A_2$ so bestimmen, daß das entstehende Polynom zweiten Grades $\varphi_2(x)$ auch noch für $x = x_2$ den richtigen Wert $f_2$ erhält, usw. Wir schreiben demgemäß

$$\varphi(x)=\varphi_n(x)=A_0+A_1(x-x_0)+A_2(x-x_0)(x-x_1)+\cdots+A_n(x-x_0)\cdots(x-x_{n-1})$$

und, wie ich gleich hinzufüge, daher

$$f(x) = \varphi_n(x) + R_n(x),$$

wobei $R_n(x)$ ein Rest ist, der jedenfalls an den Stellen $x = x_i$ $(i = 0, 1, \ldots, n)$ verschwindet. Allgemein setzen wir für $\nu \leqq n$

$$f(x) = \varphi_\nu(x) + R_\nu(x),$$

wo dann der Rest $R_\nu(x)$ an den Stellen $x_0, x_1, \ldots, x_\nu$ verschwindet.

Um den Koeffizienten $A_i$ und den Rest übersichtlich auszudrücken, denken wir uns zunächst aus $R_\nu$ den Faktor $(x - x_0)(x - x_1)\cdots(x - x_\nu)$ herausgezogen und schreiben demgemäß

$$R_\nu = (x - x_0)(x - x_1)\cdots(x - x_\nu)\, f(x, x_\nu, \ldots, x_0)$$

oder

$$f(x) = \varphi_\nu(x) + (x - x_0)\cdots(x - x_\nu)\, f(x, x_\nu, \ldots, x_0).$$

Dabei ist $f(x, x_\nu, \ldots, x_0)$ ein Polynom in $x$, dessen Koeffizienten von der Lage der Stellen $x_\nu, x_{\nu-1}, \ldots, x_0$ und den Funktionswerten $f(x_\nu), \ldots, f(x_0)$ abhängen. Es ist zunächst lediglich indirekt durch die obige Gleichung definiert; wir werden es jedoch gleich direkt berechnen.

Zunächst beachten wir, daß für $\nu < n$ nach Definition

$$f(x) = \varphi_\nu(x) + A_{\nu+1}(x - x_0)\cdots(x - x_\nu) + \cdots$$

ist, wo die letzten Punkte Glieder bezeichnen, welche für $x = x_i$ $(i = 0, 1, \ldots, \nu + 1)$ verschwinden. Setzen wir also $x = x_{\nu+1}$ ein, so erhalten wir durch Vergleichen mit der obigen Definitionsgleichung

$$A_{\nu+1} = f(x_{\nu+1}, x_\nu, \ldots, x_0).$$

Man erkennt somit, daß allgemein die Gleichungen

$$A_0 = f(x_0),\ A_1 = f(x_1, x_0),\ A_2 = f(x_2, x_1, x_0),\ \ldots,\ A_n = f(x_n, x_{n-1}, \ldots, x_0)$$

bestehen. Die Lösung unserer Interpolationsaufgabe wird also durch die folgende nach *Newton* genannte Interpolationsformel gegeben:

$$\varphi_n(x) = f(x_0) + f(x_1, x_0)(x - x_0) + \cdots + f(x_n, \ldots, x_0)(x - x_0)(x - x_1)\cdots(x - x_{n-1}).$$

Um nun die Koeffizienten $A_\nu$ sukzessive zu berechnen, schreiben wir einerseits

$$f(x) = \varphi_{\nu-1}(x) + f(x_\nu, x_{\nu-1}, \ldots, x_0)(x - x_0)\cdots(x - x_{\nu-1}) + f(x, x_\nu, \ldots, x_0)(x - x_0)\cdots(x - x_\nu),$$

andererseits

$$f(x) = \varphi_{\nu-1}(x) + f(x, x_{\nu-1}, \ldots, x_0)(x-x_0)\cdots(x - x_{\nu-1})$$

und erhalten durch Vergleichung sofort

$$f(x, x_\nu, \ldots, x_0) = \frac{1}{x - x_\nu}\{f(x, x_{\nu-1}, \ldots, x_0) - f(x_\nu, x_{\nu-1}, \ldots, x_0)\}.$$

Somit ergibt sich für unsere Koeffizienten, die sog. *Steigungen der Funktion* $f(x)$ für die Stellen $x_0, \ldots, x_n$, das folgende System von Rekursionsformeln:

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} = f(x_1, x_0) \qquad \text{erste Steigung}$$

$$\frac{f(x_2, x_0) - f(x_1, x_0)}{x_2 - x_1} = f(x_2, x_1, x_0) \qquad \text{zweite Steigung}$$

$$\frac{f(x_3, x_1, x_0) - f(x_2, x_1, x_0)}{x_3 - x_2} = f(x_3, x_2, x_1, x_0) \qquad \text{dritte Steigung}$$

. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

$$\frac{f(x_n, x_{n-2}, x_{n-3}, \ldots, x_0) - f(x_{n-1}, x_{n-2}, \ldots, x_0)}{x_n - x_{n-1}} = f(x_n, x_{n-1}, \ldots, x_0) \qquad n\text{-te Steigung}$$

Die Steigungen besitzen eine wichtige Eigenschaft: *Die $n$-te Steigung $f(x_n, \ldots, x_0)$ ändert sich nicht, wenn man die Argumente $x_0, \ldots, x_n$ irgendwie untereinander vertauscht.* Oder, wie man auch sagt, die $n$-te Steigung ist ein *symmetrischer Ausdruck* in ihren $n + 1$ Indizes. Der Beweis hierfür ergibt sich sofort aus unserem in Nr. 1 bewiesenen Eindeutigkeitssatz. Denn der Ausdruck $f(x_n, \ldots, x_0)$ erscheint als der Koeffizient der höchsten Potenz $x^n$ in unserem Interpolationspolynom, und dieses Interpolationspolynom, also auch sein höchster Koeffizient, bleibt nach dem Eindeutigkeitssatz immer dasselbe, gleichgültig, in welcher Reihenfolge man die Stellen $x_0, x_1, \ldots, x_n$ hinschreibt.

Um die Steigungen übersichtlich explizit auszudrücken, beachten wir, daß die $\nu$-te Steigung offenbar eine lineare Kombination der Funktionswerte $f_0, f_1, \ldots, f_\nu$ ist und daß insbesondere

$$f(x_n, x_{n-1}, \ldots, x_0) = c_0 f_0 + c_1 f_1 + \cdots + c_n f_n$$

gelten muß, wo die Koeffizienten $c_i$ nur noch von den Stellen $x_0, x_1, \ldots, x_n$, aber nicht von den $f_0, \ldots, f_n$ abhängen; wegen der Symmetrie muß sich $c_i$ mit $c_k$ vertauschen, wenn man $x_i$ mit $x_k$ vertauscht. Nun ist, wenn wir die Interpolationsformel für $x = x_n$ anwenden,

$$f_n = A_0 + \cdots + A_{n-1}(x_n - x_0)\cdots(x_n - x_{n-2}) + f(x_n, x_{n-1}, \ldots, x_0)(x_n - x_0)\cdots(x_n - x_{n-1}).$$

Beachten wir, daß in den Koeffizienten $A_0, \ldots, A_{n-1}$ die Größe $f_n$ nirgends auftritt, so folgt hieraus

$$f(x_n, \ldots, x_0) = \frac{f_n}{(x_n - x_0)\cdots(x_n - x_{n-1})} + \cdots,$$

wo die Punkte Glieder bedeuten, in denen $f_n$ nicht vorkommt. Also ist $\frac{1}{c_n} = (x_n - x_0) \cdots (x_n - x_{n-1})$, und wir haben somit wegen der Symmetrie in

$$f(x_n, \ldots, x_0) = \frac{f_n}{(x_n - x_0) \cdots (x_n - x_{n-1})} + \cdots + \frac{f_0}{(x_0 - x_1) \cdots (x_0 - x_n)}$$

den gesuchten expliziten Ausdruck für die $n$-te Steigung gefunden.

Ein einfacher Spezialfall der Newtonschen Formel ergibt sich, wenn die Punkte $x_0, \ldots, x_n$ äquidistant liegen, d. h. wenn

$$x_\nu = x_0 + \nu h, \qquad \nu = 0, 1, \ldots, n,$$

ist. Dann sind die Steigungen bis auf konstante Faktoren die Differenzenquotienten $\frac{\Delta^\nu f}{h^\nu}$ der Funktion $f(x)$.[1] Es ist nämlich, wenn man die Symmetrie der Steigungen beachtet, nach der Formel auf S. 271

$$f(x_1, x_0) = \frac{f_1 - f_0}{h} = \frac{\Delta^1 f_0}{h},$$

$$f(x_2, x_1, x_0) = f(x_2, x_0, x_1) = \frac{f(x_2, x_1) - f(x_0, x_1)}{x_2 - x_0} = \frac{\Delta^1 f_1 - \Delta^1 f_0}{2!\, h^2} = \frac{1}{2!} \frac{\Delta^2 f_0}{h^2},$$

. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

$$f(x_n, x_{n-1}, \ldots, x_0) = f(x_n, x_0, \ldots, x_{n-1}) = \frac{f(x_n, x_1, \ldots, x_{n-1}) - f(x_0, x_1, \ldots, x_{n-1})}{x_n - x_0}$$

$$= \frac{f(x_n, x_{n-1}, \ldots, x_1) - f(x_{n-1}, x_{n-2}, \ldots, x_0)}{n \cdot h}$$

$$= \frac{\Delta^{n-1} f_1 - \Delta^{n-1} f_0}{n \cdot h \cdot (n-1)! \cdot h^{n-1}} = \frac{\Delta^n f_0}{n! \cdot h^n}.$$

Setzen wir, indem wir statt $x$ durch die Gleichung $x - x_0 = th$ eine neue Variable $t$ einführen, zur Abkürzung wie früher

$$\binom{t}{\nu} = \frac{t(t-1) \cdots (t - \nu + 1)}{\nu!}$$

(*allgemeine Binomialkoeffizienten*), so nimmt unser Interpolationspolynom die einfache Gestalt

$$\varphi(x) = \varphi_n(x) = f_0 + \binom{t}{1} \frac{\Delta^1 f_0}{h} h + \binom{t}{2} \frac{\Delta^2 f_0}{h^2} h^2 + \cdots + \binom{t}{n} \frac{\Delta^n f_0}{h^n} h^n$$

an.

## 3. Zusammenhang zwischen Steigungen und Ableitungen. Restabschätzungen.

Bisher war es für unsere Überlegungen im Grunde gleichgültig, wie die Werte $f_0, f_1, \ldots, f_n$ gegeben waren. Sind diese Werte z. B.

[1] Mit Hilfe der Binomialkoeffizienten läßt sich die $\nu$-te Differenz $\Delta^\nu f_0$ folgendermaßen schreiben:

$$\Delta^\nu f_0 = f_\nu - \binom{\nu}{1} f_{\nu-1} + \binom{\nu}{2} f_{\nu-2} - + \cdots + (-1)^\nu f_0,$$

wie man leicht durch Schluß von $\nu$ auf $\nu + 1$ beweist.

durch physikalische Messungen gewonnen, so wird durch die Konstruktion des Polynomes $\varphi(x)$ unsere Interpolationsaufgabe völlig gelöst sein; wir haben dann in $\varphi(x)$ eine möglichst einfache Funktion, welche an den vorgegebenen Stellen die vorgegebenen Werte annimmt. Ist jedoch von vornherein die Funktion $f(x)$ gegeben, so entsteht ein neues Problem, nämlich das Problem, die Differenz $R(x) = f(x) - \varphi(x)$, den „Fehler" bei der Interpolation, abzuschätzen. Zunächst wissen wir nur, daß die $n+1$ Werte $R(x_0)$, $R(x_1), \ldots, R(x_n)$ sämtlich verschwinden. Um mehr aussagen zu können, müssen wir von der Funktion $f(x)$ und damit von $R(x)$ weitere Voraussetzungen machen. Wir nehmen an, daß $f(x)$ in dem betrachteten Intervalle stetige Ableitungen bis mindestens zur $(n+1)$-ten Ordnung besitzt.

Unter dieser Voraussetzung beweisen wir zunächst einen auch an sich wichtigen Hilfssatz über den *Zusammenhang zwischen Steigungen und Ableitungen:* Es ist

$$f(x_0, \ldots, x_n) = \frac{1}{n!} f^{(n)}(\xi),$$

wo $\xi$ einen Zwischenwert zwischen dem größten und dem kleinsten unter den Werten $x_\nu$ bedeutet. Nach dem in Nr. 1 bewiesenen Satz von Rolle gibt es nämlich einen solchen Zwischenwert $\xi$, für welchen

$$R^{(n)}(\xi) = f^{(n)}(\xi) - \varphi^{(n)}(\xi) = 0.$$

Nun ist die $n$-te Ableitung des Polynomes $\varphi(x)$ gerade gleich $n!\, f(x_0, \ldots, x_n)$; es ergibt sich daher sofort $f^{(n)}(\xi) - n!\, f(x_0, \ldots, x_n) = 0$.

Der gefundene Zusammenhang zeigt uns sofort weiter: *Wenn alle $n+1$ Stellen $x_0, \ldots, x_n$ in eine Stelle $a$ zusammenrücken, so strebt die durch $f(x_0, \ldots, x_n)$ definierte Steigung gegen den durch $n!$ dividierten $n$-ten Differentialquotienten im Punkte $a$, und insbesondere strebt der $n$-te Differenzenquotient gegen den $n$-ten Differentialquotienten.*

Mit Hilfe des gewonnenen Resultates erhalten wir nunmehr sehr leicht auch eine Restabschätzung für den bei der Interpolation einer beliebigen Funktion $f(x)$ begangenen Fehler $R(x)$. Da dieser Rest an den Stellen $x_0, x_1, \ldots, x_n$ verschwindet, so lag es nahe, ihn in der Form

$$R(x) = A(x - x_0)(x - x_1) \cdots (x - x_n)$$

zu schreiben; die Aufmerksamkeit ist nunmehr auf die Bestimmung des von den Größen $x$ und $x_0, x_1, \ldots, x_n$ abhängenden Koeffizienten $A$ zu richten. Die Betrachtungen von Nr. 2 ergaben für $A$ den Ausdruck

$$A = f(x, x_n, x_{n-1}, \ldots, x_0).$$

Der Koeffizient $A$ ist also eine Steigung für die $n+2$ Argumente $x, x_0, x_1, \ldots, x_n$. Da wir voraussetzen, daß unsere Funktion $f(x)$ stetige Ableitungen bis mindestens zur $(n+1)$-ten Ordnung besitzt, so ergibt sich nunmehr nach unserem gefundenen Satz über den Zusammenhang zwischen der Steigung und dem Differentialquotienten die Restdarstellung

$$A = \frac{1}{(n+1)!} f^{(n+1)}(\xi),$$

$$R(x) = \frac{(x-x_0)(x-x_1)\cdots(x-x_n)}{(n+1)!} f^{(n+1)}(\xi),$$

wo $\xi$ einen im übrigen nicht näher festzulegenden Zwischenwert zwischen der größten und der kleinsten der Stellen $x, x_0, x_1, \ldots, x_n$ bedeutet.

Damit ist das allgemeine Problem der Interpolation einer gegebenen Funktion vollständig gelöst. Zugleich erkennen wir, daß unsere Interpolationsformel, wenn alle Stellen $x_0, x_1, \ldots, x_n$ äquidistant in eine und dieselbe Stelle, etwa in den Nullpunkt hineinrücken, Glied für Glied in die Taylorsche Formel mit der Lagrangeschen Form des Restgliedes übergeht. *Die Taylorsche Formel ist also als ein Grenzfall der Newtonschen Interpolationsformel anzusehen.*

Durch diese Formel erhält die in der Geometrie übliche Sprechweise einen präzisen Sinn: Die Schmiegungsparabel, welche eine gegebene Kurve in einem Punkte von der $n$-ten Ordnung berührt, hat in diesem Punkte $n+1$ *„zusammenfallende" Schnittpunkte* mit der gegebenen Kurve gemeinsam. Wir erhalten nämlich diese Schmiegungsparabel ohne weiteres, indem wir die Parabel zunächst durch $n+1$ verschiedene Punkte legen und diese Punkte dann alle zusammenrücken lassen. Ganz Ähnliches gilt bei der Oskulation beliebiger Kurven. Z. B. ist der *Krümmungskreis gerade derjenige Kreis, welcher mit einer gegebenen Kurve drei zusammenfallende Schnittpunkte besitzt.*

Man wird die Interpolationsformeln stets dann anzuwenden haben, wenn man eine Funktion, deren Werte man in gewissen Stellen kennt, in dem Zwischengebiete dieser Stellen mit einigermaßen überall gleich guter Annäherung darstellen will. Wenn die Stelle $x$ außerhalb des Zwischengebietes der Stellen $x_0, x_1, \ldots, x_n$ liegt, so spricht man von einer *Extrapolation.* Man wird bei einer solchen Extrapolation um so weniger auf gute Übereinstimmung rechnen können, je weiter sich die Stelle $x$ von dem Zwischengebiet entfernt. Bei der Taylorschen Formel haben wir es gewissermaßen mit einer vollständigen Extrapolation zu tun, und dies ist der Grund dafür, daß die Taylorsche Formel tatsächlich häufig nur zu einer Darstellung der Funktion in der unmittelbaren Umgebung einer Stelle geeignet ist.

## 4. Die Interpolationsformel von Lagrange.

Zum Schluß möchte ich die Interpolationsformel noch in eine etwas andere Gestalt bringen, die gewöhnlich nach *Lagrange* bezeichnet wird und welche sich von der Newtonschen Interpolationsformel nur dadurch unterscheidet, daß die einzelnen Glieder nicht nach den Differenzenprodukten, sondern nach den Funktionswerten $f_\nu$ selbst geordnet sind. Wir gelangen zu dieser umgeformten Interpolationsformel, wenn wir von unserem expliziten Ausdruck für die Steigung ausgehen:

$$f(x_0, x_1, \ldots, x_n) = \frac{f(x_0)}{(x_0 - x_1) \cdots (x_0 - x_n)} + \frac{f(x_1)}{(x_1 - x_0)(x_1 - x_2) \cdots (x_1 - x_n)} + \cdots + \frac{f(x_n)}{(x_n - x_0) \cdots (x_n - x_{n-1})}.$$

Wenden wir diese Darstellung auf die Steigung $f(x, x_0, x_1, \ldots, x_n)$ mit einer veränderlichen Stelle $x$ an, so erhalten wir

$$f(x, x_0, \ldots, x_n) = \frac{f(x)}{(x - x_0) \cdots (x - x_n)} + \frac{f(x_0)}{(x_0 - x)(x_0 - x_1) \cdots (x_0 - x_n)} + \cdots + \frac{f(x_n)}{(x_n - x)(x_n - x_0) \cdots (x_n - x_{n-1})}.$$

Zur Abkürzung führen wir nun den Ausdruck

$$\psi(x) = (x - x_0)(x - x_1) \cdots (x - x_n)$$

ein; er ist eine zugleich mit den Stellen $x_\nu$ gegebene ganze rationale Funktion $(n+1)$-ten Grades. Differenzieren wir nach der Produktregel und setzen dann für $x$ einen der Werte $x_0, \ldots, x_n$ ein, so erhalten wir die Beziehungen

$$\psi'(x_0) = (x_0 - x_1)(x_0 - x_2) \cdots (x_0 - x_n),$$
$$\psi'(x_\nu) = (x_\nu - x_0) \cdots (x_\nu - x_{\nu-1})(x_\nu - x_{\nu+1}) \cdots (x_\nu - x_n) \quad (1 \leqq \nu \leqq n-1),$$
$$\psi'(x_n) = (x_n - x_0)(x_n - x_1) \cdots (x_n - x_{n-1}).$$

Nunmehr können wir unsere obige Darstellungsformel für $f(x, x_0, \ldots, x_n)$ nach $f(x)$ auflösen und gewinnen sofort

$$f(x) = \psi(x) \left\{ \frac{f(x_0)}{(x - x_0)\,\psi'(x_0)} + \frac{f(x_1)}{(x - x_1)\,\psi'(x_1)} + \cdots + \frac{f(x_n)}{(x - x_n)\,\psi'(x_n)} \right\} + R,$$

wo das Restglied

$$R = \psi(x)\, f(x, x_0, x_1, \ldots, x_n)$$

dasselbe wie in der Newtonschen Formel ist. Diese Formel ist die *Interpolationsformel von Lagrange*. Sie stellt uns bis auf das Restglied $R$ die Funktion $f(x)$ als eine lineare Kombination der Funktionswerte $f(x_0), f(x_1), \ldots$ dar, mit Koeffizienten, die wir aus der Kenntnis der Stellen $x_0, x_1, \ldots, x_n$ sofort berechnen können.

Zu dieser Lagrangeschen Darstellung des $n$-ten Interpolationspolynomes können wir auch ohne den Umweg über die Newtonsche Formel, sogar leichter, gelangen. Wir brauchen nur von der Bemerkung auszugehen, daß $\frac{\psi(x)}{(x - x_\nu)\,\psi'(x_\nu)}$ ein Polynom $n$-ten Grades ist, welches im Punkte $x = x_\nu$ den Wert 1, in den übrigen vorgegebenen Punkten $x_i$ dagegen den Wert Null annimmt. Dann ist sofort klar, daß der obige Ausdruck das gewünschte Interpolationspolynom darstellt.

Siebentes Kapitel.

# Exkurs über numerische Methoden.

## Vorbemerkungen.

Wer die Analysis als Instrument zur Behandlung physikalischer oder technischer Erscheinungen verwenden soll, steht vor der Frage, ob und wie sich aus den theoretischen Einsichten praktische Hilfsmittel zur wirklichen numerischen Ausführung der Rechnungen ergeben. Aber diese Frage besitzt auch vom Standpunkt des Theoretikers, der nicht die Natur beherrschen, sondern Zusammenhänge erkennen will, ein kaum geringeres Interesse. Hinsichtlich einer systematischen Behandlung numerischer Methoden muß ich auf spezielle Darstellungen verweisen[1]). Hier kann ich nur nebeneinander einige besonders wichtige mehr oder weniger unmittelbar an das Vorangehende anknüpfende Punkte behandeln. Dabei hebe ich grundsätzlich hervor, daß jede genäherte Berechnung erst dann einen präzisen Sinn besitzt, wenn sie durch eine Abschätzung des begangenen Fehlers ergänzt wird, wenn man also bei ihr eine Sicherheit für den Grad der erreichten Genauigkeit gewonnen hat.

## § 1. Numerische Integration.

Wir haben gesehen, daß sich schon verhältnismäßig einfache Funktionen nicht mehr mit Hilfe der elementaren Funktionen integrieren lassen und daß das Bestreben der Integralrechnung nicht auf ein solches prinzipiell unerreichbares Ziel gerichtet sein kann. Andererseits existiert doch das bestimmte Integral einer stetigen Funktion, und es ergibt sich daher die Aufgabe, Methoden zu seiner numerischen Berechnung zu finden. Die einfachsten und nächstliegenden dieser Methoden will ich hier an Hand der geometrischen Anschauung besprechen und dann auf die Fehlerabschätzung eingehen.

Es handelt sich um die Berechnung des Integrales $J = \int_a^b f(x)\,dx$, wobei $a < b$ sei. Wir denken uns das Integrationsintervall in $n$ gleiche Teile von der Länge $h = \frac{b-a}{n}$ eingeteilt und bezeichnen die Teilpunkte mit $x_0 = a$, $x_1 = a + h, \ldots, x_n = b$, die Funktionswerte in den Teilpunkten mit $f_0, f_1, \ldots, f_n$ und entsprechend auch die Funktionswerte in den Mittelpunkten der Intervalle mit $f_{\frac{1}{2}}, f_{\frac{3}{2}}, \ldots, f_{\frac{2n-1}{2}}$.

[1]) Vgl. etwa *Runge-König*, Vorlesungen über numerisches Rechnen, Berlin 1924, und *Whittaker-Robinson*, The Calculus of Observations, London 1926.

Unser Integral deuten wir als Flächeninhalt und zerschneiden das Flächenstück nach der üblichen Art in Streifen der Breite $h$. Es kommt jetzt nur noch darauf an, für jeden solchen Streifeninhalt eine Annäherung zu erhalten, d. h. die Integrale

$$J_\nu = \int_{x_\nu}^{x_\nu + h} f(x)\,dx$$

angenähert zu berechnen.

### 1. Rechtecksregel.

Die roheste und nächstliegende Methode zur angenäherten Berechnung von $J$ knüpft unmittelbar an die Integraldefinition an; wir ersetzen nämlich einfach den Inhalt des Streifens $J_\nu$ durch den Rechtecksinhalt $f_\nu h$ und erhalten dann für das Integral $J$ die angenäherte Darstellung[1])

$$J \approx h(f_0 + f_1 + \cdots + f_{n-1}).$$

### 2. Trapezformel und Tangentenformel.

Eine feinere Annäherung bei demselben Aufwand an Rechnung erhalten wir, wenn wir den Streifeninhalt $J_\nu$ nicht durch den obigen Rechtecksinhalt, sondern durch das in Figur 100 gezeichnete Trapez mit dem Flächeninhalt $\frac{1}{2}(f_\nu + f_{\nu+1})h$ ersetzen. Für das ganze Integral ergibt sich dann die angenäherte Darstellung (*Trapezformel*)

$$J \approx h(f_1 + f_2 + \cdots + f_{n-1}) + \frac{h}{2}(f_0 + f_n),$$

da bei Addition der Trapezinhalte jeder Funktionswert außer dem ersten und letzten zweimal vorkommt.

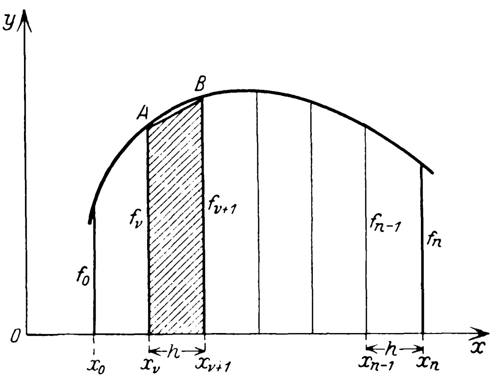



Fig. 100. Trapezformel.

Noch etwas besser wird die Annäherung im allgemeinen, wenn wir nicht das von der Sehne $AB$ begrenzte Trapez als Annäherung des Streifeninhalts $J_\nu$ wählen, sondern dasjenige Trapez, welches von der Tangente an die Kurve im Punkte mit der Abszisse $x = x_\nu + \frac{h}{2}$ begrenzt wird. Der Inhalt dieses Trapezes ist einfach $h f_{\nu+\frac{1}{2}}$, und

[1]) Das Zeichen $\approx$ bedeutet hier und im folgenden: „angenähert gleich".

wir erhalten also für das gesamte Integral die Annäherung

$$J \approx h\left(f_{\frac{1}{2}} + f_{\frac{3}{2}} + \cdots + f_{\frac{2n-1}{2}}\right),$$

die *Tangentenformel.*

### 3. Die Simpsonsche Regel.

Zu einer im allgemeinen noch sehr viel genaueren numerischen Berechnung bei kaum größerer Mühe gelangen wir durch die *Simpsonsche Regel.* Diese beruht darauf, daß man den Inhalt $J_\nu + J_{\nu+1}$ eines Doppelstreifens zwischen den Abszissen $x = x_\nu$ und $x = x_\nu + 2h = x_{\nu+2}$ berechnet, indem man ihn nach oben zu nicht mehr geradlinig, sondern durch eine Parabel begrenzt; und zwar durch diejenige Parabel, welche durch die drei Kurvenpunkte mit den Abszissen $x_\nu$, $x_{\nu+1} = x_\nu + h$, $x_{\nu+2} = x_\nu + 2h$ geht (vgl. Fig. 101). Die Gleichung dieser Parabel ist (nach der Newtonschen Interpolationsformel, S. 272)

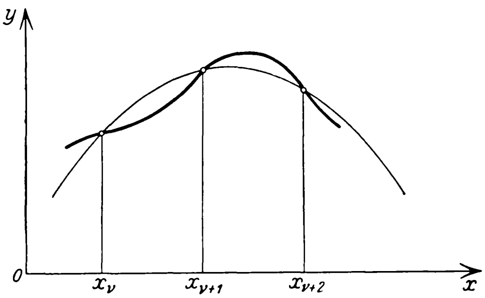


Fig. 101. Simpsonsche Regel.

$$y = f_\nu + (x - x_\nu)\cdot\frac{f_{\nu+1} - f_\nu}{h} + \frac{(x - x_\nu)(x - x_\nu - h)}{2}\cdot\frac{f_{\nu+2} - 2f_{\nu+1} + f_\nu}{h^2}.$$

Integrieren wir diese ganze Funktion zweiten Grades zwischen den Grenzen $x_\nu$ und $x_\nu + 2h$, so ergibt sich nach kurzer Rechnung als Flächeninhalt unterhalb des Parabelstückes der Ausdruck

$$\int\limits_{x_\nu}^{x_\nu+2h} y\,dx = 2hf_\nu + 2h(f_{\nu+1} - f_\nu) + \frac{\frac{8}{3}h - 2h}{2}(f_{\nu+2} - 2f_{\nu+1} + f_\nu)$$

$$= \frac{h}{3}(f_\nu + 4f_{\nu+1} + f_{\nu+2}).$$

Dieser stellt also die gesuchte Annäherung an unseren Streifeninhalt $J_\nu + J_{\nu+1}$ dar.

Setzen wir nun voraus, daß $n = 2m$, d. h. eine gerade Zahl ist, so erhalten wir durch Addition solcher Streifeninhalte für das ganze Intervall die *Simpsonsche* Näherungsdarstellung

$$J \approx \frac{4h}{3}(f_1 + f_3 + \cdots + f_{2m-1}) + \frac{2h}{3}(f_2 + f_4 + \cdots + f_{2m-2}) + \frac{h}{3}(f_0 + f_{2m}).$$

### 4. Beispiele.

Wir wollen diese Methoden anwenden, um $\log 2 = \int\limits_1^2 \frac{dx}{x}$ zu berechnen. Zerlegen wir das Intervall von 1 bis 2 in zehn gleiche Teile, so wird $h = \frac{1}{10}$, und wir erhalten zunächst nach der Trapezformel

$$\begin{array}{ll}
x_1 = 1{,}1 & f_1 = 0{,}90909 \\
x_2 = 1{,}2 & f_2 = 0{,}83333 \\
x_3 = 1{,}3 & f_3 = 0{,}76923 \\
x_4 = 1{,}4 & f_4 = 0{,}71429 \\
x_5 = 1{,}5 & f_5 = 0{,}66667 \\
x_6 = 1{,}6 & f_6 = 0{,}625 \\
x_7 = 1{,}7 & f_7 = 0{,}58824 \\
x_8 = 1{,}8 & f_8 = 0{,}55556 \\
x_9 = 1{,}9 & f_9 = 0{,}52632 \\
\hline
 & \text{Summe } 6{,}18773 \\
x_0 = 1{,}0 & \frac{1}{2} f_0 = 0{,}5 \\
x_{10} = 2{,}0 & \frac{1}{2} f_{10} = 0{,}25 \\
\hline
 & 6{,}93773 \cdot \frac{1}{10} \\
\hline
 & \log 2 \approx 0{,}69377
\end{array}$$

Dieser Wert ist zu groß, wie es auch der Umstand erwarten ließ, daß die Kurve der $x$-Achse ihre konvexe Seite zuwendet.

Nach der Tangentenregel folgt

$$\begin{array}{ll}
x_0 + \frac{1}{2}h = 1{,}05 & f_{\frac{1}{2}} = 0{,}95238 \\
x_1 + \frac{1}{2}h = 1{,}15 & f_{\frac{3}{2}} = 0{,}86957 \\
x_2 + \frac{1}{2}h = 1{,}25 & f_{\frac{5}{2}} = 0{,}8 \\
x_3 + \frac{1}{2}h = 1{,}35 & f_{\frac{7}{2}} = 0{,}74074 \\
x_4 + \frac{1}{2}h = 1{,}45 & f_{\frac{9}{2}} = 0{,}68966 \\
x_5 + \frac{1}{2}h = 1{,}55 & f_{\frac{11}{2}} = 0{,}64516 \\
x_6 + \frac{1}{2}h = 1{,}65 & f_{\frac{13}{2}} = 0{,}60606 \\
x_7 + \frac{1}{2}h = 1{,}75 & f_{\frac{15}{2}} = 0{,}57143 \\
x_8 + \frac{1}{2}h = 1{,}85 & f_{\frac{17}{2}} = 0{,}54054 \\
x_9 + \frac{1}{2}h = 1{,}95 & f_{\frac{19}{2}} = 0{,}51282 \\
\hline
 & 6{,}92836 \cdot \frac{1}{10} \\
\hline
 & \log 2 \approx 0{,}69284
\end{array}$$

Wegen der Konvexität der Kurve ist dieser Wert zu klein.

Das genaueste Resultat erzielen wir bei der gleichen Intervalleinteilung vermöge der Simpsonschen Regel. Wir erhalten:

| | | | |
|---|---|---|---|
| $x_1 = 1{,}1$ | $f_1 = 0{,}90909$ | $x_2 = 1{,}2$ | $f_2 = 0{,}83333$ |
| $x_3 = 1{,}3$ | $f_3 = 0{,}76923$ | $x_4 = 1{,}4$ | $f_4 = 0{,}71429$ |
| $x_5 = 1{,}5$ | $f_5 = 0{,}66667$ | $x_6 = 1{,}6$ | $f_6 = 0{,}625$ |
| $x_7 = 1{,}7$ | $f_7 = 0{,}58824$ | $x_8 = 1{,}8$ | $f_8 = 0{,}55556$ |
| $x_9 = 1{,}9$ | $f_9 = 0{,}52632$ | | Summe $2{,}72818 \cdot 2$ |
| | Summe $3{,}45955 \cdot 4$ | | $5{,}45636$ |
| | $13{,}83820$ | | $13{,}83820$ |
| | | $x_0 = 1{,}0$ | $f_0 = 1{,}0$ |
| | | $x_{10} = 2{,}0$ | $f_{10} = 0{,}5$ |
| | | | $20{,}79456 \cdot \frac{1}{30}$ |
| | | | $\log 2 \approx 0{,}69315$ |

Es ist tatsächlich

$$\log 2 = 0{,}693147 \ldots .$$

## 5. Fehlerabschätzung.

In allen unseren Beispielen von Integrationsmethoden ist es leicht, eine Fehlerabschätzung zu geben, wenn die Ableitungen der Funktion $f(x)$ in ihrem Verlauf bekannt sind. Wir verstehen unter $M_1, M_2, \ldots$ obere Schranken für den absoluten Betrag der ersten bzw. zweiten Ableitung usw.; d. h. wir nehmen an, daß in dem ganzen Intervall $|f^{(\nu)}(x)| < M_\nu$ gilt. Dann lauten die Abschätzungsformeln folgendermaßen:

Für die Rechtecksregel

$$|J_\nu - h f_\nu| < \frac{1}{2} M_1 h^2 \quad \text{oder} \quad \Big|J - h \sum_{\nu=0}^{n-1} f_\nu\Big| < \frac{1}{2} M_1 n h^2 = \frac{1}{2} M_1 (b-a) h.$$

Für die Tangentenregel

$$|J_\nu - h f_{\nu+\frac{1}{2}}| < \frac{M_2}{24} h^3 \quad \text{oder} \quad \Big|J - h \sum_{\nu=0}^{n-1} f_{\nu+\frac{1}{2}}\Big| < \frac{M_2}{24} (b-a) h^2.$$

Für die Trapezregel

$$\Big|J_\nu - \frac{h}{2}(f_\nu + f_{\nu+1})\Big| < \frac{M_2}{12} h^3.$$

Für die Simpsonsche Regel

$$\Big|J_\nu + J_{\nu+1} - \frac{h}{3}(f_\nu + 4 f_{\nu+1} + f_{\nu+2})\Big| < \frac{M_4}{90} h^5.$$

Aus den beiden letzten Abschätzungen folgen noch solche für das gesamte Integral $J$. Wir sehen, daß die Simpsonsche Regel einen in $h$ von viel höherer Ordnung kleinen Fehler gibt als die andern Regeln, so daß sie, falls nur nicht $M_4$ zu groß wird, für die wirkliche Rechnung sehr vorteilhaft erscheint.

Um den Leser nicht mit der Durchführung der an sich sehr einfachen Beweise für unsere Abschätzungen zu ermüden, begnüge ich mich mit dem Beweis für den Fall der Tangentenformel. Zu diesem Zwecke entwickeln wir die Funktion $f(x)$ im $(\nu+1)$-ten Streifen nach der Taylorschen Formel:

$$f(x) = f_{\nu+\frac{1}{2}} + \left(x - x_\nu - \frac{h}{2}\right) f'\left(x_\nu + \frac{h}{2}\right) + \frac{1}{2}\left(x - x_\nu - \frac{h}{2}\right)^2 f''(\xi),$$

wo $\xi$ ein gewisser Zwischenwert in dem Streifen ist. Integrieren wir die rechte Seite über das Intervall $x_\nu \leqq x \leqq x_\nu + h$, so gibt das Integral des mittleren Gliedes Null. Da, wie man leicht nachrechnet,

$$\frac{1}{2}\int\limits_{x_\nu}^{x_\nu+h} \left(x - x_\nu - \frac{h}{2}\right)^2 dx = \frac{h^3}{24}$$

ist, so folgt nunmehr sofort

$$\left|\int\limits_{x_\nu}^{x_\nu+h} f(x)\,dx - h f_{\nu+\frac{1}{2}}\right| < M_2 \frac{h^3}{24},$$

d. h. unsere Behauptung.

## § 2. Anwendungen des Mittelwertsatzes und des Taylorschen Satzes.

### 1. Die „Fehlerrechnung".

In ganz anderer Richtung liegen die numerischen Anwendungen, die man von dem Mittelwertsatz oder allgemeiner dem Taylorschen Satz mit dem Restgliede oder schließlich auch von der unendlichen Taylorschen Reihe macht. Ich betrachte zunächst als ganz einfaches, aber für die Praxis recht wichtiges Beispiel die Fehlerrechnung. Diese beruht auf dem —, für die ganze Differentialrechnung grundlegenden — Gedanken, daß wir eine hinreichend oft differenzierbare Funktion $f(x)$ in der Umgebung einer Stelle durch eine lineare Funktion bis auf einen Fehler von kleinerer als erster Ordnung oder durch eine quadratische Funktion bis auf einen Fehler von kleinerer als zweiter Ordnung ersetzen können usw. Betrachten wir die lineare Annäherung für eine Funktion $y = f(x)$. Ist $y + \Delta y = f(x + \Delta x) = f(x + h)$, so haben wir nach der Taylorschen Formel

$$\Delta y = h f'(x) + \frac{h^2}{2} f''(\xi),$$

wo $\xi = x + \vartheta h$ $(0 < \vartheta < 1)$ ein Zwischenwert ist, den wir nicht näher zu kennen brauchen. Ist $h = \Delta x$ eine kleine Größe, so erhalten wir als praktische Näherung

$$\Delta y \approx f'(x)\, h.$$

Mit anderen Worten: Man ersetzt den Differenzenquotienten angenähert durch den Differentialquotienten, die Differenz angenähert durch ihren in $h$ linearen Anteil.

Man wendet diese fast selbstverständliche Betrachtung praktisch folgendermaßen an. Es seien zwei physikalische Größen $x$ und $y$ durch eine Beziehung $y = f(x)$ miteinander verknüpft. Dann ist die Frage, welchen Einfluß eine Ungenauigkeit in der Messung von $x$ auf die Bestimmung von $y$ hat. Beobachtet man statt des „wahren" Wertes $x$ den ungenauen Wert $x + h$, so wird der entsprechende $y$-Wert sich von dem wahren Wert $y = f(x)$ um $\Delta y = f(x + h) - f(x)$ unterscheiden. Der Fehler wird also angenähert durch die obigen Beziehungen gegeben.

Am besten verstehen wir die Verwendung dieser Dinge an Hand einiger Beispiele.

1. Beispiel: Tangentenbussole. Bei einer Tangentenbussole handelt es sich um die Funktion $y = c \cdot \operatorname{tg} \alpha$, wobei $\alpha$ der Ablenkungswinkel der Magnetnadel, $c$ eine Apparatkonstante und $y = J$ die Stromstärke ist. Es ergibt sich

$$\frac{dy}{d\alpha} = \frac{c}{\cos^2 \alpha}$$

und daher $\Delta y \approx \frac{c}{\cos^2 \alpha} \Delta \alpha$. Die prozentuale Genauigkeit der Messung wird durch

$$\frac{100 \Delta y}{y} \approx \frac{100\, c\, \Delta \alpha}{c \cdot \cos^2 \alpha \cdot \operatorname{tg} \alpha} = \frac{200}{\sin 2\alpha} \Delta \alpha$$

gegeben. Man erkennt hieraus, daß die prozentuale Genauigkeit möglichst groß wird, d. h. daß einem Fehler der Winkelmessung ein prozentual möglichst kleiner Fehler der Stromstärkebestimmung entspricht, wenn der Winkel $\alpha$ gleich $\frac{\pi}{4}$, d. h. gleich $45^0$ wird.

Beispielsweise sei es möglich, die Tangentenbussole auf halbe Grade abzulesen; dann ist im Bogenmaß gemessen $|\Delta \alpha| < \frac{1}{2} \cdot 0{,}01745\ldots$, und die prozentuale Genauigkeit wird $\frac{1{,}745}{\sin 2\alpha}$. Wird etwa der Winkel $\alpha = 30^0$ abgelesen, wobei $\sin 2\alpha = \frac{1}{2}\sqrt{3} = \frac{1}{2} \cdot 1{,}73205\ldots$ wird, so ergibt sich eine prozentuale Genauigkeit von $2\frac{1{,}745}{1{,}732}$, d. h. etwa von 2%.

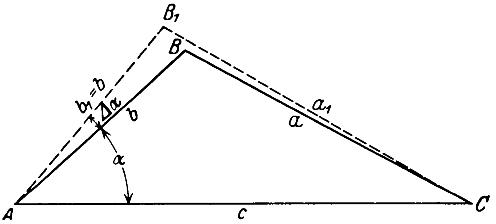


Fig. 102.

2. Beispiel. In einem Dreieck $ABC$ (vgl. Figur 102) seien die Seiten $b$ und $c$ genau gemessen, während der Winkel $\alpha = x$ nur innerhalb einer Fehlergrenze $|\Delta x| < \delta$ genau gemessen werden kann. In welchen Fehlergrenzen bewegt sich der Wert von $y = a = \sqrt{b^2 + c^2 - 2bc \cos \alpha}$?

Es wird

$$\Delta a \approx \frac{1}{a} b c \sin \alpha \Delta \alpha;$$

die prozentuale Genauigkeit ist also $\frac{100 \Delta a}{a} \approx \frac{100 b c}{a^2} \sin \alpha \Delta \alpha$. Nehmen wir als Zahlenbeispiel $b = 400$ m, $c = 500$ m und $\alpha = 60^0$, so ergibt sich nach dem Kosinussatze $y = a = 458{,}2576$ m und weiter

$$\Delta a \approx \frac{200000}{458{,}2576} \cdot \frac{1}{2} \sqrt{3}\, \Delta \alpha;$$

bei zehn Bogensekunden Genauigkeit der Messung, d. h. bei $\Delta \alpha = 10''$ $= 4848 \cdot 10^{-8}$ im Bogenmaß ergibt sich

$$\Delta a \approx 1{,}83 \text{ cm},$$

d. h. eine Genauigkeit von etwa 0,004%.

Wollen wir die Genauigkeit weitertreiben, so brauchen wir nur statt der Annäherung durch lineare Funktionen nach dem Taylorschen Satz durch ein Polynom zweiten oder höheren Grades anzunähern und erhalten dadurch ohne weiteres Korrektionen höherer Ordnung und Fehlerabschätzungen, wie der Leser selbst des Näheren durchdenken mag.

## 2. Berechnung von $\pi$.

Die Leibnizsche Reihe $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + - \cdots$, welche wir (sechstes Kapitel, § 1, Nr. 2) aus der Entwicklung des Arcustangens erhalten hatten, ist zur Berechnung von $\pi$ wegen ihrer langsamen Konvergenz ungeeignet. Man kann jedoch durch folgenden Kunstgriff zu einer verhältnismäßig mühelosen Berechnung von $\pi$ gelangen. Aus dem Additionsgesetz des Tangens $\operatorname{tg}(\alpha + \beta) = \frac{\operatorname{tg}\alpha + \operatorname{tg}\beta}{1 - \operatorname{tg}\alpha \operatorname{tg}\beta}$ folgt durch Übergang zu den Umkehrfunktionen $\alpha = \operatorname{arc\,tg} u$, $\beta = \operatorname{arc\,tg} v$ die Formel

$$\operatorname{arc\,tg} u + \operatorname{arc\,tg} v = \operatorname{arc\,tg} \frac{u+v}{1-uv}.$$

Wählt man nun $u$ und $v$ so, daß $\frac{u+v}{1-uv} = 1$ wird, so erhält man rechts den Wert $\frac{\pi}{4}$ und kann, wenn $u$ und $v$ kleine Zahlen werden, die linke Seite mit Hilfe der uns bekannten Reihenentwicklung leicht berechnen. Setzt man z. B. mit Euler $u = \frac{1}{2}$, $v = \frac{1}{3}$, so erhält man

$$\frac{\pi}{4} = \operatorname{arc\,tg} \frac{1}{2} + \operatorname{arc\,tg} \frac{1}{3}.$$

Indem man weiter die Gleichung $\dfrac{\frac{1}{3} + \frac{1}{7}}{1 - \frac{1}{21}} = \frac{1}{2}$ beachtet, ergibt sich

$\operatorname{arc\,tg}\frac{1}{2} = \operatorname{arc\,tg}\frac{1}{3} + \operatorname{arc\,tg}\frac{1}{7}$, also

$$\frac{\pi}{4} = 2\operatorname{arc\,tg}\frac{1}{3} + \operatorname{arc\,tg}\frac{1}{7},$$

eine Formel, mit welcher Vega die Zahl $\pi$ auf 140 Stellen genau berechnet hat.

Mit Hilfe der Gleichung $\dfrac{\frac{1}{5}+\frac{1}{8}}{1-\frac{1}{40}} = \frac{1}{3}$ erhalten wir weiter

$$\operatorname{arc\,tg}\frac{1}{3} = \operatorname{arc\,tg}\frac{1}{5} + \operatorname{arc\,tg}\frac{1}{8}$$

oder

$$\frac{\pi}{4} = 2\operatorname{arc\,tg}\frac{1}{5} + \operatorname{arc\,tg}\frac{1}{7} + 2\operatorname{arc\,tg}\frac{1}{8}.$$

Diese Darstellung ist außerordentlich geeignet zur Berechnung von $\pi$ mit Hilfe der Reihe $\operatorname{arc\,tg} x = x - \frac{x^3}{3} + \frac{x^5}{5} - + \cdots$; denn wenn wir für $x$ den Wert $\frac{1}{5}$, $\frac{1}{7}$ oder $\frac{1}{8}$ einsetzen, erhalten wir schon mit wenigen Gliedern eine große Genauigkeit, da die Glieder sehr rasch abnehmen. Man kann aber die Rechnung noch bequemer gestalten, wenn man von der Formel

$$\frac{\pi}{4} = \operatorname{arc\,tg}\frac{120}{119} - \operatorname{arc\,tg}\frac{1}{239} = 4\operatorname{arc\,tg}\frac{1}{5} - \operatorname{arc\,tg}\frac{1}{239}$$

ausgeht, zu der man durch ähnliche Betrachtungen wie oben gelangt.

### 3. Berechnung der Logarithmen.

Zur numerischen Berechnung der Logarithmen formt man zweckmäßig die logarithmische Reihe $\frac{1}{2}\log\frac{1+x}{1-x} = x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots$ $(|x| < 1)$ für $0 < x < 1$ durch die Substitution

$$\frac{1+x}{1-x} = \frac{p^2}{p^2-1}, \quad x = \frac{1}{2p^2-1}$$

um in die Reihe

$$\log p = \frac{1}{2}\log(p-1) + \frac{1}{2}\log(p+1) + \frac{1}{2p^2-1} + \frac{1}{3(2p^2-1)^3} + \cdots,$$

wo dann $2p^2 - 1 > 1$, d. h. $p^2 > 1$ ist. Diese letztere Reihe erlaubt uns, wenn $p$ eine ganze Zahl ist und $p+1$ sich in kleinere ganze Faktoren zerlegen läßt, den Logarithmus von $p$ durch Logarithmen kleinerer Zahlen und eine Reihe darzustellen, deren Glieder sehr rasch abnehmen, die sich also mit wenigen Gliedern schon genau genug berechnen läßt. Wir können daher aus dieser Reihe sukzessive die

Logarithmen aller Primzahlen und somit aller Zahlen berechnen, wenn wir nur schon den Wert für log 2 zuvor berechnet haben.

Was die Genauigkeitsbestimmung bei unseren Reihen anbetrifft, so wird man zweckmäßigerweise nicht auf die allgemeine Restformel zurückgreifen, sondern vielmehr besser mit Hilfe der geometrischen Reihe abschätzen. Für den Rest $R_n$ der Reihe, d. h. die Summe der Glieder, die auf das Glied $\frac{1}{n(2p^2-1)^n}$ folgen, ergibt sich

$$R_n < \frac{1}{(n+2)(2p^2-1)^{n+2}}\left[1 + \frac{1}{(2p^2-1)^2} + \frac{1}{(2p^2-1)^4} + \cdots\right]$$

$$= \frac{1}{(n+2)(2p^2-1)^n} \cdot \frac{1}{(2p^2-1)^2 - 1},$$

und diese Formel gibt uns sofort eine Fehlerabschätzung der gewünschten Art. Wir berechnen z. B. log 7, indem wir die ersten vier Glieder der Reihe benutzen. Wir finden

$$p = 7, \quad 2p^2 - 1 = 97,$$

$$\log 7 = 2 \log 2 + \frac{1}{2} \log 3 + \frac{1}{97} + \frac{1}{3 \cdot 97^3} + \cdots;$$

$$\frac{1}{97} \approx 0{,}01030928, \quad \frac{1}{3 \cdot 97^3} \approx 0{,}00000037,$$

$$2 \log 2 \approx 1{,}38629436, \quad \frac{1}{2} \log 3 \approx 0{,}54930614,$$

also

$$\log 7 \approx 1{,}94591015.$$

Die Abschätzung des Fehlers ergibt

$$R_n < \frac{1}{5 \cdot 97^3} \cdot \frac{1}{97^2 - 1} < \frac{1}{36 \cdot 10^9}.$$

Wir müssen noch beachten, daß jede der vier Zahlen, die wir addiert haben, nur bis auf einen Fehler von $\frac{5}{10^9}$ genau angegeben ist, so daß die letzte in dem Wert von log 7 angegebene Stelle sich noch um 2 ändern kann. In Wirklichkeit ist aber auch die letzte Stelle richtig.

## § 3. Numerische Auflösung von Gleichungen.

Zum Schluß möchte ich einige Bemerkungen über numerische Auflösung von Gleichungen $f(x) = 0$ hinzufügen, wobei $f(x)$ nicht notwendig eine ganze rationale Funktion zu sein braucht[1]. Jedes solche numerische Verfahren beruht darauf, daß man von einer schon irgendwie bekannten Annäherung $x_0$ an eine der Wurzeln ausgeht und diese weiter verbessert. Woher man eine solche erste Annäherung für die gesuchte

[1] Es handelt sich natürlich hier immer nur um die Bestimmung reeller Wurzeln von $f(x) = 0$.

Wurzel der Gleichung gefunden hat und wie gut diese Annäherung ist, kann dabei dahingestellt bleiben. Man wird sich häufig eine erste Annäherung durch eine rohe Überschlagsbetrachtung verschaffen oder besser an Hand einer graphischen Darstellung der Funktion $y = f(x)$ durch eine Kurve, deren Schnittpunkte mit der $x$-Achse die gesuchten Wurzeln liefern (natürlich mit einer durch den Maßstab und die Feinheit der Zeichnung bedingten Ungenauigkeit).

## 1. Das Verfahren von Newton.

Auf dem Grundgedanken der Differentialrechnung — Ersetzung einer krummen Linie durch eine gerade Linie, die Tangente, in der unmittelbaren Umgebung des Berührungspunktes — beruht das folgende von *Newton* herrührende Verfahren: Hat man einen Näherungswert $x_0$ für eine Wurzel der Gleichung $f(x) = 0$, so betrachte man auf der Kurve für die Funktion $y = f(x)$ den Punkt mit den Koordinaten $x = x_0$ und $y = f(x_0)$; man versuche sodann, indem man möglichst nahe an der Kurve bleibt, in Richtung der Kurve auf den Schnittpunkt mit der $x$-Achse zuzugehen. Dies wird man, ohne die Kurve einzeln in ihrem Verlauf berücksichtigen zu müssen, am einfachsten dadurch erreichen, daß man in Richtung der Tangente wandert. Der Schnittpunkt dieser Tangente mit der $x$-Achse wird dann in seiner Abszisse $x_1$ einen neuen, unter Umständen verbesserten Wert für die gesuchte Gleichungswurzel darstellen.

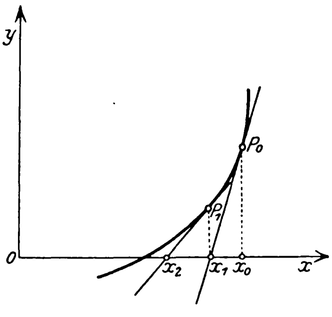



Fig. 103. Newtonsches Näherungsverfahren.

Aus Figur 103 ergibt sich vermöge der Bedeutung des Differentialquotienten unmittelbar die Beziehung

$$\frac{f(x_0)}{x_0 - x_1} = f'(x_0)$$

und daraus als Formel zur Berechnung des neuen Näherungswertes $x_1$ die Gleichung

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Hat man einen Wert mit Hilfe dieses Verfahrens verbessert, so kann man diesen verbesserten Wert wieder in einen Wert $x_2$ verbessern usw., und das Verfahren wird, wenn die Kurve die in Figur 103 gezeichnete Gestalt hat, immer genauer gegen die gesuchte Lösung konvergieren.

Die Brauchbarkeit dieses Verfahrens hängt wesentlich von dem Verlaufe der Kurve $y = f(x)$ ab. In Figur 103 erkennen wir, daß das Verfahren, wenn man es öfter wiederholt, mit immer größerer Genauigkeit gegen die gesuchte Wurzel konvergieren wird. Dies beruht darauf, daß die Kurve ihre konvexe Seite der $x$-Achse zuwendet. In Figur 104 jedoch zeigt sich, daß wir bei ungeschickter Wahl des Anfangswertes $x_0$ uns keineswegs mit unserer Konstruktion der gesuchten Wurzel nähern werden. Man erkennt hieraus, daß man auch bei dieser Newtonschen Methode in jedem Einzelfall besonders prüfen muß, ob bzw. mit welcher Genauigkeit man die Gleichung schließlich tatsächlich gelöst hat.

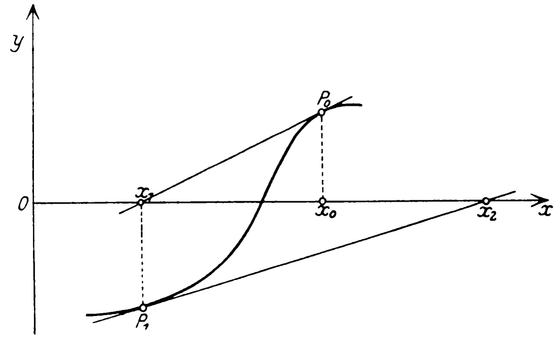


Fig. 104.

## 2. Regula falsi.

Die Newtonsche Methode, bei welcher die Tangente an die Kurve eine entscheidende Rolle spielte, ist nur der Grenzfall einer älteren Methode, die man als Regula falsi bezeichnet und bei welcher die Sekante an Stelle der Tangente tritt. Nehmen wir an, wir kennen zwei Punkte $x_0, y_0$ und $x_1, y_1$ der Kurve $y = f(x)$ in der Nähe des gesuchten Durchschnittspunktes mit der $x$-Achse. Ersetzen wir nun die Kurve durch die Kurvensekante, welche diese beiden Punkte verbindet, so werden wir in dem Schnittpunkt dieser Kurvensekante mit der $x$-Achse einen unter Umständen verbesserten Näherungswert für die gesuchte Wurzel der Gleichung zu erblicken haben. Ist $\xi$ die Abszisse dieses Schnittpunktes, so ergibt sich aus Figur 105 die Gleichung

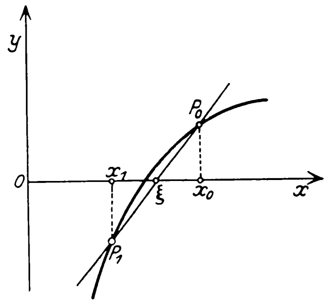


Fig. 105. Regula falsi.

$$\frac{\xi - x_0}{f(x_0)} = \frac{\xi - x_1}{f(x_1)},$$

und hieraus berechnet man:

$$\xi = \frac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)} = \frac{x_0 f(x_1) - x_0 f(x_0) + x_0 f(x_0) - x_1 f(x_0)}{f(x_1) - f(x_0)}$$

oder

$$\xi = x_0 - \frac{f(x_0)}{\dfrac{f(x_1) - f(x_0)}{x_1 - x_0}}.$$

Diese Formel, welche aus $x_0$ und $x_1$ den weiteren Näherungswert $\xi$ bestimmt, nennt man die *Regula falsi.* Man wird sie mit Vorteil anwenden, wenn ein Funktionswert positiv und der andere negativ ist, etwa, wie in Figur 105, $y_0 > 0$ und $y_1 < 0$. Sie wird im übrigen stets bei Wiederholung zum Ziele führen, wenn bei jedem Schritt je ein positiver und ein negativer Funktionswert benutzt wird, zwischen denen dann notwendig die gesuchte Wurzel eingegrenzt ist.

Die obige Newtonsche Formel ergibt sich aus dieser Regula falsi als Grenzfall, wenn wir $x_1$ gegen $x_0$ streben lassen. Denn der Nenner des zweiten Gliedes auf der rechten Seite der Regula falsi strebt, wenn $x_1$ gegen $x_0$ rückt, tatsächlich gegen $f'(x_0)$.

### 3. Beispiel.

Als Beispiel behandeln wir die Gleichung

$$f(x) = x^3 - 2x - 5 = 0.$$

Für $x_0 = 2$ wird $f(x_0) = -1$, für $x_1 = 2{,}1$ wird $f(x_1) = 0{,}061$; als weiteren Näherungswert erhalten wir

$$\xi = \frac{2 \cdot 0{,}061 - 2{,}1 \cdot (-1)}{0{,}061 - (-1)} \approx 2{,}0943.$$

Nun ergibt sich $f(\xi) \approx -0{,}0028$, und wir setzen daher weiter in unserer Formel $x_0 = 2{,}0943$ und $x_1 = 2{,}1$. Es wird dann

$$\xi_1 \approx 2{,}0946,$$

und es ist $f(\xi_1) \approx 0{,}00054$. Setzen wir jetzt $x_0 = 2{,}0943$, $x_1 = 2{,}0946$, so erhalten wir

$$\xi_2 \approx 2{,}09455,$$

und dieser Wert ist schon recht genau.

# Anhang zum siebenten Kapitel.

## Die Stirlingsche Formel.

In sehr vielen Anwendungen, vor allem in der Statistik und der Wahrscheinlichkeitsrechnung, tritt uns die Notwendigkeit entgegen, für den Ausdruck $n!$ eine einfache angenäherte Darstellung durch eine elementare Funktion von $n$ zu besitzen. Einen solchen Ausdruck liefert uns der folgende, nach *Stirling* benannte Satz: Es gilt für $n \to \infty$

$$\frac{n!}{\sqrt{2\pi}\, n^{n+\frac{1}{2}} e^{-n}} \to 1,$$

und zwar ist genauer

$$\sqrt{2\pi}\, n^{n+\frac{1}{2}} e^{-n} < n! \leqq \sqrt{2\pi}\, n^{n+\frac{1}{2}} e^{-n} e^{\frac{1}{4n}}.$$

Mit anderen Worten, da $e^{\frac{1}{4n}}$ mit wachsendem $n$ gegen 1 strebt: der Ausdruck $n!$ unterscheidet sich prozentual immer weniger von $\sqrt{2\pi n}\, n^n e^{-n}$, je größer $n$ ist, oder $n!$ ist „asymptotisch" gleich diesem Ausdruck, und zugleich gibt uns der Faktor $e^{\frac{1}{4n}}$ eine Genauigkeitsschätzung für den erreichten Grad der Annäherung.

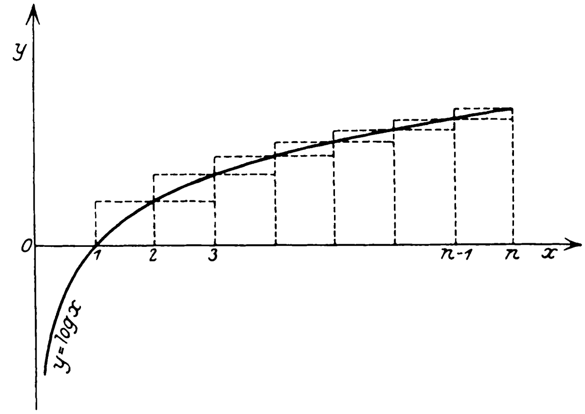


Fig. 106.

Man wird auf diese merkwürdige Formel geführt, wenn man den Ausdruck $\log n! = \log 1 + \log 2 + \cdots + \log n$ unter Heranziehung der geometrischen Deutung mit einem Integral vergleicht. Wir betrachten nämlich die Kurve $y = \log x$ für $x \geqq 1$. Aus Figur 106 entnehmen wir sofort wegen der geometrischen Bedeutung des Integrales für $n \geqq 3$ die doppelte Ungleichung

$$\log(n-1)! = \log 2 + \log 3 + \cdots + \log(n-1) < \int_1^n \log x\, dx$$
$$< \log 2 + \cdots + \log n = \log n!.$$

Da nun

$$\int_1^n \log x\, dx = x(\log x - 1)\Big|_1^n = n \log n - n + 1$$

ist, so ergibt sich

$$\log(n-1)! < n \log n - n + 1 < \log n!,$$

woraus weiter unmittelbar

$$n \log n - n + 1 < \log n! < (n+1)\log(n+1) - n$$

oder

$$e\, n^n e^{-n} < n! < (n+1)^{n+1} e^{-n} = n^{n+1}\left(1 + \frac{1}{n}\right)^{n+1} e^{-n}$$

folgt. Es liegt daher, weil $\left(1 + \frac{1}{n}\right)^{n+1}$ gegen $e$ strebt, die Vermutung nahe, daß $n!$ von der Größenordnung $n^{n+\frac{1}{2}} e^{-n}$ sein wird. Daß dies in der Tat der Fall ist, darin besteht die wesentliche Aussage des Stirlingschen Satzes.

Um ihn nun nach dieser Vorbetrachtung wirklich zu beweisen, haben wir vor allem zu zeigen, daß die Folge der Zahlen

$$a_n = \frac{n!}{n^{n+\frac{1}{2}} e^{-n}}$$

einem Grenzwert zustrebt. Wir bilden zu diesem Zwecke den Quotienten

$$\frac{a_n}{a_{n+1}} = \frac{1}{e}\left(1 + \frac{1}{n}\right)^{n+\frac{1}{2}}$$

und von diesem den Logarithmus

$$\log \frac{a_n}{a_{n+1}} = \left(n + \frac{1}{2}\right) \log\left(1 + \frac{1}{n}\right) - 1 .$$

Zur Abschätzung dieser Größe betrachten wir die Hyperbel $\eta = \frac{1}{\xi}$, von der in Figur 107 das Stück zwischen den Abszissen $\xi = n$ und $\xi = n + 1$ in etwas nach der $\eta$-Richtung vergrößertem Maßstabe gezeichnet ist. Die Kurve ist nach unten hin konvex. Der unter der Kurve liegende, von den Ordinaten $\xi = n$, $\xi = n + 1$ und der $\xi$-Achse begrenzte Flächeninhalt ist also größer als der entsprechende Flächeninhalt des Trapezes unter der im Punkte $\xi = n + \frac{1}{2}$, $\eta = \frac{1}{n + \frac{1}{2}}$ gezogenen Tangente und kleiner als der Flächeninhalt unter der in der Figur gezeichneten Sekante. Da der Flächeninhalt unter der Kurve gleich $\log(n+1) - \log n = \log\left(1 + \frac{1}{n}\right)$ ist, so ergibt dies sofort die Beziehung

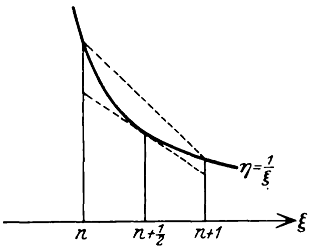


Fig. 107.

$$0 < \log\left(1 + \frac{1}{n}\right) - \frac{1}{n + \frac{1}{2}} < \frac{1}{2}\left(\frac{1}{n} + \frac{1}{n+1}\right) - \frac{1}{n + \frac{1}{2}}$$

oder nach Multiplikation mit $n + \frac{1}{2}$

$$0 < \left(n + \frac{1}{2}\right) \log\left(1 + \frac{1}{n}\right) - 1 < \frac{n + \frac{1}{2}}{2}\left(\frac{1}{n} + \frac{1}{n+1}\right) - 1$$

und nach Umformung der rechten Seite

$$0 < \log \frac{a_n}{a_{n+1}} < \frac{1}{4}\left(\frac{1}{n} - \frac{1}{n+1}\right).$$

Es ist also

$$1 < \frac{a_n}{a_{n+1}} < e^{\frac{1}{4}\left(\frac{1}{n} - \frac{1}{n+1}\right)}.$$

Hieraus aber ergibt sich unmittelbar, daß $a_n$ mit wachsendem $n$ gegen einen Grenzwert $\alpha = \lim_{n\to\infty} a_n$ konvergiert, da die Zahlen $a_n$ eine monoton abnehmende Folge bilden und positiv sind. Schreiben wir noch die entsprechenden Beziehungen für die Indizes $n+1, \ldots, n+k-1$ hin und multiplizieren, so folgt sogleich

$$1 < \frac{a_n}{a_{n+k}} < e^{\frac{1}{4}\left(\frac{1}{n} - \frac{1}{n+k}\right)}$$

und erst recht

$$1 < \frac{a_n}{a_{n+k}} < e^{\frac{1}{4n}},$$

für jedes positive, beliebig groß gewählte $k$. Es gilt also wegen des monotonen Charakters der Folge

$$1 < \frac{a_n}{\alpha} \leqq e^{\frac{1}{4n}}.$$

Um endlich den Grenzwert $\alpha$, von dem wir zunächst nur die Existenz bewiesen haben, wirklich zu bestimmen, greifen wir auf die im vierten Kapitel bewiesene Wallissche Formel bzw. auf die aus ihr gefolgerte Formel

$$\sqrt{\pi} = \lim_{n\to\infty} \frac{(n!)^2\, 2^{2n}}{(2n)!\,\sqrt{n}}$$

zurück. Ersetzen wir hier $n!$ durch $a_n n^{n+\frac{1}{2}} e^{-n}$ und $(2n)!$ durch $a_{2n} 2^{2n+\frac{1}{2}} n^{2n+\frac{1}{2}} e^{-2n}$, so ergibt sich sofort

$$\sqrt{\pi} = \lim_{n\to\infty} \frac{a_n^2}{a_{2n}\sqrt{2}} = \frac{\alpha^2}{\alpha\sqrt{2}}$$

oder $\alpha = \sqrt{2\pi}$. Damit ist die zu Beginn des Paragraphen aufgestellte Behauptung vollständig bewiesen.

Die Stirlingsche Formel ist — abgesehen von ihrer großen theoretischen Bedeutung — ein sehr nützliches Hilfsmittel, wenn es sich um die numerische Berechnung von $n!$ für große Werte von $n$ handelt. Man braucht dann nicht tatsächlich die ganzen Zahlen miteinander zu multiplizieren, sondern lediglich mit Hilfe der Logarithmentafel den Stirlingschen Ausdruck zu berechnen. So wird z. B. für $n = 10$, mit siebenstelligen Logarithmentafeln berechnet, $n! \approx 3598696$, während der genaue Wert $n! = 3628800$ ist. Der prozentuale Fehler ist kaum $^5/_6\,\%$.

Achtes Kapitel.

# Unendliche Reihen und andere Grenzprozesse.

## Vorbemerkungen.

Die geometrische Reihe, die Taylorsche Reihenentwicklung und eine Anzahl spezieller Beispiele, die uns in diesem Buche bisher begegnet sind, legen es nahe, von einem etwas allgemeineren Standpunkte aus diejenigen besonderen Grenzwertbildungen zu studieren, die man als ***unendliche Reihen*** bezeichnet. Im Prinzip läßt sich jeder Grenzwert

$$S = \lim_{n \to \infty} s_n$$

als unendliche Reihe schreiben; wir brauchen, wenn $n$ etwa von 1 an läuft, nur $s_n = s_{n-1} + a_n$ (für $n > 1$) zu setzen und $s_1 = a_1$ zu wählen, dann ist

$$s_n = a_1 + a_2 + \cdots + a_n,$$

und der Wert $S$ erscheint als Grenzwert der Summe $s_n$ aus $n$ Gliedern. Man drückt diese Tatsache aus, indem man sagt: $S$ ist die „Summe der unendlichen Reihe"

$$a_1 + a_2 + a_3 + \cdots.$$

Es ist also eine unendliche Reihe nur eine Form der Darstellung eines Grenzwertes derart, daß jeder folgende Näherungswert aus dem vorangehenden durch den einfachen Prozeß der Addition eines weiteren Gliedes entsteht. Das Prinzip der Dezimalbruchentwicklung z. B. ist nichts anderes als die Darstellung einer Zahl $a$ in der Form einer unendlichen Reihe $a = a_1 + a_2 + a_3 + \cdots$, wobei z. B., wenn $0 \leqq a \leqq 1$ war, $a_n = \alpha_n \cdot 10^{-n}$ gesetzt ist und $\alpha_n$ eine der Zahlen zwischen 0 und 9 bedeutet. Da sich jeder Grenzwert in der Form einer unendlichen Reihe schreiben läßt, könnte man eine besondere Behandlung der Reihen für überflüssig halten; aber es zeigt sich, daß in vielen Fällen ganz von selbst Grenzwerte sich als unendliche Reihen darbieten und daß durch solche Darstellungen besonders einfache Gesetzmäßigkeiten hervortreten. Natürlich wird nicht *jede* Reihenentwicklung übersichtliche Gesetzmäßigkeiten zeigen. Z. B. läßt sich zwar die Zahl $\pi$ an und für sich durch einen unendlichen Dezimalbruch darstellen; wir kennen jedoch kein einfaches Gesetz, welches uns ermöglicht, eine beliebige, etwa die 7000. Ziffer dieser Dezimalbruchentwicklung anzugeben. Verzichten wir jedoch auf die Darstellung von $\pi$ durch einen Dezimalbruch und nehmen wir dafür etwa die Leibnizsche Reihe, so erhalten wir eine Darstellung mit vollständig übersichtlichem allgemeinem Bildungsgesetz.

Ganz ähnlich wie für unendliche Reihen, bei denen die Annäherung an den Grenzwert durch fortwährendes Addieren neuer Glieder geschieht, liegt der Sachverhalt bei ***unendlichen Produkten***, wo die Annäherung an den Grenzwert durch fortgesetztes Multiplizieren mit weiteren Faktoren erfolgt. Im übrigen werden wir im folgenden nur nebenbei auf unendliche Produkte eingehen können. Der Hauptgegenstand dieses und des folgenden Kapitels werden unendliche Reihen sein.

## § 1. Die Begriffe Konvergenz und Divergenz.

### 1. Grundbegriffe.

Wir betrachten eine unendliche Reihe, deren „*allgemeines Glied*" wir mit $a_n$ bezeichnen[1]; die Reihe hat dann formal die Gestalt

$$a_1 + a_2 + \cdots = \sum_{\nu=1}^{\infty} a_\nu .$$

Das Summenzeichen bedeutet ähnlich wie früher, daß man für $\nu$ der Reihe nach alle Werte $1, 2, 3, \ldots$ einsetzen und dann summieren soll.

Je nachdem nun die „*n-te Partialsumme*"

$$s_n = a_1 + a_2 + \cdots + a_n = \sum_{\nu=1}^{n} a_\nu$$

mit wachsendem $n$ einem Grenzwert

$$S = \lim_{n\to\infty} s_n$$

zustrebt oder nicht, heißt die Reihe *konvergent* oder *divergent*, und im ersten Falle nennen wir $S$ die *Summe der Reihe*.

Beispiele für konvergente Reihen haben wir vielfach kennengelernt; etwa die geometrische Reihe $1 + q + q^2 + \cdots$ mit der Summe $\frac{1}{1-q}$, falls $|q| < 1$ ist, die Leibnizsche Reihe, die Reihe für $\log 2$ oder für $e$ und andere.

Das Konvergenzprinzip von Cauchy (vgl. Kap. I, § 6) sagt in der Sprache der unendlichen Reihen aus: *Für die Konvergenz der Reihe ist notwendige und hinreichende Bedingung, daß die Zahl*

$$|s_m - s_n| = |a_{n+1} + a_{n+2} + \cdots + a_m|$$

*beliebig klein wird, wenn nur $m$ und $n$ beide hinreichend groß gewählt werden* ($m > n$). Mit anderen Worten: *Konvergenz der Reihe besteht dann und nur dann, wenn folgende Bedingung erfüllt ist: Wie klein auch immer man eine Genauigkeitsschranke $\varepsilon > 0$ vorgibt, es läßt sich stets zu ihr*

[1] Aus formalen Gründen lassen wir dabei auch zu, daß gewisse der Zahlen $a_n$ Null sind. Falls alle $a_n$ von einer Zahl $N$ ab, d. h. für $n > N$, verschwinden, sprechen wir von einer abbrechenden Reihe.

*ein Index $N = N(\varepsilon)$ — der im allgemeinen für $\varepsilon \to 0$ über alle Grenzen streben wird — derart wählen, daß der obige Ausdruck $|s_m - s_n|$ kleiner als $\varepsilon$ wird, wenn nur zugleich $m > N$ und $n > N$ ist; gleichgültig, wie wir sonst diese beiden Zahlen $m$ und $n$ wählen.*

Machen wir uns die Bedeutung des Konvergenzkriteriums z. B. an der geometrischen Reihe klar, indem wir speziell $q = \frac{1}{2}$ setzen. Wählen wir $\varepsilon = \frac{1}{10}$, so brauchen wir nur $N = 4$ zu wählen. Es ist nämlich

$$|s_m - s_n| = \frac{1}{2^n} + \cdots + \frac{1}{2^{m-1}} = \frac{1}{2^{n-1}}\left(\frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{m-n}}\right) < \frac{1}{2^{n-1}}$$

und
$$\frac{1}{2^{n-1}} < \frac{1}{10}, \text{ sobald } n > N = 4 \text{ ist.}$$

Geben wir als Genauigkeitsschranke $\varepsilon = \frac{1}{100}$ an, so genügt es, entsprechend für $N$ die Zahl 7 zu nehmen, wie man leicht nachprüfen kann.

Selbstverständlich ist es für die Konvergenz einer Reihe eine *notwendige Bedingung*, daß

$$\lim_{n \to \infty} a_n = 0$$

ist. Denn sonst kann das Konvergenzkriterium gewiß nicht erfüllt sein. Aber diese notwendige Bedingung ist *keineswegs hinreichend* für die Konvergenz; vielmehr kann man sehr leicht unendliche Reihen angeben, deren allgemeines Glied $a_n$ mit wachsendem $n$ gegen 0 strebt und deren Summe nicht existiert, indem die Partialsumme $s_n$ mit wachsendem $n$ über alle Grenzen strebt. Ein Beispiel dafür ist die Reihe

$$1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{n}} + \cdots,$$

bei welcher also das allgemeine Glied durch $\frac{1}{\sqrt{n}}$ gegeben wird. Offenbar ist

$$s_n > \frac{1}{\sqrt{n}} + \cdots + \frac{1}{\sqrt{n}} = \frac{n}{\sqrt{n}} = \sqrt{n}.$$

Es strebt also bei wachsendem $n$ die $n$-te Partialsumme über alle Grenzen, und die Reihe divergiert daher.

Dasselbe gilt für das klassische Beispiel der *harmonischen Reihe*

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots.$$

Es ist $a_{n+1} + \cdots + a_{2n} = \frac{1}{n+1} + \cdots + \frac{1}{2n} > \frac{1}{2n} + \cdots + \frac{1}{2n} = \frac{1}{2}$. Da $n$ und $m = 2n$ beliebig groß genommen werden können, divergiert also die Reihe, weil das Kriterium von Cauchy nicht erfüllt ist, und zwar strebt die $n$-te Partialsumme offenbar gegen Unendlich, da alle

Glieder positiv sind. Hingegen *konvergiert* die mit abwechselndem Vorzeichen aus denselben Zahlen gebildete Reihe

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - + \cdots + \frac{(-1)^{n-1}}{n} + \cdots$$

nach dem sechsten Kapitel, S. 251, und besitzt die Summe log 2.

Die Divergenz einer Reihe braucht sich keineswegs immer dadurch auszudrücken, daß $s_n$ mit wachsendem $n$ gegen $+\infty$ oder $-\infty$ strebt. So sehen wir bei der Reihe

$$1 - 1 + 1 - 1 + 1 - 1 + - \cdots,$$

daß die Partialsumme $s_n$ abwechselnd die Werte 1 und 0 besitzt und wegen dieses Hin- und Herpendelns weder einem bestimmten Grenzwert zustrebt noch absolut genommen über alle Grenzen wächst.

Hinsichtlich der Konvergenz und Divergenz einer unendlichen Reihe mache ich noch folgende zwar selbstverständliche, aber doch prinzipiell wichtige Bemerkung: An der Tatsache der Konvergenz und Divergenz wird nichts geändert, wenn man eine endliche Anzahl von Gliedern zu der Reihe hinzufügt oder aus ihr fortnimmt. Es ist also hinsichtlich der Frage der Konvergenz und Divergenz ganz gleichgültig, ob man die Reihe mit einem Glied $a_0$ oder erst mit dem Glied $a_1$ oder $a_5$ oder einem beliebigen anderen beginnt.

## 2. Absolute und bedingte Konvergenz.

Die Reihe mit den Gliedern $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ divergiert, sobald wir sie mit positiven, konvergiert, sobald wir sie mit abwechselnden Vorzeichen versehen. Anders liegt es bei der geometrischen Reihe, wo für $|q| < 1$ gleichzeitig mit der Reihe $\sum\limits_{\nu=0}^{\infty} (-1)^\nu q^\nu = \frac{1}{1+q}$ die Reihe $\sum\limits_{\nu=0}^{\infty} q^\nu$ konvergent ist und die Zahl $\frac{1}{1-q}$ darstellt.

Hier kommt ein Unterschied zum Vorschein, den wir ein wenig näher betrachten müssen. Bei einer Reihe, deren Glieder sämtlich positiv sind, sind offenbar nur zwei Fälle möglich: entweder sie konvergiert, oder ihre Partialsumme $s_n$ strebt mit wachsendem $n$ über alle Grenzen. Denn die Partialsummen müssen als monoton wachsende Folge konvergieren, sobald sie beschränkt bleiben. Konvergenz herrscht dann, wenn die Glieder mit wachsendem $n$ hinreichend rasch gegen 0 streben, Divergenz dagegen, wenn sie dies überhaupt nicht oder zu langsam tun. Bei Reihen aber, deren Glieder teils positives, teils negatives Vorzeichen haben, kann die Konvergenz durch den Wechsel dieser Vorzeichen erzeugt werden, indem ein zu starkes Anwachsen der Partialsummen, welches von den positiven Gliedern herrührt, durch die negativen Glieder kompensiert wird, so daß schließlich doch ein bestimmter Grenzwert zustande kommt.

Um diese Tatsache besser zu erfassen, ordnen wir einer Reihe $\sum_{\nu=1}^{\infty} a_\nu$ mit positiven und negativen Gliedern die *Reihe der Absolutbeträge* ihrer Glieder zu, d. h. die Reihe

$$|a_1| + |a_2| + \cdots = \sum_{\nu=1}^{\infty} |a_\nu|.$$

Konvergiert diese Reihe, so wird bei hinreichend großem $n$ und $m > n$ sicherlich der Ausdruck

$$|a_{n+1}| + |a_{n+2}| + \cdots + |a_m|$$

beliebig klein sein; es wird daher wegen der Beziehung

$$|a_{n+1} + \cdots + a_m| \leqq |a_{n+1}| + \cdots + |a_m|$$

auch der Ausdruck hier links beliebig klein, und somit konvergiert sicherlich auch die ursprüngliche Reihe $\sum_{\nu=1}^{\infty} a_\nu$. Die ursprüngliche Reihe heißt in diesem Falle *absolut konvergent.* Ihre Konvergenz ist durch die absolute Kleinheit ihrer Glieder gesichert und besteht unabhängig von dem Vorzeichenwechsel.

Wenn dagegen die Reihe der Absolutglieder divergiert und die ursprüngliche Reihe doch konvergiert, so nennen wir die ursprüngliche Reihe *bedingt konvergent.* Die bedingte Konvergenz ist eine Folge der durch den Vorzeichenwechsel sich ergebenden Kompensationen.

Über bedingte Konvergenz sei hier nur der folgende häufig nützliche *Konvergenzsatz von Leibniz* angeführt: *Wenn die Vorzeichen der Reihenglieder abwechseln und außerdem der absolute Betrag* $|a_n|$ *monoton gegen* 0 *strebt (also* $|a_{n+1}| < |a_n|$ *ist), so konvergiert die Reihe* $\sum_{\nu=1}^{\infty} a_\nu$. (Beispiel: Die Leibnizsche Reihe.)

Zum Beweis nehmen wir ohne wesentliche Beschränkung der Allgemeinheit $a_1 > 0$ an und schreiben unsere Reihe in der Form

$$b_1 - b_2 + b_3 - + \cdots,$$

wo nunmehr die Glieder $b_n$ sämtlich positiv sind, gegen 0 streben und der Bedingung $b_{n+1} < b_n$ genügen. Indem wir die Glieder auf die folgenden beiden Arten zusammenfassen:

$$b_1 - (b_2 - b_3) - (b_4 - b_5) - \cdots$$

und

$$(b_1 - b_2) + (b_3 - b_4) + (b_5 - b_6) + \cdots,$$

erkennen wir sofort, daß für die Teilsummen folgende Beziehungen bestehen:

$$s_1 > s_3 > s_5 > \cdots > s_{2m+1} > \cdots,$$
$$s_2 < s_4 < s_6 < \cdots < s_{2m} < \cdots,$$

während andererseits $s_{2n} < s_{2n+1}$ und $s_{2n} < s_{2n-1}$ ist. Es bilden also die ungeraden Teilsummen $s_1, s_3, \ldots$ eine monoton absteigende Folge, die jedenfalls nicht unter den Wert $s_2$ herabsinkt; daher muß diese Folge einen Grenzwert $G$ besitzen (vgl. erstes Kapitel, § 6). Ebenso bilden die geraden Teilsummen $s_2, s_4, \ldots$ eine monoton aufsteigende Zahlenfolge, deren Glieder jedenfalls nicht größer als die feste Zahl $s_1$ werden; diese Folge muß also ebenfalls einen Grenzwert $G'$ haben. Da sich nun die Zahlen $s_{2n}$ und $s_{2n+1}$ nur um die bei wachsendem $n$ gegen 0 strebende Zahl $b_{2n+1}$ voneinander unterscheiden, so müssen die beiden Grenzwerte $G$ und $G'$ einander gleich sein. D. h. die geraden und die ungeraden Teilsummen streben demselben Grenzwert zu, den wir nunmehr mit $S$ bezeichnen (vgl. Fig. 108). Das aber bedeutet nichts anderes als die behauptete Konvergenz unserer Reihe; ihre Summe ist gleich $S$.

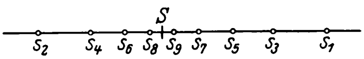


Fig. 108. Konvergenz der „alternierenden" Reihen.

Zum Schluß noch eine allgemeine Bemerkung über den grundsätzlichen Unterschied von absolut konvergenten und bedingt konvergenten Reihen: Wir bezeichnen die positiven Glieder der von uns betrachteten Reihe $\sum_{\nu=1}^{\infty} a_\nu$ der Reihe nach mit $p_1, p_2, p_3, \ldots$, die negativen Glieder mit $-q_1, -q_2, -q_3, \ldots$. Bilden wir dann die $n$-te Partialsumme $s_n = \sum_{\nu=1}^{n} a_\nu$ der gegebenen Reihe, so müssen in deren Summanden eine gewisse Anzahl, etwa $n'$, von positiven und eine gewisse Anzahl, etwa $n''$, von negativen Gliedern auftreten, und es muß dabei $n' + n'' = n$ sein; ferner werden, wenn sowohl die Anzahl der positiven Glieder als auch die Anzahl der negativen in der Reihe unendlich ist, mit $n$ zugleich beide Zahlen $n'$ und $n''$ über alle Grenzen wachsen. Die Partialsumme $s_n$ ist nun, wie wir sofort sehen, einfach gleich der Partialsumme $\sum_{\nu=1}^{n'} p_\nu$ der Reihe der positiven Glieder, vermehrt um die Partialsumme $-\sum_{\nu=1}^{n''} q_\nu$ der Reihe der negativen Glieder. Wenn die gegebene Reihe absolut konvergiert, so konvergieren sicher auch die Reihe ihrer positiven Glieder $\sum_{\nu=1}^{\infty} p_\nu$ und die Reihe der absoluten Beträge ihrer negativen Glieder $\sum_{\nu=1}^{\infty} q_\nu$ jede für sich[1]); *die Summe der gegebenen Reihe ist dann einfach gleich der Summe der nur aus den positiven und der nur aus den negativen Gliedern gebildeten Reihe,* mit anderen Worten: *die Differenz zweier*

---

[1]) Denn die Teilsummen $\sum_{\nu=1}^{m} p_\nu$ und $\sum_{\nu=1}^{m} q_\nu$ sind bei wachsendem $m$ monoton nicht abnehmende Zahlenfolgen mit der oberen Schranke $\sum_{\nu=1}^{\infty} |a_\nu|$.

*Reihen mit positiven Gliedern*[1]). Besitzt die Reihe nur endlich viele Glieder von einem der beiden Vorzeichen, so vereinfacht sich der Sachverhalt entsprechend.

Ist dagegen die Reihe nicht mehr absolut, sondern nur noch bedingt konvergent, so müssen notwendig die beiden Reihen $\sum_{\nu=1}^{\infty} p_\nu$ und $\sum_{\nu=1}^{\infty} q_\nu$ divergent sein. Denn wären beide konvergent, so würde die gegebene Reihe absolut konvergieren, entgegen der Voraussetzung. Wäre nur eine divergent, etwa die Reihe $\sum_{\nu=1}^{\infty} p_\nu$, die andere aber konvergent, so zeigt die Zerlegung $s_n = \sum_{\nu=1}^{n'} p_\nu - \sum_{\nu=1}^{n''} q_\nu$, daß die gegebene Reihe überhaupt nicht konvergieren könnte; denn bei wachsendem $n$ streben $n'$ und $\sum_{\nu=1}^{n'} p_\nu$ über alle Grenzen, während der Bestandteil $\sum_{\nu=1}^{n''} q_\nu$ einem bestimmten Grenzwert zustrebt, so daß die Teilsumme $s_n$ mit wachsendem $n$ über alle Grenzen wachsen müßte.

Wir sehen also, daß eine *bedingt konvergente Reihe* sich *nicht* als *Differenz zweier aus ihren Reihengliedern gebildeten konvergenten Reihen mit jeweils nur positiven Gliedern* auffassen läßt. Mit dieser Tatsache hängt eng ein weiterer Unterschied zwischen absoluter und bedingter Konvergenz zusammen, den ich jetzt noch kurz erörtern möchte.

### 3. Umordnung der Reihenglieder.

Es ist eine Eigenschaft endlicher Summen, daß wir die Reihenfolge der Glieder beliebig ändern oder, wie man sich ausdrückt, die endliche Summe beliebig *umordnen* können. Es entsteht die Frage, welchen Sinn der Begriff der Umordnung für eine unendliche Reihe besitzt und ob bei einer solchen Umordnung der Summenwert der Reihe geändert wird. Während es bei einer endlichen Summe ohne weiteres einen Sinn hat, die Summanden z. B. in umgekehrter Reihenfolge zu addieren, fällt eine solche Möglichkeit bei einer unendlichen Reihe fort; es gibt eben kein letztes Glied, mit dem man anfangen könnte. Vielmehr wird eine Umordnung nur folgendes bedeuten können: Wir sagen, eine Reihe $a_1 + a_2 + a_3 + \cdots$ geht durch Umordnung in eine Reihe $b_1 + b_2 + b_3 + \cdots$ über, wenn jedes Glied $a_n$ der ersten Reihe auch genau einmal in der zweiten vorkommt und umgekehrt. Es kann z. B. dabei von seinem Platze um so weiter verschoben werden, je größer $n$ ist. Nur muß es irgend einmal in der umgeordneten Reihe wieder erscheinen, während andererseits in dieser Reihe Glieder aus

---

[1]) Denn es ist $\sum_{\nu=1}^{n} a_\nu = \sum_{\nu=1}^{n'} p_\nu - \sum_{\nu=1}^{n''} q_\nu$; bei wachsendem $n$ wachsen auch $n'$ und $n''$ über alle Grenzen, und der Grenzwert der linken Seite muß also gleich der Differenz der Grenzwerte der beiden Summen rechts sein.

der ersten Reihe an einer früheren Stelle als an der ursprünglichen auftreten. So z. B. stellt die Reihe

$$1 + q + q^2 + q^4 + q^3 + q^8 + q^7 + q^6 + q^5 + q^{16} + \cdots$$

eine Umordnung der geometrischen Reihe $1 + q + q^2 + \cdots$ dar.

Es besteht nun hinsichtlich dieser Umordnung ein grundlegender Unterschied zwischen absolut und bedingt konvergenten Reihen.

*Bei absolut konvergenten Reihen wird durch eine Umordnung die Konvergenz nicht gestört, und die Summe der Reihe behält ihren Wert bei, genau so wie das für eine endliche Summe gilt.*

*Bei einer bedingt konvergenten Reihe hingegen kann man stets durch eine geeignete Umordnung den Summenwert der Reihe beliebig ändern und, wenn man will, auch die Reihe divergent machen.*

Die erste, auf absolut konvergente Reihen bezügliche Tatsache ist sehr leicht einzusehen. Setzen wir zunächst voraus, daß unsere Reihe nur positive Glieder hat, und betrachten die $n$-te Teilsumme $s_n = \sum_{\nu=1}^{n} a_\nu$. Dann kommen gewiß alle Summanden dieser Teilsumme auch in der $m$-ten Teilsumme $t_m = \sum_{\nu=1}^{m} b_\nu$ der umgeordneten Reihe vor, wenn nur $m$ hinreichend groß genommen wird. Es ist dann sicherlich $t_m \geqq s_n$. Andererseits aber kann man zu dem Index $m$ wiederum einen Index $n'$ hinzubestimmen, so daß in der Teilsumme $s_{n'} = \sum_{\nu=1}^{n'} a_\nu$ der ersten Reihe alle Summanden $b_1, b_2, \ldots, b_m$ enthalten sind. Es ist also dann sicherlich $s_{n'} \geqq t_m$. Jede Teilsumme der einen Reihe liegt also zwischen zwei Teilsummen der anderen Reihe, und daraus geht unmittelbar hervor, daß die umgeordnete Reihe konvergiert und denselben Grenzwert hat wie die Ausgangsreihe.

Hat die absolut konvergente Reihe positive und negative Glieder, so ist sie als Differenz zweier Reihen mit jeweils nur positiven Gliedern aufzufassen. Da bei der Umordnung auch jede dieser beiden Reihen nichts anderes als eine Umordnung erfährt, also den Wert beibehält, gilt dasselbe auch für die umgeordnete Ausgangsreihe; denn diese konvergiert nach dem schon erledigten Fall wieder absolut, ist also die Differenz der beiden umgeordneten Teilreihen. Damit ist der Umordnungssatz auch für solche Reihen bewiesen.

Dem Anfänger mag die eben bewiesene Tatsache als eine Selbstverständlichkeit erscheinen. Daß sie eines Beweises bedarf und daß bei diesem Beweis die absolute Konvergenz wesentlich benutzt werden muß, soll uns zunächst ein Beispiel des entgegengesetzten Verhaltens bei bedingt konvergenten Reihen zeigen. Wir nehmen die uns schon bekannte Reihe

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + - \cdots = \log 2,$$

schreiben, indem wir mit dem Faktor $\frac{1}{2}$ multiplizieren, darunter

$$\frac{1}{2} \qquad -\frac{1}{4} \qquad +\frac{1}{6} \qquad -\frac{1}{8}+-\cdots=\frac{1}{2}\log 2$$

und addieren, indem wir die untereinanderstehenden Glieder zusammenziehen[1]). So erhalten wir

$$1+\frac{1}{3}-\frac{1}{2}+\frac{1}{5}+\frac{1}{7}-\frac{1}{4}+\frac{1}{9}+\frac{1}{11}-\frac{1}{6}+\cdots=\frac{3}{2}\log 2.$$

Diese letzte Reihe läßt sich aber offenbar durch eine Umordnung aus der ursprünglichen herstellen, und der Wert der Reihe hat sich dabei mit dem Faktor $\frac{3}{2}$ multipliziert. Man kann sich leicht vorstellen, wie die Entdeckung dieses scheinbaren Paradoxons auf die Mathematiker des 18. Jahrhunderts gewirkt haben muß, welche gewohnt waren, mit unendlichen Reihen ohne Rücksicht auf ihr Konvergenzverhalten zu operieren.

Ich will für den oben formulierten Satz von der Wertänderung bedingt konvergenter Reihen bei Umordnung den Beweis hier durchführen, obwohl wir von dieser Tatsache später keinen Gebrauch zu machen haben. Es seien $p_1, p_2, \ldots$ die positiven Glieder und $-q_1, -q_2, \ldots$ die negativen Glieder der Reihe. Da der absolute Wert $|a_n|$ mit wachsendem $n$ gegen 0 strebt, so müssen auch die Zahlen $p_n$ und $q_n$ mit wachsendem $n$ gegen 0 streben. Ferner muß, wie wir oben sahen, $\sum\limits_1^\infty p_\nu$ [2]) divergieren und ebenso $\sum\limits_1^\infty q_\nu$.

Nunmehr können wir leicht eine Umordnung der ursprünglichen Reihe finden, bei der ein beliebiger Grenzwert $a$ herauskommt. Wir addieren zunächst so viele positive Glieder, bis die Summe $\sum\limits_1^{n_1} p_\nu$ gerade den Wert $a$ übersteigt. Da die Summe $\sum\limits_1^{n_1} p_\nu$ mit wachsendem $n_1$ über alle Grenzen wächst, werden wir sicher nach Benutzung einer gewissen Anzahl von Gliedern wirklich über den Wert $a$ hinauskommen. Von dem genauen Werte $a$ weicht sie dann höchstens um $p_{n_1}$ ab. Dann addieren wir so viele negative Glieder: $-\sum\limits_1^{m_1} q_\nu$, bis die Summe gerade unterhalb des Wertes $a$ liegt; daß dies möglich ist, folgt wiederum aus der Divergenz der Reihe $\sum\limits_1^\infty q_\nu$. Der Fehler gegenüber $a$ beträgt jetzt höchstens $q_{m_1}$. Nunmehr addieren wir von den weiteren positiven Gliedern so viele: $\sum\limits_{n_1+1}^{n_2} p_\nu$, daß die Summe

[1]) Zur Addition von Reihen vgl. Nr. 4.

[2]) Diese abkürzende Schreibweise für $\sum\limits_{\nu=1}^\infty p_\nu$ wird mutatis mutandis auch im folgenden öfters angewandt werden.

wieder gerade oberhalb des Wertes $a$ liegt, was wiederum infolge der Divergenz der Reihe der positiven Glieder möglich ist. Der Fehler gegenüber $a$ beträgt höchstens $p_{n_2}$. Durch Addieren weiterer negativer Glieder: $-\sum_{m_1+1}^{m_2} q_\nu$, von dem zuletzt benutzten angefangen, verkleinern wir die Summe nunmehr so lange, bis sie gerade unterhalb $a$ liegt, und fahren in dieser Weise ins Unbegrenzte fort. Die Summenwerte, die wir so erhalten, werden um den Wert $a$ herumschwanken; und zwar wird, wenn wir nur den Prozeß hinreichend weit geführt haben, diese Schwankung nur noch in beliebig kleinen Grenzen vor sich gehen können; denn da die Glieder $p_\nu$ und $q_\nu$ bei hinreichend großem $\nu$ selbst gegen 0 streben, wird der Spielraum für diese Oszillationen ebenfalls gegen 0 konvergieren. Damit ist der gewünschte Beweis geführt.

Durch ganz dieselbe Methode kann man die Umordnung so vornehmen, daß Divergenz eintritt; man braucht nur z. B. die positiven Glieder gegenüber den negativen so häufig zu machen, daß eine Kompensation nicht mehr eintritt.

### 4. Das Rechnen mit unendlichen Reihen.

Es ist selbstverständlich, daß man zwei konvergente unendliche Reihen $a_1 + a_2 + \cdots = S$ und $b_1 + b_2 + \cdots = T$ gliedweise addieren darf, d. h. daß die Reihe, welche aus den Gliedern $c_n = a_n + b_n$ gebildet ist, konvergiert und als Summenwert die Summe $S + T$ der beiden ersten Reihen besitzt. Denn es ist $\sum_{\nu=1}^{n} c_n = \sum_{\nu=1}^{n} a_n + \sum_{\nu=1}^{n} b_n \to S + T$[1]).

Ferner ist ohne weiteres klar, daß man eine unendliche Reihe mit einem Faktor multiplizieren kann, indem man jedes Glied mit diesem Faktor multipliziert.

Diese Tatsachen sind unabhängig davon, ob die Konvergenz der Reihen absolut oder bedingt ist. Dagegen zeigt ein genaueres, für uns hier nicht nötiges Studium, daß die Multiplikation zweier unendlicher Reihen miteinander nach demselben Schema, nach dem man die Multiplikation endlicher Summen auszuführen pflegt, allgemein nur dann vorgenommen werden darf, wenn mindestens eine der beiden Reihen absolut konvergiert (vgl. Anhang, § 1).

## § 2. Untersuchung der Konvergenz und Divergenz.

Wir haben soeben ein Kriterium von allgemeinem Charakter für die Konvergenz von Reihen kennengelernt, welches sich auf Reihen mit abwechselnden Vorzeichen und abnehmenden Absolutbeträgen der

[1]) Dieser Reihensatz ist eigentlich nur eine andere Formulierung der schon im ersten Kapitel, § 6, Nr. 2, vorgekommenen Tatsache, daß der Grenzwert einer Summe aus zwei Summanden gleich der Summe der beiden entsprechenden Grenzwerte ist.

Glieder bezog und zum mindesten bedingte Konvergenz solcher Reihen aussprach. Nunmehr wollen wir uns lediglich mit Kriterien befassen, welche sich auf absolute Konvergenz beziehen.

## 1. Das Prinzip der Reihenvergleichung.

Alle solchen Konvergenzbetrachtungen beruhen darauf, daß man die betreffende Reihe mit einer zweiten Reihe vergleicht, deren Glieder absolut genommen größer sind als die Glieder der ursprünglichen; man versucht dann, die zweite Reihe geeignet so zu wählen, daß sie hinsichtlich ihrer Konvergenzverhältnisse übersichtlich wird. Das allgemeine *Prinzip der Reihenvergleichung* können wir folgendermaßen aussprechen: *Sind die Zahlen $b_1, b_2, \ldots$ sämtlich positiv und gilt für jedes $n$*

$$|a_n| \leqq b_n,$$

*so ist die Reihe $\sum_{n=1}^{\infty} a_n$ absolut konvergent, wenn die Reihe $\sum_{n=1}^{\infty} b_n$ konvergiert.*

Der Beweis ist bei Verwendung des Cauchyschen Kriteriums fast selbstverständlich. Für $m \geqq n$ ist nämlich

$$|a_n + \cdots + a_m| \leqq |a_n| + \cdots + |a_m| \leqq b_n + \cdots + b_m,$$

und da bei hinreichend großem $n$ und $m$ die rechte Seite beliebig klein wird, vorausgesetzt, daß die Reihe $\sum_{n=1}^{\infty} b_n$ konvergiert, so folgt auch die beliebige Kleinheit der linken Seite und damit die Konvergenz der gegebenen Reihe; daß diese Konvergenz eine absolute ist, ergibt sich aus der Tatsache, daß unsere Betrachtung zugleich die Konvergenz der Reihe aus den Absolutbeträgen $|a_n|$ zeigt.

Den analogen Beweis für die folgende Tatsache kann ich dem Leser überlassen. *Ist*

$$|a_n| \geqq b_n > 0,$$

*so ist die Reihe $\sum_{n=1}^{\infty} a_n$ sicher nicht absolut konvergent, wenn die Reihe $\sum_{n=1}^{\infty} b_n$ divergiert.*

## 2. Vergleichung mit der geometrischen Reihe.

Die häufigste Anwendung findet unser Prinzip, indem wir als Vergleichsreihe eine geometrische Reihe wählen. Es folgt dann sofort folgender Satz: Die Reihe $\sum_{n=1}^{\infty} a_n$ ist absolut konvergent, wenn — von einem gewissen Gliede an — ständig eine Beziehung der Form

$$\text{I} \qquad |a_n| < c \cdot q^n$$

gilt, wobei $c$ eine von $n$ unabhängige positive Zahl und $q$ einen positiven „echten Bruch", d. h. *irgend* eine Zahl zwischen 0 und 1, bedeutet. Gewöhnlich bringt man dieses Kriterium in eine der beiden folgenden, schwächeren Formen: Die Reihe $\sum_{n=1}^{\infty} a_n$ konvergiert absolut, wenn von einem gewissen Gliede an eine Beziehung der Form

$$\text{IIa} \qquad \left|\frac{a_{n+1}}{a_n}\right| < q$$

gilt, wobei wieder $q$ ein positiver, von $n$ unabhängiger echter Bruch ist, oder: wenn von einer gewissen Stelle ab mit einem positiven echten Bruch $q$ eine Beziehung der Form

$$\text{IIb} \qquad \sqrt[n]{|a_n|} < q$$

besteht. Die Bedingungen dieser Kriterien sind gewiß z. B. dann erfüllt, wenn eine Beziehung der Form

$$\text{IIIa} \qquad \lim_{n \to \infty} \left|\frac{a_{n+1}}{a_n}\right| = k < 1$$

oder der Form

$$\text{IIIb} \qquad \lim_{n \to \infty} \sqrt[n]{|a_n|} = k < 1$$

vorliegt. Der Beweis für unsere Behauptungen ergibt sich sehr leicht folgendermaßen.

Das Kriterium IIa, das *Quotientenkriterium*, möge etwa von dem Index $n_0$ ab, d. h. für $n > n_0$, erfüllt sein. Dann setzen wir zur Abkürzung $a_{n_0+m+1} = b_m$ und haben

$$|b_1| < q\,|b_0|, \qquad |b_2| < q\,|b_1| < q^2\,|b_0|, \qquad |b_3| < q\,|b_2| < q^3\,|b_0|,$$

usw., also

$$|b_m| < q^m\,|b_0|,$$

womit unsere Behauptung erwiesen ist. Bei dem Kriterium IIb, dem *Wurzelkriterium*, ergibt sich sofort $|a_n| < q^n$ und damit die Richtigkeit unserer Behauptung.

Um endlich die Kriterien III zu beweisen, betrachten wir eine beliebige Zahl $q$ derart, daß $k < q < 1$ ist. Dann wird von einem gewissen $n_0$ ab, d. h. für $n > n_0$, sicherlich auch $\left|\frac{a_{n+1}}{a_n}\right| < q$ bzw. $\sqrt[n]{|a_n|} < q$ sein, da der Wert der Zahlen $\left|\frac{a_{n+1}}{a_n}\right|$ bzw. $\sqrt[n]{|a_n|}$ sich bei genügend großem $n$ beliebig wenig von $k$ unterscheidet. Somit ist der Beweis durch Zurückführung auf das soeben gewonnene allgemeinere Ergebnis geführt.

Ich mache darauf aufmerksam, daß die vier aus dem ursprünglichen Kriterium $|a_n| < cq^n$ gewonnenen Kriterien nicht miteinander oder mit dem ursprünglichen gleichwertig sind, d. h. daß sie sich nicht wechselseitig auseinander ableiten lassen. Wir werden sogleich an Beispielen sehen, daß bei einer Reihe, bei der das eine von ihnen erfüllt ist, keineswegs jedes der anderen erfüllt zu sein braucht[1]).

Zur Ergänzung bemerke ich noch, daß eine Reihe sicherlich divergiert, wenn von einem gewissen Gliede ab bei passend gewähltem positiven $c$

$$|a_n| > c$$

oder wenn von einem gewissen Gliede ab

$$\sqrt[n]{|a_n|} > 1$$

oder wenn

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = k \quad \text{oder} \quad \lim_{n\to\infty} \sqrt[n]{|a_n|} = k$$

ist, wo $k$ eine Zahl oberhalb 1 bedeutet. Denn, wie man sofort erkennt, können bei einer solchen Reihe die Glieder mit wachsendem $n$ nicht gegen Null streben; die Reihe muß also divergieren. (Auch bedingte Konvergenz kommt also nicht in Frage.)

Unsere Kriterien stellen im übrigen lediglich hinreichende Bedingungen für die absolute Konvergenz einer Reihe dar; d. h. man kann aus ihrer Erfüllung die absolute Konvergenz schließen. Jedoch sind sie durchaus keine notwendigen Kriterien; d. h. es kann sehr wohl absolut konvergente Reihen geben, bei denen sie nicht erfüllt sind.

Es läßt sich z. B. keine allgemeine Aussage mehr machen, wenn

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 1 \quad \text{oder} \quad \lim_{n\to\infty} \sqrt[n]{|a_n|} = 1$$

ist. Derartige Reihen können konvergent oder divergent sein. Z. B. ist die Reihe

$$\sum_{n=1}^{\infty} \frac{1}{\nu},$$

bei welcher $\lim\limits_{n\to\infty} \sqrt{|a_n|} = 1$ und $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 1$ ist, nach § 1, Nr. 1 divergent. Dagegen werden wir sogleich sehen, daß die Reihe $\sum\limits_{\nu=1}^{\infty} \frac{1}{\nu^2}$, für welche dieselben Beziehungen bestehen, konvergiert.

Als Beispiel für die Anwendung unserer Kriterien nenne ich zunächst die Reihe

$$q + 2q^2 + 3q^3 + \cdots + n q^n + \cdots.$$

Es ist

$$\lim_{n\to\infty} \sqrt[n]{|a_n|} = |q| \cdot \lim_{n\to\infty} \sqrt[n]{n} = |q|,$$

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = |q| \cdot \lim_{n\to\infty} \frac{n+1}{n} = |q|.$$

[1]) Genauer: Wir können zwar aus der Erfüllung von IIIa bzw. IIIb auf die von IIa bzw. IIb und I schließen, aber nicht umgekehrt.

Daß die Reihe konvergiert, sobald $|q| < 1$ ist, folgt sowohl aus dem Quotienten- als auch aus dem Wurzelkriterium, sogar in der schwächeren Form III.

Betrachten wir dagegen die Reihe

$$1 + 2q + q^2 + 2q^3 + \cdots + q^{2n} + 2q^{2n+1} + \cdots,$$

so können wir ihre Konvergenz für $\frac{1}{2} < |q| < 1$ nach dem Quotientenkriterium nicht mehr beweisen; denn es ist ja $\left|\frac{2q^{2n+1}}{q^{2n}}\right| = 2|q| > 1$. Dagegen liefert das Wurzelkriterium sofort $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} = |q|$ und zeigt die Konvergenz der Reihe für $|q| < 1$, was wir natürlich auch direkt hätten einsehen können.

### 3. Vergleichung mit einem Integral[1]).

Neben den eben durchgeführten Konvergenzbetrachtungen hat eine andere, auf sie nicht zurückführbare eine selbständige Bedeutung. Wir führen sie an dem einfachsten und wichtigsten Beispiel der Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} = 1 + \frac{1}{2^\alpha} + \frac{1}{3^\alpha} + \cdots$$

durch, wo also das allgemeine Glied $a_n$ gleich $\frac{1}{n^\alpha}$ ist; unter $\alpha$ soll dabei eine positive Zahl verstanden werden. Um die Konvergenz bzw. Divergenz dieser Reihen zu untersuchen, stellen wir uns die Kurve $y = \frac{1}{x^\alpha}$ vor und markieren auf der $x$-Achse die ganzzahligen Abszissen $x = 1$, $x = 2, \ldots$. Wir legen einmal über das Stück $n - 1 \leqq x \leqq n$ der $x$-Achse ($n > 1$), das andere Mal über die Strecke $n \leqq x \leqq n + 1$ das Rechteck von der Höhe $\frac{1}{n^\alpha}$ und vergleichen es mit dem (in der Figur 109 schräg bzw. gekreuzt schraffierten) Gebiet, das von demselben Stück der $x$-Achse, den beiden Ordinaten in den Endpunkten und dem von ihnen ausgeschnittenen Stück der Kurve $y = \frac{1}{x^\alpha}$ begrenzt

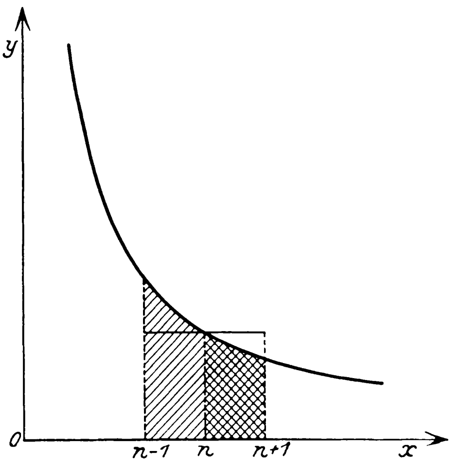


Fig. 109. Vergleichung einer Reihe mit einem Integral.

wird. Offenbar ist der Flächeninhalt des genannten Gebietes im ersten Falle größer, im zweiten kleiner als der Inhalt des Rechtecks, der

---

[1]) Vgl. hierzu auch den Anhang zum siebenten Kapitel.

gerade $\frac{1}{n^\alpha}$ beträgt. Es ist mit anderen Worten

$$\int_n^{n+1} \frac{dx}{x^\alpha} < \frac{1}{n^\alpha} = a_n < \int_{n-1}^{n} \frac{dx}{x^\alpha},$$

wie man natürlich auch direkt an den Integralen erkennt. (Vgl. zweites Kapitel, § 7, Nr. 1.) Somit erhalten wir für die $n$-te Partialsumme $s_m = \sum_{n=1}^{m} \frac{1}{n^\alpha}$, indem wir die Ungleichungen für $n = 2, n = 3, \ldots, n = m$ summieren, sofort die Beziehung[1])

$$1 + \int_2^{m+1} \frac{1}{x^\alpha} dx < s_m < 1 + \int_1^{m} \frac{1}{x^\alpha} dx.$$

Nun strebt (vgl. viertes Kapitel, § 8) bei wachsendem $m$ das Integral $\int_1^m \frac{1}{x^\alpha} dx$ einem endlichen Grenzwert zu oder nicht, je nachdem, ob $\alpha > 1$ oder $\alpha \leqq 1$ ist. Die monotone Folge der Zahlen $s_m$ ist also bzw. beschränkt oder über alle Grenzen wachsend, und wir entnehmen hieraus sofort den folgenden Satz: Die Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} = \frac{1}{1^\alpha} + \frac{1}{2^\alpha} + \frac{1}{3^\alpha} + \cdots$$

ist dann und nur dann konvergent — und zwar natürlich absolut konvergent —, wenn $\alpha > 1$ ist. Die oben auf andere Art bewiesene Divergenz der harmonischen Reihe ist, wie wir sehen, eine unmittelbare Folge hiervon. Speziell folgt im übrigen die Tatsache, daß die Reihen

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots,$$
$$\frac{1}{1^3} + \frac{1}{2^3} + \frac{1}{3^3} + \cdots,$$
$$\cdots\cdots\cdots\cdots$$

konvergieren.

---

1) Aus dieser Beziehung für $\alpha = 1$ folgt übrigens unmittelbar, daß die Zahlenfolge $C_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} - \log n$ nach unten beschränkt ist. Da sie überdies wegen $\frac{1}{n+1} < \int_n^{n+1} \frac{dx}{x} = \log(n+1) - \log n$ bei wachsendem $n$ monoton abnimmt, so existiert

$$\lim_{n\to\infty} C_n = \lim_{n\to\infty} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n} - \log n\right) = C.$$

Diese Zahl $C$, deren Wert 0,5772 ... ist, heißt die *Eulersche Konstante.* Im Gegensatz zu anderen wichtigen besonderen Zahlen der Analysis wie $e$ oder $\pi$ ist es bei der Eulerschen Konstanten nicht gelungen, neben der Definitionsgleichung andere Darstellungen mit einfachen arithmetischen Bildungsgesetzen zu finden.

Die eben hinsichtlich ihrer Konvergenz untersuchten Reihen $\sum_{\nu=1}^{\infty} \frac{1}{\nu^\alpha}$ dienen nun wiederum häufig als Vergleichsreihen bei Konvergenzuntersuchungen. Z. B. sehen wir sofort, daß für $\alpha > 1$ die Reihe $\sum_{\nu=1}^{\infty} \frac{c_\nu}{\nu^\alpha}$ absolut konvergiert, wenn der absolute Betrag $|c_\nu|$ der Koeffizienten unterhalb einer festen, von $\nu$ unabhängigen Schranke bleibt.

## § 3. Grenzübergänge und Reihen von Funktionen einer Veränderlichen.

### 1. Allgemeines.

Die Glieder der unendlichen Reihen, die wir bisher betrachtet haben, waren konstant; demgemäß stellten diese Reihen auch immer bestimmte Zahlen dar. Für die Theorie und für die Anwendungen sind aber vor allem solche Reihen von Bedeutung, bei welchen die Reihenglieder Funktionen einer Veränderlichen sind und demgemäß auch die Reihensumme eine solche Funktion darstellt; wie z. B. die in Kap. VI behandelten Taylorschen Reihen.

Wir betrachten jetzt also allgemein eine Reihe

$$g_1(x) + g_2(x) + g_3(x) + \cdots,$$

bei der $g_n(x)$ eine in einem festen Intervall $a \leqq x \leqq b$ definierte Funktion von $x$ ist. Die $n$-te Partialsumme dieser Reihe, $g_1(x) + \cdots + g_n(x)$, bezeichnen wir mit $f_n(x)$. Dann ist die Summe $f(x)$ unserer Reihe, falls sie existiert, nichts anderes als der Grenzwert $\lim\limits_{n \to \infty} f_n(x)$.

Wir können also eine *unendliche Reihe von Funktionen* auffassen als *Grenzwert einer Folge Funktionen* $f_1(x), f_2(x), \ldots, f_n(x), \ldots$, ebenso wie wir auch *umgekehrt* zu jeder solchen Funktionenfolge $f_n(x)$ eine mit ihr gleichwertige Reihe bilden können, indem wir $g_n(x) = f_n(x) - f_{n-1}(x)$ (für $n > 1$) und $g_1(x) = f_1(x)$ setzen. Wir dürfen daher, wie es uns bequem ist, an Stelle der Betrachtung von unendlichen Reihen die von Funktionenfolgen setzen und umgekehrt.

### 2. Grenzübergänge mit Funktionen und Kurven.

Wir wollen nun genau formulieren, wann in einem bestimmten Intervall eine Funktion $f(x)$ die Grenzfunktion einer Funktionenfolge $f_1(x), f_2(x), \ldots, f_n(x), \ldots$ ist. Wir sagen: die Funktionenfolge $f_1(x)$, $f_2(x), \ldots$ konvergiert in diesem Intervalle gegen die Grenzfunktion $f(x)$, wenn an jeder Stelle $x$ des Intervalles im gewöhnlichen Sinne der Wert $f_n(x)$ gegen den Wert $f(x)$ strebt, und schreiben $\lim\limits_{n \to \infty} f_n(x) = f(x)$ oder $f_n(x) \underset{n \to \infty}{\longrightarrow} f(x)$. Nach dem Konvergenzprinzip von Cauchy (vgl.

erstes Kapitel, § 6) können wir die Konvergenz der Funktionenfolge auch zum Ausdruck bringen, ohne von vornherein die Grenzfunktion $f(x)$ kennen oder erwähnen zu müssen. Es wird nämlich unsere Funktionenfolge sicherlich dann und nur dann gegen eine Grenzfunktion $f(x)$ konvergieren, wenn an jeder Stelle $x$ in unserem Intervall zu jedem noch so klein vorgegebenen positiven $\varepsilon$ die Größe $|f_n(x) - f_m(x)|$ kleiner als $\varepsilon$ wird, sobald wir nur die Zahlen $n$ und $m$ hinreichend groß — d. h. größer als eine im allgemeinen von $\varepsilon$ abhängende und mit unbegrenzt abnehmendem $\varepsilon$ unbegrenzt wachsende Zahl $N = N(\varepsilon)$ — wählen.

Beispiele für solche Grenzwerte von Funktionen sind uns vielfach begegnet. Ich nenne nur die Definition der Potenz $x^\alpha$ für irrationales $\alpha$ durch die Gleichung

$$x^\alpha = \lim_{n \to \infty} x^{r_n},$$

wobei unter $r_1, r_2, \ldots, r_n, \ldots$ eine Folge von rationalen gegen $\alpha$ konvergierenden Zahlen verstanden war; oder die Gleichung

$$e^x = \lim_{n \to \infty} \left(1 + \frac{x}{n}\right)^n,$$

wo rechts die Funktionen $f_n(x)$ ganze rationale Funktionen $n$-ten Grades sind.

Die Darstellung der Funktionen durch Kurven legt es nahe, auch von einem Grenzübergang bei Kurven zu sprechen und zu sagen, daß die Kurven für die obigen Grenzfunktionen $x^\alpha$ und $e^x$ als Grenzkurven der Kurven für die Funktionen $x^{r_n}$ und $\left(1 + \frac{x}{n}\right)^n$ aufzufassen sind. Es besteht jedoch zwischen den Grenzübergängen mit Funktionen und mit Kurven ein feiner Unterschied, dessen Bedeutung bis in die Mitte des 19. Jahrhunderts nicht genügend beachtet worden ist und dessen klare Erfassung dennoch allein gewisse scheinbare Paradoxien zu vermeiden gestattet.

Betrachten wir als Beispiel die Funktionen

$$f_n(x) = x^n \qquad (n = 1, 2, \ldots)$$

im Intervalle $0 \leqq x \leqq 1$. Alle diese Funktionen sind stetig, und es existiert die Grenzfunktion $\lim\limits_{n \to \infty} f_n(x) = f(x)$. Diese Funktion ist aber nicht mehr stetig. Vielmehr ist, da ja für jedes $n$ der Funktionswert $f_n(1) = 1$ ist,

$$f(1) = 1,$$

während andererseits für $0 \leqq x < 1$, wie wir schon im ersten Kapitel, § 5, S. 24, sahen, $f(x) = 0$ wird; die Grenzfunktion $f(x)$ ist also in unserem Intervalle eine unstetige Funktion, welche für $x = 1$ den Wert 1 und sonst überall den Wert 0 besitzt.

Diese schon früher betrachtete Unstetigkeit wird uns anschaulich sofort verständlich, wenn wir die Kurven $C_n$ betrachten, welche zu den Funktionen $y = f_n(x)$ gehören. Sie sind (vgl. Fig. 11 S. 15) stetige Kurvenzüge, die alle durch den Nullpunkt und den Punkt $x = 1$, $y = 1$ gehen und die sich um so mehr der $x$-Achse anschmiegen, je größer $n$ ist. Diese Kurven besitzen eine Grenzkurve $C$, welche keineswegs unstetig ist, sondern (vgl. Fig. 110) aus dem Stück der $x$-Achse von $x = 0$ bis $x = 1$ und dem dazu senkrechten Stück der Geraden $x = 1$ von $y = 0$ bis $y = 1$ besteht. Also die Kurven konvergieren gegen eine stetige Grenzkurve mit einem senkrecht verlaufenden Stück, die Funktionen dagegen gegen eine unstetige Grenzfunktion. So erkennen wir, daß diese Unstetigkeit der Grenzfunktion sich geometrisch durch das Auftreten eines zur $x$-Achse senkrechten Stückes der Grenzkurve ausprägt. Ein solches Stück muß in der Funktion immer eine Unstetigkeit bedeuten; denn eine Funktion $f(x)$ ordnet jedem $x$ einen bestimmten Wert zu, und diese Bestimmtheit kann bei der Veranschaulichung einer Funktion durch eine Kurve nur vorliegen, wo keine solchen senkrechten Stücke vorhanden sind, vielmehr die Kurve von jeder solchen senkrechten Linie genau in einem Punkte geschnitten wird.

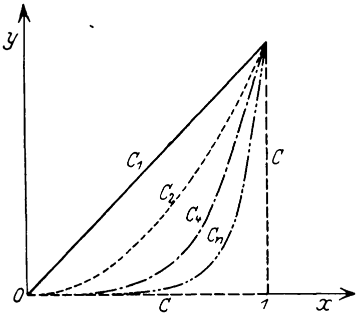


Fig. 110. Grenzkurve und Grenzfunktion.

Was bisher für Funktionenfolgen gezeigt wurde, gilt natürlich entsprechend für unendliche Reihen von Funktionen (vgl. S. 307).

## § 4. Gleichmäßige und ungleichmäßige Konvergenz.

### 1. Allgemeines und Beispiele.

Die Unstimmigkeit zwischen dem Begriff der Konvergenz von Funktionen und von Kurven bildet den Kern einer Erscheinung, deren klare Erkenntnis für das Studium der Konvergenz unentbehrlich ist: der sog. *ungleichmäßigen Konvergenz* von Funktionenfolgen oder unendlichen Reihen von Funktionen. Da erfahrungsgemäß gerade hier der Anfänger Schwierigkeiten zu finden pflegt, so möchte ich diesen Punkt ein wenig ausführlicher behandeln.

Daß eine Funktion $f(x)$ der Limes einer Funktionenfolge $f_n(x)$ im Intervalle $a \leqq x \leqq b$ ist, besagt nach der Definition nur, daß die gewöhnliche Limesbeziehung $f(x) = \lim\limits_{n \to \infty} f_n(x)$ an jeder Stelle $x$ des Intervalles gilt. Von einem naiven Standpunkte aus wird man nun vielleicht erwarten, daß mit diesem Konvergenzbegriffe ganz von selbst auch folgende Tatsache verbunden sei: Wenn man einen beliebigen Genauigkeitsgrad vorschreibt, z. B. $\varepsilon = \frac{1}{1000}$ oder $\varepsilon = \frac{1}{100}$, so werden von einem gewissen Index $N$ ab die Funktionen $f_n(x)$ für alle $x$ des Intervalles zwischen $f(x) + \varepsilon$ und $f(x) - \varepsilon$ liegen, die Kurven $y = f_n(x)$ also ganz in dem in Figur 111 gezeichneten Streifen verlaufen; d. h. es gibt eine zu $\varepsilon$ hinzubestimmbare Zahl $N = N(\varepsilon)$ — die natürlich im allgemeinen bei abnehmendem $\varepsilon$ immer größer werden wird —, so daß für $n > N$ stets $|f(x) - f_n(x)| < \varepsilon$ bleibt, wo auch immer wir im Intervall den Wert $x$ wählen. (Ist diese Forderung erfüllt, so bleibt also $|f_n(x) - f_m(x)| < 2\varepsilon$, sobald gleichzeitig $n > N$ und $m > N$ gilt.) Daß der Genauigkeitsgrad der Annäherung überall im Intervalle gleichzeitig, d. h. durch Wahl derselben von $x$ nicht abhängigen Zahl $N(\varepsilon)$, mindestens gleich einer vorgeschriebenen Zahl $\varepsilon$ gemacht werden kann, bezeichnet man als *Gleichmäßigkeit der Annäherung*. Es ist nun eine zunächst überraschende Einsicht, daß jene naive Annahme, als ob jede Konvergenz notwendig gleichmäßig sein müsse, durchaus nicht zutrifft; daß, anders gesprochen, die *Konvergenz* sehr wohl *ungleichmäßig* sein kann.

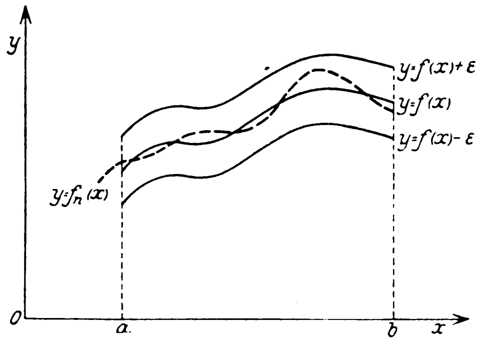


Fig. 111. Gleichmäßige Konvergenz.

1. Beispiel. Ungleichmäßige Konvergenz zeigt sich schon bei der oben behandelten Folge der Funktionen $f_n(x) = x^n$; sie konvergieren im Intervalle $0 \leqq x \leqq 1$ gegen die Grenzfunktion $f(x) = 0$ für $0 \leqq x < 1$, $f(1) = 1$. An jeder Stelle des Intervalles findet Konvergenz statt; d. h.: gibt man irgend eine noch so kleine positive Zahl $\varepsilon$ vor und wählt einen bestimmten festen Wert $x = \xi$, so braucht man nur $n$ hinreichend groß zu machen, damit $|\xi^n - f(\xi)| < \varepsilon$ gilt. Und doch ist diese Annäherung nicht gleichmäßig; denn nehmen wir etwa $\varepsilon = \frac{1}{2}$, so können wir, wie groß auch immer die Zahl $n$ genommen werden mag, immer noch Stellen $x = \eta \neq 1$ in der Nähe von $x = 1$ finden, wo $|\eta^n - f(\eta)| = \eta^n > \frac{1}{2}$ ist; nämlich alle solchen Stellen $x = \eta$, für welche $1 > \eta > \sqrt[n]{\frac{1}{2}}$ ist. Es ist also nicht möglich, die Zahl $n$ so groß zu wählen, daß im *ganzen* Intervalle der Unterschied zwischen $f(x)$ und $f_n(x)$ kleiner als $\frac{1}{2}$ wird.

Dieses Verhalten wird uns verständlich, wenn wir die Funktionen geometrisch repräsentieren und uns davon Rechenschaft geben, daß die Kurven für die Funktionen $f_n(x)$ in der Nähe des Punktes $x = 1$, $y = 1$ immer steiler werden, daß der steile Anstieg sich aber auch bei wachsendem $n$ auf eine immer kleinere Umgebung dieses Punktes beschränkt und daß schließlich diese Kurven der oben betrachteten Grenzlage zustreben.

Ein ganz ähnliches Verhalten zeigen die Funktionen

$$f_n(x) = \frac{1}{1 + x^{2n}}$$

in der Umgebung der Stellen $x = 1$ und $x = -1$, wie man sich leicht überlegt. (Man vergleiche auch die Betrachtungen im ersten Kapitel, § 8.)

2. Beispiel. Bei den ersten beiden Beispielen hing die Ungleichmäßigkeit der Konvergenz damit zusammen, daß die Grenzfunktion unstetig war. Wir können uns aber auch leicht eine Folge von stetigen Funktionen bilden, die zwar gegen eine stetige Grenzfunktion konvergieren, aber trotzdem ungleichmäßig. Wir legen das Intervall $0 \leqq x \leqq 1$ zugrunde und definieren für $n \geqq 2$:

$$f_n(x) = x\, n^{\alpha} \quad \text{für} \quad 0 \leqq x \leqq \frac{1}{n},$$

$$f_n(x) = \left(\frac{2}{n} - x\right) n^{\alpha} \quad \text{für} \quad \frac{1}{n} \leqq x \leqq \frac{2}{n},$$

$$f_n(x) = 0 \quad \text{für} \quad \frac{2}{n} \leqq x \leqq 1,$$

wobei $\alpha$ eine zunächst noch verfügbare, aber für unsere Funktionenfolge dann fest zu wählende positive Zahl ist. Geometrisch werden unsere Funktionen repräsentiert durch ein dachartiges, aus zwei geraden Linien bestehendes zur Geraden $x = \frac{1}{n}$ symmetrisches Gebilde über dem Stück zwischen der Abszisse $x = 0$ und der Abszisse $x = \frac{2}{n}$ der $x$-Achse und im übrigen die $x$-Achse selbst (vgl. Fig. 112).

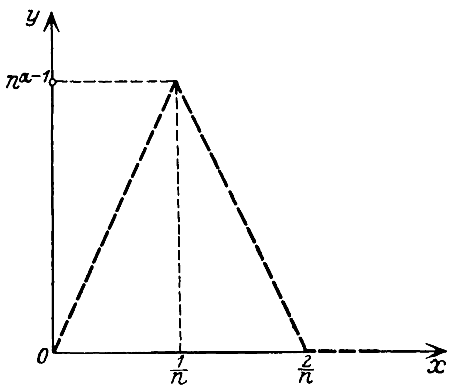


Fig. 112. Ungleichmäßige Konvergenz.

Ist $\alpha < 1$, so wird die Höhe dieses aufgesetzten Dreieckes, welche allgemein $n^{\alpha-1}$ beträgt, mit wachsendem $n$ gegen 0 streben; die Kurven werden dann gegen die $x$-Achse, die Funktionen $f_n(x)$ gleichmäßig gegen den Grenzwert 0 streben.

Ist $\alpha = 1$, so wird die aufgesetzte Zacke für jedes $n$ die Höhe 1 besitzen. Ist $\alpha > 1$, so wird die Höhe der aufgesetzten Zacke sogar mit wachsendem $n$ über alle Grenzen wachsen.

Wie aber auch $\alpha$ gewählt werden möge, die Funktionenfolge $f_n(x)$ wird doch stets der Grenzfunktion $f(x) = 0$ zustreben. Jedes positive $x$ des Intervalles bleibt nämlich bei hinreichend großem $n$ doch gewiß außerhalb der Grundlinie des aufgesetzten Daches; für $x = 0$ aber sind alle Funktionen $f_n(x)$ selbst 0, also auch ihr Grenzwert $f(0) = \lim\limits_{n\to\infty} f_n(0)$.

Die Konvergenz ist aber für $\alpha \geqq 1$ sicher ungleichmäßig; denn es ist nicht mehr möglich, durch Wahl eines hinreichend großen $n$ gleichzeitig für alle Werte von $x$ im ganzen Intervall den Ausdruck $|f(x) - f_n(x)| = f_n(x)$ kleiner als etwa $\frac{1}{2}$ zu machen.

3. Beispiel. Ein ganz ähnliches Verhalten zeigt die — im Gegensatze zum vorigen Beispiel — durch einen einheitlichen analytischen Ausdruck dargestellte Funktionenfolge

$$f_n(x) = x\,n^\alpha\,e^{-nx}$$

(vgl. Fig. 113). Auch hier wird für positives $x$ ganz gewiß $\lim\limits_{n\to\infty} f_n(x) = 0$ gelten, da die Funktion $e^{-nx}$ bei wachsendem $n$ stärker 0 wird als jede Potenz von $\frac{1}{n}$ (vgl. drittes Kapitel, § 9). Für $x = 0$ ist aber stets $f_n(x) = 0$, und so erhalten wir in dem ganzen Intervall $0 \leqq x \leqq a$ bei beliebigem positivem $a$

$$f(x) = \lim_{n\to\infty} f_n(x) = 0.$$

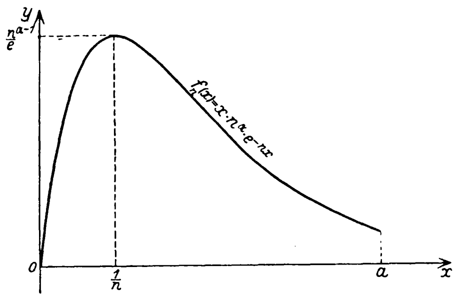


Fig. 113.

Aber auch hier ist die Konvergenz der Funktionenfolge gegen die Grenzfunktion nicht immer gleichmäßig. Es wird nämlich für die Stelle $x = \frac{1}{n}$ (die Stelle des Maximums von $f_n(x)$)

$$f_n(x) = f_n\left(\frac{1}{n}\right) = \frac{n^{\alpha-1}}{e},$$

und wir erkennen daraus, daß die durch den Buckel an der mit $n$ veränderlichen Stelle $x = \frac{1}{n}$ gemessene „Höhe" der Annäherungskurve im Falle $\alpha \geqq 1$ bei wachsendem $n$ gewiß nicht gegen 0 strebt. Im übrigen liegen die Verhältnisse bei diesem Beispiel genau wie bei dem eben betrachteten.

4. Beispiel. Die Begriffe der gleichmäßigen und ungleichmäßigen Konvergenz übertragen sich natürlich auch auf unendliche Reihen. Wir nennen eine Reihe $g_1(x) + g_2(x) + \cdots$ gleichmäßig konvergent oder nicht, je nach dem Verhalten der Folge ihrer Partialsummen $f_n(x)$. Ein

sehr einfaches Beispiel für eine nicht gleichmäßig konvergente Reihe finden wir in

$$f(x) = x^2 + \frac{x^2}{1+x^2} + \frac{x^2}{(1+x^2)^2} + \frac{x^2}{(1+x^2)^3} + \cdots.$$

Für $x = 0$ ist jede Partialsumme $f_n(x) = x^2 + \cdots + \frac{x^2}{(1+x^2)^{n-1}} = 0$; also ist auch $f(0) = 0$. Für $x \neq 0$ haben wir einfach eine geometrische Reihe mit dem positiven Quotienten $\frac{1}{1+x^2} < 1$ vor uns; also können wir nach den elementaren Regeln summieren und erhalten unter der Voraussetzung $x \neq 0$ als Summe der Reihe den Wert $\frac{x^2}{1 - \frac{1}{1+x^2}} = 1 + x^2$. Die Grenzfunktion $f(x)$ wird also überall außer für $x = 0$ durch den Ausdruck $f(x) = 1 + x^2$ gegeben, während $f(0) = 0$ ist; ihr ist mithin im Nullpunkt gewissermaßen künstlich eine Unstetigkeit aufgeprägt.

Auch hier haben wir es in jedem den Nullpunkt enthaltenden Intervall mit einer ungleichmäßigen Konvergenz zu tun. Die Differenz $f(x) - f_n(x) = r_n(x)$ ist nämlich für $x = 0$ selbst 0, während sie für jeden anderen Wert von $x$ durch den Ausdruck $r_n(x) = \frac{1}{(1+x^2)^{n-1}}$ gegeben wird. Verlangen wir, daß dieser Ausdruck etwa kleiner als $\frac{1}{2}$ wird, so können wir das bei festgehaltenem $x$ stets durch die Wahl eines hinreichend großen $n$ erreichen. Wir können aber stets für jeden beliebigen festen Wert von $n$ noch immer nahe am Nullpunkt Stellen $x = \xi$ angeben, für welche $r_n(\xi) > \frac{1}{2}$ wird; eine gleichmäßige Erzielung des vorgeschriebenen Genauigkeitsgrades ist also unmöglich. Anschaulich wird der Sachverhalt sofort wieder durch Betrachtung der Näherungskurven klar (vgl. Fig. 114). Diese Näherungskurven werden sich bei wachsendem $n$, abgesehen von der unmittelbaren Umgebung der Stelle $x = 0$, immer besser der Parabel $y = 1 + x^2$ anschmiegen; in der Umgebung der Stelle $x = 0$ aber werden diese Kurven einen immer schmaler werdenden rüsselartigen Fortsatz nach dem Nullpunkt herunterstrecken, und dieser Fortsatz

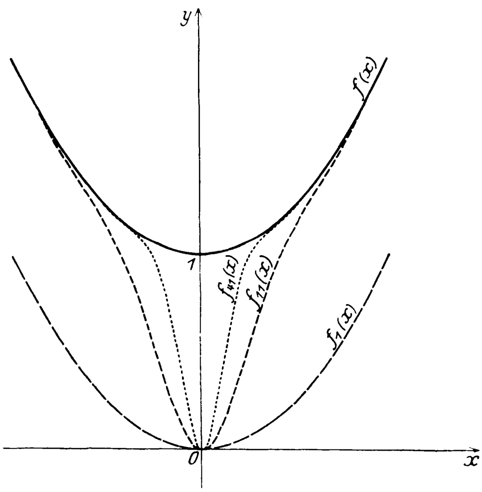


Fig. 114.

wird sich bei wachsendem $n$ immer mehr auf ein gewisses geradliniges Stück der $y$-Achse zusammenziehen, so daß als Grenzkurve die Parabel mit einem geradlinigen senkrecht nach unten bis zum Nullpunkt reichenden Fortsatz erscheint.

Als weiteres Beispiel für ungleichmäßige Konvergenz nenne ich die Reihe $\sum_{\nu=0}^{\infty} g_\nu(x)$ mit $g_\nu(x) = x^\nu - x^{\nu-1}$ für $\nu \geqq 1$, $g_0(x) = 1$ im Intervalle $0 \leqq x \leqq 1$, deren Partialsummen die oben als erstes Beispiel betrachteten Funktionen $x^\nu$ sind.

## 2. Kriterium der gleichmäßigen Konvergenz.

Alle unsere Überlegungen zeigen uns, daß die Gleichmäßigkeit der Konvergenz in einem Intervalle stets eine besondere Eigenschaft einer Funktionenfolge bzw. einer unendlichen Reihe ist. Wir wollen nun den Begriff der gleichmäßigen Konvergenz noch einmal formulieren. *Die konvergente Reihe $g_1(x) + g_2(x) + \cdots$ heißt in einem Intervalle gleichmäßig konvergent, wenn man die Summe $f(x)$ mit jeder noch so klein gegebenen Genauigkeitsgrenze $\varepsilon$ berechnen kann, sobald man eine für das ganze Intervall feste hinreichend große Zahl von Gliedern verwendet; m.a.W., wenn man zu jedem noch so kleinen $\varepsilon$ eine nur von $\varepsilon$, nicht mehr von $x$ abhängige Zahl $N$ angeben kann, derart, daß für $n > N$ stets gilt $|f(x) - f_n(x)| = |R_n(x)| < \varepsilon$, wobei mit $R_n(x)$ der Rest der Reihe für die $n$-te Partialsumme $f_n(x)$ bezeichnet wird.*

Wenn wir statt von einer unendlichen Reihe von einer Funktionenfolge $f_1(x)$, $f_2(x)$, ... mit der Grenzfunktion $f(x)$ sprechen, so werden wir demgemäß wie oben die gleichmäßige Konvergenz durch die Eigenschaft definieren, daß der Ausdruck $|f(x) - f_n(x)|$ oder auch der Ausdruck $|f_n(x) - f_m(x)|$ bei hinreichend großem $n$ bzw. bei hinreichend großem $n$ und $m$ für das ganze Intervall gleichzeitig unter jede vorgegebene Schranke $\varepsilon$ sinkt.

Wir werden sogleich sehen, daß erst durch die Forderung der Gleichmäßigkeit unendliche Reihen oder sonstige Grenzwertbildungen von Funktionen zu schmiegsamen Hilfsmitteln der Analysis gemacht werden; glücklicherweise zeigt es sich, daß die Ungleichmäßigkeit der Konvergenz eine Art Ausnahmeerscheinung ist, die uns in der Handhabung der Methoden der Analysis kaum stört.

Gewöhnlich beweist man die Gleichmäßigkeit der Konvergenz einer Reihe aus folgendem Kriterium: *Wenn die Glieder der Reihe $\sum_{\nu=1}^{\infty} g_\nu(x)$ den Bedingungen $|g_\nu(x)| \leqq a_\nu$ genügen, wo die $a_\nu$ Konstante sind und eine konvergente Reihe $\sum_{\nu=1}^{\infty} a_\nu$ bilden, so konvergiert die Reihe $\sum_{\nu=1}^{\infty} g_\nu(x)$ gleichmäßig (und, nebenbei bemerkt, auch absolut).*

In der Tat ist dann

$$\left|\sum_{\nu=n}^{m} g_\nu(x)\right| \leqq \sum_{\nu=n}^{m} |g_\nu(x)| \leqq \sum_{\nu=n}^{m} a_\nu,$$

und da nach dem Cauchyschen Kriterium $\sum\limits_{\nu=n}^{m} a_\nu$ schließlich beliebig klein wird, drückt sich hierin unsere Behauptung unmittelbar aus.

Ein erstes Beispiel bietet uns die geometrische Reihe $1 + x + x^2 + \cdots$, wenn man sich auf das Intervall $|x| \leqq q$ beschränkt, wo $q$ irgend eine positive Zahl kleiner als 1 ist. Die Glieder der Reihe sind dann absolut genommen nicht größer als die Glieder der konvergenten geometrischen Reihe $\sum q^\nu$.

Ein weiteres Beispiel gibt die „trigonometrische Reihe"

$$\frac{c_1 \sin(x - \delta_1)}{1^2} + \frac{c_2 \sin(x - \delta_2)}{2^2} + \frac{c_3 \sin(x - \delta_3)}{3^2} + \cdots,$$

sobald $|c_n| < c$ ist, unter $c$ eine von $n$ unabhängige positive Konstante verstanden. Es ist dann einfach $g_n(x) = \frac{c_n \sin(x - \delta_n)}{n^2}$ und $|g_n(x)| < \frac{c}{n^2}$, woraus wegen der Konvergenz der Reihe $\sum\limits_{\nu=1}^{\infty} \frac{c}{\nu^2}$ die gleichmäßige und absolute Konvergenz unserer trigonometrischen Reihe folgt.

### 3. Stetigkeit gleichmäßig konvergenter Reihen stetiger Funktionen.

Die Bedeutung der gleichmäßigen Konvergenz unendlicher Reihen liegt, wie schon angedeutet, vor allem darin, daß eine gleichmäßig konvergente Reihe in vieler Hinsicht sich genau so verhält wie eine Summe von endlich vielen Funktionen. So z. B. ist eine Summe endlich vieler stetiger Funktionen wieder stetig, und ganz entsprechend gilt der Satz:

*Eine in einem Intervall gleichmäßig konvergente Reihe, deren Glieder stetige Funktionen sind, stellt dort wieder eine stetige Funktion dar.*

Der Beweis ist sehr einfach: Wir zerlegen die Reihe

$$f(x) = g_1(x) + g_2(x) + \cdots$$

in die $n$-te Partialsumme $f_n(x)$ und den „Rest" $R_n(x)$. Es ist dabei $f_n(x) = g_1(x) + \cdots + g_n(x)$. Wird nun eine beliebig kleine positive Zahl $\varepsilon$ vorgegeben, so können wir zunächst wegen der vorausgesetzten Gleichmäßigkeit der Konvergenz die Zahl $n$ so groß wählen, daß der Rest im ganzen Intervall kleiner als $\frac{\varepsilon}{4}$ wird und daher erst recht

$$|R_n(x + h) - R_n(x)| < \frac{\varepsilon}{2}$$

gilt, wenn $x$ und $x + h$ irgendwelche Zahlen des Intervalles bedeuten. Die Partialsumme $f_n(x)$ besteht aber nur aus endlich vielen stetigen Summanden und ist daher gewiß wieder stetig; wir können also zu jeder Stelle $x$ des Intervalles ein so kleines positives $\delta$ wählen, daß

$$|f_n(x+h) - f_n(x)| < \frac{\varepsilon}{2}$$

ist, wenn $|h| < \delta$ ist und $x + h$ im Intervall liegt. Daher gilt dann erst recht

$$\begin{aligned} |f(x+h) - f(x)| &= |f_n(x+h) - f_n(x) + R_n(x+h) - R_n(x)| \\ &\leqq |f_n(x+h) - f_n(x)| + |R_n(x+h) - R_n(x)| < \varepsilon; \end{aligned}$$

diese Beziehung drückt aber die Stetigkeit der Funktion $f(x)$ aus.

Die Bedeutung dieses Satzes wird erst klar, wenn wir uns vergegenwärtigen, daß bei ungleichmäßig konvergenten Reihen stetiger Funktionen die Summe keineswegs stetig zu sein braucht, wie uns unsere früheren Beispiele unmittelbar zeigen. Wir können aus unserem obigen Resultat schließen: Wenn die Summe einer konvergenten Reihe stetiger Funktionen an einer Stelle unstetig ist, so konvergiert die Reihe in einer Umgebung dieser Stelle sicher ungleichmäßig. Es beruht also jede Darstellung unstetiger Funktionen durch Reihen stetiger Funktionen immer gerade darauf, daß man dabei ungleichmäßig konvergente Grenzübergänge verwendet.

## 4. Die Integration gleichmäßig konvergenter Reihen.

Eine Summe von endlich vielen stetigen Funktionen läßt sich „gliedweise integrieren"; d. h. das Integral der Summe läßt sich bilden, indem man die einzelnen Funktionen integriert und die Integrale addiert. Bei einer konvergenten Reihe von Funktionen, einer unendlichen Summe, ist nun ein solches Verfahren ebenfalls gestattet, vorausgesetzt, daß die Reihe in dem Integrationsintervall gleichmäßig konvergiert. *Eine in einem Intervall gleichmäßig konvergente Reihe stetiger Funktionen* $\sum_{\nu=1}^{\infty} g_\nu(x) = f(x)$ *darf dort gliedweise integriert werden.* Genauer: *Wenn a und x zwei Zahlen in diesem Intervalle sind, so konvergiert die Reihe* $\sum_{\nu=1}^{\infty} \int_a^x g_\nu(t)\,dt$, *und zwar gleichmäßig in x (bei festem a), und ihre Summe ist gleich* $\int_a^x f(t)\,dt$.

Zum Beweise schreiben wir mit unseren früheren Bezeichnungen

$$f(x) = \sum_{\nu=1}^{\infty} g_\nu(x) = f_n(x) + R_n(x)$$

und setzen voraus, daß die einzelnen Reihenglieder, also auch die Summe $f(x)$ in dem gegebenen Intervall stetig sind, sich also integrieren lassen.

Dann können wir zu jedem $\varepsilon$ jedenfalls $n$ so groß wählen, daß im ganzen Intervall $|R_n(x)| < \varepsilon$ wird und somit auch $|f(x) - f_n(x)| < \varepsilon$. Nach dem Mittelwertsatz der Integralrechnung ist nun

$$\int_a^x (f(t) - f_n(t))\,dt \leqq \varepsilon l,$$

wo $l$ die Länge des Integrationsintervalles bezeichnet; da wir aber das Integral über die endliche Summe $f_n(x)$ ohne weiteres gliedweise ausführen dürfen, erhalten wir

$$\left|\int_a^x f(t)\,dt - \sum_1^n \int_a^x g_\nu(t)\,dt\right| < \varepsilon l.$$

Lassen wir nun $\varepsilon$ gegen 0 streben und dementsprechend $n$ über alle Grenzen wachsen, so ergibt sich aus dieser letzten Beziehung unmittelbar unser Satz von der Möglichkeit, eine gleichmäßig konvergente Reihe gliedweise zu integrieren.

Wenn man statt von unendlichen Reihen einfach von einem Grenzwert

$$f(x) = \lim_{n\to\infty} f_n(x)$$

spricht, so kann man unser Resultat auch folgendermaßen aussprechen:

*Wenn in einem Intervalle die Beziehung* $f(x) = \lim\limits_{n\to\infty} f_n(x)$ *gleichmäßig erfüllt ist, d. h. wenn die Funktionenfolge* $f_n(x)$ *gleichmäßig gegen die Funktion* $f(x)$ *konvergiert, so gilt, falls a und b im Intervall liegen,*

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx,$$

*d. h. mit anderen Worten, wir dürfen dann die Reihenfolge der Operationen des Integrierens und des Grenzüberganges bei* $f_n(x)$ *miteinander vertauschen.*

Diese Tatsache ist keineswegs von vornherein selbstverständlich. Zwar von einem naiven Standpunkt aus, wie er bis ins 19. Jahrhundert herrschte, wird man die Vertauschbarkeit der beiden Prozesse kaum bezweifeln; aber ein Blick auf die Beispiele aus diesem Paragraphen, Nr. 1, zeigt uns, daß bei nicht mehr gleichmäßig konvergenten Grenzwertbildungen unsere Behauptung unzutreffend sein kann. Man braucht nur Beispiel 2 zu betrachten, bei dem das Integral der Grenzfunktion 0 ist, während das Integral der Funktion $f_n(x)$ über das Intervall $0 \leqq x \leqq 1$, d. h. der Inhalt des aufgesetzten Dreiecks, den Wert

$$\int_0^1 f_n(x)\,dx = n^{\alpha-2}$$

besitzt und also im Falle $\alpha \geqq 2$ nicht gegen Null strebt. Wir erkennen hier anschaulich sofort den Grund für den Unterschied zwischen $\int\limits_0^1 f(x)\,dx$ und $\lim\limits_{n\to\infty}\int\limits_0^1 f_n(x)\,dx$ in der Ungleichmäßigkeit der Konvergenz.

Dagegen zeigt sich im Falle $1 \leqq \alpha < 2$, daß trotz ungleichmäßiger Konvergenz die Gleichung $\lim\limits_{n\to\infty}\int\limits_0^1 f_n(x)\,dx = \int\limits_0^1 f(x)\,dx$ bestehen kann. Dasselbe zeigen uns auch die anderen in Nr. 1 gegebenen Beispiele. Wir dürfen z. B. die Reihe $\sum\limits_0^\infty g_n(x)$ mit $g_n(x) = x^n - x^{n-1}$ für $n \geqq 1$ und mit $g_0(x) = 1$ trotz ihrer ungleichmäßigen Konvergenz zwischen den Grenzen 0 und 1 gliedweise integrieren, d. h. wir erhalten so ein richtiges Resultat. Es ist also für die Möglichkeit, gliedweise zu integrieren, die Gleichmäßigkeit der Konvergenz zwar eine hinreichende, aber keineswegs eine notwendige Bedingung. Die Nichtbeachtung dieses Punktes gibt leicht zu Mißverständnissen Anlaß.

## 5. Differentiation unendlicher Reihen.

Anders als bei der Integration verhalten sich gleichmäßig konvergente Reihen oder Funktionenfolgen gegenüber der Differentiation. Z. B. ist für die Funktionenfolge $f_n(x) = \frac{\sin n^2 x}{n}$ sicherlich gleichmäßig $\lim\limits_{n\to\infty} f_n(x) = 0$, aber der Differentialquotient $f_n'(x) = n\cos n^2 x$ konvergiert bei wachsendem $n$ keineswegs gegen den Differentialquotienten der Grenzfunktion, d. h. gegen den Wert 0, z. B. nicht für $x = 0$. Wir dürfen also trotz der Gleichmäßigkeit des Grenzüberganges hier keineswegs die Prozesse der Differentiation und des Grenzüberganges vertauschen.

Für eine unendliche Reihe gilt natürlich Entsprechendes. Z. B. ist die Reihe

$$\sin x + \frac{\sin 2^4 x}{2^2} + \frac{\sin 3^4 x}{3^2} + \cdots$$

absolut und gleichmäßig konvergent; denn ihre Glieder sind sicher absolut genommen nicht größer als die Glieder der konvergenten Reihe $\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots$. Differenzieren wir aber unsere obere Reihe gliedweise, so entsteht die Reihe

$$\cos x + 2^2\cos 2^4 x + 3^2 \cos 3^4 x + \cdots,$$

welche ganz offenkundig nicht überall konvergiert, z. B. nicht für $x = 0$.

Das einzige brauchbare Kriterium, welches uns in speziellen Fällen Sicherheit dafür gibt, daß die Differentiation gliedweise erlaubt ist, gibt folgender Satz: *Wenn durch gliedweises Differenzieren einer konvergenten* *unendlichen Reihe* $\sum\limits_{\nu=0}^{\infty} G_\nu(x) = F(x)$ *eine gleichmäßig konvergente Reihe stetiger Funktionen* $\sum\limits_{\nu=0}^{\infty} g_\nu(x) = f(x)$ *entsteht, so ist der Wert dieser letzteren gleich dem Differentialquotienten der Summe der ersten Reihe.* Dieser Satz verlangt also ausdrücklich, daß man sich nach Ausführung der gliedweisen Differentiation noch davon überzeugt, ob das Resultat dieser Differentiation eine gleichmäßig konvergente Reihe ist.

Der Beweis des Satzes versteht sich fast von selbst. Denn man darf ja nach dem Satz von Nr. 4 die durch Differentiation entstandene Reihe in unserem Intervall gliedweise integrieren, wobei wir die obere Grenze $x$ unbestimmt lassen können. Mit Rücksicht auf $g_\nu(t) = G'_\nu(t)$ erhalten wir also

$$\int_a^x f(t)\,dt = \int_a^x \Big(\sum_{\nu=0}^{\infty} g_\nu(t)\Big)\,dt = \sum_{\nu=0}^{\infty} \int_a^x g_\nu(t)\,dt = \sum_{\nu=0}^{\infty} (G_\nu(x) - G_\nu(a)) = F(x) - F(a).$$

Daher ist

$$f(x) = F'(x),$$

wie es behauptet wurde.

## § 5. Potenzreihen.

Unter den unendlichen Reihen stehen die Potenzreihen an wichtigster Stelle. Unter einer Potenzreihe verstehen wir eine Reihe der Form

$$\mathfrak{P}(x) = c_0 + c_1 x + c_2 x^2 + \cdots = \sum_{\nu=0}^{\infty} c_\nu x^\nu$$

(„Potenzreihe in $x$") oder auch allgemeiner

$$\mathfrak{P}(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)^2 + \cdots = \sum_{\nu=0}^{\infty} c_\nu (x - x_0)^\nu$$

(„Potenzreihe in $x - x_0$"), wo $x_0$ eine feste Zahl ist. Wenn wir in der letzten Reihe $x - x_0 = x'$ als neue unabhängige Veränderliche einführen, so geht sie in die Potenzreihe $\sum\limits_{\nu=0}^{\infty} c_\nu x'^\nu$ in der Variablen $x'$ über, und wir können uns daher bei unseren Betrachtungen, ohne der Allgemeinheit Abbruch zu tun, auf Potenzreihen der spezielleren Form $\sum\limits_{\nu=0}^{\infty} c_\nu x^\nu$ beschränken.

Wir haben schon im sechsten Kapitel ausführlich die Annäherung von Funktionen durch ganze rationale und im Zusammenhang damit die Entwicklung in Taylorsche Reihen — die ja Potenzreihen sind — behandelt. In diesem Paragraphen wollen wir die Potenzreihen etwas näher studieren und werden dadurch neue und in vieler Hinsicht gegenüber dem damaligen Kapitel einfachere und bequemere Zugänge zu den Reihenentwicklungen der wichtigsten Funktionen gewinnen.

## 1. Das Konvergenzverhalten einer Potenzreihe.

Es gibt Potenzreihen, welche für keinen Wert von $x$ konvergieren, natürlich abgesehen von dem Werte $x = 0$; z. B. die Reihe

$$x + 2^2 x^2 + 3^3 x^3 + \cdots + n^n x^n + \cdots.$$

Denn ist etwa $|x| > \frac{1}{N}$, wo $N$ eine natürliche Zahl ist, so werden alle Glieder $n^n x^n$ für $n > N$ absolut genommen größer als 1 sein, sogar mit wachsendem $n$ über alle Grenzen wachsen, während sie doch bei einer konvergenten Reihe gegen 0 streben müßten.

Andererseits gibt es Potenzreihen, welche für jeden Wert von $x$ konvergieren, wie z. B. die Potenzreihe für die Exponentialfunktion

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$

wo die Konvergenz für jeden Wert von $x$ gemäß dem Kriterium IIIa aus § 2, Nr. 2 sofort folgt. Das $(n+1)$-te Glied, durch das $n$-te dividiert, ergibt nämlich $\frac{x}{n}$, und, wie auch immer wir die Zahl $x$ wählen, stets wird dieser Quotient bei wachsendem $n$ gegen 0 streben.

Das Konvergenzverhalten von Potenzreihen wird nun allgemein durch den folgenden grundlegenden Satz beherrscht: *Wenn eine Potenzreihe überhaupt für einen Wert $x = \xi$ konvergiert, so konvergiert sie absolut für jeden Wert $x$, für welchen $|x| < |\xi|$ ist, und die Konvergenz ist gleichmäßig für jedes Intervall $|x| \leqq \eta$, sobald nur $\eta$ eine positive Zahl kleiner als $|\xi|$ — aber im übrigen beliebig nahe an $|\xi|$ gelegen — ist.*

Der Beweis ist sehr einfach. Wenn die Reihe $\sum\limits_{\nu=0}^{\infty} c_\nu \xi^\nu$ konvergiert, so müssen ihre Glieder mit wachsendem $\nu$ gegen 0 streben; sie werden also gewiß, was ja viel weniger besagt, absolut genommen unterhalb einer von $\nu$ unabhängigen Schranke $M$ liegen, d. h. es wird $|c_\nu \xi^\nu| < M$ sein. Ist nun $q$ eine den Bedingungen $0 < q < 1$ genügende feste Zahl und beschränken wir $x$ auf das Intervall $|x| \leqq q|\xi|$, so wird $|c_\nu x^\nu| \leqq |c_\nu \xi^\nu| q^\nu < M q^\nu$. Die Glieder unserer Reihe $\sum\limits_{0}^{\infty} c_\nu x^\nu$ sind also absolut genommen in diesem Intervalle kleiner als die Glieder der aus positiven Konstanten bestehenden, konvergenten geometrischen Reihe $\sum M q^\nu$. Hieraus folgt nunmehr nach dem Satze von § 4, Nr. 2 sofort die behauptete absolute und gleichmäßige Konvergenz im Intervall $-q|\xi| \leqq x \leqq q|\xi|$.

Wenn eine Potenzreihe nicht überall konvergiert, wenn es vielmehr einen Wert $x = \xi$ gibt, für den sie divergiert, so muß sie für jeden Wert von $x$ divergieren, für welchen $|x| > |\xi|$ ist. Denn wäre sie für einen solchen Wert von $x$ konvergent, so müßte sie erst recht nach dem vorhin bewiesenen Satze für den absolut kleineren Wert $\xi$ konvergieren.

Man erkennt hieraus, daß es für eine Potenzreihe, die in mindestens einem von 0 verschiedenen Punkte konvergiert und in mindestens einem Punkte divergiert, eine *Konvergenzstrecke* gibt; d. h. es gibt eine ganz bestimmte positive Zahl $\varrho$, derart, daß für $|x| > \varrho$ Divergenz, für $|x| < \varrho$ Konvergenz herrscht, während man für $|x| = \varrho$ von vornherein keine allgemeinen Aussagen machen kann. Die Grenzfälle, daß die Potenzreihe nur für $x = 0$ konvergiert bzw. daß sie für alle Werte $x$ konvergiert, drücken wir symbolisch durch die Schreibweise $\varrho = 0$ bzw. $\varrho = \infty$ aus[1]).

Z. B. ist für die geometrische Reihe $1 + x + x^2 + \cdots$ der Wert $\varrho = 1$; in den Endpunkten der Konvergenzstrecke herrscht Divergenz. Ebenso ist für die uns aus dem sechsten Kapitel bekannte Arcustangens-Reihe

$$\operatorname{arc\,tg} x = x - \frac{x^3}{3} + \frac{x^5}{5} - + \cdots$$

$\varrho = 1$, und in beiden Endpunkten der Konvergenzstrecke, nämlich für die Werte $x = \pm 1$, herrscht Konvergenz, wie man sofort aus dem Leibnizschen Kriterium § 1, Nr. 2 erkennt.

Als Folge der gleichmäßigen Konvergenz nenne ich die wichtige Tatsache, daß eine Potenzreihe im Innern ihres (etwaigen) Konvergenzintervalls immer eine stetige Funktion darstellt.

### 2. Die Integration und Differentiation von Potenzreihen.

Wegen der Gleichmäßigkeit der Konvergenz ist es *ohne weiteres erlaubt, eine Potenzreihe*

$$f(x) = \sum_{\nu=0}^{\infty} c_\nu x^\nu$$

über ein ganz im Innern ihres Konvergenzintervalles gelegenes abgeschlossenes Intervall *gliedweise zu integrieren.* Wir erhalten so die Funktion

$$F(x) = c + \sum_{\nu=0}^{\infty} \frac{c_\nu}{\nu+1} x^{\nu+1},$$

für welche

$$F'(x) = f(x)$$

ist. Übrigens bemerke ich, daß stets $\left|\frac{c_\nu}{\nu+1}\right| \leqq |c_\nu|$ ist, daß also die

[1]) Es läßt sich direkt angeben, in welcher Weise die Konvergenzstrecke von den Koeffizienten $c_\nu$ der Reihe abhängt. Existiert der Grenzwert $\lim\limits_{n\to\infty} \sqrt[n]{|c_n|}$, so ist

$$\varrho = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{|c_n|}}.$$

Allgemein drückt sich $\varrho$ durch die Formel

$$\varrho = \frac{1}{\overline{\lim\limits_{n\to\infty}} \sqrt[n]{|c_n|}}$$

aus (wobei $\overline{\lim}$ das im ersten Kapitel, Anhang, definierte Symbol der oberen Häufungsstelle ist).

absolute Konvergenz durch die Integration gegenüber der ursprünglichen Reihe verbessert wird.

*Wir dürfen aber eine Potenzreihe im Innern ihres Konvergenzintervalls auch gliedweise differenzieren* und gelangen so zu der Gleichung

$$f'(x) = \sum_{\nu=1}^{\infty} \nu c_\nu x^{\nu-1}.$$

Um die Richtigkeit dieser Behauptung zu beweisen, brauchen wir nur zu zeigen, daß die rechts stehende Potenzreihe gleichmäßig konvergiert, sobald $x$ auf ein Intervall beschränkt wird, welches ganz im Innern der Konvergenzstrecke liegt. Mit den obigen Bezeichnungen aus Nr. 1 sei demgemäß $|x| \leqq q|\xi|$. Die Glieder der in Frage stehenden Reihe sind absolut genommen nicht größer als die Glieder der Reihe $\sum_{\nu=1}^{\infty} |\nu c_\nu q^{\nu-1} \xi^{\nu-1}|$; also brauchen wir, da $|c_\nu \xi^{\nu-1}| < \frac{M}{|\xi|} = N$ ist, nur die Konvergenz der Reihe $\sum_{\nu=1}^{\infty} N \nu q^{\nu-1}$ zu beweisen. In dieser Reihe ist aber der Quotient des $(n+1)$-ten Gliedes, dividiert durch das $n$-te Glied, gerade gleich $\frac{n+1}{n} q$. Er strebt bei wachsendem $n$ gegen den Wert $q$, der seinerseits ein echter Bruch ist, und daraus folgt nach dem Kriterium IIIa von § 2, Nr. 2 die Konvergenz dieser Reihe, deren Glieder nicht mehr von $x$ abhängen. Die oben betrachtete Potenzreihe konvergiert also gleichmäßig und stellt daher nach dem Satz des vorigen Paragraphen den Differentialquotienten $f'(x)$ der Funktion $f(x)$ dar, womit unsere Behauptung bewiesen ist.

Wenden wir dieses Ergebnis wiederum auf die Potenzreihe

$$f'(x) = \sum_{\nu=1}^{\infty} \nu c_\nu x^{\nu-1}$$

an, so erkennen wir, daß sich durch nochmalige gliedweise Differentiation die Potenzreihe

$$f''(x) = \sum_{\nu=2}^{\infty} \nu(\nu-1) c_\nu x^{\nu-2}$$

ergibt, und gelangen, so weiter schließend, allgemein zu dem Satze: *Jede durch eine Potenzreihe darstellbare Funktion läßt sich im Innern des Konvergenzintervalls beliebig oft differenzieren, und diese Differentiation darf gliedweise an der Potenzreihe vorgenommen werden*[1]).

---

[1]) Als expliziter Ausdruck der $k$-ten Ableitung ergibt sich

$$f^{(k)}(x) = \sum_{\nu=k}^{\infty} \nu(\nu-1)\cdots(\nu-k+1) c_\nu x^{\nu-k}$$

oder, ein wenig anders geschrieben,

$$\frac{f^{(k)}(x)}{k!} = \sum_{\nu=k}^{\infty} \binom{\nu}{k} c_\nu x^{\nu-k} = \sum_{\nu=0}^{\infty} \binom{k+\nu}{k} c_{k+\nu} x^\nu;$$

zwei häufig gebrauchte Formeln.

### 3. Das Rechnen mit Potenzreihen.

Das durch die vorigen Sätze gekennzeichnete Verhalten der Potenzreihen bildet die Grundlage für die Tatsache, daß man mit Potenzreihen ganz genau so rechnen kann wie mit ganzen rationalen Funktionen. Daß man zwei Potenzreihen addiert und subtrahiert, indem man die entsprechenden Koeffizienten addiert bzw. subtrahiert, ist selbstverständlich (siehe S. 301). Ebenso versteht es sich von selbst, daß man eine Potenzreihe — wie jede Reihe — mit einem konstanten Faktor multipliziert, indem man diese Multiplikation an jedem Gliede ausführt. Dagegen bedarf die Multiplikation und Division zweier Potenzreihen einer etwas näheren Betrachtung, für die ich auf den Anhang verweise. Hier erwähne ich ohne Beweis nur, daß man zwei Potenzreihen

$$f(x)=\sum_{\nu=0}^{\infty} a_\nu x^\nu$$

und

$$g(x)=\sum_{\nu=0}^{\infty} b_\nu x^\nu$$

miteinander multipliziert wie ganze rationale Funktionen. Es gilt nämlich der Satz: Das Produkt der beiden obigen Potenzreihen wird in dem beiden Potenzreihen gemeinsamen Konvergenzgebiete durch eine dritte Potenzreihe $\sum_{\nu=0}^{\infty} c_\nu x^\nu$ dargestellt, deren Koeffizienten durch die Formeln

$$\begin{aligned}
c_0 &= a_0 b_0,\\
c_1 &= a_0 b_1 + a_1 b_0,\\
c_2 &= a_0 b_2 + a_1 b_1 + a_2 b_0,\\
&\cdots\cdots\cdots\cdots\\
c_n &= a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}$$

gegeben werden. (Beweis siehe Anhang, § 1.)

### 4. Eindeutigkeitssatz für die Potenzreihen.

Für die Theorie der Potenzreihen ist folgende Tatsache von Wichtigkeit: Wenn man von zwei Potenzreihen $\sum_{\nu=0}^{\infty} a_\nu x^\nu$ und $\sum_{\nu=0}^{\infty} b_\nu x^\nu$ weiß, daß sie in einem gemeinsamen, den Punkt $x = 0$ im Innern enthaltenden Intervall konvergieren und dort eine und dieselbe Funktion $f(x)$ darstellen, so sind sie miteinander identisch, d. h. für jedes $n$ besteht die Gleichung $a_n = b_n$. Oder mit anderen Worten: *Eine Funktion $f(x)$ kann, wenn überhaupt, nur auf eine Art durch eine Potenzreihe in $x$ dargestellt werden.* Kurz: Die Darstellung durch eine Potenzreihe ist *„eindeutig“*.

Zum Beweise dieser Tatsache brauchen wir nur zu beachten, daß die Differenz dieser beiden Potenzreihen, d. h. die Potenzreihe $\varphi(x) = \sum_{\nu=0}^{\infty} c_\nu x^\nu$ mit den Koeffizienten $c_\nu = a_\nu - b_\nu$ für alle in dem Intervall gelegenen Werte von $x$ den Wert 0 besitzt, d. h. die Funktion $\varphi(x) = 0$ darstellt. Speziell muß also für $x = 0$ der Wert der Potenzreihe 0 sein, d. h. $c_0 = 0$, also $a_0 = b_0$. Differenzieren wir nun die Potenzreihe im Innern des Intervalls und beachten, daß der Differentialquotient $\varphi'(x)$ wieder überall 0 ist, so ergibt sich, daß auch die Potenzreihe $\sum_{\nu=1}^{\infty} \nu c_\nu x^{\nu-1}$ im ganzen Innern des Intervalls den Wert 0 besitzen muß, woraus speziell für $x = 0$ die Gleichung $c_1 = 0$ oder $a_1 = b_1$ folgt. In derselben Weise können wir fortfahren zu differenzieren und danach $x = 0$ zu setzen und finden dabei, daß sämtliche Koeffizienten $c_\nu$ verschwinden, was gerade unsere Behauptung darstellt.

Wir sehen übrigens, daß wir unseren Überlegungen auch folgende Wendung geben können: Differenzieren wir eine Reihe $f(x) = \sum a_\nu x^\nu$ $\nu$-mal und setzen dann $x = 0$, so erhalten wir sofort

$$a_\nu = \frac{1}{\nu!} f^{(\nu)}(0);$$

d. h.: *jede Potenzreihe, die nicht nur im Punkte* $x = 0$ ***konvergiert****, ist die Taylorsche Reihe der durch sie dargestellten Funktion.* Die Eindeutigkeit der Entwicklung kommt hierbei darin zum Ausdruck, daß sich die Koeffizienten in eindeutiger Weise aus dem Funktionsverlauf bestimmen.

## § 6. Entwickelung gegebener Funktionen in Potenzreihen. Methode der unbestimmten Koeffizienten. Beispiele.

Jede Potenzreihe stellt im Innern ihres Konvergenzgebiets eine stetige Funktion mit stetigen Ableitungen aller Ordnungen dar. Wir behandeln nun die umgekehrte Aufgabe, eine vorgegebene Funktion $f(x)$ in eine Potenzreihe zu entwickeln. Man kann hierzu im Prinzip stets die Taylorsche Formel heranziehen, stößt dabei aber im Einzelfalle oft auf Schwierigkeiten bei der wirklichen allgemeinen Berechnung der $n$-ten Ableitung sowie bei der Restabschätzung. Aber man kommt häufig in viel einfacherer Weise zum Ziele, wenn man folgendermaßen vorgeht: Man macht mit zunächst noch unbekannten Koeffizienten $c_\nu$ den Ansatz $f(x) = \sum_{\nu=0}^{\infty} c_\nu x^\nu$, bestimmt dann die Koeffizienten aus irgendwelchen bekannten Eigenschaften der Funktion $f(x)$ und beweist hinterher die Konvergenz der gefundenen Reihe. Diese Reihe stellt dann eine Funktion dar, und man hat sich jetzt nur noch davon zu überzeugen, daß diese mit $f(x)$ identisch ist. Wegen der bewiesenen Eindeutigkeit der Potenzreihenentwicklung ist man sicher, daß keine andere als die gefundene

Reihe die gewünschte Entwicklung darstellen kann. Für dieses Verfahren werden wir sogleich Beispiele kennen lernen. Übrigens haben wir im sechsten Kapitel schon die Reihen für $\operatorname{arc\,tg} x$ und $\log(1+x)$ auf einem Wege erhalten, der eigentlich in den Gedankenkreis des gegenwärtigen Kapitels hineingehört. Wir haben nämlich die Reihen für die Differentialquotienten dieser Funktionen, die wir als geometrische Reihen angeben konnten, einfach gliedweise integriert.

## 1. Die Exponentialfunktion.

Stellen wir uns die Aufgabe, eine Funktion $f(x)$ zu finden, für welche $f'(x) = f(x)$ und $f(0) = 1$ ist, und machen wir mit unbestimmten Koeffizienten den Potenzreihenansatz

$$f(x) = c_0 + c_1 x + c_2 x^2 + \cdots,$$

so erhalten wir durch Differentiation

$$f'(x) = c_1 + 2c_2 x + 3c_3 x^2 + \cdots.$$

Da diese beiden Potenzreihen nach Voraussetzung übereinstimmen sollen, ergibt sich sofort für jedes $n \geqq 1$ die Gleichung

$$n c_n = c_{n-1}.$$

Beachten wir noch, daß wegen $f(0) = 1$ der Koeffizient $c_0$ den Wert 1 haben muß, so können wir nun sukzessive alle Koeffizienten berechnen und erhalten die Potenzreihe

$$f(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots.$$

Diese Potenzreihe — wir machen bei diesen Betrachtungen absichtlich von dem früher über die Exponentialfunktion Gelernten keinen Gebrauch — konvergiert, wie wir mit Hilfe des Quotientenkriteriums leicht sehen, für alle Werte von $x$ und stellt daher eine Funktion $f(x)$ dar, für welche tatsächlich die Bedingungen $f'(x) = f(x)$, $f(0) = 1$ erfüllt sind.

Daß diese Funktion $f(x)$ mit der Exponentialfunktion $e^x$ übereinstimmt, folgt sofort, wenn wir beachten, daß jedenfalls die Funktion $e^x$ den obigen Bedingungen genügt. Bilden wir nun den Quotienten $\varphi(x) = \frac{f(x)}{e^x}$ und differenzieren, so ergibt sich $\varphi'(x) = \frac{e^x f'(x) - e^x f(x)}{e^{2x}} = 0$. Also ist die Funktion $\varphi(x)$ eine Konstante, und zwar muß sie, da sie für $x = 0$ den Wert 1 hat, gleich 1 sein, womit die Übereinstimmung unserer Potenzreihe mit der Exponentialfunktion bewiesen ist. (Vgl. auch die ganz analoge Betrachtung auf S. 143.)

## 2. Die binomische Reihe.

Die Herleitung der binomischen Reihe (sechstes Kapitel, § 3, Nr. 3) können wir jetzt leicht mit Hilfe der Methode der unbestimmten Koeffizienten nachholen. Wir wollen die Funktion $f(x) = (1 + x)^\alpha$ in eine Reihe entwickeln und setzen mit unbestimmten Koeffizienten an

$$f(x) = (1 + x)^\alpha = c_0 + c_1 x + c_2 x^2 + \cdots.$$

Beachten wir nun, daß für unsere Funktion offenbar die Beziehung

$$(1 + x) f'(x) = \alpha f(x) = \sum_{\nu=0}^{\infty} \alpha c_\nu x^\nu$$

gilt, und bilden wir, indem wir die obige Reihe gliedweise differenzieren und mit dem Faktor $1 + x$ multiplizieren, die Reihe

$$(1 + x) f'(x) = c_1 + (2c_2 + c_1) x + (3c_3 + 2c_2) x^2 + \cdots,$$

so erhalten wir durch Koeffizientenvergleichung die Beziehungen

$$\alpha c_0 = c_1, \quad \alpha c_1 = 2c_2 + c_1, \quad \alpha c_2 = 3c_3 + 2c_2, \quad \ldots.$$

Nun ist gewiß $c_0 = 1$, da unsere Reihe für $x = 0$ den Wert 1 darstellen muß, und so erhalten wir der Reihe nach für die Koeffizienten die Ausdrücke

$$c_1 = \alpha, \quad c_2 = \frac{(\alpha - 1)\alpha}{2}, \quad c_3 = \frac{(\alpha - 2)(\alpha - 1)\alpha}{3 \cdot 2}, \quad \ldots$$

und, wie man leicht bestätigt, allgemein

$$c_\nu = \frac{(\alpha - \nu + 1)(\alpha - \nu + 2)\cdots(\alpha - 1)\alpha}{\nu(\nu - 1)\cdots 2 \cdot 1} = \binom{\alpha}{\nu},$$

so daß wir tatsächlich die binomische Reihe

$$(1 + x)^\alpha = \sum_{\nu=0}^{\infty} \binom{\alpha}{\nu} x^\nu$$

gewinnen. Das Quotientenkriterium zeigt uns, daß diese Reihe für $|x| < 1$ konvergiert und für $|x| > 1$ divergiert, wenn $\alpha$ nicht ganzzahlig $\geqq 0$ ist; denn es ist in diesem Falle der Quotient des $(n + 1)$-ten durch das $n$-te Glied gleich $\frac{\alpha - n + 1}{n} x$, und dieser Ausdruck strebt absolut genommen gegen $|x|$, wenn $n$ über alle Grenzen wächst[1]. Unsere Reihe stellt uns also für $|x| < 1$ eine Funktion $f(x)$ dar, die, wie aus der

---

[1] Ohne Beweise gebe ich noch die genauen Konvergenzbedingungen für die Entwicklung an. Ist der Exponent $\alpha$ ganzzahlig und $\geqq 0$, so bricht die Reihe ab und gilt daher für alle $x$ (gewöhnlicher binomischer Satz). Bei allen anderen Werten von $\alpha$ ist die Reihe für $|x| < 1$ absolut konvergent, für $|x| > 1$ divergent. Für $x = +1$ liegt bei $\alpha > 0$ absolute Konvergenz, bei $-1 < \alpha < 0$ bedingte Konvergenz, bei $\alpha \leqq -1$ Divergenz vor. Für $x = -1$ schließlich ist die Reihe absolut konvergent bei $\alpha > 0$, divergent bei $\alpha < 0$.

Bildung der Koeffizienten hervorgeht, der Beziehung $(1+x)f'(x) = \alpha f(x)$ genügt. Außerdem ist $f(0) = 1$. Diese letzten Bedingungen aber charakterisieren unsere Funktion $f(x)$ als identisch mit der Funktion $(1+x)^\alpha$. Denn es ist für den Quotienten $\varphi(x) = \frac{f(x)}{(1+x)^\alpha}$

$$\varphi'(x) = \frac{(1+x)^\alpha f'(x) - \alpha(1+x)^{\alpha-1} f(x)}{(1+x)^{2\alpha}} = 0;$$

es ist $\varphi(x)$ also eine Konstante, und zwar gerade 1, da $\varphi(0) = 1$ ist.

Ich will von der binomischen Reihe die folgenden Spezialfälle besonders anführen: die geometrische Reihe

$$\frac{1}{1+x} = (1+x)^{-1} = 1 - x + x^2 - x^3 + x^4 - + \cdots = \sum_{\nu=0}^{\infty} (-1)^\nu x^\nu,$$

die negative Ableitung der geometrischen Reihe

$$\frac{1}{(1+x)^2} = (1+x)^{-2} = 1 - 2x + 3x^2 - 4x^3 + - \cdots = \sum_{\nu=0}^{\infty} (-1)^\nu (\nu+1) x^\nu$$

und die Reihen

$$\sqrt{1+x} = (1+x)^{\frac{1}{2}} = 1 + \frac{1}{2}x - \frac{1}{2\cdot 4}x^2 + \frac{1\cdot 3}{2\cdot 4\cdot 6}x^3 - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6\cdot 8}x^4 + - \cdots,$$

$$\frac{1}{\sqrt{1+x}} = (1+x)^{-\frac{1}{2}} = 1 - \frac{1}{2}x + \frac{1\cdot 3}{2\cdot 4}x^2 - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}x^3 + \frac{1\cdot 3\cdot 5\cdot 7}{2\cdot 4\cdot 6\cdot 8}x^4 - + \cdots,$$

von denen die ersten beiden oder die ersten drei Glieder beliebte Näherungsformeln geben.

### 3. Die Reihe für arc sin $x$

erhalten wir am einfachsten, indem wir den Ausdruck $\frac{1}{\sqrt{1-t^2}}$ nach dem binomischen Satz in die Reihe

$$(1-t^2)^{-\frac{1}{2}} = 1 + \frac{1}{2}t^2 + \frac{1\cdot 3}{2\cdot 4}t^4 + \cdots$$

entwickeln und diese für $|t| \leqq q < 1$ gleichmäßig konvergente Reihe zwischen 0 und $x$ integrieren; wir gewinnen auf diese Art sofort die Entwicklung

$$\arcsin x = x + \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} + \cdots,$$

welche für $|x| < 1$ konvergiert und, wie man aus dem Quotientenkriterium von S. 304 leicht sieht, für $|x| > 1$ divergiert.

Die Herleitung dieser Reihe aus dem Taylorschen Satz würde wegen der Schwierigkeit der Restabschätzung wesentlich unbequemer sein.

**4. Die Potenzreihenentwicklung von $\mathfrak{Ar}\,\mathfrak{Sin}\,x = \log(x + \sqrt{1+x^2})$** erhalten wir ganz ebenso, indem wir für ihre Ableitung die Entwicklung

$$\frac{1}{\sqrt{1+x^2}} = 1 - \frac{1}{2}x^2 + \frac{1\cdot 3}{2\cdot 4}x^4 - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}x^6 + - \cdots$$

auf Grund des binomischen Satzes hinschreiben und dann gliedweise integrieren. So gelangen wir zu der Reihenentwicklung

$$\mathfrak{Ar}\,\mathfrak{Sin}\,x = x - \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} - + \cdots,$$

deren Konvergenzstrecke das Intervall $-1 \leqq x \leqq 1$ ist.

**5. Beispiel für Reihenmultiplikation.**

Die Entwicklung der Funktion

$$\frac{\log(1+x)}{1+x}$$

gibt uns ein einfaches Beispiel für die Anwendung der Multiplikationsregel für Potenzreihen. Wir brauchen nämlich nur die logarithmische Reihe $\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + - \cdots$ mit der geometrischen Reihe $\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - + \cdots$ zu multiplizieren, dann erhalten wir, wie der Leser leicht ausrechnen kann, für $|x| < 1$ die merkwürdige Entwicklung

$$\frac{\log(1+x)}{1+x} = x - \left(1 + \frac{1}{2}\right)x^2 + \left(1 + \frac{1}{2} + \frac{1}{3}\right)x^3 - \left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}\right)x^4 + - \cdots.$$

**6. Beispiel für gliedweises Integrieren. Elliptisches Integral.**

In früheren Anwendungen ist uns das elliptische Integral

$$K = \int_0^{\frac{\pi}{2}} \frac{d\varphi}{\sqrt{1 - k^2\sin^2\varphi}} \qquad (k^2 < 1)$$

entgegengetreten (Schwingungszeit eines Pendels). Um das Integral auszuwerten, können wir so vorgehen, daß wir zuerst den Integranden nach dem binomischen Satz entwickeln:

$$\frac{1}{\sqrt{1 - k^2\sin^2\varphi}} = 1 + \frac{1}{2}k^2\sin^2\varphi + \frac{1\cdot 3}{2\cdot 4}k^4\sin^4\varphi + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}k^6\sin^6\varphi + \cdots.$$

Da die Reihe für alle Werte von $\varphi$ gleichmäßig konvergiert ($k^2\sin^2\varphi$ ist nie größer als $k^2$), so können wir sie gliedweise integrieren:

$$K = \int_0^{\frac{\pi}{2}} \frac{d\varphi}{\sqrt{1 - k^2\sin^2\varphi}} = \int_0^{\frac{\pi}{2}} d\varphi + \frac{1}{2}k^2\int_0^{\frac{\pi}{2}} \sin^2\varphi\,d\varphi + \frac{1\cdot 3}{2\cdot 4}k^4\int_0^{\frac{\pi}{2}} \sin^4\varphi\,d\varphi + \cdots.$$

Die hier auftretenden Integrale haben wir schon früher berechnet (vgl. viertes Kapitel, § 4, Nr. 4). Setzen wir ihre Werte ein, so ergibt sich

$$K = \int_0^{\frac{\pi}{2}} \frac{d\varphi}{\sqrt{1 - k^2 \sin^2 \varphi}} = \frac{\pi}{2}\left\{1 + \left(\frac{1}{2}\right)^2 k^2 + \left(\frac{1\cdot 3}{2\cdot 4}\right)^2 k^4 + \left(\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\right)^2 k^6 + \cdots\right\}.$$

Hinsichtlich weiterer Beispiele zur Reihentheorie verweise ich auf den Anhang.

## § 7. Potenzreihen mit komplexen Gliedern.

### 1. Einführung komplexer Glieder in Potenzreihen.

Die Ähnlichkeiten zwischen Potenzreihen für scheinbar ganz verschiedenartige Funktionen haben schon Euler dazu veranlaßt, einen Zusammenhang rein formal herzustellen, indem er für die Variable $x$ in den Potenzreihen auch komplexe Werte zuließ und insbesondere auch rein imaginäre Werte einsetzte. Wir wollen dies zunächst einmal unbekümmert tun und den Erfolg eines solchen Verfahrens studieren.

Die erste frappierende derartige Beziehung ergibt sich, wenn wir in der Exponentialreihe für $e^x$ die Größe $x$ durch die rein imaginäre Größe $i\varphi$ ersetzen, wo $\varphi$ wieder eine reelle Zahl bedeutet. Beachten wir die Grundbeziehung für die imaginäre Einheit $i$, nämlich $i^2 = -1$, aus der $i^3 = -i$, $i^4 = 1$, $i^5 = i, \ldots$ folgt, und trennen in der entstehenden Reihe Reelles und Imaginäres, so erhalten wir sofort

$$e^{i\varphi} = \left(1 - \frac{\varphi^2}{2!} + \frac{\varphi^4}{4!} - \frac{\varphi^6}{6!} + - \cdots\right) + i\left(\varphi - \frac{\varphi^3}{3!} + \frac{\varphi^5}{5!} - + \cdots\right)$$

oder anders geschrieben

$$e^{i\varphi} = \cos\varphi + i\sin\varphi.$$

Diese zunächst rein formale höchst wichtige „*Eulersche Relation*" steht im Einklang mit dem *Moivreschen Theorem*, welches wir schon im Anhang zum ersten Kapitel in Erinnerung gebracht haben, nämlich der Gleichung

$$(\cos\varphi + i\sin\varphi)(\cos\psi + i\sin\psi) = \cos(\varphi + \psi) + i\sin(\varphi + \psi).$$

Diese Gleichung besagt jetzt vermöge der Eulerschen Relation nur, daß auch für rein imaginäre Werte $x = i\varphi$, $y = i\psi$ die Beziehung

$$e^x \cdot e^y = e^{x+y}$$

besteht.

Ersetzt man in der Potenzreihe für $\cos x$ die Variable $x$ durch die rein imaginäre Größe $ix$, so erhält man sofort die Reihe für $\mathfrak{Cos}\, x$, eine

Beziehung, die wir durch die Gleichung

$$\mathfrak{Cof}\, x = \cos i\, x$$

ausdrücken können. Ganz ebenso erhalten wir

$$\mathfrak{Sin}\, x = \frac{1}{i} \sin i\, x\,.$$

Die Eulersche Relation liefert uns, da nach ihr auch $e^{-i\varphi} = \cos\varphi - i \sin\varphi$ gilt, für die trigonometrischen Funktionen die Exponentialdarstellungen

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i}, \qquad \cos x = \frac{e^{ix} + e^{-ix}}{2},$$

welche zu den Exponentialdarstellungen der Hyperbelfunktionen in voller Analogie stehen und in diese sofort durch die Beziehungen $\mathfrak{Cof}\, x = \cos i\, x$, $\mathfrak{Sin}\, x = \frac{1}{i} \sin i\, x$ übergehen.

Ganz entsprechende formale Beziehungen gelten nunmehr natürlich auch für die Funktionen $\operatorname{tg} x$, $\mathfrak{Tg}\, x$, $\operatorname{ctg} x$, $\mathfrak{Ctg}\, x$, welche durch die Gleichungen $\mathfrak{Tg}\, x = \frac{1}{i} \operatorname{tg} i\, x$, $\mathfrak{Ctg}\, x = i \operatorname{ctg} i\, x$ verknüpft sind.

Schließlich bestehen ähnliche Beziehungen auch für die Umkehrfunktionen der trigonometrischen bzw. hyperbolischen Funktionen.

Wir berechnen z. B. aus

$$y = \operatorname{tg} x = \frac{e^{ix} - e^{-ix}}{i\,(e^{ix} + e^{-ix})} = \frac{e^{2ix} - 1}{i\,(e^{2ix} + 1)}$$

sofort

$$e^{2ix} = \frac{1 + i\,y}{1 - i\,y}.$$

Logarithmieren wir diese Gleichung und schreiben wir $x$ anstatt $y$ und $\operatorname{arc\,tg} x$ anstatt $x$, so erhalten wir die Gleichung

$$\operatorname{arc\,tg} x = \frac{1}{2i} \log \frac{1 + i\,x}{1 - i\,x},$$

die einen merkwürdigen Zusammenhang des $\operatorname{arc\,tg} x$ mit dem Logarithmus andeutet. Setzen wir in die bekannte Potenzreihe für $\frac{1}{2} \log \frac{1+x}{1-x}$ die Größe $i\,x$ an Stelle von $x$ ein, so erhalten wir tatsächlich die Potenzreihe für $\operatorname{arc\,tg} x$

$$\operatorname{arc\,tg} x = \frac{1}{i}\left(i\,x + \frac{(i\,x)^3}{3} + \frac{(i\,x)^5}{5} + \cdots\right)$$
$$= x - \frac{x^3}{3} + \frac{x^5}{5} - + \cdots.$$

Die obigen Beziehungen tragen zunächst einen rein formalen Charakter, und es bedarf natürlich einer genaueren Präzisierung, was ihre inhaltliche Bedeutung sein soll. Wie dies mit Hilfe der Funktionentheorie geschehen kann, wird in der nächsten Nummer angedeutet werden.

Für das Folgende wird jedoch für uns lediglich die Eulersche Relation $e^{i\varphi} = \cos\varphi + i\sin\varphi$ eine Rolle spielen, und zu ihrer Begründung können wir eine solche eingehendere Analyse mit Hilfe der Funktionentheorie entbehren, indem wir hier einfach das *Symbol* $e^{i\varphi}$ *als eine formale Abkürzung* für die rechte Seite $\cos\varphi + i\sin\varphi$ auffassen, wobei dann die *Moivresche Formel* $e^{i\varphi}\cdot e^{i\psi} = e^{i(\varphi+\psi)}$ als einfache Folge der *elementaren trigonometrischen Additionstheoreme* erscheint. Weiter werden wir von diesem formalen Standpunkte aus, um die Beziehung $e^x \cdot e^y = e^{x+y}$ auch für komplexe Argumente allgemeingültig zu machen, definieren:

$$e^x = e^\xi(\cos\eta + i\sin\eta) \qquad \text{für } x = \xi + i\eta \qquad (\xi, \eta \text{ reell}).$$

## 2. Ausblick auf die allgemeine Funktionentheorie.

Trotzdem der gekennzeichnete rein formale Standpunkt in sich einwandfrei ist, bleibt doch der Wunsch bestehen, in den obigen Formeln mehr als bloß formale Verknüpfungen zu erkennen. Die Verfolgung dieses Zieles führt zu der allgemeinen Funktionentheorie, wie man abgekürzt die Theorie der sogenannten analytischen Funktionen einer komplexen Veränderlichen nennt. In ihr kann man davon ausgehen, daß man eine allgemeine Theorie der Potenzreihen mit komplexen Veränderlichen und mit komplexen Koeffizienten aufstellt. Der Aufbau einer solchen Theorie der Potenzreihen hat in der Tat keinerlei Schwierigkeiten, wenn man zunächst den Limesbegriff im Gebiet der komplexen Zahlen definiert hat, und vollzieht sich fast genau so wie bei reellen Größen. Da wir jedoch im folgenden von diesen Dingen keinen Gebrauch machen müssen, so will ich mich hier mit der Angabe einiger Tatsachen begnügen und auf die Beweise verzichten. Es zeigt sich, daß für Potenzreihen im Komplexen der folgende Satz gilt, der eine unmittelbare Verallgemeinerung des Satzes in § 5, Nr. 1 ist: *Wenn eine Potenzreihe für irgend einen komplexen Wert $x = \xi$ konvergiert, so konvergiert sie* ***absolut*** *für jeden Wert $x$, für den $|x| < |\xi|$ ist; wenn sie für einen Wert $x = \xi$ divergiert, so divergiert sie für jeden Wert $x$, für den $|x| > |\xi|$ ist. Eine Potenzreihe, die nicht überall und auch nicht bloß für $x = 0$ konvergiert, besitzt einen Konvergenzkreis, d. h. es gibt* ***eine Zahl*** *$\varrho > 0$, derart, daß die Reihe für $|x| < \varrho$ absolut konvergiert, für $|x| > \varrho$ divergiert.*

Hat man diesen Begriff der durch Potenzreihen dargestellten Funktionen einer komplexen Veränderlichen $x$ und die Regeln für das Rechnen mit ihnen einmal begründet, so kann man die Funktionen $e^x$, $\sin x$, $\cos x$, $\operatorname{arc\,tg} x$ usw. als Funktionen einer komplexen Veränderlichen $x$ einfach durch dieselben Potenzreihen definiert denken, durch die sie für reelle Werte von $x$ dargestellt werden, und alle die obigen formalen Relationen sind dann tatsächlich Selbstverständlichkeiten.

Ich möchte nur an zwei Beispielen den Nutzen der Heranziehung des Komplexen für das Verständnis der elementaren Funktionen andeuten. Die geometrische Reihe für $\frac{1}{1+x^2}$ hört beim Verlassen der Strecke $(-1 \ldots +1)$ auf zu konvergieren, ebenso die Reihe für $\operatorname{arc\,tg} x$, obwohl diese Funktionen an den Enden der Konvergenzstrecke keinerlei Besonderheiten aufweisen, sondern für alle reellen Werte von $x$ mit allen ihren Ableitungen stetig sind. (Daß die Konvergenz der Reihen für $\frac{1}{1-x^2}$ und für $\log(1-x)$ beim Überschreiten der Stelle $x=1$ aufhört, ist uns dagegen ohne weiteres verständlich, weil die Funktionen an diesen Stellen unendlich werden.) Das Aufhören der Konvergenz der Arkustangensreihe und der Reihe $\sum_{\nu=0}^{\infty}(-1)^{\nu} x^{2\nu}$ für $|x|>1$ wird aber sofort klar, wenn man auch komplexe Werte von $x$ in Betracht zieht. Für $x=i$ werden nämlich diese Funktionen unendlich, können also nicht mehr durch eine konvergente Reihe dargestellt werden; und daher verbietet von selbst der Satz vom Konvergenzkreise die Konvergenz im Reellen außerhalb der Strecke $|x| \leqq 1$.

Ein zweites Beispiel gibt uns die früher (S. 264) behandelte Funktion $f(x)=e^{-\frac{1}{x^2}}$ für $x \neq 0$, $f(0)=0$, die sich trotz ihres augenscheinlich regulären Verhaltens nicht in eine Taylorsche Reihe entwickeln ließ. In der Tat hört aber die Stetigkeit der Funktion in der Umgebung des Nullpunktes sofort auf, wenn man rein imaginäre Werte $x=i\xi$ heranzieht. Die Funktion geht dann in $e^{\frac{1}{\xi^2}}$ über und strebt für $\xi \to 0$ über alle Grenzen; es ist daher nunmehr verständlich, daß eine Potenzreihenentwicklung nach Potenzen von $x$ für sie nicht existiert.

Mit diesen Andeutungen über die Theorie der Funktionen und Potenzreihen im Komplexen muß ich mich hier begnügen.

# Anhang zum achten Kapitel.

## § 1. Multiplikation und Division von Reihen.

### 1. Multiplikation absolut konvergenter Reihen.

Es seien

$$A=\sum_{\nu=0}^{\infty} a_\nu, \qquad B=\sum_{\nu=0}^{\infty} b_\nu$$

zwei absolut konvergente Reihen, mit denen wir gleichzeitig die Reihen ihrer absoluten Beträge $\mathsf{A}=\sum_{\nu=0}^{\infty}|a_\nu|$, $\mathsf{B}=\sum_{\nu=0}^{\infty}|b_\nu|$ betrachten. Wir setzen ferner

$$A_n = \sum_{\nu=0}^{n} a_\nu, \quad B_n = \sum_{\nu=0}^{n} b_\nu, \quad \boldsymbol{A}_n = \sum_{\nu=0}^{n} |a_\nu|, \quad \boldsymbol{B}_n = \sum_{\nu=0}^{n} |b_\nu|$$

und

$$c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0.$$

Unsere Behauptung ist, daß die Reihe $\sum_{\nu=0}^{\infty} c_\nu$ absolut konvergiert und daß ihre Summe gleich $A \cdot B$ ist.

Zum Beweise denken wir uns die $n$-ten Partialsummen

$$A_n \text{ und } B_n \text{ bzw. } \boldsymbol{A}_n \text{ und } \boldsymbol{B}_n$$

miteinander multipliziert:

$$A_n B_n = a_0 b_0 + a_1 b_0 + a_0 b_1 + a_1 b_1 + \cdots + a_n b_n,$$

$$\boldsymbol{A}_n \boldsymbol{B}_n = |a_0| |b_0| + |a_1| |b_0| + |a_0| |b_1| + |a_1| |b_1| + \cdots + |a_n| |b_n|.$$

Läßt man $n$ wachsen, so erkennt man, daß rechts Reihen erscheinen, deren einzelne Glieder die Form $a_\nu b_\mu$ bzw. $|a_\nu| |b_\mu|$ haben. Diese letzte Reihe besteht aus nicht negativen Gliedern. Ihre Partialsummen wachsen also beständig. Da diese Partialsummen wegen $\boldsymbol{A}_n \leqq \boldsymbol{A}$, $\boldsymbol{B}_n \leqq \boldsymbol{B}$ unterhalb der festen Schranke $\boldsymbol{A}\boldsymbol{B}$ liegen, konvergiert also die zweite Reihe; somit konvergiert auch die erste Reihe absolut genommen. Die Summen der beiden Reihen sind offenbar $A B$ bzw. $\boldsymbol{A}\boldsymbol{B}$. Da man in konvergenten Reihen an beliebig vielen Stellen aufeinanderfolgende Glieder zusammenfassen kann, ohne ihren Wert zu ändern, da ferner absolute Konvergenz dabei offensichtlich erhalten bleibt und da die Reihe $\sum_{\nu=0}^{\infty} c_\nu$ auf solche Art aus der oben hingeschriebenen Reihe entsteht, so ist damit der gewünschte Nachweis erbracht.

## 2. Multiplikation und Division von Potenzreihen.

Die Hauptanwendung findet unser Satz in der Theorie der Potenzreihen. Man folgert aus ihm ohne weiteres die folgende Tatsache: Das Produkt der beiden Potenzreihen

$$\sum_{\nu=0}^{\infty} a_\nu x^\nu \quad \text{und} \quad \sum_{\nu=0}^{\infty} b_\nu x^\nu$$

wird im Innern jedes den beiden Potenzreihen gemeinsamen Konvergenzgebietes durch eine dritte Potenzreihe $\sum_{\nu=0}^{\infty} c_\nu x^\nu$ dargestellt, deren Koeffizienten durch die Formel

$$c_\nu = a_0 b_\nu + a_1 b_{\nu-1} + \cdots + a_\nu b_0$$

gegeben werden.

Was die Division von Potenzreihen anbetrifft, so kann man den Quotienten der obigen beiden Potenzreihen ebenfalls durch eine Potenzreihe $\sum_{\nu=0}^{\infty} q_\nu x^\nu$ darstellen, falls das konstante Glied $b_0$ des Nenners nicht verschwindet. (Im letzten Falle ist eine solche Darstellung offenbar deswegen im allgemeinen unmöglich, weil sie für $x = 0$ wegen des Verschwindens des Nenners nicht konvergieren könnte, während doch andererseits jede Potenzreihe in $x$ für $x = 0$ konvergieren muß.) Die Koeffizienten der Potenzreihe

$$\sum_{\nu=0}^{\infty} q_\nu x^\nu$$

lassen sich der Reihe nach berechnen, indem man beachtet, daß $\sum_{\nu=0}^{\infty} q_\nu x^\nu \cdot \sum_{\nu=0}^{\infty} b_\nu x^\nu = \sum_{\nu=0}^{\infty} a_\nu x^\nu$ sein muß, daß also die Gleichungen

$$\begin{aligned} a_0 &= q_0 b_0, \\ a_1 &= q_0 b_1 + q_1 b_0, \\ a_2 &= q_0 b_2 + q_1 b_1 + q_2 b_0, \\ &\cdots\cdots\cdots\cdots\cdots \\ a_\nu &= q_0 b_\nu + q_1 b_{\nu-1} + \cdots + q_\nu b_0 \end{aligned}$$

gelten müssen. Aus der ersten dieser Gleichungen ergibt sich $q_0$, aus der zweiten können wir den Wert $q_1$ ausrechnen, aus der dritten sodann, indem wir die gefundenen Werte immer mit benutzen, den Wert $q_2$ usw. Um die Darstellung des Quotienten der beiden Potenzreihen durch die dritte Potenzreihe jedoch streng zu rechtfertigen, müssen wir noch untersuchen, wo die formal gefundene Potenzreihe $\sum_{\nu=0}^{\infty} q_\nu x^\nu$ konvergiert. Ich möchte diese allgemeine Untersuchung, von deren Resultat wir keinen weiteren Gebrauch machen werden, übergehen und mich hier mit der Angabe begnügen, daß die Reihe für den Quotienten tatsächlich konvergiert, solange $x$ auf ein passend klein zu wählendes Intervall beschränkt bleibt, für welches jedenfalls der Nenner nicht verschwindet und sowohl Zähler als auch Nenner konvergiert.

## § 2. Grenzübergänge, die mit der Exponentialfunktion zusammenhängen.

### 1. Die Gleichmäßigkeit des Grenzüberganges $\left(1+\frac{x}{n}\right)^n \to e^x$.

Für jede positive Zahl $a$ ist der Grenzübergang $\left(1+\frac{x}{n}\right)^n \to e^x$ im Intervalle $0 \leqq x \leqq a$ gleichmäßig, ebenso wie der Grenzübergang $\left(1-\frac{x}{n}\right)^n \to e^{-x}$. Es genügt, den ersten Teil der Behauptung zu beweisen, da der Beweis für den zweiten Teil genau so verläuft. Zu dem Zweck gehen wir von der Grenzwertformel

$$(1+h)^{\frac{1}{h}} \to e$$

aus, deren Sinn folgender ist: Zu jeder vorgegebenen positiven, wenn auch noch so kleinen Zahl $\delta$ kann man eine positive Größe $k$ genügend klein wählen, derart, daß für $0 < h \leqq k$ die Beziehung

$$e(1-\delta) < (1+h)^{\frac{1}{h}} < e(1+\delta)$$

besteht. Nun ist für $h = \frac{x}{n}$ offenbar

$$\left(1+\frac{x}{n}\right)^n = (1+h)^{\frac{x}{h}} = \left((1+h)^{\frac{1}{h}}\right)^x.$$

Wenn $\frac{a}{n} = k$ gesetzt wird, wo $k$ nur von $n$ und $a$, nicht von $x$ abhängt und bei wachsendem $n$ gegen Null strebt, so ist für das ganze Intervall $\frac{x}{n} = h \leqq k$, und somit ist für beliebig kleines $\delta$ mit $0 < \delta < 1$ bei genügend groß gewähltem $n$, etwa für $n > N = N(\delta)$,

$$e(1-\delta) < (1+h)^{\frac{1}{h}} < e(1+\delta),$$

also auch

$$e^x(1-\delta)^x \leqq (1+h)^{\frac{x}{h}} = \left(1+\frac{x}{n}\right)^n \leqq e^x(1+\delta)^x,$$

gleichgültig, wo man in unserm Intervall den Punkt $x = nh$ wählt. Nun ist für $0 \leqq x \leqq a$

$$(1+\delta)^x \leqq (1+\delta)^a$$

und

$$(1-\delta)^x \geqq (1-\delta)^a.$$

Wir erhalten also

$$e^x(1-\delta)^a \leqq \left(1+\frac{x}{n}\right)^n \leqq e^x(1+\delta)^a,$$

$$e^x((1-\delta)^a - 1) \leqq \left(1+\frac{x}{n}\right)^n - e^x \leqq e^x((1+\delta)^a - 1) \quad \text{für} \quad n > N(\delta).$$

Die letzte Ungleichung aber drückt — da $(1-\delta)^a$ und $(1+\delta)^a$ beliebig nahe an 1 liegen, wenn $\delta$ hinreichend klein gewählt wird, und da ferner der Faktor $e^x$ in unserm Intervall beschränkt ist — die behauptete Gleichmäßigkeit aus. Wir erkennen nämlich aus ihr, daß tatsächlich, wenn man eine beliebig kleine positive Zahl vorgibt, bei jedem hinreichend groß gewählten $n$ der Ausdruck $\left|\left(1+\frac{x}{n}\right)^n - e^x\right|$ im ganzen Intervall $0 \leqq x \leqq a$ unter die betreffende Zahl hinabsinkt.

### 2. Bemerkung über Integration und Differentiation der Exponentialfunktion.

Die bewiesene Gleichmäßigkeit gestattet folgende zur Illustration unserer früheren Integrationssätze nützliche Betrachtung. Es ist

$$\int_0^x \left(1+\frac{t}{n}\right)^n dt = \frac{n}{n+1}\left(\left(1+\frac{x}{n}\right)^{n+1} - 1\right),$$

wie wir sofort durch die Substitution $u = 1 + \frac{t}{n}$ in der linken Seite der Gleichung erhalten. Nunmehr dürfen wir wegen der Gleichmäßigkeit der Konvergenz den Grenzübergang $n \to \infty$ unter dem Integralzeichen ausführen und erhalten

$$\int_0^x e^t\,dt = e^x - 1,$$

also die Integrationsformel für die Exponentialfunktion.

### 3. Beweis der Formel $\int_0^\infty e^{-x^2}\,dx = \frac{1}{2}\sqrt{\pi}$.

Eine andere, tiefer gehende Anwendung machen wir unter Zuhilfenahme des in etwas abweichender Gestalt früher (viertes Kapitel, § 4, Nr. 4) bei der Wallisschen Formel erhaltenen Resultates

$$J_n = \sqrt{n}\int_0^{\frac{\pi}{2}} \cos^{2n+1} x\,dx = \sqrt{n}\cdot\frac{2\cdot 4\cdots 2n}{3\cdot 5\cdots(2n+1)} \to \frac{1}{2}\sqrt{\pi}$$

und des Begriffes „uneigentliches Integral". Führen wir in die eben hingeschriebene Integralformel $u = \sin x$ als neue Integrationsvariable ein, so erhalten wir sofort $\cos x\,dx = du$ und

$$\lim_{n\to\infty} \sqrt{n}\int_0^1 (1-u^2)^n\,du = \frac{1}{2}\sqrt{\pi}.$$

Setzen wir $u = \frac{t}{\sqrt{n}}$, so gewinnen wir

$$J_n = \int_0^{\sqrt{n}} \left(1 - \frac{t^2}{n}\right)^n dt \to \frac{1}{2}\sqrt{\pi}.$$

Nunmehr schreiben wir, indem wir unter $A$ eine feste positive Zahl $< \sqrt{n}$ verstehen,

$$J_n = K_A + R_n,$$

wo

$$K_A = \int_0^A \left(1 - \frac{t^2}{n}\right)^n dt, \qquad R_n = \int_A^{\sqrt{n}} \left(1 - \frac{t^2}{n}\right)^n dt$$

gesetzt ist.

Wir notieren ferner die Beziehung

$$1 - \alpha < e^{-\alpha} \quad \text{für} \quad \alpha > 0,$$

welche sofort aus dem Mittelwertsatz der Differentialrechnung in der Form

$$e^{-\alpha} - 1 = e^{-\alpha} - e^0 = -\alpha e^{-\vartheta\alpha} > -\alpha$$

folgt. Insbesondere ergibt sich für $\alpha = \frac{t^2}{n} (\leqq 1)$

$$1 - \frac{t^2}{n} < e^{-\frac{t^2}{n}}, \quad \left(1 - \frac{t^2}{n}\right)^n < e^{-t^2}.$$

Ist nun, was wir voraussetzen wollen, $A > 1$, so wird für $t > A$ sicher $e^{-t^2} < e^{-t}$, und wir erhalten für den zweiten Bestandteil $R_n$ die Abschätzungen

$$0 < R_n < \int_A^{\sqrt{n}} e^{-t^2} dt < \int_A^{\sqrt{n}} e^{-t} dt = e^{-A} - e^{-\sqrt{n}},$$

d. h. jedenfalls $0 < R_n < e^{-A}$. Nunmehr führen wir den Grenzübergang $n \to \infty$ aus. Wegen der in Nr. 1 bewiesenen Gleichmäßigkeitseigenschaft geht dabei das Integral $K_A$ in $\int_0^A e^{-t^2} dt$ über, und wir erhalten also sofort

$$0 \leqq \frac{1}{2}\sqrt{\pi} - \int_0^A e^{-t^2} dt = \lim_{n\to\infty} (J_n - K_A) = \lim_{n\to\infty} R_n \leqq e^{-A}.$$

Jetzt dürfen wir die bisher festgehaltene positive Zahl $A < \sqrt{n}$ eine beliebige über alle Grenzen wachsende Folge positiver Werte durchlaufen lassen, wobei $n$ über alle Grenzen wächst und $e^{-A}$ gegen 0 strebt; es ergibt sich die Gleichung

$$\lim_{A\to\infty} \int_0^A e^{-t^2} dt = \int_0^\infty e^{-t^2} dt = \frac{1}{2}\sqrt{\pi},$$

die wir beweisen wollten und die sich übrigens, da $e^{-t^2}$ eine gerade Funktion ist, auch in der Form

$$\int_{-\infty}^{+\infty} e^{-t^2} dt = \sqrt{\pi}$$

schreiben läßt.

## § 3. Unendliche Reihen und uneigentliche Integrale.

Die unendlichen Reihen und die bei ihnen angewandten Begriffsbildungen führen zu einfachen Anwendungen und Analogien in der Theorie der uneigentlichen Integrale (vgl. viertes Kapitel, § 8). Ich beschränke mich hier auf den Fall eines konvergenten Integrales mit unendlichem Integrationsintervall, etwa eines Integrales der Form $\int_0^\infty f(x)\,dx$. Teilt man das Integrationsintervall durch die monoton gegen $\infty$

strebende Zahlenfolge $x_0 = 0, x_1, \ldots$ ein, so kann man das uneigentliche Integral in die Gestalt

$$\int_0^\infty f(x)\,dx = a_1 + a_2 + \cdots$$

setzen, wobei jedes Glied unserer unendlichen Reihe wieder ein Integral, nämlich

$$a_1 = \int_0^{x_1} f(x)\,dx, \qquad a_2 = \int_{x_1}^{x_2} f(x)\,dx, \qquad \ldots$$

ist. Dabei ist die Wahl der Punkte $x_\nu$ gleichgültig. Man kann also den Begriff eines konvergenten uneigentlichen Integrales in mannigfacher Weise auf den einer unendlichen Reihe zurückführen.

Besonders bequem ist es, die Stellen $x_\nu$ so zu wählen, daß der Integrand innerhalb jedes einzelnen Teilintervalls das Vorzeichen nicht wechselt. Der Reihe $\sum\limits_{\nu=1}^{\infty} |a_\nu|$ würde dann das Integral des absoluten Betrages unserer Funktion:

$$\int_0^\infty |f(x)|\,dx$$

entsprechen. Wir werden so ganz von selbst auf folgenden Begriff geführt: *Ein uneigentliches Integral* $\int\limits_0^\infty f(x)\,dx$ *heißt absolut konvergent, wenn das Integral* $\int\limits_0^\infty |f(x)|\,dx$ *existiert.* Andernfalls heißt unser Integral, wenn es überhaupt existiert, *bedingt konvergent.*

Die früher betrachteten Integrale

$$\int_0^\infty \frac{1}{1+x^2}\,dx, \qquad \int_0^\infty e^{-x^2}\,dx, \qquad \Gamma(x) = \int_0^\infty e^{-t}\,t^{x-1}\,dt$$

sind alle absolut konvergent. Dagegen bildet das schon auf S. 203 f. behandelte Integral

$$\int_0^\infty \frac{\sin x}{x}\,dx = \lim_{A\to\infty} \int_0^A \frac{\sin x}{x}\,dx$$

ein einfaches Beispiel für ein bedingt konvergentes Integral. Um für die Konvergenz des Integrales einen von dem früheren unabhängigen Beweis zu geben, zerlegen wir das Intervall von 0 bis $A$ durch die Stellen $x_\nu = \nu\pi$ $(\nu = 0, 1, 2, \ldots)$. Wir zerlegen somit das Integral in Summanden der Form $a_\nu = \int\limits_{(\nu-1)\pi}^{\nu\pi} \frac{\sin x}{x}\,dx$ $(\nu = 1, 2, \ldots)$, von denen wir sofort erkennen, daß sie abwechselnde Vorzeichen haben und daß $|a_{\nu+1}| < |a_\nu|$ ist, und in einen Rest $\int\limits_{\mu\pi}^{A} \frac{\sin x}{x}\,dx$ mit $0 \leqq A - \mu\pi < \pi$, der mit wachsendem $A$ gegen 0 strebt. Die Reihe $a_1 + a_2 + a_3 + \cdots$ konvergiert, da $a_\nu \underset{\nu\to\infty}{\longrightarrow} 0$ gilt, nach dem Leibnizschen Kriterium vom achten Kapitel, § 1, und hieraus folgt die Konvergenz unseres Integrales. Daß diese Konvergenz nur bedingt ist, ergibt sich durch Vergleich mit der harmonischen Reihe.

## § 4. Unendliche Produkte.

Schon in den Vorbemerkungen zu diesem Kapitel haben wir hervorgehoben, daß die unendlichen Reihen nur eine, allerdings eine besonders wichtige Art der Darstellung von Größen oder Funktionen durch unendliche Prozesse sind. Ich möchte hier ohne Beweise als Beispiel für einen anderen derartigen Prozeß die unendlichen Produkte anführen.

Im vierten Kapitel, § 4, haben wir die Wallissche Formel

$$\frac{\pi}{2} = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdots$$

kennengelernt, welche die Zahl $\frac{\pi}{2}$ durch ein „unendliches Produkt" darstellt. Als den Wert eines *unendlichen Produktes*

$$\prod_{\nu=1}^{\infty} a_\nu = a_1\cdot a_2\cdot a_3\cdot a_4\cdot\,\cdots$$

bezeichnen wir dabei den Grenzwert der Folge der Teilprodukte

$$a_1, \quad a_1\cdot a_2, \quad a_1\cdot a_2\cdot a_3, \quad a_1\cdot a_2\cdot a_3\cdot a_4, \quad \ldots,$$

falls er existiert.

Die Faktoren $a_1, a_2, a_3, \ldots$ eines solchen Produktes können natürlich auch Funktionen einer Variablen $x$ sein. Ein besonders interessantes Beispiel ist die „Produktzerlegung" der Funktion $\sin \pi x$, die wir im nächsten Kapitel, § 4, Nr. 7, ableiten werden,

$$\sin \pi x = \pi x\left(1 - \frac{x^2}{1^2}\right)\left(1 - \frac{x^2}{2^2}\right)\left(1 - \frac{x^2}{3^2}\right)\cdots.$$

Eine sehr große Rolle spielt in der Zahlentheorie die *Produktzerlegung der „Zetafunktion"*. Um mit der in der Zahlentheorie üblichen Bezeichnungsweise im Einklang zu bleiben, will ich hier die unabhängige Variable mit $s$ bezeichnen und diese Funktion für $s > 1$ durch die Darstellung

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

definieren. Wir wissen aus Kap. 8, § 2, Nr. 3, daß für $s > 1$ die Reihe rechts konvergiert. Ist nun $p$ irgend eine Zahl größer als 1, so erhalten wir durch Entwicklung in eine geometrische Reihe ohne weiteres die Gleichung

$$\frac{1}{1-\frac{1}{p^s}} = 1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \frac{1}{p^{3s}} + \cdots.$$

Denken wir uns hier für $p$ der Reihe nach alle Primzahlen $p_1, p_2, p_3, \ldots$ in wachsender Folge eingesetzt und alle die entstehenden Gleichungen miteinander multipliziert, so erhalten wir links ein Produkt der Form

$$\frac{1}{1-p_1^{-s}} \cdot \frac{1}{1-p_2^{-s}} \cdots.$$

Wenn wir die Reihen auf der rechten Seite unserer Gleichung miteinander, unbekümmert um eine nähere Rechtfertigung dieses Verfahrens, ausmultiplizieren und dabei beachten, daß sich nach einem elementaren Satze jede ganze Zahl $n > 1$ auf eine und nur eine Weise als Produkt von Potenzen verschiedener Primzahlen darstellen läßt, so bemerken wir, daß rechts gerade die Funktion $\zeta(s)$ entsteht, und wir gelangen so zu der merkwürdigen Produktzerlegung

$$\zeta(s) = \frac{1}{1-p_1^{-s}} \cdot \frac{1}{1-p_2^{-s}} \cdot \frac{1}{1-p_3^{-s}} \cdots.$$

Diese Produktzerlegung, deren Beweis ich hier nur kurz skizziert habe, ist tatsächlich eine Zerlegung der $\zeta$-Funktion in ein unendliches Produkt, da es unendlich viele Primzahlen gibt.

In der allgemeinen Theorie der unendlichen Produkte pflegt man den Fall auszuschließen, daß das Produkt $a_1 a_2 \cdots a_n$ den Grenzwert Null hat. Insbesondere darf also keiner der Faktoren $a_n$ verschwinden. Es müssen dann, damit das Produkt konvergiert, die Faktoren $a_n$ mit wachsendem $n$ gegen 1 streben. Indem wir nötigenfalls eine endliche Anzahl von Faktoren fortlassen, was für die Frage der Konvergenz unerheblich ist, können wir also erreichen, daß $a_n > 0$ ist. Auf diesen Fall bezieht sich die folgende Tatsache: Notwendig und hinreichend für die Konvergenz des Produktes $\prod_{\nu=1}^{\infty} a_\nu$ mit $a_\nu > 0$ ist die Konvergenz der unendlichen Reihe $\sum_{\nu=1}^{\infty} \log a_\nu$. Denn es ist klar, daß die Teilsummen dieser Reihe dann und nur dann einem Grenzwert zustreben, wenn die Teilprodukte $a_1 a_2 \cdots a_n$ einen positiven Grenzwert besitzen.

Gewöhnlich bedient man sich, indem man $a_n = 1 + \alpha_n$ setzt, bei Konvergenzuntersuchungen des folgenden hinreichenden Kriteriums: Das Produkt

$$\prod_{\nu=1}^{\infty} (1 + \alpha_\nu)$$

konvergiert, wenn die Reihe

$$\sum_{\nu=1}^{\infty} |\alpha_\nu|$$

konvergiert. Zum Beweise dürfen wir, nötigenfalls unter Weglassung einer endlichen Anzahl von Faktoren, annehmen, daß jedes $|\alpha_\nu| < \frac{1}{2}$ sei. Es ist dann also auch $1 - |\alpha_\nu| > \frac{1}{2}$. Nun wird nach dem Mittelwertsatz $\log(1+h) = \log(1+h) - \log 1 = h\frac{1}{1+\vartheta h}$ mit $0 < \vartheta < 1$. Somit ist

$$|\log(1+\alpha_\nu)| = \left|\frac{\alpha_\nu}{1+\vartheta\alpha_\nu}\right| \leqq \frac{|\alpha_\nu|}{1-|\alpha_\nu|} \leqq 2|\alpha_\nu|,$$

und daher folgt die Konvergenz der Reihe $\sum_{\nu=1}^{\infty} \log(1+\alpha_\nu)$ aus der Konvergenz von $\sum_{\nu=1}^{\infty} |\alpha_\nu|$.

Aus unserem Kriterium ergibt sich ohne weiteres die Konvergenz der oben für den $\sin \pi x$ angegebenen Produktdarstellung für alle Werte von $x$. — Ferner schließt man leicht für $p \geqq 2$ und $s > 1$

$$\frac{1}{1-p^{-s}} = 1 + \frac{1}{p^s - 1}, \quad 0 < \frac{1}{p^s - 1} < \frac{2}{p^s}.$$

Da nun, wenn wir $p$ die Reihe der Primzahlen durchlaufen lassen, die Reihe $\sum \frac{1}{p^s}$ konvergieren muß — ihre Glieder bilden ja nur einen Teil der Glieder der konvergenten Reihe $\sum_{\nu=1}^{\infty} \frac{1}{\nu^s}$ —, so ist die Konvergenz des Produktes $\prod \frac{1}{1-p^{-s}}$ für $s > 1$ bewiesen.

## § 5. Weitere Beispiele für unendliche Reihen.

### 1. Verschiedene Entwicklungen.

Als Beispiele für die Methode der unbestimmten Koeffizienten gebe ich die folgenden Reihenentwicklungen:

$$\frac{\arcsin x}{\sqrt{1-x^2}} = x + \frac{2}{3}x^3 + \frac{2\cdot 4}{3\cdot 5}x^5 + \cdots, \tag{1}$$

$$(\arcsin x)^2 = \frac{x^2}{1} + \frac{2}{3}\frac{x^4}{2} + \frac{2\cdot 4}{3\cdot 5}\frac{x^6}{3} + \cdots, \tag{2}$$

$$\cos(\mu \arcsin x) = 1 - \frac{\mu^2}{2!}x^2 + \frac{\mu^2(\mu^2-2^2)}{4!}x^4 - \frac{\mu^2(\mu^2-2^2)(\mu^2-4^2)}{6!}x^6 + - \cdots, \tag{3}$$

aus deren letzter durch Differentiation die Reihe

$$\frac{\sin(\mu \arcsin x)}{\sqrt{1-x^2}} = \frac{\mu}{1!}x - \frac{\mu(\mu^2-2^2)}{3!}x^3 + \frac{\mu(\mu^2-2^2)(\mu^2-4^2)}{5!}x^5 - + \cdots$$

folgt. Ferner die Reihe

(4) $$\sin(\mu \arcsin x) = \frac{\mu}{1!}x - \frac{\mu(\mu^2-1^2)}{3!}x^3 + \frac{\mu(\mu^2-1^2)(\mu^2-3^2)}{5!}x^5 - + \cdots,$$

aus der sich

$$\frac{\cos(\mu \arcsin x)}{\sqrt{1-x^2}} = 1 - \frac{\mu^2-1^2}{2!}x^2 + \frac{(\mu^2-1^2)(\mu^2-3^2)}{4!}x^4 - + \cdots$$

ableitet. Alle diese Reihenentwicklungen sind für $|x| < 1$ gültig.

Aus den Reihen (3) und (4) ergeben sich, indem wir $x = \sin\varphi$ setzen, die Entwicklungen der Funktionen $\cos\mu\varphi$ und $\sin\mu\varphi$ nach Potenzen von $\sin\varphi$. Diese Potenzreihen brechen ab, d. h. aus ihnen werden Polynome, wenn $\mu$ eine ganze Zahl, und zwar in der ersten Reihe gerade, in der zweiten ungerade oder auch Null ist.

Wir gewinnen diese Reihen folgendermaßen: Zunächst machen wir für die zu entwickelnde Funktion $y = f(x)$ einen Ansatz mit unbestimmten Koeffizienten: $y = \sum_{\nu=0}^{\infty} c_\nu x^\nu$, stellen dann — in der Form von Differentialgleichungen und Bedingungen für die Stelle $x = 0$ — Forderungen für die Funktion $f(x)$ auf, aus denen sich die Koeffizienten bestimmen, beweisen die Konvergenz der so gefundenen Potenzreihe und haben schließlich zu zeigen, daß deren Summe $\varphi(x)$ mit $f(x)$ identisch ist.

Zunächst behandeln wir die Funktion $y = f(x) = \frac{\arcsin x}{\sqrt{1-x^2}}$. Für sie ist $y' = \frac{1}{1-x^2} + \frac{x \arcsin x}{(\sqrt{1-x^2})^3}$, also $y'(1-x^2) - xy - 1 = 0$ und ferner $f(0) = 0$. Setzen wir die Potenzreihe $\sum_{\nu=0}^{\infty} c_\nu x^\nu$ links in die obige Gleichung ein und ordnen die linke Seite nach Potenzen von $x$, so finden wir, daß die Potenzreihe

$$\begin{aligned}(c_1 - 1) &+ (2c_2 - c_0)x + (3c_3 - 2c_1)x^2 \\ &+ \cdots + ((\nu+2)c_{\nu+2} - (\nu+1)c_\nu)x^{\nu+1} + \cdots\end{aligned}$$

für alle Werte verschwinden muß; d. h. (nach S. 324) es müssen alle ihre Koeffizienten verschwinden, woraus wir der Reihe nach sämtliche Werte $c_0, c_1, c_2, \ldots$ bestimmen können. Beachten wir nämlich, daß wegen $f(0) = 0$ auch $c_0 = 0$ sein muß, so ergibt sich $c_0 = c_2 = c_4 = \cdots = 0$ und

$$c_1 = 1, \quad c_3 = \frac{2}{3}, \quad c_5 = \frac{2}{3}\cdot\frac{4}{5}, \quad c_7 = \frac{2}{3}\cdot\frac{4}{5}\cdot\frac{6}{7}, \ \cdots;$$

so erhalten wir die Reihe

$$\varphi(x) = x + \frac{2}{3}x^3 + \frac{2\cdot 4}{3\cdot 5}x^5 + \cdots,$$

deren Konvergenz für $|x| < 1$ man sofort dem Quotientenkriterium entnimmt.

Wir haben noch zu zeigen, daß die so gefundene Funktion $\varphi(x)$ mit $f(x) = \frac{\arc\sin x}{\sqrt{1-x^2}}$ übereinstimmt. Jedenfalls genügt $\varphi(x)$ den Bedingungen $\varphi(0) = 0$ und $(1 - x^2)\,\varphi'(x) - x\varphi(x) - 1 = 0$, wie man aus der Reihendarstellung von $\varphi(x)$ entnimmt. Für die Funktion $u(x) = \sqrt{1-x^2}\,\varphi(x)$ ist $u'(x) = \sqrt{1-x^2}\,\varphi' - \frac{x\varphi}{\sqrt{1-x^2}}$, und hieraus folgt leicht, daß $u$ den beiden Bedingungen $u(0) = 0$ und $u'(x) = \frac{1}{\sqrt{1-x^2}}$ genügt, aus deren zweiter durch Integration $u(x) = u(0) + \arc\sin x$ und somit wegen der ersten $u = \arc\sin x$ folgt. Damit ist die Gleichung $\varphi(x) = \frac{\arc\sin x}{\sqrt{1-x^2}}$ und somit die Reihenentwicklung (1) bewiesen. Da $\frac{\arc\sin x}{\sqrt{1-x^2}} = \frac{1}{2}\frac{d}{dx}(\arc\sin x)^2$ ist, so folgt die Reihenentwicklung (2) durch gliedweises Integrieren zwischen den Grenzen 0 und $x$.

Für die Funktion $f(x) = \cos(\mu \arc\sin x)$ findet man

$$f'(x) = -\frac{\mu}{\sqrt{1-x^2}}\sin(\mu \arc\sin x), \quad \text{also} \quad (1-x^2)\,f'(x)^2 = \mu^2(1 - f(x)^2)$$

und durch nochmalige Differentiation und Weglassung des Faktors $2f'(x)$

$$(1-x^2)\,f'' - x f' + \mu^2 f = 0.$$

Außerdem ist $f(0) = 1$ und $f'(0) = 0$. Wissen wir umgekehrt von einer Funktion $f(x)$, daß sie diese drei Bedingungen erfüllt, dann können wir rückwärts schließen, daß $f(x) = \cos(\mu \arc\sin x)$ sein muß. Denn zunächst ergibt sich durch Multiplikation mit $2f'(x)$ die Differentialgleichung $2(1-x^2)\,f'f'' - 2xf'^2 + 2\mu^2 ff' = 0$, die wir auch in der Form

$$\frac{d}{dx}\{(1-x^2)\,f'^2 - \mu^2(1-f^2)\} = 0$$

schreiben können; hieraus folgt, daß $(1-x^2)\,f'^2 - \mu^2(1-f^2) = \text{const.}$, und zwar, da für $x = 0$ die linke Seite wegen $f'(0) = 0$ und $f(0) = 1$ verschwindet, $(1-x^2)\,f'^2 - \mu^2(1-f^2) = 0$ sein muß. Führen wir statt $x$ eine neue Veränderliche $t$ durch die Gleichung $\cos t = f(x)$ oder $t = \arccos f(x)$ ein, so geht nach kurzer Rechnung unsere Differentialgleichung in $(1-x^2)\left(\frac{dt}{dx}\right)^2 - \mu^2 = 0$ über, woraus wir unter Berücksichtigung von $f(0) = 1$, d. h. $t(0) = 0$ das gewünschte Resultat $t(x) = \pm\mu \arc\sin x$, also $f(x) = \cos(\mu \arc\sin x)$ erhalten.

Wenn wir jetzt in der Gleichung $(1-x^2)\,f'' - xf' + \mu^2 f = 0$ den Ansatz $f(x) = c_0 + c_1 x + c_2 x^2 + \cdots$ machen, den Koeffizienten von $x^n$ links sammeln und gleich Null setzen, so finden wir, daß sie mit dem System der folgenden Relationen gleichbedeutend ist:

$$
\begin{aligned}
2c_2 + \mu^2 c_0 \qquad\quad &= 0\,,\\
3\cdot 2 c_3 + (\mu^2 - 1^2)\, c_1 &= 0\,,\\
4\cdot 3 c_4 + (\mu^2 - 2^2)\, c_2 &= 0\,,\\
\dots\dots\dots\dots\dots\dots\dots\dots&\dots\\
(n+2)(n+1)\, c_{n+2} + (\mu^2 - n^2)\, c_n &= 0\,,\\
\dots\dots\dots\dots\dots\dots\dots\dots&\dots
\end{aligned}
$$

Aus dieser Rekursionsformel und den Bedingungen $c_0 = 1$ und $c_1 = 0$ bestimmen wir nunmehr der Reihe nach die Koeffizienten: $c_{2m-1} = 0$, $c_0 = 1$, $c_2 = -\frac{\mu^2}{2}$, $c_4 = \frac{\mu^2(\mu^2-2^2)}{4!}$, .... Wir erhalten also die in (3) angegebene Reihe, deren Konvergenz für $|x| < 1$ aus dem Quotientenkriterium folgt. Nach dem Bewiesenen muß diese Reihe mit $\cos(\mu \arcsin x)$ identisch sein.

Genau ebenso folgt mit Hilfe derselben Differentialgleichung unter Berücksichtigung der anderen Bedingungen $f(0) = 0$ und $f'(0) = \mu$ die Reihenentwicklung (4).

## 2. Reihen, in denen die Bernoullischen Zahlen auftreten.

Wir haben bisher für einige unter den elementaren Funktionen keine Potenzreihenentwicklungen angegeben, z. B. nicht für die Funktion $\operatorname{tg} x$. Der Grund hierfür ist, daß die hier auftretenden Zahlenkoeffizienten sich nicht mehr ganz leicht übersehen lassen. Man kann nun diese Koeffizienten ebenso wie die Entwickelungskoeffizienten einer Anzahl anderer Funktionen durch die sogenannten *Bernoullischen Zahlen* ausdrücken, gewisse rationale, in ihrem Bildungsgesetz nicht einfache Zahlen, die an vielen Stellen der Analysis auftreten. Zu diesen Zahlen gelangt man am einfachsten, wenn man versucht, die Funktion

$$\frac{x}{e^x - 1} = \frac{1}{1 + \frac{x}{2!} + \frac{x^2}{3!} + \cdots}$$

in eine Potenzreihe der Form

$$\frac{x}{e^x - 1} = \sum_{\nu=0}^{\infty} \frac{B_\nu}{\nu!} x^\nu$$

zu entwickeln. Schreibt man diese Gleichung in der Form

$$x = (e^x - 1) \sum_{\nu=0}^{\infty} \frac{B_\nu}{\nu!} x^\nu$$

und trägt rechts für $e^x - 1$ die Potenzreihe ein, so erhält man in der auf S. 334 geschilderten Weise für die Zahlen $B_\nu$ eine Rekursionsformel, welche die Berechnung aller $B_\nu$ gestattet. Diese Zahlen nennt man die Bernoullischen Zahlen. Es sind, da bei ihrer Bildung nur rationale Rechenoperationen auftreten, rationale Zahlen, welche, wie man leicht erkennt, für ungeraden Index außer für $\nu = 1$ verschwinden; die ersten werden durch die Ausdrücke

$$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_4 = -\frac{1}{30}, \quad B_6 = \frac{1}{42},$$

$$B_8 = -\frac{1}{30}, \quad B_{10} = \frac{5}{66}, \quad \ldots$$

gegeben.

Ich muß mich mit einer kurzen Andeutung begnügen, wie diese Zahlen in die fraglichen Potenzreihenentwicklungen eingehen. Man erhält zunächst, indem man von der Umformung

$$1 + \frac{B_2}{2!} x^2 + \cdots = \frac{x}{e^x - 1} + \frac{x}{2} = \frac{x}{2} \cdot \frac{e^x + 1}{e^x - 1} = \frac{x}{2} \cdot \frac{e^{\frac{x}{2}} + e^{-\frac{x}{2}}}{e^{\frac{x}{2}} - e^{-\frac{x}{2}}}$$

Gebrauch macht,

$$\frac{x}{2} \mathfrak{Ctg} \frac{x}{2} = \sum_{\nu=0}^{\infty} \frac{B_{2\nu}}{(2\nu)!} x^{2\nu}.$$

Ersetzt man hierin $x$ durch $2x$, so entsteht die für $|x| < \pi$ gültige Reihenentwicklung

$$x \mathfrak{Ctg}\, x = \sum_{\nu=0}^{\infty} \frac{2^{2\nu} B_{2\nu}}{(2\nu)!} x^{2\nu},$$

aus welcher wir, indem wir $x$ durch $-ix$ ersetzen, die Entwicklung

$$x \operatorname{ctg} x = \sum_{\nu=0}^{\infty} (-1)^\nu \frac{2^{2\nu} B_{2\nu}}{(2\nu)!} x^{2\nu}, \qquad |x| < \pi,$$

gewinnen.

Mit Hilfe der Gleichung $2 \operatorname{ctg} 2x = \operatorname{ctg} x - \operatorname{tg} x$ gelangt man nun zu der Reihenentwicklung

$$\operatorname{tg} x = \sum_{\nu=1}^{\infty} (-1)^{\nu-1} \frac{2^{2\nu}(2^{2\nu} - 1) B_{2\nu}}{(2\nu)!} x^{2\nu-1},$$

die für $|x| < \frac{\pi}{2}$ gültig ist.

Im übrigen muß ich wegen der Durchführung der Beweise auf die spezielle Literatur verweisen[1]).

[1]) Vgl. z. B. *K. Knopp*: Theorie und Anwendung der unendlichen Reihen, 2. Aufl., Berlin 1924, S. 181ff.

Neuntes Kapitel.

# Fouriersche Reihen.

Neben den Potenzreihen spielt sowohl in rein mathematischen Disziplinen als auch bei den Anwendungen noch eine andere Klasse von unendlichen Reihen eine besonders wichtige Rolle, die Fourierschen Reihen, bei welchen die einzelnen Glieder trigonometrische Funktionen sind und die Summe der Reihe eine periodische Funktion wird.

## § 1. Die periodischen Funktionen.

### 1. Allgemeines.

Funktionen der Zeit, welche periodisch sind, d. h. sich in ihrem Verlaufe mit bestimmten Zeitabständen regelmäßig wiederholen, treffen wir in vielen Anwendungen. Bei den meisten Maschinen wird ein periodischer Vorgang im Rhythmus der Umdrehungen eines Schwungrades erzeugt, wie z. B. der Wechselstrom einer Dynamomaschine. Periodische Funktionen treten weiter bei allen Schwingungsvorgängen auf.

Eine periodische Funktion mit der Periode $2l$ ist durch die für alle Werte von $x$ bestehende Gleichung

$$f(x+2l) = f(x)$$

charakterisiert, wo $2l$ die *Periode* heißt[1]). Haben wir nur in einem bestimmten Intervall, etwa dem Intervalle $-l \leqq x \leqq l$ eine beliebige Funktion $f(x)$ gegeben, so läßt sie sich stets als periodische Funktion fortsetzen, indem man außerhalb des Intervalles die Funktionswerte durch die Gleichungen $f(x+2nl) = f(x)$ festlegt ($n$ bedeutet dabei eine beliebige ganze positive oder negative Zahl). Dabei ist allerdings zu beachten, daß durch die Fortsetzung in den Punkten $x = (2n \pm 1)l$ eine Unstetigkeit der Funktion erzeugt wird, falls nicht ganz von selbst schon bei der ursprünglich gegebenen Funktion Anfangs- und Endwert miteinander übereinstimmten, d. h. $f(-l) = f(l)$ war; und Entsprechendes gilt natürlich auch für die Ableitungen der Funktion. Wir wollen, um ein anschauliches Bild bei der Hand zu haben, die unabhängige Variable $x$ als die Zeit deuten — demgemäß auch gelegent-

---

[1]) Für die Darstellung periodischer Funktionen ist es manchmal bequem, die unabhängige Variable $x$ nicht durch einen Punkt auf einer Zahlengeraden, sondern unter Verwendung der Peripherie eines Kreises zu deuten. Hat eine Funktion $f(x)$ etwa die Periode $2\pi$, d. h. besteht für jedes $x$ die Gleichung

$$f(x+2\pi) = f(x),$$

und deuten wir $x$ als den von irgend einer Anfangslage aus gerechneten Zentriwinkel eines Kreises vom Radius 1, so drückt sich die Periodizität der Funktion $f(x)$ einfach darin aus, daß die Funktionswerte eindeutig den Punkten der Kreisperipherie zugeordnet sind (z. B. den Stellungen des Schwungrades einer Maschine).

lich $t$ statt $x$ schreiben — und durch die Funktion $f(x)$ dann einen periodischen Vorgang oder, wie wir auch sagen wollen, eine *Schwingung* repräsentieren. Die Periode $2l = T$ heißt dann die *Schwingungsdauer*.

Gleich hier merke ich eine durch die Gleichung

$$\int_{-l-a}^{l-a} f(x)\,dx = \int_{-l}^{l} f(x)\,dx$$

für eine periodische Funktion ausgedrückte allgemeine Tatsache an, welche in Worten besagt, daß das Integral einer periodischen Funktion über ein Intervall von der Periodenlänge $T = 2l$ stets denselben Wert hat, wo auch das Intervall hingeschoben werden mag. Der Beweis ergibt sich, wenn wir beachten, daß wegen $f(\xi - 2l) = f(\xi)$ bei beliebigem $\alpha$ und $\beta$ durch die Substitution $x = \xi - 2l$

$$\int_{\alpha}^{\beta} f(x)\,dx = \int_{\alpha+2l}^{\beta+2l} f(\xi)\,d\xi = \int_{\alpha+2l}^{\beta+2l} f(x)\,dx$$

folgt. Speziell wird für $\alpha = -l - a$ und $\beta = -l$

$$\int_{-l-a}^{-l} f(x)\,dx = \int_{l-a}^{l} f(x)\,dx$$

und daher

$$\int_{-l-a}^{l-a} f(x)\,dx = \int_{-l-a}^{-l} f(x)\,dx + \int_{-l}^{l-a} f(x)\,dx = \int_{l-a}^{l} f(x)\,dx + \int_{-l}^{l-a} f(x)\,dx = \int_{-l}^{l} f(x)\,dx,$$

was unsere Behauptung ausdrückt. Im übrigen veranschaulicht sich dies an Hand der beistehenden Figur 115 wegen der geometrischen Bedeutung des Integrales.

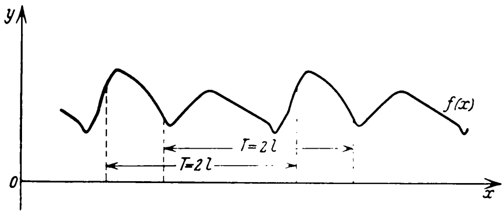



Fig. 115. Integral über eine volle Periode.

Die einfachsten periodischen Funktionen, in denen wir die Bausteine der allgemeinsten periodischen Funktionen erkennen werden, sind die Funktionen $a \sin \omega x$ und $a \cos \omega x$ oder allgemeiner $a \sin \omega (x - \xi)$ und $a \cos \omega (x - \xi)$, wobei $a \geqq 0$, $\omega > 0$, $\xi$ Konstanten bedeuten. Die durch sie dargestellten Vorgänge nennen wir *reine Schwingungen*. Die Schwingungsdauer ist $T = \frac{2\pi}{\omega}$; die Zahl $\omega$ heißt die *Frequenz der Schwingung*, sie ist, da $\frac{1}{T}$ die Anzahl der Schwingungen in der Zeiteinheit ist, gleich der *Schwingungszahl in der Zeit* $2\pi$.[1]) Die Zahl $a$ heißt

[1]) Vielfach, besonders in der Technik, nennt man die hier eingeführte Anzahl der Schwingungen in der Zeit $2\pi$ die *Kreis*frequenz, während man die Zahl $\frac{1}{T}$,

die *Amplitude* der Schwingung; sie gibt uns den maximalen Wert, welchen die Funktion $a \sin \omega (x - \xi)$ oder $a \cos \omega (x - \xi)$ annehmen kann. Die Zahl $\omega (x - \xi)$ nennen wir die *Phase* und die Zahl $\omega \xi$ die *Phasenverschiebung*.

Graphisch erhalten wir diese Funktionen, indem wir die Sinuskurven (vgl. Fig. 116) nach den Koordinatenrichtungen im Maßstab $1 : \omega$ bzw. $a : 1$ dilatieren und noch eine Verschiebung um $\omega \xi$ in der $x$-Richtung vornehmen.

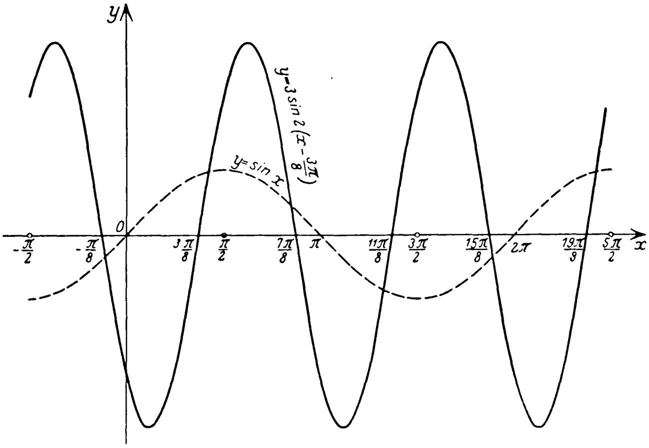



Fig. 116. Reine Schwingungen.

Nach den Additionsformeln der trigonometrischen Funktionen kann man unsere reinen Schwingungen auch in der Gestalt

$$\alpha \cos \omega x + \beta \sin \omega x \quad \text{bzw.} \quad \beta \cos \omega x - \alpha \sin \omega x$$

ausdrücken, wo $\alpha = -a \sin \omega \xi$, $\beta = a \cos \omega \xi$ ist. Umgekehrt repräsentiert jede Funktion der Form

$$\alpha \cos \omega x + \beta \sin \omega x$$

eine reine Schwingung $a \sin \omega (x - \xi)$ mit der Amplitude $a = \sqrt{\alpha^2 + \beta^2}$ und der durch die Gleichungen $\alpha = -a \sin \omega \xi$, $\beta = a \cos \omega \xi$ gegebenen Phasenverschiebung. Wir erkennen hieraus, daß die Summe solcher Funktionen mit derselben Frequenz $\omega$ stets wieder eine reine Schwingung mit der Frequenz $\omega$ darstellt.

## 2. Zusammensetzung von reinen Schwingungen. Obertöne. Schwebungen.

Wenn auch oft Bewegungen sich als reine Schwingungen erweisen (vgl. fünftes Kapitel, § 4, und zehntes Kapitel), so werden doch meistens

die Anzahl der Schwingungen in der Zeiteinheit, schlechthin als Frequenz bezeichnet. Im folgenden ist mit dem Wort Frequenz jedoch stets die Kreisfrequenz gemeint.

die periodischen Bewegungen einen komplizierteren Charakter tragen, nämlich sich als „*Superposition*“ von reinen Schwingungen darstellen. Mathematisch bedeutet dies einfach folgendes: Die Bewegung, z. B. die Entfernung eines Punktes aus seiner Ruhelage als Funktion der Zeit, wird durch eine Funktion gegeben, welche die Summe einer Anzahl von rein periodischen Funktionen der obigen Art ist. Die Sinuswellen der Funktion lagern sich dann geometrisch übereinander, oder, wie man sagt, sie superponieren sich. Bei einer solchen Übereinanderlagerung werden wir die Perioden bzw. Frequenzen der zusammenzusetzenden Schwingungen als voneinander verschieden annehmen; denn die Übereinanderlagerung zweier reiner Schwingungen mit derselben Frequenz $\omega$ liefert wiederum eine reine Schwingung mit derselben Frequenz, nur mit veränderter Amplitude und Phase, wie zum Schluß von Nr. 1 ausgeführt wurde.

Betrachten wir als einfachsten Fall die Superposition von zwei Schwingungen mit den Frequenzen $\omega_1$ und $\omega_2$, so finden wir, daß ein fundamentaler Unterschied besteht, je nachdem die beiden Frequenzen in einem rationalen Verhältnis zueinander stehen oder nicht oder, wie man sagt, je nachdem die Frequenzen kommensurabel oder inkommensurabel sind. Behandeln wir zunächst den ersten Fall und nehmen als Beispiel für die zweite Frequenz den Wert $\omega_2 = 2\omega_1$ an. Dann wird die zweite Schwingung genau die halbe Schwingungsdauer $\frac{2\pi}{2\omega_1} = T_2 = \frac{T_1}{2}$ der ersten besitzen, und sie wird ganz von selbst nicht nur die Periode $T_2$, sondern auch die doppelte Periode $T_1$ haben, da sich der Funktionsverlauf erst recht auch nach dieser doppelten Periode wiederholt. Man nennt eine solche Schwingung mit der doppelten Schwingungszahl und der halben Schwingungsdauer eine *erste harmonische Oberschwingung* zu der ursprünglichen.

Ganz Entsprechendes gilt, wenn wir nun noch eine weitere Schwingung mit der Frequenz $\omega_3 = 3\omega_1$ hinzufügen. Auch hier wird die Schwingungsfunktion $\sin 3\omega_1 x$ ganz von selbst sich mit der Periode $\frac{2\pi}{\omega_1} = T_1$ wiederholen. Eine solche Schwingung heißt *zweite harmonische Oberschwingung* zu der gegebenen; ebenso können wir eine dritte, vierte, ..., $(n-1)$-te Oberschwingung betrachten mit den Frequenzen $\omega_4 = 4\omega_1$, $\omega_5 = 5\omega_1$, ..., $\omega_n = n\omega_1$ und im übrigen beliebigen Phasenverschiebungen. Jede solche Oberschwingung wird sich auch ganz von selbst nach der Schwingungsdauer $T_1 = \frac{2\pi}{\omega_1}$ periodisch wiederholen, und daher wird jede Superposition einer Anzahl von Schwingungen, die sämtlich harmonische Oberschwingungen zu einer gegebenen Grundfrequenz $\omega_1$ sind, wiederum eine periodische Funktion mit der Periode $\frac{2\pi}{\omega_1} = T_1$ sein. Indem wir Schwingungen von der Grundschwingung bis zur $(n-1)$-ten Oberschwingung superponieren, erhalten wir eine periodische Funktion der Form

$$S(x) = \alpha + \sum_{\nu=1}^{n} (a_\nu \cos \nu \omega x + b_\nu \sin \nu \omega x)$$

— die hinzugefügte Konstante $\alpha$ ändert an der Periodizität nichts, sie ist eine periodische Funktion für jede Periode —. Durch die in dieser Funktion enthaltenen $2n+1$ Konstanten, über die wir frei

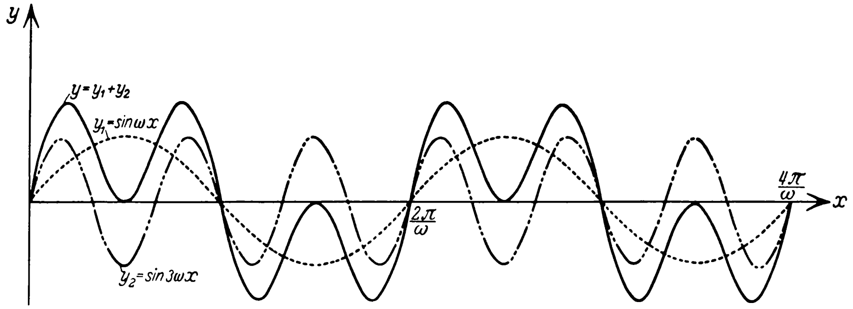
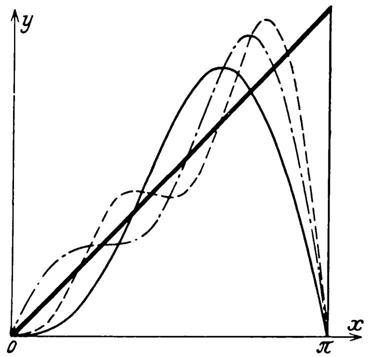



Fig. 117 [1]).




——— $\sin x - \frac{\sin 2x}{2}$

—·—·— $\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3}$

------ $\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \frac{\sin 4x}{4}$

Fig. 118.

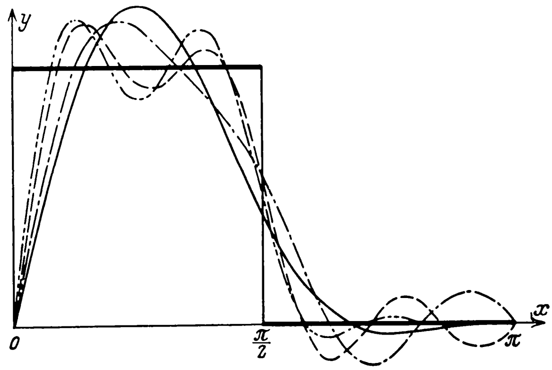



Fig. 119 [2]).

Zusammensetzung von Schwingungen.

verfügen können, gewinnen wir die Möglichkeit, recht komplizierte und an die ursprüngliche Sinusform kaum noch erinnernde Kurven-

[1]) Die Maßverhältnisse der Figur entsprechen der Annahme $\omega = 1$.

[2]) Die in der Figur eingezeichneten Kurven entsprechen den durch Abbrechen bei 3, 5, 6 bzw. 7 aus der Reihe

$$\frac{\sin x}{1} + 2\frac{\sin 2x}{2} + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + 2\frac{\sin 6x}{6} + \frac{\sin 7x}{7} + \frac{\sin 9x}{9} + \cdots$$

entstehenden trigonometrischen Polynomen.

bilder zu erzeugen. Die Figuren **117** bis **119** geben davon eine anschaulichere Vorstellung als eine Schilderung.

Die Bezeichnung „*harmonische Oberschwingungen*" stammt aus der Akustik, wo sich zeigt, daß, wenn einer Grundschwingung mit der Frequenz $\omega$ ein bestimmter Ton zugeordnet ist, dann der ersten, zweiten, dritten usw. Oberschwingung die Reihe der harmonischen Obertöne zum Grundton entspricht, d. h. die Oktave, Oktave + Quinte, Doppeloktave usw.

Ganz allgemein lassen sich bei der Zusammensetzung von Schwingungen, bei denen die Frequenzen in rationalen Verhältnissen stehen, die Perioden als ganzzahlige Vielfache einer gemeinsamen Grundperiode darstellen. Einen grundsätzlich anderen Typus von Erscheinungen stellt jedoch die Zusammensetzung zweier Schwingungen mit inkommensurablen Frequenzen $\omega_1$ und $\omega_2$ dar. Hier wird der durch Superposition reiner Schwingungen entstehende Vorgang nicht mehr rein periodisch werden. Ich kann auf die hiermit zusammenhängenden mathematischen Gesichtspunkte nicht eingehen. Nur möchte ich bemerken, daß solche Funktionen doch immer noch einen nahezu periodischen Charakter tragen, d. h. daß sie sich „beinahe" periodisch wiederholen oder, wie man sagt, *fastperiodisch* sind. Derartige Funktionen sind gerade in neuerer Zeit besonders eingehend untersucht worden.

Eine letzte Bemerkung über die Zusammensetzung von reinen Schwingungen betrifft die sog. *Schwebungserscheinungen.* Setzen wir zwei Schwingungen mit derselben Amplitude 1 und zwei verschiedenen Frequenzen $\omega_1$ und $\omega_2$ zusammen und nehmen wir der Einfachheit halber für beide (im Sinne von S. **347**) dasselbe $\xi$ an — die Verallgemeinerung auf beliebige Phasen kann dem Leser überlassen bleiben —, so handelt es sich einfach um das Studium der Funktion

$$y = \sin \omega_1 x + \sin \omega_2 x \qquad (\omega_1 > \omega_2 > 0).$$

Nach einer bekannten trigonometrischen Formel ergibt sich sofort

$$y = 2 \cos \frac{\omega_1 - \omega_2}{2} x \cdot \sin \frac{\omega_1 + \omega_2}{2} x,$$

und diese Gleichung stellt uns einen Vorgang dar, den wir folgendermaßen auffassen können: Wir haben eine Schwingung mit der Frequenz $\frac{\omega_1 + \omega_2}{2}$ und der Periode $\frac{4\pi}{\omega_1 + \omega_2}$. Diese Schwingung aber hat keine konstante Amplitude; die „Amplitude" wird vielmehr durch den Ausdruck $2 \cos \frac{\omega_1 - \omega_2}{2} x$ gegeben, sie ändert sich mit einer längeren Periode, nämlich der Periode $\frac{4\pi}{\omega_1 - \omega_2}$. Diese Auffassung wird besonders dann nützlich und anschaulich sein, wenn die Frequenzen $\omega_1$ und $\omega_2$ beide verhältnismäßig groß sind, ihre Differenz $\omega_1 - \omega_2$ aber eine im Vergleich dazu sehr kleine Zahl wird. Es wird dann die Schwingung mit der Schwingungsdauer $\frac{4\pi}{\omega_1 + \omega_2}$ ihre Amplitude $2\cos\frac{\omega_1 - \omega_2}{2}x$ nur langsam im Verhältnis zu dieser Schwingungsdauer ändern, und diese Änderungen werden sich wieder selbst periodisch vollziehen, eben mit der langen Periode $\frac{4\pi}{\omega_1 - \omega_2}$. Diese rhythmischen Amplitudenschwingungen nennt man *Schwebungen.* Jedermann kennt insbesondere aus der Akustik und vielleicht auch aus der drahtlosen Telegraphie diese Erscheinung. In der drahtlosen Telegraphie pflegt man die Frequenzen $\omega_1$ und $\omega_2$ weit oberhalb der Frequenzen der

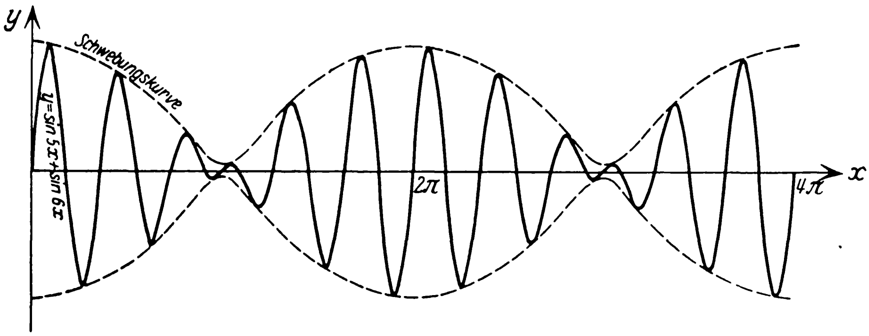



Fig. 120. Schwebung.

akustisch wahrnehmbaren Töne zu legen, während die Differenz $\omega_1 - \omega_2$ in den Bereich der akustisch hörbaren Töne fällt. Die Schwebung wird dann in einen akustisch hörbaren Ton umgesetzt werden können, während die eigentlichen Schwingungen für das Ohr unwahrnehmbar bleiben.

In Figur 120 ist ein Beispiel für eine Schwebung graphisch dargestellt.

## § 2. Die Verwendung der komplexen Schreibweise.

### 1. Allgemeine Bemerkungen.

Die Behandlung von Schwingungsvorgängen und periodischen Funktionen gewinnt an formaler Einfachheit, wenn wir uns der komplexen Zahlen bedienen, indem wir je zwei trigonometrische Funktionen $\cos\omega x$ und $\sin\omega x$ durch Ausdrücke der Gestalt $\cos\omega x + i\sin\omega x = e^{i\omega x}$ zusammenfassen (vgl. hierzu achtes Kapitel, § 7). Dabei müssen wir uns dessen bewußt bleiben, daß eine Gleichung zwischen komplexen Größen mit zwei Gleichungen zwischen reellen Größen gleichbedeutend ist und daß wir unsere Resultate stets im Reellen zu deuten und zu verstehen haben.

Indem wir gemäß den Formeln

$$2\cos\vartheta = e^{i\vartheta} + e^{-i\vartheta}, \qquad 2i\sin\vartheta = e^{i\vartheta} - e^{-i\vartheta}$$

überall die trigonometrischen Funktionen durch Exponentialfunktionen ersetzen, gelangen wir zu der Darstellung reiner Schwingungen durch die komplexen Größen $e^{i\omega x}$, $e^{-i\omega x}$ bzw.

$$a\,e^{i\omega(x-\xi)}, \qquad a\,e^{-i\omega(x-\xi)},$$

wobei wiederum $a, \omega$ und $\omega\xi$ reelle Größen: Amplitude, Frequenz und Phasenverschiebung sind. Die reellen Schwingungen ergeben sich aus dieser komplexen Darstellung einfach als der reelle und der imaginäre Teil.

Die Annehmlichkeit einer solchen Darstellung für viele Anwendungen beruht darauf, daß man die Differentialquotienten der reellen Schwingungen nach der Zeit $x$ formal erhält, indem man die komplexe Exponentialfunktion genau so differenziert, als ob $i$ eine reelle Konstante wäre, was durch die Formel

$$\begin{aligned}\frac{d}{dx}a\{\cos\omega(x-\xi)+i\sin\omega(x-\xi)\} &= a\,\omega\{-\sin\omega(x-\xi)+i\cos\omega(x-\xi)\}\\ &= i\,a\,\omega\{\cos\omega(x-\xi)+i\sin\omega(x-\xi)\}\end{aligned}$$

oder

$$\frac{d}{dx}\,a\,e^{i\omega(x-\xi)} = i\,a\,\omega\,e^{i\omega(x-\xi)}$$

seinen Ausdruck findet.

## 2. Anwendung in der Lehre vom Wechselstrom.

Ich erläutere diese Dinge an Hand eines wichtigen Beispieles. Wir wollen dabei die unabhängige Veränderliche, die Zeit, nicht mehr mit $x$, sondern mit $t$ bezeichnen.

Stellen wir uns einen Stromkreis mit dem Ohmschen Widerstande $R$ und der Selbstinduktion $L$ vor, dem wir von außen her eine elektromotorische Kraft $E$ aufprägen. Im Falle des idealen Gleichstroms ist $E$ konstant, und es gilt für die Stromstärke $J$ das Ohmsche Gesetz:

$$E = RJ.$$

Handelt es sich jedoch um Wechselstrom, so ist $E$ und demgemäß auch $J$ eine Funktion der Zeit $t$, und das Ohmsche Gesetz nimmt die Gestalt

$$E - L\frac{dJ}{dt} = RJ$$

an.

Im einfachsten Falle, den wir als vorliegend voraussetzen, wird nun die äußere elektromotorische Kraft $E$ eine reine harmonische Schwingung von der Frequenz $\omega$ sein. Statt nun aber diese Schwingung in der Form $a\cos\omega t$ oder $a\sin\omega t$ anzunehmen, fassen wir beide Möglichkeiten in komplexer Gestalt formal folgendermaßen zusammen:

$$E = \varepsilon e^{i\omega t} = \varepsilon \cos \omega t + i \varepsilon \sin \omega t,$$

wobei $\varepsilon\,(>0)$ die Amplitude darstellt. Wir rechnen nun mit dieser „komplexen Spannung", als ob $i$ ein reeller Parameter wäre, und werden aus ihr eine komplexe Stromstärke $J$ gewinnen. Die Bedeutung der so zwischen den komplexen Größen $E$ und $J$ entstehenden Beziehungen ist dann, daß einer elektromotorischen Kraft $\varepsilon \cos \omega t$ der reelle Teil von $J$, einer elektromotorischen Kraft $\varepsilon \sin \omega t$ der imaginäre Teil von $J$ als Stromstärke entspricht. Die komplexe Stromstärke $J$ können wir sofort berechnen, wenn wir $J$ in der Form

$$J = \alpha e^{i\omega t} = \alpha (\cos \omega t + i \sin \omega t)$$

ansetzen; d. h. wenn wir zunächst einmal hypothetisch voraussetzen, daß $J$ wieder durch eine harmonische Schwingung mit der Frequenz $\omega$ gegeben wird. Es wird dann der Differentialquotient von $J$ formal durch

$$\frac{dJ}{dt} = i \alpha \omega e^{i\omega t} = \alpha \omega (-\sin \omega t + i \cos \omega t)$$

dargestellt, und wir erhalten aus dem verallgemeinerten Ohmschen Gesetz durch Einführung unserer Größen und Weglassung des überall auftretenden Faktors $e^{i\omega t}$ sofort die Gleichung $\varepsilon - \alpha L i \omega = R\alpha$ oder $\alpha = \frac{\varepsilon}{R + i\omega L}$, also

$$E = (R + i\omega L) J = WJ.$$

Diese letzte Gleichung können wir als das Ohmsche Gesetz für den Wechselstrom in komplexer Form ansehen, wenn wir die Größe

$$W = R + i\omega L$$

als den *komplexen Widerstand* des Stromkreises bezeichnen. Das Ohmsche Gesetz lautet dann genau so wie im Falle des Gleichstromes: Stromstärke gleich Spannung durch Widerstand.

Stellen wir den komplexen Widerstand in der Form

$$W = w e^{i\delta} = w \cos \delta + i w \sin \delta$$

dar, wo

$$w = \sqrt{R^2 + L^2 \omega^2}, \qquad \operatorname{tg} \delta = \frac{\omega L}{R}$$

ist, so erhalten wir

$$J = \frac{\varepsilon}{w} e^{i(\omega t - \delta)}.$$

Nach dieser Formel hat der Strom dieselbe Schwingungsdauer (und Frequenz) wie die Spannung; die Amplitude $a$ des Stromes hängt mit der Amplitude $\varepsilon$ der Spannung durch die Beziehung

$$a = \frac{\varepsilon}{w}$$

zusammen, und ferner tritt eine *Phasenverschiebung* des Stromes gegen die Spannung ein. Der Strom erreicht sein Maximum nicht zu derselben Zeit wie die Spannung, sondern erst um $\frac{\delta}{\omega}$ später, und dasselbe gilt natürlich auch für das Minimum. Die Größe $w = \sqrt{R^2 + L^2 \omega^2}$ nennt man in der Elektrotechnik häufig den *Wechselstromwiderstand* oder die *Impedanz* des Stromkreises für die Frequenz $\omega$.

### 3. Komplexe Darstellung der Superposition von reinen Schwingungen.

Bisher bedeutete die komplexe Schreibweise die Zusammenfassung zweier reiner Schwingungen. Wir können aber auch eine einzelne reine Schwingung oder auch eine Superposition von reinen Schwingungen (der Einfachheit halber werde $\omega = 1$ angenommen) in der Gestalt

$$S(x) = \alpha + \sum_{\nu=1}^{n} (a_\nu \cos \nu x + b_\nu \sin \nu x)$$

sofort in komplexe Schreibweise überführen, indem wir

$$\cos \nu x = \frac{1}{2}\left(e^{i\nu x} + e^{-i\nu x}\right) \quad \text{und} \quad \sin \nu x = \frac{1}{2i}\left(e^{i\nu x} - e^{-i\nu x}\right)$$

einsetzen. Unser Ausdruck geht dabei über in einen Ausdruck der Form

$$S(x) = \sum_{\nu=-n}^{n} \alpha_\nu e^{i\nu x},$$

wobei die komplexen Zahlen $\alpha_\nu$ mit den reellen Zahlen $\alpha$, $a_\nu$ und $b_\nu$ durch die Gleichungen

$$a_\nu = \alpha_\nu + \alpha_{-\nu}, \qquad \alpha = \alpha_0, \qquad b_\nu = i(\alpha_\nu - \alpha_{-\nu})$$

verknüpft sind. Damit die Gleichung $a_\nu = \alpha_\nu + \alpha_{-\nu}$ auch den Fall $\nu = 0$ formal umfaßt, setzen wir gelegentlich $\alpha = \alpha_0 = \frac{a_0}{2}$.

Umgekehrt können wir jeden beliebigen Ausdruck der Form

$$\sum_{\nu=-n}^{n} \alpha_\nu e^{i\nu x}$$

als Superposition komplex geschriebener Schwingungen ansehen; damit das Resultat dieser Superposition reell ist, muß nur $\alpha_\nu + \alpha_{-\nu}$ reell und $\alpha_\nu - \alpha_{-\nu}$ rein imaginär sein, d. h. $\alpha_\nu$ und $\alpha_{-\nu}$ müssen konjugiert komplex sein.

### 4. Ableitung einer trigonometrischen Formel.

Durch die komplexe Schreibweise erhalten wir einen sehr einfachen Beweis für die folgende später zu verwendende, für $\alpha \neq 0, \pm 2\pi,$

$\pm 4\pi, \ldots$ gültige *trigonometrische Summationsformel*:

$$\sigma_n(\alpha) = \frac{1}{2} + \cos\alpha + \cos 2\alpha + \cdots + \cos n\alpha = \frac{\sin\left(n + \frac{1}{2}\right)\alpha}{2\sin\frac{\alpha}{2}}.$$

Zum Beweise ersetzen wir die Kosinusfunktionen durch ihre Exponentialausdrücke und bringen dadurch die Summe $\sigma_n(\alpha)$ in die Gestalt

$$\sigma_n(\alpha) = \frac{1}{2}\sum_{\nu=-n}^{n} e^{i\nu\alpha}.$$

Hier steht rechts eine geometrische Reihe mit dem Quotienten $q = e^{i\alpha} \neq 1$, die wir sofort summieren können. Wir erhalten

$$\sigma_n(\alpha) = \frac{1}{2} e^{-in\alpha} \frac{1 - q^{2n+1}}{1 - q} = \frac{1}{2} \frac{e^{-in\alpha} - e^{(n+1)i\alpha}}{1 - e^{i\alpha}}$$

und, indem wir Zähler und Nenner mit $e^{-\frac{i\alpha}{2}}$ multiplizieren,

$$\sigma_n(\alpha) = \frac{\sin\left(n + \frac{1}{2}\right)\alpha}{2\sin\frac{\alpha}{2}},$$

wie behauptet wurde.

## § 3. Trigonometrische Interpolation.

### 1. Lösung des Interpolationsproblems.

Wir verlassen jetzt die einleitenden Betrachtungen und wenden uns einer Aufgabe zu, welche den natürlichen Ausgangspunkt sowohl für die Theorie als auch für praktische Anwendungen der trigonometrischen Reihen bildet, einem *Interpolationsproblem*. Wenn wir eine periodische Funktion $f(x)$ etwa mit der Periode $2l$ gegeben haben, so liegt es nach den Vorbemerkungen aus § 1 nahe, zu fragen, ob und wie wir eine solche Funktion durch Superposition von reinen Schwingungen darstellen können, und zwar genauer von einer Schwingung mit der Grundfrequenz $\frac{\pi}{l}$ und dazu harmonischen Oberschwingungen mit den Frequenzen $\frac{2\pi}{l}, \frac{3\pi}{l}, \ldots$.

Eine solche Darstellung wird im allgemeinen nur eine Annäherung sein, die wir allerdings hernach durch einen Grenzübergang zu einer exakten Darstellung auszugestalten trachten werden. Es entsteht die Frage, in welcher Weise wir eine solche Annäherung versuchen können.

Um zunächst alles Überflüssige in der Schreibweise zu vermeiden, nehmen wir als das Intervall, für das die Funktion $f(x)$ gegeben sein soll, das Intervall $-\pi < x < \pi$ an, setzen also die Zahl $l$ gleich $\pi$, eine Einschränkung, die wir zum Schluß leicht aufheben können, indem wir $x$ durch $\frac{\pi}{l}x$ ersetzen. Wir betrachten nun rein periodische Schwingungen von der Form

$$S_n(x) = \alpha + \sum_{\nu=1}^{n}(a_\nu \cos \nu x + b_\nu \sin \nu x)$$

und wollen eine Annäherung an unsere Funktion $f(x)$ durch eine solche Funktion vornehmen, wobei wir die Zahl $n$ als gegeben betrachten.

In den Funktionen $S_n(x)$ sind im ganzen $2n+1$ Konstanten verfügbar, nämlich die Zahlen $\alpha, a_1, b_1, \ldots, a_n, b_n$. Wir werden also erwarten, daß wir bei der Annäherung der Funktion $2n+1$ Bedingungen erfüllen können. Die nächstliegende Art solcher Bedingungen wird durch die Interpolationsaufgabe formuliert: Man soll die verfügbaren Konstanten so bestimmen, daß die Funktion $S_n(x)$ an $2n+1$ gegebenen Stellen des Intervalles mit der gegebenen Funktion $f(x)$ übereinstimmt. Dies ist das allgemeine Problem der trigonometrischen Interpolation. Es stellt eine genaue Analogie zu dem schon im sechsten Kapitel, Anhang, § 4, behandelten Problem der Interpolation durch ganze rationale Funktionen dar. Das Interpolationsproblem führt zu dem einfachsten Ergebnis, wenn die $2n+1$ Stellen, an welchen wir Übereinstimmung von $f(x)$ und $S_n(x)$ fordern, in gleichen Abständen und zwar im Abstande $\lambda = \frac{2\pi}{2n+1}$ voneinander liegen und das Intervall symmetrisch bedecken. Wir werden demgemäß die Strecke $\lambda = \frac{2\pi}{2n+1}$, welche gerade den $(2n+1)$-ten Teil des Gesamtintervalles darstellt, vom Nullpunkt aus nach rechts und links $n$-mal hintereinander abtragen, so daß wir $2n+1$ Teilpunkte

$$x = -n\lambda,\ -(n-1)\lambda,\ \ldots,\ 0,\ \ldots,\ (n-1)\lambda,\ n\lambda$$

erhalten, die wir der Reihe nach mit

$$x_{-n} = -n\lambda, \ldots, x_\varkappa = \varkappa\lambda, \ldots, x_n = n\lambda$$

bezeichnen. Die beiden Endpunkte $-\pi$ und $\pi$ des Gesamtintervalles sind nicht selbst Teilpunkte, liegen vielmehr gerade um die Länge $\frac{1}{2}\lambda$ von dem ersten bzw. letzten Teilpunkt entfernt. Die vorgegebenen Funktionswerte in diesen Teilpunkten wollen wir der Reihe nach mit

$$f_{-n}, \ldots, f_\varkappa, \ldots, f_n$$

bezeichnen, sodaß also $f_\varkappa = f(x_\varkappa)$ wird. (Im übrigen wird von unserer vorgegebenen Funktion $f(x)$ nichts als eben diese $2n+1$ Werte $f_\varkappa$ bei unseren Betrachtungen in Erscheinung treten.)

In formaler Hinsicht erleichtern wir uns die Durchführung der Interpolationsaufgabe, wenn wir uns der komplexen Schreibweise bedienen.

Wir setzen also unsere Interpolationsfunktion gemäß § 2, Nr. 3 in der Form

$$S_n(x) = \sum_{\nu=-n}^{n} \alpha_\nu e^{i\nu x}$$

an, wobei die Größen $\alpha_\nu$ komplexe Zahlen sind. Unsere Aufgabe ist, die $2n+1$ Größen $\alpha_{-n}, \alpha_{-n+1}, \ldots, \alpha_0, \ldots, \alpha_n$ so zu bestimmen, daß die $2n+1$ Gleichungen

$$\begin{aligned}
\sum_{\nu=-n}^{n} \alpha_\nu e^{-\nu i\lambda n} &= f_{-n},\\
\sum_{\nu=-n}^{n} \alpha_\nu e^{-\nu i\lambda(n-1)} &= f_{-n+1},\\
&\ldots\ldots\\
\sum_{\nu=-n}^{n} \alpha_\nu e^{\nu i\lambda\varkappa} &= f_\varkappa,\\
&\ldots\ldots\\
\sum_{\nu=-n}^{n} \alpha_\nu e^{\nu i\lambda n} &= f_n
\end{aligned}$$

oder kurz geschrieben

$$S_n(x_\varkappa) = f_\varkappa, \qquad \varkappa = -n, \ldots, 0, \ldots, n,$$

erfüllt sind. (Diese Gleichungen besagen einfach, daß die Funktion $S_n(x)$ an den Stellen $x_\varkappa$ die vorgegebenen Werte $f_\varkappa$ besitzen soll.)

Sie stellen ein System von $2n+1$ linearen Gleichungen für die $2n+1$ Unbekannten $\alpha_{-n}, \ldots, \alpha_0, \ldots, \alpha_n$ dar und können sehr leicht explizite aufgelöst werden. Wir wählen irgend eine ganze Zahl $\mu$ unter den Zahlen $-n, -n+1, \ldots, n$, multiplizieren die erste unserer Gleichungen mit $e^{i\mu\lambda n}$, die zweite mit $e^{i\mu\lambda(n-1)}, \ldots$, die letzte mit $e^{-i\mu\lambda n}$ und addieren sodann die sämtlichen Gleichungen. Hierdurch erhalten wir auf der rechten Seite den Ausdruck $\sum_{\varkappa=-n}^{n} f_\varkappa e^{-i\mu\lambda\varkappa}$; auf der linken Seite ordnen wir das entstehende Aggregat

$$\sum_{\varkappa=-n}^{n} \sum_{\nu=-n}^{n} \alpha_\nu e^{i\lambda\varkappa(\nu-\mu)},$$

indem wir jeweils die Faktoren von $\alpha_\nu$ sammeln, d. h. die Doppelsumme in der Form

$$\sum_{\nu=-n}^{n} \alpha_\nu \sum_{\varkappa=-n}^{n} e^{i\lambda\varkappa(\nu-\mu)}$$

schreiben. Der Faktor von $\alpha_\nu$ ist gerade die Summe

$$e^{-i\lambda(\nu-\mu)n} + e^{-i\lambda(\nu-\mu)(n-1)} + \cdots + e^{i\lambda(\nu-\mu)n},$$

und diese Summe können wir wie auf S. 356 berechnen. Wir setzen zur Abkürzung $\nu - \mu = \delta$ und $e^{i\lambda\delta} = \varrho$ und beachten, daß

$$\varrho^{2n+1} = e^{2i\pi\delta} = 1$$

ist oder, wie man sich auch ausdrückt, daß die Zahl $\varrho$ eine $(2n+1)$-te Einheitswurzel ist.

Ist $\nu = \mu$, so werden alle Ausdrücke $e^{-i\lambda(\nu-\mu)n}, \ldots, e^{i\lambda(\nu-\mu)n}$ gleich 1, und wir erhalten als Faktor von $\alpha_\nu$ die Zahl $2n+1$.

Ist aber $\nu \neq \mu$, so stellt unsere Summe eine geometrische Reihe mit einem von 1 verschiedenen Quotienten $\varrho$ dar, die wir mit unserer Abkürzung folgendermaßen schreiben und summieren können:

$$\varrho^{-n} + \varrho^{-n+1} + \cdots + \varrho^{n} = \varrho^{-n}\,\frac{\varrho^{2n+1}-1}{\varrho-1}.$$

Nun brauchen wir nur zu beachten, daß $\varrho^{2n+1} = 1$ ist, und haben damit als Wert unserer Summe den Wert Null gefunden.

Somit fallen bei unserer Zusammenfassung der Gleichungen auf der linken Seite alle Glieder fort bis auf das Glied mit $\alpha_\mu$, dessen Faktor $2n+1$ ist, und wir erhalten als Auflösung die einfache Gleichung

$$\alpha_\mu = \frac{1}{2n+1}\sum_{\varkappa=-n}^{n} f_\varkappa\, e^{-i\mu\lambda\varkappa}.$$

Mit dieser Gleichung ist das Interpolationsproblem gelöst; denn dies besteht ja nach dem Obigen in nichts anderem als der Bestimmung dieser Größen $\alpha_\mu$.

Mit Hilfe des doppelten Summenzeichens können wir die Lösung des Interpolationsproblemes in der folgenden eleganten Form schreiben:

$$S_n(x) = \frac{1}{2n+1}\sum_{\nu=-n}^{n}\sum_{\varkappa=-n}^{n} f_\varkappa\, e^{i\nu(x-\varkappa\lambda)}.$$

Was an unseren obigen Formeln vor allem in die Augen fällt, ist ihre Symmetrie, die wir noch einmal zusammenfassend zum Ausdruck bringen: *Damit die Gleichungen*

$$f_\varkappa = \sum_{\nu=-n}^{n} \alpha_\nu\, e^{i\nu\varkappa\lambda}, \qquad \lambda = \frac{2\pi}{2n+1}$$

*für $\varkappa = -n, \ldots, 0, \ldots, n$ bestehen, müssen die Größen $\alpha_\nu$ durch die Gleichungen*

$$\alpha_\nu = \frac{1}{2n+1}\sum_{\varkappa=-n}^{n} f_\varkappa\, e^{-i\varkappa\nu\lambda}$$

*gegeben werden und umgekehrt.* Die Symmetrie dieser Gleichungen können wir noch vollkommener machen, wenn wir $\alpha_\nu = c_\nu \frac{1}{\sqrt{2n+1}}$

setzen. Wir erhalten dann das Resultat in der folgenden, völlig symmetrischen Gestalt: *Damit die Gleichungen*

$$f_\varkappa = \frac{1}{\sqrt{2n+1}} \sum_{\nu=-n}^{n} c_\nu e^{i\nu\varkappa\lambda}$$

*bestehen, ist es notwendig und hinreichend, daß das andere Gleichungssystem*

$$c_\nu = \frac{1}{\sqrt{2n+1}} \sum_{\varkappa=-n}^{n} f_\varkappa e^{-i\varkappa\nu\lambda}$$

*erfüllt ist.*

Gewöhnlich ist es nötig, unser Ergebnis in einer reellen Form zu schreiben. Indem wir für die Exponentialfunktion die trigonometrischen Funktionen einführen, geht unser Interpolationsausdruck $S_n(x) = \sum_{\nu=-n}^{n} \alpha_\nu e^{i\nu x}$ in die Gestalt

$$S_n(x) = \frac{a_0}{2} + \sum_{\nu=1}^{n} (a_\nu \cos \nu x + b_\nu \sin \nu x)$$

über, die Koeffizienten $a_\nu$ und $b_\nu$ ergeben sich durch die Formeln

$$a_\nu = \alpha_\nu + \alpha_{-\nu}, \qquad b_\nu = i(\alpha_\nu - \alpha_{-\nu}),$$

und wir erhalten also die Lösung unseres Interpolationsproblems in reeller Gestalt durch folgenden Satz: *Die Koeffizienten des trigonometrischen Ausdruckes n-ten Grades*

$$\frac{a_0}{2} + \sum_{\nu=1}^{n} (a_\nu \cos \nu x + b_\nu \sin \nu x),$$

*welcher an den vorgeschriebenen Stellen* $x_\varkappa = \varkappa\lambda$ $(\varkappa = -n, \ldots, n)$ *die vorgeschriebenen Werte* $f(x_\varkappa) = f_\varkappa$ *annimmt, werden durch die Summen*

$$a_\nu = \frac{2}{2n+1} \sum_{\varkappa=-n}^{n} f_\varkappa \cos \nu\varkappa\lambda, \qquad b_\nu = \frac{2}{2n+1} \sum_{\varkappa=-n}^{n} f_\varkappa \sin \nu\varkappa\lambda$$

*gegeben.*

Natürlich kann man das eben rein im Reellen formulierte Resultat auch gewinnen, ohne die komplexe Schreibweise zu benutzen. Man hat an Stelle der einfachen Rechnungen mit der Exponentialfunktion und der geometrischen Reihe sich dann entsprechender trigonometrischer Formeln zu bedienen. Die Durchführung dieser Betrachtung mag dem Leser überlassen bleiben.

## 2. Grenzübergang zur Fourierschen Reihe.

Ähnlich wie wir bei der Interpolation rationaler Funktionen durch einen Grenzübergang zu der Darstellung von Funktionen durch die

Taylorsche Reihe kommen, werden wir hier von einer Annäherung der Funktion durch das trigonometrische Interpolationspolynom $n$-ten Grades zu einer exakten Darstellung der Funktion durch eine unendliche trigonometrische Reihe gelangen, indem wir die Anzahl der Interpolationsstellen über alle Grenzen wachsen lassen. Ich begnüge mich damit, diesen Grenzübergang hier rein formal auszuführen, indem ich die strenge Begründung für die Reihenentwicklung auf den § 5 verweise.

Bei dem Grenzübergang knüpfen wir zunächst an die reellen eben abgeleiteten Formeln an. Lassen wir $n$ über alle Grenzen wachsen, so werden die Zahlen $a_\nu$ und $b_\nu$ in der Grenze übergehen in die Größen

$$a_\nu = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos\nu x\,dx\,, \qquad b_\nu = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin\nu x\,dx\,,$$

da

$$\frac{2}{2n+1} = \frac{\lambda}{\pi} = \frac{\Delta x}{\pi}$$

gegen Null strebt, wobei nur etwa vorausgesetzt zu werden braucht, daß die Funktion $f(x)$ in dem Intervalle $-\pi < x < \pi$ stetig ist. Wir werden also rein formal von der Interpolation zu der Darstellung einer willkürlich im Intervalle $-\pi < x < \pi$ vorgegebenen stetigen Funktion $f(x)$ durch die Reihe

$$f(x) = \frac{a_0}{2} + \sum_{\nu=1}^{\infty}(a_\nu\cos\nu x + b_\nu\sin\nu x)$$

mit den Koeffizienten

$$a_\nu = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos\nu t\,dt, \qquad b_\nu = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin\nu t\,dt$$

gelangen. Diese einer Funktion $f(x)$ zugeordnete Reihe heißt ihre *Fouriersche Reihe,* nach dem großen französischen Mathematiker Fourier, der im Zusammenhang mit Fragen der mathematischen Physik solche Reihen als Erster systematisch zur Darstellung willkürlicher Funktionen benutzt hat. Nach der vorangegangenen Lösung des Interpolationsproblemes ist es zwar plausibel, daß die so gefundene unendliche Reihe konvergiert und die Funktion überall darstellt. In der Tat jedoch müssen wir unserer Funktion gewisse Einschränkungen auferlegen, damit dies der Fall ist.

Wir werden finden, daß sich in eine Fouriersche Reihe sicherlich jede Funktion entwickeln läßt, für welche der Funktionswert und die erste Ableitung bis auf endlich viele Sprungstellen stetige Funktionen in dem Grundintervall sind. Dabei braucht die Funktion in den verschiedenen Intervallen keineswegs demselben analytischen oder geometrischen Bildungsgesetz zu gehorchen; sie darf „willkürlich“ sein.

In dieser Tatsache, ebenso wie in der Tatsache, daß die Koeffizienten der Fourierschen Reihe von dem gesamten Verlauf der Funktion abhängen, zeigt sich ein tiefgehender Unterschied gegen die Potenzreihen. Man kann nämlich in einer Potenzreihe die Koeffizienten durch die Ableitungen der Funktion in einer bestimmten Stelle ausdrücken. Die Potenzreihe ist uns also gegeben, wenn wir die Funktion in einer noch so kleinen Umgebung einer Stelle kennen. Die durch Potenzreihen darstellbaren Funktionen, die „analytischen" Funktionen, tragen daher einen ganz speziellen Charakter: Ihr Gesamtverlauf ist bekannt, sobald der Verlauf in einem wenn auch noch so kleinen Intervalle gegeben ist. Analytische Funktionen stellen also gewissermaßen ein organisches einheitliches Gebilde dar. Bei einer Fourierschen Reihe dagegen ist nach dem eben Gesagten hiervon nicht die Rede. Der Funktionsverlauf in der Umgebung einer Stelle kann ganz unabhängig von dem Funktionsverlauf in der Umgebung einer anderen willkürlich abgeändert werden, ohne daß die Entwickelbarkeit in eine Fouriersche Reihe gestört wird.

Nehmen wir den Grenzübergang von der trigonometrischen Interpolation zur Fourierschen Reihe in der komplexen Schreibweise vor, so gelangen wir formal zur Fourierschen Reihe in der Gestalt

$$f(x) = \sum_{\nu=-\infty}^{+\infty} \alpha_\nu e^{i\nu x}$$

mit den Koeffizienten

$$\alpha_\nu = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\, e^{-i\nu t}\, dt.$$

Dabei bedeutet die Summation von $-\infty$ bis $+\infty$ nichts anderes als einen Übergang, bei dem zunächst von $-n$ bis $+n$ summiert und dann ein Grenzübergang $n \to \infty$ gemacht wird.

Bevor ich nun dazu übergehe, unseren Grenzübergang von der Interpolation zur Fourierschen Reihe streng zu begründen, möchte ich — indem ich den erwähnten Satz von der Entwickelbarkeit in eine Fouriersche Reihe hypothetisch voraussetze — an einer Reihe von Beispielen die außerordentliche Fruchtbarkeit und Eleganz der Fourierschen Reihenentwicklung dartun.

## § 4. Beispiele für die Fouriersche Reihe.

### 1. Vorbemerkungen.

Wir nehmen als Periode unserer Funktionen $f(x)$ die Größe $2\pi$ und als das betrachtete Intervall dasjenige zwischen $-\pi$ und $+\pi$. Die Funktionen $f(x)$ sind dann zunächst nur für das Intervall

$-\pi < x < \pi$ definiert und müssen gemäß § 1 periodisch über dieses Intervall hinaus nach rechts und links fortgesetzt werden.

Ich schicke unseren Rechnungen eine einfache Bemerkung voraus: Ist $f(x)$ eine **gerade Funktion** (vgl. S. 12f.), so wird offenbar $f(x) \sin \nu x$ ungerade, $f(x) \cos \nu x$ gerade und daher

$$b_\nu = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \sin \nu x \, dx = 0; \qquad a_\nu = \frac{2}{\pi}\int_0^{\pi} f(x) \cos \nu x \, dx.$$

Wir erhalten also eine „Kosinusreihe". Ist dagegen die Funktion $f(x)$ eine **ungerade Funktion**, so wird

$$a_\nu = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \cos \nu x \, dx = 0; \qquad b_\nu = \frac{2}{\pi}\int_0^{\pi} f(x) \sin \nu x \, dx.$$

Wir erhalten also eine „Sinusreihe"[1].

## 2. Entwicklung der Funktionen $\psi(x) = x$ und $\varphi(x) = x^2$.

Für die ungerade Funktion $x$ wird $b_\nu = \frac{2}{\pi}\int_0^{\pi} x \sin \nu x \, dx$ und durch Anwendung der Produktintegration

$$\frac{\pi}{2} b_\nu = -\frac{x \cos \nu x}{\nu}\Bigg|_0^{\pi} + \frac{1}{\nu}\int_0^{\pi} \cos \nu x \, dx = (-1)^{\nu+1}\frac{\pi}{\nu}.$$

Wir erhalten also für die periodische im Intervalle $-\pi < x < \pi$ durch die Größe $x$ selbst dargestellte Funktion (vgl. Fig. 121), die wir mit $\psi(x)$ bezeichnen, die Reihenentwicklung

$$\psi(x) = 2\left(\frac{\sin x}{1} - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - + \cdots\right).$$

Setzt man $x = \frac{\pi}{2}$, so ergibt sich die uns schon von früher bekannte Leibnizsche Reihe

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - + \cdots.$$

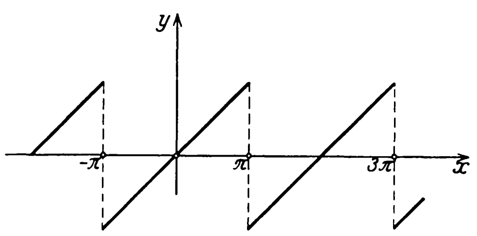


Fig. 121.

Die durch unsere obige Reihe dargestellte Funktion $\psi(x)$ ist als Ganzes keineswegs stetig; sie springt vielmehr an den Stellen $x = k\pi$, $k = \pm 1, \pm 3, \pm 5, \ldots$ um den Wert $2\pi$. In den Sprungstellen selbst, d. h. für $x = k\pi$,

[1] Ist also die Funktion $f(x)$ von vornherein nur im Intervalle $0 < x < \pi$ gegeben, so können wir sie in das Intervall $-\pi < x < 0$ entweder als gerade oder als ungerade Funktion fortsetzen und demgemäß im Intervalle $0 < x < \pi$ entweder in eine bloße Kosinusreihe oder in eine bloße Sinusreihe entwickeln.

$k = \pm 1, \pm 3, \ldots$, wird jedes Reihenglied Null und daher der Funktionswert selbst Null. In den Sprungstellen wird also das arithmetische Mittel zwischen dem Grenzwert von links und dem Grenzwert von rechts durch die Reihe dargestellt.

Ist $\xi$ irgend eine feste Zahl zwischen $-\pi$ und $\pi$ und ersetzt man in der obigen Reihe $x$ durch $x - \xi$, so erhält man die Reihe

$$\begin{aligned}\psi(x-\xi) &= 2\left(\frac{\sin(x-\xi)}{1} - \frac{\sin 2(x-\xi)}{2} + \frac{\sin 3(x-\xi)}{3} - + \cdots\right)\\ &= -\frac{2}{1}\sin\xi\cos x + \frac{2}{1}\cos\xi\sin x + \frac{2}{2}\sin 2\xi\cos 2x\\ &\quad -\frac{2}{2}\cos 2\xi\sin 2x - \frac{2}{3}\sin 3\xi\cos 3x + \frac{2}{3}\cos 3\xi\sin 3x + \cdots,\end{aligned}$$

die man ebenfalls in Form einer Fourierschen Reihe mit den gegen Null strebenden Koeffizienten

$$a_0 = 0, \qquad a_n = 2\frac{(-1)^n}{n}\sin n\xi, \qquad b_n = 2\frac{(-1)^{n-1}}{n}\cos n\xi$$

schreiben kann und welche eine Funktion darstellt, die lediglich an den Stellen $x = \xi \pm \pi$, $x = \xi \pm 3\pi, \ldots$ die eben beschriebene Unstetigkeit aufweist.

Für die gerade Funktion $\varphi(x) = x^2$ ergibt sich durch zweimalige partielle Integration

$$a_\nu = \frac{2}{\pi}\int_0^\pi x^2\cos\nu x\,dx = (-1)^\nu\frac{4}{\nu^2} \qquad (\nu > 0),$$

$$a_0 = \frac{2\pi^2}{3},$$

so daß die Entwicklung

$$\varphi(x) = \frac{\pi^2}{3} - 4\left(\frac{\cos x}{1^2} - \frac{\cos 2x}{2^2} + \frac{\cos 3x}{3^2} - + \cdots\right)$$

entsteht, aus der man formal durch gliedweise Differentiation und Division durch die Zahl 2 die Reihe für $\psi(x) = x$ zurückerhält.

### 3. Entwicklung der Funktion $x \cos$

Für diese ebenfalls ungerade Funktion ergibt sich

$$a_\nu = 0, \qquad b_\nu = \frac{2}{\pi}\int_0^\pi x\cos x\sin\nu x\,dx$$

und, unter Benutzung der in Nr. 2 gewonnenen Formel

$$\int_0^\pi x\sin\mu x\,dx = (-1)^{\mu+1}\frac{\pi}{\mu} \qquad (\mu = 1, 2, \ldots),$$

$$b_\nu = \frac{2}{\pi}\int_0^\pi x\cos x \sin\nu x\,dx = \frac{1}{\pi}\int_0^\pi x\,(\sin(\nu+1)\,x + \sin(\nu-1)\,x)\,dx$$
$$= \frac{(-1)^{\nu+2}}{\nu+1} + \frac{(-1)^\nu}{\nu-1} = (-1)^\nu \frac{2\nu}{\nu^2-1} \qquad (\nu = 2, 3, \ldots),$$
$$b_1 = -\frac{1}{2}.$$

Wir erhalten also die Reihenentwicklung

$$x\cos x = -\frac{1}{2}\sin x + 2\sum_{\nu=2}^{\infty} \frac{(-1)^\nu \nu}{\nu^2-1}\sin\nu x$$

und durch Addition der in Nr. 2 gefundenen Reihe die Reihenentwicklung

$$x\,(1+\cos x) = \frac{3}{2}\sin x + 2\left(\frac{\sin 2x}{1\cdot 2\cdot 3} - \frac{\sin 3x}{2\cdot 3\cdot 4} + \frac{\sin 4x}{3\cdot 4\cdot 5} - + \cdots\right).$$

Die Funktion $x\cos x$ im Intervall $-\pi < x < \pi$ (vgl. Fig. 122) besitzt, wenn sie periodisch über die Intervallendpunkte hinaus fortgesetzt

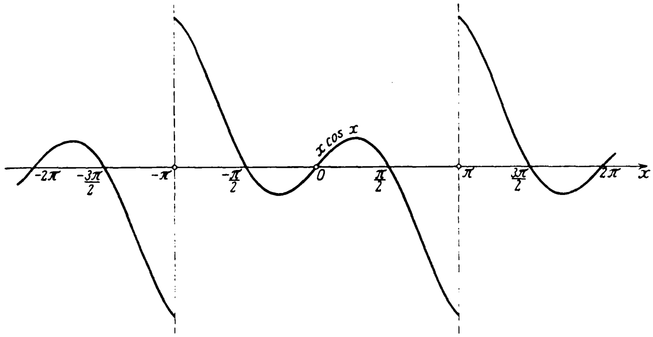



Fig. 122.

wird, beim Überschreiten der Endpunkte des Intervalles dieselbe Unstetigkeit wie die in Nr. 2 betrachtete Funktion $\psi(x)$; dagegen wird die Funktion $x(1+\cos x)$, wenn man sie periodisch fortsetzt, auch beim Überschreiten der Endpunkte sogar mit ihrer Ableitung stetig bleiben, da die Unstetigkeit durch den in den Endpunkten mit seiner Ableitung verschwindenden Faktor $1+\cos x$ beseitigt wird.

### 4. $f(x) = |x|$.

Diese Funktion ist gerade; es wird also $b_\nu = 0$, $a_\nu = \frac{2}{\pi}\int_0^\pi x\cos\nu x\,dx$,

und man erhält leicht mit Hilfe der Produktintegration

$$\int_0^\pi x\cos\nu x\,dx = \frac{1}{\nu}\,x\sin\nu x\Big|_0^\pi - \frac{1}{\nu}\int_0^\pi \sin\nu x\,dx = \begin{cases} 0, \text{ wenn } \nu \text{ gerade und } \neq 0, \\ -\frac{2}{\nu^2}, \text{ wenn } \nu \text{ ungerade,} \end{cases}$$

also

$$f(x) = \frac{\pi}{2} - \frac{4}{\pi}\left(\cos x + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \cdots\right).$$

Setzt man hier $x = 0$ ein, so gewinnt man die merkwürdige Formel

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots.$$

**5. Beispiel.**

Die Fouriersche Reihe für die durch die Gleichungen

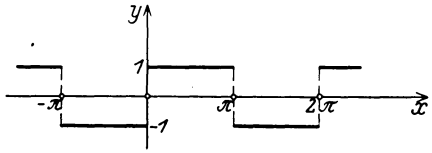


Fig. 123.

$$f(x) = \begin{cases} -1, & \text{wenn } -\pi < x < 0, \\ 0, & \text{wenn } \quad x = 0, \\ +1, & \text{wenn } \quad 0 < x < \pi, \end{cases}$$

bzw. durch Fig. 123 definierte Funktion lautet, da $a_\nu = 0$,

$$b_\nu = \frac{2}{\pi}\int_0^\pi \sin \nu x \, dx = \begin{cases} 0, & \text{wenn } \nu \text{ gerade}, \\ \dfrac{4}{\pi \nu}, & \text{wenn } \nu \text{ ungerade}, \end{cases}$$

ist,

$$f(x) = \frac{4}{\pi}\left(\frac{\sin x}{1} + \frac{\sin 3x}{3} + \cdots\right).$$

Speziell ergibt sich hieraus für $x = \frac{\pi}{2}$ wieder die Leibnizsche Reihe.

**6. $f(x) = |\sin x|$.**

Die gerade Funktion $f(x) = |\sin x|$ können wir in eine Kosinusreihe entwickeln, wobei sich der Wert von $a_\nu$ durch die folgende Rechnung ergibt:

$$\begin{aligned} \frac{\pi}{2} a_\nu &= \int_0^\pi \sin x \cos \nu x \, dx = \frac{1}{2}\int_0^\pi \{\sin(\nu+1)x - \sin(\nu-1)x\}\, dx \\ &= \begin{cases} 0, & \text{wenn } \nu \text{ ungerade}, \\ \dfrac{-2}{\nu^2 - 1}, & \text{wenn } \nu \text{ gerade}. \end{cases} \end{aligned}$$

Wir erhalten also

$$f(x) = |\sin x| = \frac{2}{\pi} - \frac{4}{\pi}\sum_{\mu=1}^{\infty} \frac{\cos 2\mu x}{4\mu^2 - 1}.$$

**7. Entwicklung der Funktion $\cos \mu x$. Partialbruchzerlegung des Kotangens. Produktzerlegung des Sinus.**

Für $-\pi < x < \pi$ sei $f(x) = \cos \mu x$, wobei $\mu$ nicht ganzzahlig ist. Man erhält, da $f(x)$ gerade ist, wieder $b_\nu = 0$ und

$$\frac{\pi}{2} a_\nu = \int_0^\pi \cos \mu x \cos \nu x \, dx = \frac{1}{2} \int_0^\pi \{\cos (\mu + \nu) x + \cos (\mu - \nu) x\} \, dx$$
$$= \frac{1}{2} \left( \frac{\sin (\mu + \nu) \pi}{\mu + \nu} + \frac{\sin (\mu - \nu) \pi}{\mu - \nu} \right)$$
$$= \frac{\mu (-1)^\nu}{\mu^2 - \nu^2} \sin \mu \pi .$$

Also ergibt sich

$$\cos \mu x = \frac{2 \mu \sin \mu \pi}{\pi} \left( \frac{1}{2 \mu^2} - \frac{\cos x}{\mu^2 - 1^2} + \frac{\cos 2x}{\mu^2 - 2^2} - + \cdots \right).$$

Diese Funktion bleibt beim Überschreiten der Stellen $x = \pm \pi$ stetig. Setzt man $x = \pi$ ein, dividiert die Gleichung durch $\sin \mu \pi$ und schreibt dann $x$ statt $\mu$, so erhält man die Gleichung

$$\operatorname{ctg} \pi x = \frac{2x}{\pi} \left( \frac{1}{2x^2} + \frac{1}{x^2 - 1^2} + \frac{1}{x^2 - 2^2} + \cdots \right),$$

welche die sog. *Partialbruchzerlegung des Kotangens* darstellt, eine in der Analysis häufig diskutierte sehr wichtige Formel. Wir schreiben diese Reihe nun in die Form

$$\operatorname{ctg} \pi x - \frac{1}{\pi x} = - \frac{2x}{\pi} \left\{ \frac{1}{1^2 - x^2} + \frac{1}{2^2 - x^2} + \cdots \right\}.$$

Das $n$-te Glied der Reihe auf der rechten Seite ist, wenn $x$ in einem Intervall $0 \leqq x \leqq q < 1$ liegt, absolut genommen kleiner als $\frac{2}{\pi} \frac{1}{n^2 - q^2}$. Die Reihe konvergiert also gleichmäßig in diesem Intervall, und wir dürfen sie gliedweise integrieren. Wir erhalten so

$$\pi \int_0^x \left( \operatorname{ctg} \pi t - \frac{1}{\pi t} \right) dt = \log \frac{\sin \pi x}{\pi x} - \lim_{a \to 0} \log \frac{\sin \pi a}{\pi a} = \log \frac{\sin \pi x}{\pi x}$$
$$= \log \left( 1 - \frac{x^2}{1^2} \right) + \log \left( 1 - \frac{x^2}{2^2} \right) + \cdots = \lim_{n \to \infty} \sum_{\nu = 1}^{n} \log \left( 1 - \frac{x^2}{\nu^2} \right).$$

Gehen wir vom Logarithmus zur Exponentialfunktion über, so ergibt sich

$$\frac{\sin \pi x}{\pi x} = e^{\lim\limits_{n \to \infty} \sum\limits_{\nu = 1}^{n} \log \left( 1 - \frac{x^2}{\nu^2} \right)} = \lim_{n \to \infty} e^{\sum\limits_{\nu = 1}^{n} \log \left( 1 - \frac{x^2}{\nu^2} \right)} = \lim_{n \to \infty} \prod_{\nu = 1}^{n} \left( 1 - \frac{x^2}{\nu^2} \right).$$

Es ist also

$$\sin \pi x = \pi x \left( 1 - \frac{x^2}{1^2} \right) \left( 1 - \frac{x^2}{2^2} \right) \left( 1 - \frac{x^2}{3^2} \right) \cdots .$$

Wir haben so die berühmte Zerlegung des Sinus in ein unendliches

Produkt gewonnen[1]). Aus ihr erhalten wir für $x = \frac{1}{2}$ das Wallissche Produkt

$$\frac{\pi}{2} = \prod_{\nu=1}^{\infty} \frac{2\nu}{2\nu - 1} \cdot \frac{2\nu}{2\nu + 1} = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdots,$$

in Übereinstimmung mit S. 181.

### 8. Weitere Beispiele.

Durch kleine Rechnungen nach den hier durchgeführten Mustern erhält man die folgenden weiteren Beispiele für Reihenentwicklungen.

Die für $-\pi < x < \pi$ durch $f(x) = \sin \mu x$ definierte Funktion $f(x)$ läßt sich in die Reihe entwickeln

$$f(x) = \sin \mu x = -\frac{2 \sin \mu \pi}{\pi} \cdot \left( \frac{\sin x}{\mu^2 - 1} - \frac{2 \sin 2x}{\mu^2 - 2^2} + \frac{3 \sin 3x}{\mu^2 - 3^2} - + \cdots \right),$$

aus der für $x = \frac{\pi}{2}$ $\left(\text{unter Benutzung von } \sin \mu \pi = 2 \sin \mu \frac{\pi}{2} \cos \mu \frac{\pi}{2}\right)$ die Partialbruchzerlegung des Sekans, d. h. von $\dfrac{1}{\cos \mu \frac{\pi}{2}}$ hergeleitet werden kann, die, wenn man für $\frac{\mu}{2}$ wieder $x$ schreibt, lautet:

$$\pi \sec \pi x = \frac{\pi}{\cos \pi x} = 4 \sum_{\nu=1}^{\infty} \frac{(-1)^{\nu} (2\nu - 1)}{4x^2 - (2\nu - 1)^2}.$$

Die Reihen für die Hyperbelfunktionen $\mathfrak{Cof}\, \mu x$ und $\mathfrak{Sin}\, \mu x$ $(-\pi < x < \pi)$ lauten

$$\mathfrak{Cof}\, \mu x = \frac{2\mu}{\pi} \mathfrak{Sin}\, \mu \pi \left( \frac{1}{2\mu^2} - \frac{\cos x}{\mu^2 + 1^2} + \frac{\cos 2x}{\mu^2 + 2^2} - \frac{\cos 3x}{\mu^2 + 3^2} + - \cdots \right),$$

$$\mathfrak{Sin}\, \mu x = \frac{2}{\pi} \mathfrak{Sin}\, \mu \pi \left( \frac{\sin x}{\mu^2 + 1^2} - \frac{2 \sin 2x}{\mu^2 + 2^2} + \frac{3 \sin 3x}{\mu^2 + 3^2} - + \cdots \right).$$

## § 5. Strenge Begründung der Fourierschen Reihenentwicklung.

Man kann zu einer strengen Begründung der Fourierschen Reihe bei einer genaueren Untersuchung des im § 3 nur formal ausgeführten Grenzüberganges gelangen. Aber dieser Weg ist recht mühsam, und wir wollen daher hier anders vorgehen.

---

[1]) Diese Formel ist vor allem dadurch interessant, daß sie unmittelbar das Verschwinden der Funktion $\sin \pi x$ an den Stellen $x = 0, \pm 1, \pm 2, \ldots$ erkennen läßt; sie entspricht in dieser Hinsicht der elementaren Produktzerlegung einer ganzen rationalen Funktion mittels ihrer Nullstellen.

## 1. Die Konvergenz der Fourierschen Reihe einer stückweise glatten Funktion.

Wir bilden zunächst zu einer gegebenen Funktion $f(x)$ gemäß den Formeln

$$a_\nu = \frac{1}{\pi}\int_{-\pi}^{+\pi} f(t)\cos \nu t\, dt, \quad b_\nu = \frac{1}{\pi}\int_{-\pi}^{+\pi} f(t)\sin \nu t\, dt$$

die sog. *Fourierschen Koeffizienten*, was immer möglich ist, sobald die Funktion $f(x)$ stetig ist oder höchstens an einer endlichen Anzahl von Stellen eine Unterbrechung der Stetigkeit durch endliche Sprünge („Sprungstellen") erfährt. Wir können dann zweitens die formal gebildete unendliche Reihe

$$\frac{a_0}{2} + \sum_{\nu=1}^{\infty}(a_\nu \cos \nu x + b_\nu \sin \nu x)$$

auf ihre Konvergenz hin untersuchen.

Um das zu beweisende Resultat zu formulieren, stellen wir die folgenden Definitionen auf. Eine Funktion $f(x)$ heißt in einem Intervall *stückweise glatt*, wenn sie selbst *stückweise stetig*, d. h. im Intervalle bis auf endlich viele Sprungstellen stetig ist und wenn ferner auch die erste Ableitung $f'(x)$ stückweise stetig ist.

Von vornherein denken wir uns die Funktion $f(x)$ periodisch über das Intervall $-\pi \leqq x \leqq \pi$ hinaus fortgesetzt.

An einer Sprungstelle wollen wir der Funktion $f(x)$ den Wert zuschreiben, der gleich dem arithmetischen Mittel aus den Grenzwerten von rechts und links ist; wir setzen also in unmittelbar verständlicher Schreibweise

$$f(x) = \frac{1}{2}\,(f(x-0) + f(x+0)),$$

eine selbstverständlich auch für Stetigkeitsstellen bestehende Beziehung.

Unser Ziel ist der folgende Satz: *Die zu einer stückweise glatten Funktion $f(x)$ gehörige Fouriersche Reihe konvergiert an jeder Stelle $x$ und stellt die Funktion dar*[1].

Zum Beweise dieses Satzes betrachten wir die Partialsummen

$$S_n(x) = \frac{a_0}{2} + \sum_{\nu=1}^{n}(a_\nu \cos \nu x + b_\nu \sin \nu x)$$

und formen sie um, indem wir für die Koeffizienten die obigen Integralausdrücke einsetzen; vertauschen wir danach Summation und Inte-

[1] Beiläufig sei bemerkt, daß man diesen Satz auch noch für allgemeinere Funktionenklassen nachweisen kann. Jedoch reicht das hier formulierte Resultat für alle Anwendungen aus.

gration, so gewinnen wir den Ausdruck

$$S_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left\{\frac{1}{2} + \sum_{\nu=1}^{n}(\cos\nu t\cos\nu x + \sin\nu t\sin\nu x)\right\}dt$$

oder auf Grund des Additionstheorems des Kosinus

$$S_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left\{\frac{1}{2} + \sum_{\nu=1}^{n}\cos\nu(t-x)\right\}dt.$$

Wenden wir nun die in § 2, Nr. 4 abgeleitete Summationsformel an, so ergibt sich

$$S_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,\frac{\sin\left(n+\frac{1}{2}\right)(t-x)}{\sin\frac{t-x}{2}}\,dt;$$

und endlich, wenn wir die Transformation $\tau = t - x$ ausführen und die Periodizität des Integranden beachten:

$$S_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x+\tau)\,\frac{\sin\left(n+\frac{1}{2}\right)\tau}{\sin\frac{\tau}{2}}\,d\tau.$$

Wir werden die Konvergenz von $S_n(x)$ gegen $f(x)$ an Hand dieser Gestalt der Partialsummen $S_n(x)$ auf Grund zweier Hilfssätze leicht erkennen.

*Hilfssatz 1. Ist $s(x)$ eine im Intervall $a \leqq x \leqq b$ stückweise stetige Funktion, so strebt das Integral*

$$J = \int_a^b s(t)\sin\lambda t\,dt$$

*mit wachsendem $\lambda$ gegen Null.*

Beim Beweise dürfen wir $s(x)$ im ganzen Intervall als stetig voraussetzen, da wir andernfalls die Betrachtung für jedes Teilintervall, in dem $s(x)$ stetig ist, gesondert durchführen können.

Wir beachten nun ganz analog wie in der ähnlichen Betrachtung auf S. **338** f., daß bei positivem $\lambda$ die Funktion $\sin\lambda t$ jeweils im Abstande $h = \frac{\pi}{\lambda}$ ihr Vorzeichen wechselt, so daß sich bei genügend großem $\lambda$ die von zwei benachbarten Intervallen herrührenden Bestandteile des Integrales nahezu zerstören werden, weil die Werte von $s(x)$ in zwei solchen Intervallen wegen der Stetigkeit nur wenig voneinander abweichen. Wir nutzen diesen Umstand aus, indem wir den Ausdruck $J$ durch die Substitution $t = \tau + h$ mit $h = \frac{\pi}{\lambda}$ transformieren, wobei $\sin\lambda t = -\sin\lambda\tau$ wird und die Relation

$$J = -\int_{a-h}^{b-h} s(\tau+h)\sin\lambda\tau\,d\tau$$

entsteht. Schreiben wir hier für die Integrationsvariable wieder $t$ statt $\tau$ und addieren die beiden verschiedenen Ausdrücke für $J$, so ergibt sich

$$2J = -\int_{a-h}^{a} s(t+h)\sin\lambda t\,dt + \int_{a}^{b-h}(s(t) - s(t+h))\sin\lambda t\,dt + \int_{b-h}^{b} s(t)\sin\lambda t\,dt.$$

Verstehen wir unter $M$ eine obere Schranke für den absoluten Betrag von $s(x)$, d. h. gilt

$$|s(x)| \leqq M,$$

so folgt aus dieser Darstellung von $J$ sofort die Ungleichung

$$2|J| \leqq 2Mh + \int_{a}^{b-h} |s(t) - s(t+h)|\,dt.$$

Es sei nun irgend eine positive Größe $\varepsilon$ gegeben; wählen wir dann $\lambda$ so groß, daß im ganzen Intervall $a \leqq t \leqq b-h$ der Ausdruck $|s(t) - s(t+h)|$ kleiner als $\frac{\varepsilon}{b-a}$ bleibt und ferner $Mh = \frac{M\pi}{\lambda} < \frac{\varepsilon}{2}$ gilt, so ergibt sich

$$|J| < \varepsilon$$

und daher — weil $\varepsilon$ beliebig klein gewählt werden kann — $\lim\limits_{\lambda\to\infty} J = 0$[1].

*Hilfssatz 2. Bei beliebigem $a > 0$ konvergiert das Integral $\int_0^a \frac{\sin\lambda t}{t}\,dt$ mit wachsendem $\lambda$ gegen den Wert $\frac{\pi}{2}$.*

Daß der Limes existiert und von $a$ nicht abhängt, erkennen wir, wenn wir vermöge der Substitution $\tau = \lambda t$ schreiben

$$\int_0^a \frac{\sin\lambda t}{t}\,dt = \int_0^{a\lambda} \frac{\sin\tau}{\tau}\,d\tau;$$

denn das Integral rechts besitzt nach S. **203** f. und **338** f. einen von $a$ unabhängigen Limes, den wir mit $\int_0^\infty \frac{\sin\tau}{\tau}\,d\tau$ bezeichnet haben. Wir beweisen hier über das frühere Resultat hinaus, daß dieses uneigentliche Integral den Wert $\frac{\pi}{2}$ besitzt.

---

[1]) Unter der Voraussetzung, daß $s(x)$ auch noch eine stückweise stetige Ableitung $s'(x)$ besitzt, ergibt sich der Beweis von Hilfssatz 1 einfach durch Produktintegration. Es ist nämlich

$$\int_a^b s(t)\sin\lambda t\,dt = \frac{1}{\lambda}\left\{s(a)\cos\lambda a - s(b)\cos\lambda b + \int_a^b s'(t)\cos\lambda t\,dt\right\}.$$

Hier erkennt man aber sofort, daß die rechte Seite mit wachsendem $\lambda$ gegen Null konvergiert.

Zu dem Zwecke schließen wir aus Hilfssatz 1 auf das Bestehen der Relation

$$\lim_{\lambda \to \infty} \int_0^a \sin \lambda t \left( \frac{1}{t} - \frac{1}{2 \sin \frac{t}{2}} \right) dt = 0,$$

wo $a$ eine beliebige positive Zahl unterhalb $2\pi$ sein kann, so daß der Faktor $\frac{1}{t} - \frac{1}{2 \sin \frac{t}{2}}$ im Integrationsgebiet stetig bleibt[1]).

Speziell dürfen wir $a = \pi$ setzen und erkennen auf diese Weise, daß die beiden Ausdrücke

$$\int_0^\pi \frac{\sin \lambda t}{t} dt \quad \text{und} \quad \int_0^\pi \frac{\sin \lambda t}{2 \sin \frac{t}{2}} dt$$

bei wachsendem $\lambda$ demselben Grenzwert zustreben. Wir können diesen bestimmen, indem wir für $\lambda$ die Zahlen $n + \frac{1}{2}$ einsetzen, wobei $n$ die Folge der ganzen Zahlen durchläuft.

Nunmehr können wir auf Grund unserer Summationsformel für jeden Wert $\lambda = n + \frac{1}{2}$ das Integral

$$\int_0^\pi \frac{\sin \left( n + \frac{1}{2} \right) t}{2 \sin \frac{t}{2}} dt = \int_0^\pi \left( \frac{1}{2} + \sum_1^n \cos \nu t \right) dt$$

ausführen. Wir erhalten

$$\int_0^\pi \frac{\sin \left( n + \frac{1}{2} \right) t}{2 \sin \frac{t}{2}} dt = \frac{\pi}{2}$$

---

[1]) An der Stelle $t = 0$ ist dabei für diesen Faktor der Grenzwert von rechts, d. h., wie man leicht aus der Ungleichung

$$0 < \frac{\frac{t}{2} - \sin \frac{t}{2}}{\frac{t}{2} \sin \frac{t}{2}} < \frac{\operatorname{tg} \frac{t}{2} - \sin \frac{t}{2}}{\frac{t}{2} \sin \frac{t}{2}} = \frac{1 - \cos \frac{t}{2}}{\frac{t}{2} \cos \frac{t}{2}} \quad \text{für} \quad 0 < t < \pi$$

erkennt, der Wert Null zu nehmen. Denn nach Seite 37 strebt der Ausdruck

$$\frac{1 - \cos \frac{t}{2}}{\frac{t}{2} \cos \frac{t}{2}}$$

mit $t$ gegen Null. (Vgl. auch S. 268.)

und daher allgemein

$$\lim_{\lambda\to\infty}\int_0^a \frac{\sin\lambda t}{t}\,dt = \lim_{\lambda\to\infty}\int_0^\pi \frac{\sin\lambda t}{2\sin\frac{t}{2}}\,dt = \frac{\pi}{2}.$$

*Beweis des Hauptsatzes.* Mittels dieser beiden Hilfssätze folgt unser Hauptsatz, d. h. die Formel

$$\lim_{n\to\infty} S_n(x) = \lim_{n\to\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x+t)\,\frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\frac{t}{2}}\,dt = f(x)$$

leicht.

Wir zerlegen zunächst das Integrationsgebiet durch den Nullpunkt. Nun ist bei festem $x$ die Funktion

$$s(t) = \frac{f(x+t)-f(x+0)}{2\sin\frac{t}{2}}\ ^{1)}$$

im Intervall $0 \leqq t \leqq \pi$ stückweise stetig. Denn diese Aussage gilt jedenfalls für das Gebiet $0 < t \leqq \pi$, während die Stetigkeit für $t = 0$ aus der vorausgesetzten Existenz des vorderen Differentialquotienten

$$\lim_{t\to 0,\,t>0}\frac{f(x+t)-f(x+0)}{t} = \lim_{t\to+0}\frac{f(x+t)-f(x+0)}{2\sin\frac{t}{2}}\cdot\frac{2\sin\frac{t}{2}}{t}$$
$$= \lim_{t\to+0}\frac{f(x+t)-f(x+0)}{2\sin\frac{t}{2}}$$

folgt. Folglich strebt das Integral

$$\frac{1}{\pi}\int_0^\pi s(t)\sin\lambda t\,dt = \frac{1}{2\pi}\int_0^\pi f(x+t)\,\frac{\sin\lambda t}{\sin\frac{t}{2}}\,dt - \frac{1}{2\pi}\int_0^\pi f(x+0)\,\frac{\sin\lambda t}{\sin\frac{t}{2}}\,dt$$

mit wachsendem $\lambda = n + \frac{1}{2}$ gegen Null.

Da aber aus dem zweiten Bestandteil $f(x+0)$ als Faktor heraustritt und $\int_0^\pi \frac{\sin\lambda t}{2\sin\frac{t}{2}}\,dt$ nach dem Beweis von Hilfssatz 2 gegen $\frac{\pi}{2}$ konvergiert, so ergibt sich sofort die Gleichung

$$\lim_{\lambda\to\infty}\frac{1}{2\pi}\int_0^\pi f(x+t)\,\frac{\sin\lambda t}{\sin\frac{t}{2}}\,dt = \frac{1}{2}f(x+0).$$

---

1) Zu dieser Schreibweise vgl. S. 369.

Ebenso bekommen wir für das Intervall $-\pi \leqq t \leqq 0$

$$\lim_{\lambda\to\infty} \frac{1}{2\pi} \int_{-\pi}^{0} f(x+t) \frac{\sin \lambda t}{\sin \frac{t}{2}} \, dt = \frac{1}{2} f(x-0)$$

und durch Addition

$$\lim_{\lambda\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x+t) \frac{\sin \lambda t}{\sin \frac{t}{2}} \, dt = f(x) \, .$$

## 2. Genauere Untersuchung der Konvergenz.

In der Umgebung derjenigen Stellen, wo die Funktion $f(x)$ unstetig wird, konvergiert die Fouriersche Reihe ungleichmäßig; denn nach Kapitel VIII, § 4 besitzt eine gleichmäßig konvergente Reihe stetiger Funktionen eine stetige Summe. Es gilt aber der folgende wichtige Satz: *Besitzt eine stückweise glatte periodische Funktion keinerlei Unstetigkeitsstellen, so konvergiert ihre Fouriersche Reihe absolut und gleichmäßig. Die Konvergenz der Reihe für eine beliebige stückweise glatte Funktion ist gleichmäßig in jedem abgeschlossenen Intervall, welches keinen Unstetigkeitspunkt enthält.*

Um diesen Satz zu beweisen, gehen wir von einer fundamentalen Ungleichung aus, welcher die Fourierschen Koeffizienten einer beliebigen stückweise stetigen Funktion $f(x)$ genügen — es ist also hierbei nicht einmal stückweise Glattheit vorausgesetzt —. Diese sog. *Besselsche Ungleichung* lautet mit beliebigem $n$

$$\frac{a_0^2}{2} + \sum_{\nu=1}^{n} (a_\nu^2 + b_\nu^2) \leqq \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)^2 \, dx \, .$$

Der Beweis folgt aus der Tatsache, daß der Ausdruck

$$\int_{-\pi}^{\pi} \left\{ f(x) - \frac{a_0}{2} - \sum_{\nu=1}^{n} (a_\nu \cos \nu x + b_\nu \sin \nu x) \right\}^2 dx$$

stets positiv oder Null ist. Führt man das Integral aus, indem man das Quadrat unter dem Integralzeichen bildet und die Orthogonalitätsrelationen sowie die Definition der Fourierschen Entwicklungskoeffizienten berücksichtigt, so erhält man sofort die Besselsche Ungleichung in der Form

$$\int_{-\pi}^{\pi} f(x)^2 \, dx - \pi \left\{ \frac{a_0^2}{2} + \sum_{\nu=1}^{n} (a_\nu^2 + b_\nu^2) \right\} \geqq 0 \, .$$

Neben der Besselschen Ungleichung benutzen wir noch eine einfache algebraische *Ungleichung von Schwarz*. Sind $u_1, u_2, \ldots, u_n$ und $v_1, v_2, \ldots, v_n$ beliebige Zahlen, so gilt stets

$$\left(\sum_{\nu=1}^{n} u_\nu v_\nu\right)^2 \leqq \sum_{\nu=1}^{n} u_\nu^2 \cdot \sum_{\nu=1}^{n} v_\nu^2,$$

wobei das Gleichheitszeichen nur dann eintritt, wenn die Folge der $u_\nu$ der Folge der $v_\nu$ proportional ist. Der Beweis folgt sofort aus der Identität

$$\frac{1}{2}\sum_{\mu=1}^{n}\sum_{\nu=1}^{n}(a_\nu b_\mu - a_\mu b_\nu)^2 = \sum_{\nu=1}^{n} a_\nu^2 \cdot \sum_{\nu=1}^{n} b_\nu^2 - \left(\sum_{\nu=1}^{n} a_\nu b_\nu\right)^2,$$

wo links eine Summe von Quadraten steht, die nur bei Proportionalität der $a_\nu$ und $b_\nu$ verschwindet.

Nunmehr setzen wir zunächst voraus, daß die stückweise glatte periodische Funktion $f(x)$ stetig ist. Die Ableitung $g(x) = f'(x)$ ist stückweise stetig, und für die Fourierschen Entwicklungskoeffizienten $c_\nu$ und $d_\nu$ von $g(x)$ gelten die Relationen

$$\begin{aligned} c_0 &= 0, \\ c_\nu &= \nu b_\nu, \\ d_\nu &= -\nu a_\nu \end{aligned} \quad \Bigg\} \; (\nu \geqq 1).$$

Die Besselsche Ungleichung, angewandt auf die Funktion $g(x)$, lautet infolgedessen

$$\sum_{\nu=1}^{n} \nu^2 (a_\nu^2 + b_\nu^2) = \sum_{\nu=1}^{n} (c_\nu^2 + d_\nu^2) \leqq \frac{1}{\pi}\int_{-\pi}^{\pi} g(x)^2\, dx.$$

Schreiben wir für die rechte Seite dieser Ungleichung zur Abkürzung $M^2$, so erhalten wir, wenn $m > n$ ist, durch Anwendung der Schwarzschen Ungleichung

$$\sum_{\nu=n+1}^{m} |a_\nu \cos \nu x + b_\nu \sin \nu x| \leqq \sum_{\nu=n+1}^{m} (|a_\nu| + |b_\nu|) \leqq \sqrt{2M^2}\sqrt{\sum_{\nu=n+1}^{m} \frac{1}{\nu^2}},$$

und da wegen der Konvergenz von $\sum_{\nu=1}^{\infty} \frac{1}{\nu^2}$ die von $x$ unabhängige rechte Seite bei hinreichend großen $n$ und $m$ beliebig klein wird, so ist damit die absolute und gleichmäßige Konvergenz der Reihe bewiesen[1]).

---

[1]) Nebenbei bemerkt zeigt dieselbe Betrachtung, daß für periodische Funktionen mit stetiger Ableitung $(h-1)$-ter Ordnung und wenigstens stückweise stetiger Ableitung $h$-ter Ordnung die Summe $\sum_{1}^{n} \nu^{2h}(a_\nu^2 + b_\nu^2)$ unterhalb einer festen Schranke bleibt, eine Tatsache, die eine charakteristische Aussage über die Art des Kleinwerdens der Fourierschen Koeffizienten gibt. Für eine solche Funktion konvergieren sodann auch die Fourierschen Reihen der Ableitungen bis zur Ordnung $h - 1$ absolut und gleichmäßig.

Um nun auch für unstetige stückweise glatte Funktionen die Formulierung des obigen Satzes zu beweisen, betrachten wir zunächst eine spezielle Funktion $\psi(x)$ dieser Art.

Im Intervall $-\pi < x < \pi$ sei $\psi(x)$ mit der Größe $x$ identisch und werde außerhalb dieses Intervalles periodisch fortgesetzt. Nach § 4 lautet dann ihre Fouriersche Reihe

$$2\left(\frac{\sin x}{1} - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - + \cdots\right).$$

Diese Reihe kann nicht gleichmäßig konvergieren, da ihre Summe die unstetige Funktion $\psi(x)$ darstellt. Wir werden jedoch zeigen, daß die Konvergenz in jedem Intervall $-l \leqq x \leqq l$, für welches $0 < l < \pi$ gilt, gleichmäßig ist.

Den Beweis führen wir durch einen Kunstgriff[1]). Wir beachten, daß die Funktion $\cos\frac{x}{2}$ im Intervalle $-l \leqq x \leqq l$ nirgends unterhalb der positiven Größe $\cos\frac{l}{2} = \varkappa$ liegt. Multiplizieren wir den Betrag der Differenz der $m$-ten und der $n$-ten Partialsumme der obigen Reihe $(m > n)$, d. h. den Ausdruck

$$|S_m(x) - S_n(x)| = 2\left|\frac{\sin(n+1)x}{n+1} - \frac{\sin(n+2)x}{n+2} + - \cdots \pm \frac{\sin mx}{m}\right|$$

mit der Funktion $\cos\frac{x}{2}$, so erhalten wir infolge der bekannten trigonometrischen Formel $2\sin u \cos v = \sin(u+v) + \sin(u-v)$ den Betrag des Ausdrucks

$$\begin{aligned} 2\cos\frac{x}{2}\left(\frac{\sin(n+1)x}{n+1} - \frac{\sin(n+2)x}{n+2} + - \cdots \pm \frac{\sin mx}{m}\right) \\ = \frac{\sin\left(n+\frac{3}{2}\right)x}{n+1} - \frac{\sin\left(n+\frac{5}{2}\right)x}{n+2} + - \cdots \pm \frac{\sin\left(m+\frac{1}{2}\right)x}{m} \\ + \frac{\sin\left(n+\frac{1}{2}\right)x}{n+1} - \frac{\sin\left(n+\frac{3}{2}\right)x}{n+2} + \frac{\sin\left(n+\frac{5}{2}\right)x}{n+3} - + \cdots. \end{aligned}$$

Fassen wir die Glieder rechts in der Art zusammen, wie sie untereinanderstehen, so ergibt sich der Ausdruck

---

[1]) Auf diesen wird man ganz naturgemäß geführt, wenn man bedenkt, daß die Funktion $2y\cos y$, aus dem Intervalle $-\frac{\pi}{2} \leqq y \leqq \frac{\pi}{2}$ periodisch fortgesetzt, stetig bleibt, daß also ihre Fouriersche Reihe nach dem ersten Teil des Satzes gleichmäßig konvergieren und die Funktion darstellen muß. Diese Reihe wird aber aus der Fourierschen Reihe für $2y$ durch Multiplikation mit $\cos y$ entstehen. Setzen wir nun $y = \frac{x}{2}$, so führt diese Multiplikation gerade auf die Betrachtungen des Textes.

$$\frac{\sin\left(n+\frac{1}{2}\right)x}{n+1} \pm \frac{\sin\left(m+\frac{1}{2}\right)x}{m}$$

$$+\frac{\sin\left(n+\frac{3}{2}\right)x}{(n+1)(n+2)} - \frac{\sin\left(n+\frac{5}{2}\right)x}{(n+2)(n+3)} + - \cdots \mp \frac{\sin\left(m-\frac{1}{2}\right)x}{(m-1)m},$$

und wir erhalten daher wegen $\cos\frac{x}{2} \geqq \varkappa$ und $|\sin u| \leqq 1$ die Abschätzung

$$|S_m - S_n| \leqq \frac{1}{\varkappa}\left[\frac{1}{n+1} + \frac{1}{m} + \frac{1}{(n+1)(n+2)} + \cdots + \frac{1}{(m-1)m}\right].$$

Der Ausdruck rechts aber hängt nicht mehr von $x$ ab und wird, wenn nur $n$ und $m$ beide hinreichend groß genommen sind, wegen der Konvergenz der Reihe $\sum_{\nu=1}^{\infty} \frac{1}{\nu(\nu+1)}$ unter jede vorgeschriebene Grenze herabsinken. Das aber ist gleichbedeutend mit der behaupteten gleichmäßigen Konvergenz unserer Fourierschen Reihe.

Nachdem wir so die Reihenentwicklung für eine spezielle unstetige Funktion gewonnen haben, können wir (vgl. § 4) durch Parallelverschiebung des Kurvenbildes bzw. des Koordinatensystemes diese Unstetigkeit an eine beliebige Stelle des Intervalles verlegen. Es wird nämlich die Funktion

$$\psi(x-\xi) = 2\left(\frac{\sin(x-\xi)}{1} - \frac{\sin 2(x-\xi)}{2} + \frac{\sin 3(x-\xi)}{3} - + \cdots\right)$$

bis auf die Stellen $(2k-1)\pi + \xi$ ($k$ eine ganze Zahl) stetig sein. An diesen Stellen aber wird die Funktion beim Überschreiten von links nach rechts einen Sprung „von der Größe $2\pi$“, nämlich von dem Werte $\pi$ auf den Wert $-\pi$ machen und in dem Punkte selbst den Wert 0 annehmen.

Ist nun $f(x)$ eine beliebige stückweise glatte Funktion, die im Intervall $-\pi \leqq x \leqq \pi$ nur die Unstetigkeitsstellen $\xi_1, \xi_2, \ldots, \xi_m$ besitzt, und sind die Sprünge, welche die Funktion dort beim Durchgange von links nach rechts erleidet, $\delta_1, \delta_2, \ldots, \delta_m$, so wird die Funktion

$$f(x) - \frac{\delta_1}{2\pi}\psi(x+\pi-\xi_1) - \frac{\delta_2}{2\pi}\psi(x+\pi-\xi_2) - \cdots - \frac{\delta_m}{2\pi}\psi(x+\pi-\xi_m)$$

stetig und stückweise glatt sein, sich also nach dem bisherigen Resultat in eine gleichmäßig konvergente Fouriersche Reihe entwickeln lassen. Wir erhalten nunmehr ohne weiteres die Fouriersche Reihe der Funktion $f(x)$, indem wir die endlich vielen Fourierschen Reihen der Funktionen $\frac{\delta_1}{2\pi}\psi(x+\pi-\xi_1), \ldots, \frac{\delta_m}{2\pi}\psi(x+\pi-\xi_m)$ gliedweise hinzufügen. Damit aber ist der oben formulierte Hauptsatz bewiesen.

Dieses Resultat ist für die meisten mathematischen Untersuchungen und für die Anwendungen voll ausreichend. Ich möchte aber zum Schluß doch noch darauf hinweisen, daß die Untersuchung der Fourierschen Reihen sehr viel weiter getrieben worden ist. Die Bedingungen, die wir für die Entwickelbarkeit hier aufgestellt haben, sind hinreichende Bedingungen, aber keineswegs notwendige; es ist möglich, Funktionenklassen mit sehr viel weitergehenden Unstetigkeiten durch Fouriersche Reihen darzustellen, und es hat sich an diese Fragen und überhaupt an die Frage der Entwickelbarkeit einer Funktion in eine Fouriersche Reihe eine große spezielle Literatur angeknüpft. Als merkwürdiges Ergebnis solcher Untersuchungen nenne ich nur die Tatsache, daß es stetige Funktionen gibt, deren Fouriersche Reihe in keinem auch noch so kleinen Intervalle konvergiert. Ein solches Resultat besagt natürlich nichts gegen die Brauchbarkeit der Fourierschen Reihen; man wird es im Gegenteil als einen Beleg dafür ansehen müssen, daß der bloße abstrakte Begriff der stetigen Funktion unmittelbar gar nicht zu übersehende Möglichkeiten in sich schließt, wie ja auch das Beispiel stetiger, aber nirgends differenzierbarer Funktionen zeigt.

## § 6. Die mittlere Approximation durch trigonometrische Polynome.

Bei der Fourierschen Reihe wie überhaupt bei allen unendlichen Reihen und sonstigen unendlichen Prozessen muß man sich immer wieder vor Augen halten, daß der Sinn jeder solchen Entwicklung in der Tatsache der Approximation der zu entwickelnden Funktion durch einen endlichen Ausdruck besteht; Entwicklung in eine unendliche Reihe heißt, daß die endliche Summe, die durch Abbrechen nach dem $n$-ten Gliede entsteht, eine Annäherung an die entwickelte Funktion darstellt, und zwar eine Annäherung, die durch Wahl eines hinreichend großen $n$ beliebig gut gemacht werden kann.

Die Annäherung einer Funktion $f(x)$ durch das trigonometrische Interpolationspolynom aus § 3 ist nur eine unter vielen möglichen Approximationen. Das Prinzip der Annäherung ist hier Übereinstimmung mit der gegebenen Funktion an $2n+1$ gegebenen Stellen, während für andere Stellen nichts mehr verlangt wird. Approximiert man dagegen eine Funktion $f(x)$ durch dasjenige trigonometrische Polynom, welches entsteht, wenn man die Fouriersche Reihe nach dem Gliede $a_n \cos nx + b_n \sin nx$ abbricht, so hat man eine Annäherung anderer Art; denn die Koeffizienten stimmen jetzt nicht mehr exakt mit den Koeffizienten des Interpolationspolynomes überein. Diese Annäherung durch die Partialsummen der Fourierschen Reihe hat eine einfache Bedeutung, die ich jetzt kurz entwickeln will.

Man kann sich folgende Aufgabe stellen: Unter allen trigonometrischen Polynomen $n$-ten Grades

$$s_n(x) = \frac{\alpha_0}{2} + \sum_{\nu=1}^{n} (\alpha_\nu \cos \nu x + \beta_\nu \sin \nu x)$$

soll durch geeignete Wahl der Koeffizienten $\alpha_\nu$ und $\beta_\nu$ dasjenige gefunden werden, für welches der Ausdruck

$$\frac{1}{2\pi}\int_{-\pi}^{+\pi}\{f(x)-s_n(x)\}^2 dx = \frac{1}{2\pi}\int_{-\pi}^{+\pi}\left\{f(x)-\frac{\alpha_0}{2}-\sum_{\nu=1}^{n}(\alpha_\nu\cos\nu x+\beta_\nu\sin\nu x)\right\}^2 dx,$$

das „*mittlere Fehlerquadrat*", möglichst klein wird. Wir sagen, daß das so bestimmte Polynom die beste *Approximation im Mittel* für die Funktion $f(x)$ liefert.

Diese Forderung bzw. diese Bezeichnung rechtfertigt sich unmittelbar durch folgende Erwägung. Je kleiner das zum Minimum zu machende Integral wird, desto kleiner muß durchschnittlich auch der Integrand, d. h. die Abweichung von $f(x)$ und $s_n(x)$ sein, wenn auch die Kleinheit dieses Integrales es nicht prinzipiell ausschließt, daß in der engen Umgebung einzelner Stellen noch beträchtliche Abweichungen der Funktionen vorkommen. Diese Aufgabe, die beste mittlere Approximation zu finden, läßt sich nun unmittelbar lösen: Führen wir das Quadrat unter dem Integralzeichen aus, so erhalten wir für den Integranden den folgenden Ausdruck:

$$f(x)^2-\alpha_0 f(x)-2\sum_{\nu=1}^{n}(\alpha_\nu f(x)\cos\nu x+\beta_\nu f(x)\sin\nu x)$$
$$+\left(\frac{\alpha_0}{2}+\sum_{\nu=1}^{n}(\alpha_\nu\cos\nu x+\beta_\nu\sin\nu x)\right)^2.$$

Beachten wir nun, daß die Ausdrücke

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos\nu x\,dx = a_\nu, \qquad \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin\nu x\,dx = b_\nu$$

gerade die in dem vorigen Paragraphen eingeführten Fourierschen Koeffizienten der Funktion $f(x)$ sind, und führen die Integration gliedweise aus, so gelangen wir unter Beachtung der Orthogonalitätsrelationen (vgl. § 5, Nr. 2, S. **374**) sofort zu der Gleichung

$$\int_{-\pi}^{\pi}\{f(x)-s_n(x)\}^2 dx$$
$$=\int_{-\pi}^{\pi} f(x)^2 dx+\pi\left\{\frac{\alpha_0^2}{2}+\sum_{\nu=1}^{n}(\alpha_\nu^2+\beta_\nu^2)-\alpha_0 a_0-2\sum_{\nu=1}^{n}(\alpha_\nu a_\nu+\beta_\nu b_\nu)\right\}$$
$$=\int_{-\pi}^{\pi} f(x)^2 dx+\pi\left\{\frac{1}{2}(\alpha_0-a_0)^2+\sum_{\nu=1}^{n}\left((\alpha_\nu-a_\nu)^2+(\beta_\nu-b_\nu)^2\right)\right\}$$
$$-\pi\left\{\frac{a_0^2}{2}+\sum_{\nu=1}^{n}(a_\nu^2+b_\nu^2)\right\}.$$

Diese Gleichung macht es evident, daß das Integral der linken Seite seinen kleinsten Wert erhält, wenn wir gerade

$$\alpha_\nu = a_\nu, \qquad \beta_\nu = b_\nu$$

setzen. Wir erhalten so das Resultat: Die Fouriersche Partialsumme $S_n(x) = \frac{a_0}{2} + \sum_{\nu=1}^{n}(a_\nu \cos\nu x + b_\nu \sin\nu x)$ liefert die beste mittlere Approximation einer Funktion $f(x)$ durch ein trigonometrisches Polynom $n$-ten Grades; für den Minimumwert des mittleren Fehlerquadrats gilt

$$(*) \qquad \int_{-\pi}^{\pi}(f(x) - S_n(x))^2\,dx = \int_{-\pi}^{\pi} f(x)^2\,dx - \pi\left\{\frac{a_0^2}{2} + \sum_{\nu=1}^{n}(a_\nu^2 + b_\nu^2)\right\}.$$

Dieses Resultat zeigt uns, daß, wenn man die Genauigkeit der mittleren Approximation durch Steigerung des Grades $n$ erhöhen will, eine Modifikation der einmal gefundenen Koeffizienten nicht mehr nötig ist, daß vielmehr nur noch die weiteren Koeffizienten des Polynomes neu hinzuzubestimmen sind. Aus der obigen Gleichung haben wir bereits früher die Besselsche Ungleichheitsbeziehung

$$\frac{a_0^2}{2} + \sum_{\nu=1}^{n}(a_\nu^2 + b_\nu^2) \leqq \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)^2\,dx$$

hergeleitet. Diese lehrt uns, daß die aus nichtnegativen Summanden bestehende linke Seite höchstens gleich der von $n$ nicht mehr abhängigen Zahl $\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)^2\,dx$ ist. Daher konvergiert die linke Seite mit wachsendem $n$ gegen eine bestimmte Grenze, die selbst nicht größer als jener Integralausdruck ist; mit anderen Worten: Wir erhalten die Beziehung

$$\frac{a_0^2}{2} + \sum_{\nu=1}^{\infty}(a_\nu^2 + b_\nu^2) \leqq \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)^2\,dx.$$

Aber es gilt noch mehr. Setzen wir der Einfachheit halber voraus, daß die Funktion $f(x)$ in eine gleichmäßig konvergente Fouriersche Reihe entwickelbar ist, so wird bei hinreichend großem $n$ der Integrand auf der linken Seite von $(*)$ gleichmäßig in $x$ beliebig klein werden; also wird auch das Integral beliebig klein, und wir erhalten daher an Stelle der Besselschen Ungleichung die Gleichung

$$\frac{a_0^2}{2} + \sum_{\nu=1}^{\infty}(a_\nu^2 + b_\nu^2) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)^2\,dx.$$

Diese Gleichung, die man auch als die *Vollständigkeitsrelation* der trigonometrischen Funktionen bezeichnet, gilt übrigens, wie ich ebenfalls hier nicht beweisen will, auch für viel allgemeinere Klassen von Funktionen $f(x)$, z. B. für beliebige stückweise stetige Funktionen[1]).

Man kann dieser Vollständigkeitsrelation noch leicht eine etwas allgemeinere Form geben. Sind $f(x)$ und $\varphi(x)$ zwei Funktionen mit den Fourierschen Koeffizienten $a_\nu$, $b_\nu$ bzw. $\alpha_\nu$, $\beta_\nu$, so lautet die Vollständigkeitsrelation für $f(x)+\varphi(x)$

$$\frac{(a_0+\alpha_0)^2}{2}+\sum_{\nu=1}^{\infty}\{(a_\nu+\alpha_\nu)^2+(b_\nu+\beta_\nu)^2\}=\frac{1}{\pi}\int_{-\pi}^{+\pi}\{f(x)^2+2f(x)\,\varphi(x)+\varphi(x)^2\}\,dx,$$

während die für $f(x)$ und $\varphi(x)$ lauten

$$\frac{a_0^2}{2}+\sum_{\nu=1}^{\infty}(a_\nu^2+b_\nu^2)=\frac{1}{\pi}\int_{-\pi}^{+\pi}f(x)^2\,dx,$$

$$\frac{\alpha_0^2}{2}+\sum_{\nu=1}^{\infty}(\alpha_\nu^2+\beta_\nu^2)=\frac{1}{\pi}\int_{-\pi}^{+\pi}\varphi(x)^2\,dx.$$

Ziehen wir die beiden letzten von der ersten ab, so erhalten wir die Beziehung

$$\frac{a_0\,\alpha_0}{2}+\sum_{\nu=1}^{\infty}(a_\nu\,\alpha_\nu+b_\nu\,\beta_\nu)=\frac{1}{\pi}\int_{-\pi}^{+\pi}f(x)\,\varphi(x)\,dx,$$

welche man als die verallgemeinerte Vollständigkeitsrelation für ein Funktionenpaar bezeichnet.

# Anhang zum neunten Kapitel.

## Beispiele zur trigonometrischen Interpolation.

### 1. Vorbemerkungen.

Während wir bisher nur Beispiele für die Fouriersche Reihenentwicklung betrachtet haben, möchte ich jetzt zeigen, daß auch die trigonometrische Interpolation zu interessanten Formeln führt[2]). Demgemäß will ich in diesem Paragraphen für einige der einfachsten Funktionen die im neunten Kapitel, § 3, gestellten Interpolationsaufgaben kurz behandeln. Dabei ist es nützlich, unserer Formel aus § 3, welche sich auf die Einteilung des Intervalles von $-\pi$ bis $+\pi$ in $2n+1$ Teile

---

[1]) Vgl. *R. Courant* und *D. Hilbert*: Methoden der mathem. Physik. I. S. 36.

[2]) Ich glaube, auf die kurze Angabe der folgenden — in der Literatur sonst fehlenden — Rechnungen nicht verzichten zu sollen, wenn sie auch vielleicht dem Anfänger Schwierigkeiten bereiten werden. Im übrigen wird von den Ergebnissen dieses Paragraphen nirgends in diesem Buche Gebrauch gemacht.

bezog, noch eine andere zur Seite zu stellen, welche nicht so symmetrisch erscheint, weil sie den Anfangspunkt $x = -\pi$ des Intervalles auszeichnet. Denken wir uns das Intervall durch die $2n$ Teilpunkte

$$x_{-n} = -n\lambda = -\pi, \ldots, x_\varkappa = \varkappa\lambda, \ldots, x_{n-1} = (n-1)\lambda,$$

wo $\lambda = \frac{\pi}{n}$ ist, in $2n$ Teile geteilt und bezeichnen wir wieder den Funktionswert einer Funktion $f(x)$ im Teilpunkte $x_\varkappa$ mit $f_\varkappa = f(x_\varkappa)$, so erhalten wir ganz wie früher die Koeffizienten eines Interpolationspolynomes

$$S_n(x) = \sum_{\nu=-n}^{n-1} \alpha_\nu e^{i\nu x},$$

welches in diesen Teilpunkten die vorgeschriebenen Werte $f_\varkappa$ besitzt, in der Form

$$\alpha_\nu = \frac{1}{2n} \sum_{\varkappa=-n}^{n-1} f_\varkappa e^{-i\varkappa\nu\lambda}.$$

Diese unsymmetrische Darstellung wird immer dann zweckmäßig verwendet werden, wenn es sich um eine wirklich periodische Funktion $f(x)$ handelt, wenn also $f(\pi) = f(-\pi)$ ist. Da nämlich $S_n(\pi) = S_n(-\pi)$ ist, so stellt die Funktion $S_n(x)$ dann von selbst auch den richtigen Funktionswert an der Stelle $x = \pi$ dar, so daß die formale Unsymmetrie nicht stört.

Ist $f(-\pi) \neq f(\pi)$, so werden wir die ursprünglichen Formeln aus § 3 verwenden. Da in die Interpolationsformeln von der Funktion nur die $2n+1$ bzw. $2n$ Funktionswerte $f_\varkappa$ eingehen, so können wir unsere Interpolationsformeln in der Gestalt folgender *reziproker Beziehungen* aussprechen:

I. Wenn zwischen $2n$ Zahlenwerten $f_{-n}, \ldots, f_\varkappa, \ldots, f_{n-1}$ und $2n$ Zahlenwerten $c_{-n}, \ldots, c_\varkappa, \ldots, c_{n-1}$ die Gleichungen

$$f_\varkappa = \frac{1}{\sqrt{2n}} \sum_{\nu=-n}^{n-1} c_\nu e^{i\nu\varkappa\lambda}, \qquad \lambda = \frac{\pi}{n}$$

bestehen, so gelten auch die Gleichungen

$$c_\nu = \frac{1}{\sqrt{2n}} \sum_{\varkappa=-n}^{n-1} f_\varkappa e^{-i\varkappa\nu\lambda}$$

und umgekehrt.

II. Wenn zwischen $2n+1$ Werten $f_{-n}, \ldots, f_\varkappa, \ldots, f_n$ und $2n+1$ Werten $c_{-n}, \ldots, c_\varkappa, \ldots, c_n$ die Gleichungen

$$f_\varkappa = \frac{1}{\sqrt{2n+1}} \sum_{\nu=-n}^{n} c_\nu e^{i\nu\varkappa\lambda}, \qquad \lambda = \frac{2\pi}{2n+1}$$

bestehen, so folgen daraus die Gleichungen

$$c_\nu = \frac{1}{\sqrt{2n+1}} \sum_{\varkappa=-n}^{n} f_\varkappa e^{-i\varkappa\nu\lambda}$$

und umgekehrt.

Aus diesen Darstellungen der Werte $f_\varkappa$ erhalten wir das Interpolationspolynom, indem wir rechts $\varkappa\lambda$ durch $x$ ersetzen.

## 2. Einzelne Beispiele.

Wir behandeln nun einzelne Beispiele, indem wir von den einfachsten früher in Fouriersche Reihen entwickelten Funktionen $f(x)$ ausgehen und diesen die Werte $f_\varkappa$ entnehmen.

1. Wir setzen $f_\varkappa = \varkappa$, was der Funktion $f(x) = x$ entspricht, und wenden die Formel II an. Dann ist

$$\sqrt{2n+1}\, c_\nu = \sum_{\mu=-n}^{n} \mu\, e^{-i\mu\nu\lambda}.$$

Der Wert dieser Summe für $\nu = 0$ ist $\sqrt{2n+1}\, c_0 = 0$. Im Falle $\nu \neq 0$ erhalten wir, indem wir $\varrho = e^{-i\nu\lambda}$ setzen,

$$\sqrt{2n+1}\, c_\nu = \sum_{\mu=-n}^{n} \mu\, \varrho^\mu.$$

Um diese Summe auszuwerten, multiplizieren wir sie mit $1-\varrho$. Wir erhalten, da $\varrho \neq 1$ ist,

$$(1-\varrho)\sum_{\mu=-n}^{n} \mu\varrho^\mu = \sum_{\mu=-n}^{n} \varrho^\mu - n\varrho^{n+1} - (n+1)\varrho^{-n}$$
$$= \varrho^{-n}\left[\frac{1-\varrho^{2n+1}}{1-\varrho} - n\varrho^{2n+1} - (n+1)\right].$$

Beachten wir nun, daß $\varrho^{2n+1} = e^{-i\nu 2\pi} = 1$ und $\varrho^{n+\frac{1}{2}} = e^{-i\nu\pi} = (-1)^\nu$ ist, so ergibt sich

$$\sqrt{2n+1}\, c_\nu = -(2n+1)\frac{\varrho^{n+1}}{1-\varrho} = -(2n+1)(-1)^\nu \frac{e^{-i\nu\frac{\lambda}{2}}}{1-e^{-i\nu\lambda}}$$

oder

$$\alpha_\nu = \frac{c_\nu}{\sqrt{2n+1}} = \frac{(-1)^\nu i}{2\sin\frac{\nu\lambda}{2}}.$$

Wir erhalten also für $f_\varkappa = \varkappa$ die Interpolation

$$\varkappa = \sum_{\nu=-n}^{n} \frac{(-1)^\nu i e^{i\nu\varkappa\lambda}}{2\sin\frac{\nu\lambda}{2}}$$

(wobei $\nu$ den Wert 0 nicht annehmen soll) oder in reeller Schreibweise

$$\varkappa = -\sum_{\nu=1}^{n} (-1)^{\nu} \frac{\sin \nu \varkappa \lambda}{\sin \frac{\nu\lambda}{2}}.$$

2. Zur Darstellung der Werte $f_{\varkappa} = |\varkappa|$, die der Funktion $f(x) = |x|$ entsprechen, verwenden wir die erste Interpolationsformel. Wir erhalten

$$\sqrt{2n}\, c_{\nu} = \sum_{\mu=-n}^{n-1} |\mu| e^{-i\mu\nu\lambda}$$

und finden sofort

$$\sqrt{2n}\, c_0 = 2\sum_{\mu=1}^{n-1} \mu + n = n^2, \text{ d. h. } \alpha_0 = \frac{n}{2}.$$

Für $\nu \neq 0$ ergibt sich

$$\sqrt{2n}\, c_{\nu} = \sum_{\mu=1}^{n-1} \mu (e^{-i\mu\nu\lambda} + e^{i\mu\nu\lambda}) + (-1)^{\nu} n.$$

Setzen wir $e^{-i\nu\lambda} = \varrho$, so ist stets $\varrho \neq 1$, und die hier auftretende Summe

$$\sum_{\mu=1}^{n-1} \mu \varrho^{\mu} + \sum_{\mu=1}^{n-1} \mu \varrho^{-\mu}$$

läßt sich ähnlich wie im ersten Beispiel auswerten; wir erhalten

$$\frac{1}{1-\varrho}\left[\frac{\varrho - \varrho^n}{1-\varrho} - (n-1)\varrho^n - \frac{\varrho^{-1} - \varrho^{-n}}{1-\varrho^{-1}} - 1 + n\varrho^{-n+1}\right].$$

Ist $\nu$ gerade, so nimmt dieser Ausdruck wegen $\varrho^n = \varrho^{-n} = e^{-i\nu\pi} = 1$ (man beachte $\lambda n = \pi$) den Wert $-n$ an; also ist

$$\sqrt{2n}\, c_{\nu} = -n + (-1)^{\nu} n = 0.$$

Ist $\nu$ ungerade, also $\varrho^n = \varrho^{-n} = -1$, so hat unsere Summe den Wert

$$n + \frac{4\varrho}{(1-\varrho)^2} = n + \frac{4}{\left(e^{\frac{i\nu\lambda}{2}} - e^{-\frac{i\nu\lambda}{2}}\right)^2};$$

also ist in diesem Falle

$$\sqrt{2n}\, c_{\nu} = -\frac{1}{\sin^2 \frac{\nu\lambda}{2}}.$$

Als Interpolationsformel erhalten wir, wenn wir gleich auf die reelle Darstellung übergehen und uns auf den Fall beschränken, daß $n$

gerade ist,

$$|\varkappa| = \frac{n}{2} - \frac{1}{n}\sum_{\nu=1}^{n-1} \frac{\cos \varkappa\nu\lambda}{\sin^2\frac{\nu\lambda}{2}},$$

wobei nur über ungerade Werte von $\nu$ summiert werden soll.

Die beiden Interpolationsformeln für $f_\varkappa = \varkappa$ und $f_\varkappa = |\varkappa|$ gehen rein formal in die in § 4, Nr. 2 und 4 aufgestellten Fourierschen Reihen für $f(x) = x$ und $f(x) = |x|$ über, wenn wir $\varkappa\lambda$ mit $x$ identifizieren, $n$ und damit $\frac{1}{\lambda}$ über alle Grenzen wachsen lassen und $\frac{1}{\lambda}\sin\frac{\nu\lambda}{2}$ durch $\frac{\nu}{2}$ ersetzen. Entsprechendes gilt für die folgenden Interpolationsformeln, die ich ohne Beweis angebe.

3. Bei Verwendung der Formel II erhalten die Werte

$$\begin{aligned} f_\varkappa &= -1 \quad \text{für} \quad -n \leqq \varkappa \leqq -1,\\ f_0 &= 0,\\ f_\varkappa &= 1 \quad \text{für} \quad 1 \leqq \varkappa \leqq n, \end{aligned}$$

die der in § 4, Nr. 5 behandelten Funktion entsprechen, die Darstellung

$$f_\varkappa = \frac{2}{2n+1}\sum_{\nu=1}^{n} \frac{\cos\frac{\nu\lambda}{2} - (-1)^\nu}{\sin\frac{\nu\lambda}{2}} \sin\varkappa\nu\lambda.$$

4. Für $f_\varkappa = \cos\alpha\varkappa\lambda$, wo $\alpha$ nicht ganz ist, erhalten wir unter Verwendung der Formel I

$$\cos\alpha\varkappa\lambda = \frac{1}{2n}\sin\alpha\pi\left\{\operatorname{ctg}\frac{\alpha\lambda}{2} + \sum_{\nu=1}^{n-1}(-1)^\nu \frac{\sin\alpha\lambda\cos\varkappa\nu\lambda}{\sin^2\frac{\alpha\lambda}{2} - \sin^2\frac{\nu\lambda}{2}} + (-1)^\varkappa(-1)^n \operatorname{tg}\frac{\alpha\lambda}{2}\right\}.$$

Setzen wir $\varkappa = -n$ und dividieren durch $\sin\alpha\pi$, so entsteht hieraus

$$\operatorname{ctg}\alpha\pi = \frac{1}{2n}\left\{\operatorname{ctg}\frac{\alpha\lambda}{2} + \sum_{\nu=1}^{n-1} \frac{\sin\alpha\lambda}{\sin^2\frac{\alpha\lambda}{2} - \sin^2\frac{\nu\lambda}{2}} + \operatorname{tg}\frac{\alpha\lambda}{2}\right\}.$$

Integrieren wir diese Formel nach $\alpha$ von $\alpha = 0$ bis $\alpha = x$, wo $0 < x < 1$ ist, so erhalten wir unter Beachtung der Integrationsformel

$$\pi\int_0^x \left(\operatorname{ctg}\alpha\pi - \frac{1}{2n}\operatorname{ctg}\frac{\alpha\pi}{2n}\right)d\alpha = \log\frac{\sin x\pi}{\sin\frac{x\pi}{2n}} - \log 2n$$

die Gleichung

$$\log\sin x\pi = \log\left(2n\sin\frac{x\lambda}{2}\right) + \sum_{\nu=1}^{n-1}\log\left(1 - \frac{\sin^2\frac{x\lambda}{2}}{\sin^2\frac{\nu\lambda}{2}}\right) - \log\cos\frac{x\lambda}{2} \quad \left[\lambda = \frac{\pi}{n}\right],$$

aus der wir sofort

$$\sin \varkappa\pi = \sin \varkappa n\lambda = 2n \operatorname{tg}\frac{\varkappa\lambda}{2} \prod_{\nu=1}^{n-1}\left(1 - \frac{\sin^2\frac{\varkappa\lambda}{2}}{\sin^2\frac{\nu\lambda}{2}}\right)$$

gewinnen.

Diese Formel stellt eine Produktzerlegung des Sinus dar, das Analogon zu der Zerlegung in ein unendliches Produkt, die wir in § 4, Nr. 7 aufgestellt haben.

5. Weitere interessante Beispiele bieten die Darstellungen der Größen

$$f_\varkappa = \mathfrak{Cos}\, \alpha\varkappa\lambda$$

und

$$f_\varkappa = \mathfrak{Sin}\, \alpha\varkappa\lambda,$$

wobei $\alpha$ eine beliebige Zahl ist.

Die Benutzung der Formel I führt nach analogen Rechnungen wie die unter Nr. 4 zu den Darstellungen

$$\mathfrak{Cos}\, \alpha\varkappa\lambda = \frac{\mathfrak{Sin}\, \alpha\pi}{2n}\left[\mathfrak{Ctg}\frac{\alpha\lambda}{2} + \sum_{\nu=1}^{n-1}(-1)^\nu \frac{\mathfrak{Sin}\, \alpha\lambda \cos\nu\varkappa\lambda}{\sin^2\frac{\nu\lambda}{2} + \mathfrak{Sin}^2\frac{\alpha\lambda}{2}} + (-1)^{\varkappa+n}\mathfrak{Tg}\frac{\alpha\lambda}{2}\right]$$

bzw.

$$\mathfrak{Sin}\, \alpha\varkappa\lambda = -\frac{\mathfrak{Sin}\, \alpha\pi}{2n}\left[\sum_{\nu=1}^{n-1}(-1)^\nu \frac{\sin\nu\lambda \sin\nu\varkappa\lambda}{\sin^2\frac{\nu\lambda}{2} + \mathfrak{Sin}^2\frac{\alpha\lambda}{2}}\right.$$
$$\left. + 1 + 2\sum_{\nu=1}^{n-1}(-1)^\nu \cos\nu\varkappa\lambda + (-1)^{\varkappa+n}\right],$$

wobei die Summe der letzten drei Ausdrücke für alle Werte von $\varkappa$ außer $\varkappa = -n$ verschwindet.

Zum Schlusse dieser Betrachtungen mag hervorgehoben werden, wieviel komplizierter die Lösung des algebraischen Interpolationsproblemes sich ausnimmt als die Lösung des entsprechenden Problemes der Fourierschen Reihenentwicklung, die man durch Grenzübergang aus den Interpolationsformeln formal erhält. Es bestätigt sich damit, daß Grenzübergänge, wenn sie auch prinzipiell aus dem Bereiche der algebraischen elementaren Betrachtungen herausführen, doch tatsächlich in sehr vielen Fällen das Hilfsmittel zu einer wesentlichen Vereinfachung der formalen Beziehungen darstellen.

Zehntes Kapitel.

# Die Differentialgleichungen der einfachsten Schwingungsvorgänge.

Schon bei verschiedenen Gelegenheiten sind uns *Differentialgleichungen* begegnet, d. h. Gleichungen, aus welchen eine zunächst unbekannte Funktion zu bestimmen ist und in welchen nicht nur diese Funktion selbst, sondern auch ihre Differentialquotienten auftreten.

Das einfachste derartige Problem bietet uns die unbestimmte Integration einer gegebenen Funktion $f(x)$; dieses Integrationsproblem verlangt, eine Funktion $y = F(x)$ zu bestimmen, für welche die Differentialgleichung $y' - f(x) = 0$ erfüllt ist. Weiter hatten wir im dritten Kapitel, § 7, erkannt, daß eine Gleichung der Form $y' = \alpha y$ durch Exponentialfunktionen der Form $y = c\,e^{\alpha x}$ gelöst wird. Sodann haben wir im fünften Kapitel, § 4, gesehen, daß die Probleme der Mechanik auf Differentialgleichungen führen, und ganz allgemein zeigt es sich, daß viele Teile der reinen Mathematik und die meisten Anwendungsgebiete von Differentialgleichungen beherrscht werden. Ohne auf eine allgemeine Theorie der Differentialgleichungen einzugehen, will ich in diesem Kapitel als theoretisch instruktives und für die Anwendungen überaus wichtiges B. ispiel die Differentialgleichungen der einfachsten Schwingungsvorgänge behandeln.

Dabei wird es nützlich sein, folgende allgemeinen Begriffe und Bezeichnungen sich vor Augen zu halten. Unter *Lösung* einer Differentialgleichung verstehen wir eine Funktion, welche, in die Differentialgleichung eingesetzt, diese für alle in Betracht kommenden Werte der unabhängigen Veränderlichen erfüllt. Man pflegt statt Lösung einer Differentialgleichung auch häufig *Integral* der Differentialgleichung zu sagen, einmal, weil es sich gewissermaßen um die Verallgemeinerung des gewöhnlichen Integrationsproblemes handelt, und zweitens, weil man in vielen Fällen die Auflösung tatsächlich auf die Ausführung von Integrationen zurückführt.

## § 1. Schwingungsprobleme der Mechanik und Physik.

### 1. Einfachste mechanische Schwingungen.

Den einfachsten Typus mechanischer Schwingungen haben wir schon im fünften Kapitel, § 4, betrachtet. Es handelt sich um einen Massenpunkt der Masse $m$, welcher auf der $x$-Achse beweglich und durch eine elastische Kraft an seine Ruhelage im Punkte $x = 0$ gefesselt ist. Die Größe dieser elastischen Kraft haben wir proportional der Elongation $x$ angenommen, nämlich gleich $-kx$ gesetzt, wobei $k$ eine positive Konstante ist und das negative Vorzeichen der Tatsache Ausdruck verleiht, daß die Kraft immer auf den Nullpunkt zu gerichtet ist. Wir wollen nunmehr weiter das Vorhandensein einer Reibungskraft voraussetzen und annehmen, daß diese Reibungskraft proportional der Geschwindigkeit $\frac{dx}{dt} = \dot{x}$ des Punktes und ihr entgegengesetzt ist, d. h. durch einen Ausdruck der Form $-r\dot{x}$ mit einer positiven Reibungskonstanten $r$ gemessen wird. Nehmen wir endlich an, daß auf den Punkt eine als Funktion $f(t)$ der Zeit $t$ gegebene äußere Kraft wirkt, so muß nach dem Newtonschen Grundgesetz (vgl. fünftes Kapitel, § 4) das Produkt von Masse $m$ und Beschleunigung $\ddot{x}$ gleich der elastischen Kraft vermehrt um die Reibungskraft und die äußere Kraft sein, und wir erhalten somit als Ausdruck des Newtonschen Grundgesetzes die Gleichung

$$m\ddot{x} + r\dot{x} + kx = f(t).$$

Diese *Schwingungsgleichung* regelt den Ablauf des mechanischen Vorganges. Wie uns die schon früher behandelten speziellen Beispiele für Differentialgleichungen zeigen (z. B. das Integrationsproblem $\dot{x} = \frac{dx}{dt} = f(t)$ mit seiner Lösung $x = \int f(t)\,dt + c$ oder die Auflösung der speziellen Differentialgleichung $m\ddot{x} + kx = 0$ im fünften Kapitel, § 4), wird eine solche Differentialgleichung keineswegs nur eine einzige oder endlich viele voneinander verschiedene Lösungen besitzen; vielmehr wird es, wie wir sogleich im einzelnen sehen werden, unendlich viele Lösungen geben, die sich folgendermaßen darstellen. Es läßt sich eine *„allgemeine Lösung“* $x(t)$ der Schwingungsgleichung angeben, welche außer von der unabhängigen Veränderlichen $t$ noch von zwei gänzlich beliebigen Parametern $c_1$ und $c_2$, den sog. unbestimmten *Integrationskonstanten*, abhängt; setzen wir für $c_1$ und $c_2$ bestimmte Werte ein, so erhalten wir eine bestimmte spezielle Lösung, und zwar können wir in dieser Weise jede solche Lösung gewinnen. Die „allgemeine Lösung“ ist also der Inbegriff aller speziellen Lösungen.

Diese Tatsache ist durchaus verständlich (vgl. auch fünftes Kapitel, § 4). Wir können nämlich nicht erwarten, daß unsere Schwingungsgleichung allein den Vorgang vollständig bestimmt. Vielmehr ist es plausibel, daß man zu einem gegebenen Zeitpunkt, etwa zum Zeitpunkt $t = 0$ die Anfangslage $x(0) = x_0$ und die Anfangsgeschwindigkeit $\dot{x}(0) = \dot{x}_0$ — kurz gesagt, den *Anfangszustand* — noch willkürlich wählen darf, d. h. daß man zur Zeit $t = 0$ den Punkt von irgend einer Anfangslage aus mit gegebener Geschwindigkeit in Bewegung setzen kann; erst dann dürfen wir den zwangsläufigen Ablauf der Bewegung erwarten. Die beiden verfügbaren Konstanten $c_1$ und $c_2$ in der allgemeinen Lösung werden dann gerade dazu dienen, eine bestimmte, diesen speziellen Anfangsbedingungen angepaßte Lösung herauszugreifen; wir werden im nächsten Paragraphen erkennen, daß und wie dies auf eindeutige Weise möglich ist.

Ist keine äußere Kraft vorhanden, d. h. ist $f(t) = 0$, so sprechen wir von einer *freien Bewegung*. Die Differentialgleichung heißt dann *homogen*. Ist dagegen $f(t)$ nicht für alle $t$ gleich Null, so sprechen wir von einer *erzwungenen Bewegung* bzw. einer *unhomogenen* Differentialgleichung. $f(t)$ heißt gelegentlich auch das *Störungsglied*.

## 2. Elektrische Schwingungen.

Ein mechanisches System der geschilderten einfachen Art wird in Wirklichkeit stets nur angenähert vorliegen. Eine Annäherung bietet z. B. das Pendel, solange die Ausschläge klein sind; auch die Schwingungen einer Magnetnadel, die Schwingungen des Mittelpunktes einer Telephon- oder einer Mikrophonmembran und andere mechanische Schwingungsvorgänge kann man mit einer gewissen Näherung durch ein System des beschriebenen Schemas ersetzen. Vorgänge jedoch, welche der obigen Differentialgleichung viel exakter entsprechen, zeigt uns der elektrische Schwingungskreis.

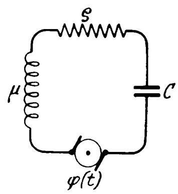


Fig. 124. Elektrischer Schwingungskreis.

Betrachten wir den in der Figur **124** angedeuteten Schwingungskreis, dessen Selbstinduktion $\mu$, dessen Widerstand $\varrho$ und dessen Kapazität $C = \frac{1}{\varkappa}$ sei. Es möge außerdem eine von außen wirkende, als Funktion der Zeit $t$ bekannte elektromotorische Kraft $\varphi(t)$ auf den Schwingungskreis einwirken, z. B. die Spannung, die eine Maschine liefert, oder die Spannungen, welche von elektrischen Wellen erzeugt werden. Um den Vorgang in unserem Stromkreise wiederum durch eine Differentialgleichung zu beschreiben, bezeichnen wir die am Kondensator herrschende Spannung mit $E$, die im Kondensator vorhandene Gesamtladung mit $Q$, so daß zwischen diesen Größen die Gleichung $CE = \frac{1}{\varkappa} E = Q$ besteht. Die Stromstärke $J$, welche ebenso wie die Spannung $E$ eine Funktion der Zeit sein wird, ist definiert als Größe der Änderung der Ladung pro Zeiteinheit, d. h. als die Geschwindigkeit, mit welcher die Kondensatorladung sich vermindert: $J = -\dot{Q} = -\frac{dQ}{dt} = -\frac{1}{\varkappa}\dot{E}$. Das Ohmsche Gesetz sagt ferner, daß Stromstärke mal Ohmschem Widerstand gleich der wirkenden elektromotorischen Kraft ist, d. h. also jetzt gleich der Kondensatorspannung $E$ vermindert um die Gegenspannung der Selbstinduktion, d. h. um $\mu\dot{J}$, und vermehrt um die äußere Spannung $\varphi(t)$. Wir gelangen so zu der Gleichung $J\varrho = E - \mu\dot{J} + \varphi(t)$ oder $-\frac{\varrho}{\varkappa}\dot{E} = E + \frac{\mu}{\varkappa}\ddot{E} + \varphi(t)$, d. h. $\mu\ddot{E} + \varrho\dot{E} + \varkappa E = -\varkappa\varphi(t)$, welcher die Spannung in dem Stromkreise genügt. Wir sehen, daß wir eine Differentialgleichung ganz von dem Typus der unter Nr. 1 betrachteten Art erhalten haben. Es entspricht dabei der Masse die Selbstinduktion, der Reibung der Widerstand und dem Elastizitätskoeffizienten die reziproke Kapazität; der äußeren Kraft entspricht (bis auf einen konstanten Faktor) die von außen aufgeprägte Spannung. Ist diese letztere Null, so haben wir es mit einer homogenen Differentialgleichung zu tun.

Multiplizieren wir unsere Differentialgleichung mit $-\frac{1}{\varkappa}$ und differenzieren nach der Zeit, so erhalten wir sofort für die Stromstärke $J$ die entsprechende Differentialgleichung

$$\mu \ddot{J} + \varrho \dot{J} + \varkappa J = \dot{\varphi}(t),$$

welche sich nur durch die rechte Seite von der Gleichung für die Spannung unterscheidet und bei freier Bewegung ($\varphi = 0$) vollständig mit ihr übereinstimmt.

## § 2. Lösung der homogenen Gleichung. Freie Bewegungen.

### 1. Formale Auflösung.

Man kann eine Lösung der homogenen Differentialgleichung $m\ddot{x} + r\dot{x} + kx = 0$ aus § 1, Nr. 1 leicht in der Form einer Exponentialfunktion erhalten, indem man versucht, eine Konstante $\lambda$ so zu bestimmen, daß der Ausdruck $e^{\lambda t} = x$ eine Lösung wird. Setzt man diesen Ausdruck und die daraus folgenden $\dot{x} = \lambda e^{\lambda t}$, $\ddot{x} = \lambda^2 e^{\lambda t}$ in die Differentialgleichung ein und läßt man dann den überall auftretenden Faktor $e^{\lambda t}$ fort, so entsteht für $\lambda$ die quadratische Gleichung

$$m\lambda^2 + r\lambda + k = 0,$$

deren beide Wurzeln durch

$$\lambda_1 = -\frac{r}{2m} + \frac{1}{2m}\sqrt{r^2 - 4mk}, \quad \lambda_2 = -\frac{r}{2m} - \frac{1}{2m}\sqrt{r^2 - 4mk}$$

gegeben werden. Jeder der beiden Ausdrücke $x = e^{\lambda_1 t}$ und $x = e^{\lambda_2 t}$ ist sodann zumindest formal eine spezielle Lösung der Differentialgleichung, wie man erkennt, wenn man die letzte Rechnung rückwärts durchgeht. Es sind nun drei verschiedene Fälle möglich:

1. Es ist $r^2 - 4mk > 0$. Dann sind die beiden Wurzeln $\lambda_1$ und $\lambda_2$ reell, negativ und voneinander verschieden, und wir erhalten zunächst zwei Lösungen $x = u_1 = e^{\lambda_1 t}$ und $x = u_2 = e^{\lambda_2 t}$ der Differentialgleichung. Mit Hilfe dieser beiden Lösungen kann man sich nun sofort eine Lösung verschaffen, in welcher zwei willkürliche Konstanten $c_1$ und $c_2$ vorkommen. Man erkennt nämlich durch Differenzieren sofort, daß auch

$$x = c_1 u_1 + c_2 u_2$$

eine Lösung der Differentialgleichung ist, und wir werden in Nr. 3 sogleich weiter beweisen, daß dieser Ausdruck die allgemeinste Lösung unserer Differentialgleichung darstellt, d. h. daß wir jede Lösung erhalten, indem wir für $c_1$ und $c_2$ geeignete Zahlenwerte einsetzen.

2. Ist $r^2 - 4mk = 0$, so hat unsere quadratische Gleichung eine Doppelwurzel. Es ergibt sich also zunächst bis auf einen konstanten Faktor nur die eine Lösung $x = w_1 = e^{-\frac{r}{2m}t}$. Man bestätigt aber sehr leicht, daß in diesem Falle auch noch die Funktion

$$x = w_2 = t e^{-\frac{r}{2m}t}$$

eine Lösung der Differentialgleichung ist[1]). Es wird nämlich für diesen Ausdruck $\dot{x} = \left(1 - \frac{r}{2m}t\right)e^{-\frac{r}{2m}t}$, $\ddot{x} = \left(\frac{r^2}{4m^2}t - \frac{r}{m}\right)e^{-\frac{r}{2m}t}$, woraus wir durch Einsetzen sofort finden, daß die Differentialgleichung

$$m\ddot{x} + r\dot{x} + \frac{r^2}{4m}x = m\ddot{x} + r\dot{x} + kx = 0$$

erfüllt ist. Wir gewinnen jetzt in dem Ausdruck

$$x = c_1 e^{-\frac{r}{2m}t} + c_2 t e^{-\frac{r}{2m}t}$$

wiederum eine Lösung der homogenen Differentialgleichung mit zwei willkürlichen Integrationskonstanten $c_1$ und $c_2$.

3. Ist endlich $r^2 - 4mk < 0$, so setzen wir $r^2 - 4mk = -4m^2\nu^2$ und erhalten nunmehr zunächst zwei Lösungen der Differentialgleichung in komplexer Form, nämlich die Ausdrücke $x = u_1 = e^{-\frac{r}{2m}t + i\nu t}$ und $x = u_2 = e^{-\frac{r}{2m}t - i\nu t}$. Die Zerlegungen

$$e^{\pm i\nu t} = \cos\nu t \pm i\sin\nu t$$

geben uns für den reellen und den imaginären Teil der komplexen Lösung $u_1$ einerseits die Ausdrücke

$$v_1 = e^{-\frac{r}{2m}t}\cos\nu t, \qquad v_2 = e^{-\frac{r}{2m}t}\sin\nu t,$$

andrerseits die Darstellungen

$$v_1 = \frac{u_1 + u_2}{2}, \qquad v_2 = \frac{u_1 - u_2}{2i}.$$

An der zweiten Darstellungsweise sieht man, daß $v_1$ und $v_2$ (reelle)

[1]) Auf diese Lösung wird man ganz von selbst durch folgenden Grenzübergang geführt: Ist $\lambda_1 \neq \lambda_2$, so stellt auch speziell der Ausdruck $\frac{e^{\lambda_1 t} - e^{\lambda_2 t}}{\lambda_1 - \lambda_2}$ eine Lösung dar. Lassen wir nun $\lambda_1$ gegen $\lambda_2$ rücken und schreiben $\lambda$ statt $\lambda_1, \lambda_2$, so geht unser Ausdruck in $\frac{d}{d\lambda}e^{\lambda t} = te^{\lambda t}$ über.

Lösungen der Differentialgleichung sind. Es ist eine nützliche kleine Übungsaufgabe, dies direkt ohne Bezugnahme auf das Komplexe durch Differenzieren und Einsetzen zu bestätigen.

Wiederum können wir aus unseren beiden speziellen Lösungen mit zwei willkürlichen Konstanten $c_1$ und $c_2$ eine allgemeine Lösung

$$x = c_1 v_1 + c_2 v_2 = (c_1 \cos \nu t + c_2 \sin \nu t)\, e^{-\frac{r}{2m} t}$$

zusammensetzen, die wir auch in die Form

$$x = a \cos \nu (t - \delta) \cdot e^{-\frac{r}{2m} t}$$

schreiben können, wo $c_1 = a \cos \nu \delta$, $c_2 = a \sin \nu \delta$ gesetzt ist und $a$, $\delta$ zwei neue Konstanten bedeuten.

Ich erinnere daran, daß uns diese Lösung für den speziellen Fall $r = 0$ schon aus dem fünften Kapitel, § 4, bekannt ist.

## 2. Physikalische Deutung der Lösung.

In den beiden Fällen $r > 2\sqrt{mk}$ und $r = 2\sqrt{mk}$ wird die Lösung durch Exponentialkurven oder durch die der Exponentialkurve bei großem $t$ ähnliche Kurve der Funktion $t e^{-\frac{r}{2m} t}$ bzw. durch Superposition solcher Kurven dargestellt. In diesen Fällen haben wir es mit *aperiodisch* abklingenden Vorgängen zu tun, d. h. mit Vorgängen, bei denen mit wachsender Zeit die „Elongation" $x$ sich asymptotisch dem Werte 0 nähert, ohne daß Oszillationen um den Wert $x = 0$ herum stattfinden. Es ist also von Schwingungen keine Rede; der Einfluß der Reibung $r$ oder die *Dämpfung* ist so groß geworden, daß ihr gegenüber die elastischen Kräfte nicht mehr schwingende Bewegungen durchsetzen können.

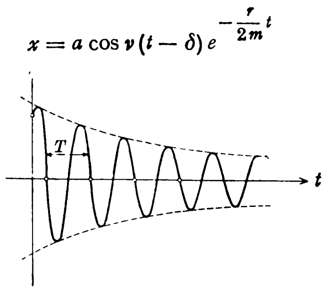



Fig. 125. Gedämpfte harmonische Schwingungen.

Ganz anders ist es im dritten Fall bei $r < 2\sqrt{mk}$, wo die Dämpfung so klein ist, daß komplexe Wurzeln $\lambda_1$, $\lambda_2$ auftreten. Wir erhalten hier durch den Ausdruck $x = a \cos \nu (t - \delta)\, e^{-\frac{r}{2m} t}$ sog. *gedämpfte harmonische Schwingungen*; d. h. Schwingungen, welche nach einem Sinusgesetz verlaufen, deren Frequenz[1] $\nu = \sqrt{\frac{k}{m} - \frac{r^2}{4m^2}}$ ist, deren „Amplitude" jedoch nicht wie bei den reinen Schwingungen konstant bleibt, sondern durch den Ausdruck $a e^{-\frac{r}{2m} t}$ gegeben ist, d. h. exponentiell *abklingt*, und zwar um so schneller, je größer der „Dämpfungsfaktor" $\frac{r}{2m}$ ist. Diesen

[1]) Unter Frequenz ist überall wie in Kap. IX die Kreisfrequenz verstanden.

Dämpfungsfaktor nennt man in der Physik häufig auch das *logarithmische Dekrement* der gedämpften Schwingung, womit man andeuten will, daß der Logarithmus der Amplitude mit der Geschwindigkeit $\frac{r}{2m}$ abnimmt. In Figur 125 ist der Verlauf solcher gedämpfter Schwingungen gezeichnet. Wie früher nennen wir die Größe $T = \frac{2\pi}{\nu}$ die Schwingungsdauer und die Größe $\nu\delta$ die Phasenverschiebung der Schwingung. Für den speziellen Fall $r = 0$ erhalten wir die schon früher behandelten reinen Schwingungen mit der Frequenz $\nu_0 = \sqrt{\frac{k}{m}}$, der *Eigenfrequenz des ungedämpft schwingenden Systemes.*

### 3. Anpassung an gegebene Anfangsbedingungen. Eindeutigkeit der Lösung.

Ich habe noch den Beweis zu erbringen, daß wir die gefundene, zwei Konstanten $c_1$ und $c_2$ enthaltende Lösung wirklich auch jedem vorgegebenen Anfangszustand anpassen können, und ferner, daß wir in ihr die Gesamtheit aller Lösungen gewonnen haben. Verlangen wir, daß eine Lösung gefunden werden soll, für welche zur Zeit $t = 0$ die Anfangswerte $x(0) = x_0$, $\dot{x}(0) = \dot{x}_0$ vorliegen, wobei die Werte $x_0$ und $\dot{x}_0$ beliebig vorgegeben sind, dann haben wir in dem Falle 1 aus Nr. 1 zu setzen

$$c_1 + c_2 = x_0,$$
$$c_1 \lambda_1 + c_2 \lambda_2 = \dot{x}_0.$$

Wir erhalten also für die beiden Konstanten $c_1$ und $c_2$ zwei lineare Gleichungen, die sich sofort in eindeutiger Weise auflösen lassen:

$$c_1 = \frac{\dot{x}_0 - \lambda_2 x_0}{\lambda_1 - \lambda_2}, \qquad c_2 = \frac{\dot{x}_0 - \lambda_1 x_0}{\lambda_2 - \lambda_1}.$$

Im Falle 2 ergibt dasselbe Verfahren die beiden linearen Gleichungen

$$c_1 = x_0,$$
$$\lambda c_1 + c_2 = \dot{x}_0 \qquad \left(\lambda = -\frac{r}{2m}\right),$$

aus denen sich ebenfalls $c_1$ und $c_2$ eindeutig bestimmen. Im Falle 3 endlich nehmen die Gleichungen zur Bestimmung von $a$ und $\delta$ die Gestalt

$$a \cos \nu\delta = x_0,$$
$$a\left(\nu \sin \nu\delta - \frac{r}{2m} \cos \nu\delta\right) = \dot{x}_0$$

an und lassen sich durch die Formeln

$$\delta = \frac{1}{\nu} \operatorname{arc\,cos} \frac{x_0}{a}, \qquad a = \frac{1}{\nu} \sqrt{\nu^2 x_0^2 + \left(\dot{x}_0 + \frac{r}{2m} x_0\right)^2}$$

auflösen.

Hiermit ist tatsächlich gezeigt, daß die gewonnenen allgemeinen Lösungen jedem Anfangszustand angepaßt werden können. Es bleibt zu zeigen, daß es keine anderen Lösungen geben kann, und hierzu genügt der Nachweis, daß für denselben Anfangszustand niemals zwei verschiedene Lösungen bestehen können.

Gäbe es zwei solche Lösungen $u(t)$ und $v(t)$, für welche $u(0) = x_0$, $\dot{u}(0) = \dot{x}_0$ und $v(0) = x_0$, $\dot{v}(0) = \dot{x}_0$ ist, so würde die Differenz $w = u - v$ wiederum eine Lösung der Differentialgleichung sein, und zwar wäre $w(0) = 0$, $\dot{w}(0) = 0$. Diese Lösung würde also einem anfänglichen Ruhezustand entsprechen, d. h. einem Zustand, in welchem zur Zeit $t = 0$ der Punkt sich mit der Geschwindigkeit 0 in seiner Ruhelage befindet. Es ist zu zeigen, daß er sich niemals in Bewegung setzen kann. Zu diesem Zwecke multiplizieren wir die Differentialgleichung $m\ddot{w} + r\dot{w} + kw = 0$ mit $2\dot{w}$, beachten, daß $2\dot{w}\ddot{w} = \frac{d}{dt}\dot{w}^2$, $2w\dot{w} = \frac{d}{dt}w^2$ ist, und erhalten daher

$$\frac{d}{dt}(m\dot{w}^2) + \frac{d}{dt}(kw^2) + 2r\dot{w}^2 = 0.$$

Integrieren wir jetzt zwischen den Zeitmomenten $t = 0$ und $t = \tau$ und berücksichtigen dabei die Anfangsbedingungen $w(0) = 0$, $\dot{w}(0) = 0$, so erhalten wir

$$m\dot{w}^2(\tau) + kw^2(\tau) + 2r\int_0^\tau \left(\frac{dw}{dt}\right)^2 dt = 0.$$

Diese Gleichung bedeutet aber einen Widerspruch, sobald $w$ für irgend ein $\tau > 0$ von Null verschieden ist; denn dann würde sicherlich links eine positive Größe stehen, da wir $m$, $k$ und $r$ als positiv angenommen haben, während die rechte Seite der Gleichung Null ist. Also ist ständig $w = 0$, womit wir unseren Beweis erbracht haben.

## § 3. Unhomogene Gleichung. Erzwungene Bewegungen.

### 1. Allgemeine Bemerkungen.

Der Lösung unseres Problemes bei Vorhandensein einer äußeren Kraft $f(t)$, d. h. der Integration der unhomogenen Differentialgleichung schicke ich folgende Bemerkung voraus:

Wenn $w$ und $v$ zwei Lösungen der unhomogenen Differentialgleichung sind, so genügt die Differenz $u = w - v$ der homogenen Differentialgleichung, wie man sofort durch Einsetzen in die Differentialgleichung erkennt. Umgekehrt, wenn $u$ eine Lösung der homogenen, $v$ eine Lösung der unhomogenen Differentialgleichung ist, so genügt auch $w = u + v$ der unhomogenen Differentialgleichung. Man erhält also aus einer einzigen Lösung der unhomogenen Differentialgleichung sämtliche Lösungen der unhomogenen, indem man zu ihr die allgemeinste Lösung der homogenen Differentialgleichung addiert, und es kommt daher für uns nur noch darauf an, eine einzige Lösung der unhomogenen Differentialgleichung zu finden. Physikalisch bedeutet dies: Überlagert sich einem durch eine äußere Kraft erzwungenen Vorgang ein beliebiger freier Vorgang, repräsentiert durch eine additiv hinzutretende Lösung der homogenen Gleichung, so entsteht ein Vorgang, welcher derselben unhomogenen Gleichung genügt wie der ursprüngliche. — Die freie Bewegung wird stets bei Vorhandensein von Reibung (z. B. wegen des Dämpfungsfaktors $e^{-\frac{r}{2m}t}$) mit der Zeit abklingen; es wird sich also bei einer gegebenen erzwungenen Bewegung mit Reibung, gleichgültig, was für freie Bewegungen sich überlagern, mit der Zeit mehr und mehr derselbe Endzustand herausbilden.

Zweitens bemerke ich, daß man die Wirkung einer Kraft $f(t)$ in derselben Weise zerlegen kann wie die Kraft selbst. Damit ist folgendes gemeint: Sind $f_1(t)$, $f_2(t)$ und $f(t)$ drei Funktionen, für welche

$$f_1(t) + f_2(t) = f(t)$$

gilt, ist ferner $x_1 = x_1(t)$ eine Lösung der Differentialgleichung $m\ddot{x} + r\dot{x} + kx = f_1(t)$ und $x_2 = x_2(t)$ eine Lösung der Differentialgleichung $m\ddot{x} + r\dot{x} + kx = f_2(t)$, so ist $x(t) = x_1(t) + x_2(t)$ eine Lösung der Differentialgleichung $m\ddot{x} + r\dot{x} + kx = f(t)$. Genau das Entsprechende gilt natürlich bei der Zusammensetzung der Funktion $f(t)$ aus einer beliebigen Anzahl von Summanden. Der Beweis dieser einfachen, aber wichtigen, als *„Superpositionsprinzip"* bezeichneten Tatsache ergibt sich durch einen Blick auf die Gleichungen von selbst. Wir gewinnen damit die Freiheit, durch Zerlegung einer Funktion $f(t)$ in zwei oder mehr Summanden die Differentialgleichung in andere zu spalten, die wir unter Umständen viel leichter behandeln können.

Der wichtigste Fall ist der einer periodischen äußeren Kraft $f(t)$. Eine solche periodische äußere Kraftfunktion $f(t)$ können wir nämlich nach dem vorigen Kapitel in reine harmonische periodische Funktionen auflösen, d. h. sie durch eine Summe endlich vieler solcher Funktionen beliebig genau annähern, und es genügt daher für uns, die Lösung unserer Differentialgleichung unter der Voraussetzung durchzuführen, daß die rechte Seite die Form hat

$$a \cos \omega t \quad \text{oder} \quad b \sin \omega t,$$

wo $a$, $b$ und $\omega$ irgendwelche Konstanten sind.

Anstatt mit diesen trigonometrischen Funktionen zu rechnen, können wir einfacher und übersichtlicher zu den Lösungen unserer Differentialgleichungen gelangen, wenn wir uns der komplexen Schreibweise bedienen. Setzen wir $f(t) = c\,e^{i\omega t}$, so werden wir durch unsere obige Bemerkung darauf geführt, die Differentialgleichung

$$m\ddot{x} + r\dot{x} + kx = c\,e^{i\omega t}$$

zu behandeln, wobei wir unter $c$ eine beliebige reelle oder komplexe Konstante verstehen. Der Sinn einer solchen Differentialgleichung ist der, daß sie zwei reelle Differentialgleichungen repräsentiert. Zerlegen wir nämlich die rechte Seite in zwei Summanden, nehmen wir etwa $c = 1$ und schreiben $e^{i\omega t} = \cos\omega t + i \sin\omega t$, so werden sich die Lösungen $x_1$ und $x_2$ der beiden reellen Differentialgleichungen $m\ddot{x} + r\dot{x} + kx = \cos\omega t$ und $m\ddot{x} + r\dot{x} + kx = \sin\omega t$ zu der Lösung $x = x_1 + i x_2$ unserer komplexen Differentialgleichung verbinden; umgekehrt erhalten wir, wenn wir zunächst die Differentialgleichung in der komplexen Form lösen, in dem reellen Teil dieser Lösung die Funktion $x_1$, im imaginären Teil die Funktion $x_2$.

## 2. Lösung der unhomogenen Gleichung.

Unsere Gleichung $m\ddot{x} + r\dot{x} + kx = c e^{i\omega t}$ lösen wir wiederum durch einen naheliegenden Ansatz, wobei wir $c$ als reell und zunächst $r \neq 0$ voraussetzen. Wir dürfen nämlich vermuten, daß eine Bewegung existieren wird, die der periodischen äußern Kraft in demselben Rhythmus folgt, und wir werden demgemäß versuchen, der Differentialgleichung durch eine Funktion der Form

$$x = \sigma e^{i\omega t}$$

zu genügen, wobei lediglich der von der Zeit nicht abhängige Faktor $\sigma$ zu bestimmen ist. Gehen wir mit diesem Ansatz in die Differentialgleichung ein, so ergibt sich wegen $\dot{x} = i\omega\sigma e^{i\omega t}$, $\ddot{x} = -\omega^2 \sigma e^{i\omega t}$ nach Unterdrückung des Faktors $e^{i\omega t}$ die Gleichung

$$-m\omega^2\sigma + i r\omega\sigma + k\sigma = c$$

oder

$$\sigma = \frac{c}{-m\omega^2 + i r\omega + k},$$

und umgekehrt sehen wir, daß für diesen Wert der Konstanten $\sigma$ tatsächlich der Ausdruck $\sigma e^{i\omega t}$ eine Lösung der unhomogenen Differentialgleichung ist. Wir haben also eine solche Lösung in komplexer Form gefunden. Um jedoch den Sinn unseres Resultates klarzustellen, müssen wir einige Umformungen vornehmen.

Zunächst schreiben wir den komplexen Faktor $\sigma$ in der Form

$$\sigma = c\frac{k - m\omega^2 - i r\omega}{(k - m\omega^2)^2 + r^2\omega^2} = c\alpha e^{-i\omega\delta},$$

wobei der positive „*Verzerrungsfaktor*" $\alpha$ und die „*Phasenverschiebung*" $\omega\delta$ durch die Gleichungen

$$\alpha^2 = \frac{1}{(k - m\omega^2)^2 + r^2\omega^2}, \quad \sin\omega\delta = \frac{r\omega}{\alpha}, \quad \cos\omega\delta = \frac{k - m\omega^2}{\alpha}$$

aus den gegebenen Größen $m, r, k, \omega$ gewonnen werden. Mit diesen Ausdrücken nimmt unsere Lösung die Form

$$x = c\alpha e^{i\omega(t-\delta)}$$

an, und die Bedeutung unseres Ergebnisses ist die folgende: Der Kraft $c \cos \omega t$ entspricht die „Wirkung“ $c\,\alpha \cos \omega (t - \delta)$, der Kraft $c \sin \omega t$ entspricht die Wirkung $c\,\alpha \sin \omega (t - \delta)$.

Das Ergebnis ist also: Die Wirkung ist eine Funktion genau derselben Art wie die Kraft, d. h. eine ungedämpfte Schwingung. Diese Schwingung unterscheidet sich von der die Kraft repräsentierenden Schwingung durch eine Verzerrung der Amplitude mit dem Verzerrungsfaktor $\alpha$ und ferner durch eine Verschiebung der Phase um den Winkel $\omega \delta$. Es ist natürlich leicht, das Ergebnis ohne Benutzung der komplexen Zahlen mit etwas mehr Aufwand an Rechnung zu gewinnen.

Nach unseren Bemerkungen zu Beginn des Paragraphen haben wir mit der Gewinnung einer Lösung unser Problem vollständig gelöst, da wir aus ihr durch Überlagerung einer freien Bewegung den allgemeinsten erzwungenen Bewegungszustand erhalten.

Ich fasse das Resultat noch einmal zusammen: *Die allgemeinste Lösung der Differentialgleichung* $m \ddot{x} + r \dot{x} + k x = c e^{i \omega t}$ *lautet*

$$x = c\,\alpha\, e^{i \omega (t - \delta)} + u,$$

*wobei* $u$ *die allgemeinste Lösung der homogenen Differentialgleichung* $m \ddot{x} + r \dot{x} + k x = 0$ *ist und die Größen* $\alpha$ *und* $\delta$ *durch die Gleichungen*

$$\alpha^2 = \frac{1}{(k - m \omega^2)^2 + r^2 \omega^2}, \quad \sin \omega \delta = \frac{r \omega}{\alpha}, \quad \cos \omega \delta = \frac{k - m \omega^2}{\alpha}$$

*definiert werden.* Die Konstanten $c_1$ und $c_2$ in dieser allgemeinen Lösung geben die Möglichkeit, die Lösung einem willkürlich vorgegebenen Anfangszustand anzupassen, d. h. zu erreichen, daß bei willkürlich vorgegebenen Werten von $x_0$ und $\dot{x}_0$ die Gleichungen $x(0) = x_0$ und $\dot{x}(0) = \dot{x}_0$ bestehen.

### 3. Die Resonanzkurve.

Um uns von der gewonnenen Lösung und ihrer Bedeutung in den Anwendungen eine anschauliche Vorstellung zu machen, studieren wir den Verzerrungsfaktor $\alpha$ als Funktion der „erregenden Frequenz“ $\omega$, d. h. die Funktion

$$\varphi(\omega) = \frac{1}{\sqrt{(k - m \omega^2)^2 + r^2 \omega^2}}.$$

Das Motiv zu einer solchen näheren Untersuchung liegt in der Tatsache, daß man sich bei gegebenen konstanten $k$, $m$, $r$ oder, wie wir sagen, bei einem gegebenen „schwingungsfähigen System“ periodische „erregende Kräfte“ $e^{i \omega t}$ von sehr verschiedenartigen Frequenzen $\omega$ vorgegeben denken kann und daß es wichtig ist, bei solchen verschiedenen erregenden Kräften die Lösung der Differentialgleichung zu übersehen. Um den Verlauf unserer Funktion bequem zu beschreiben, führe ich zur Abkürzung die Größe $\omega_0 = \sqrt{\frac{k}{m}}$ ein; $\omega_0$ ist die Schwingungszahl oder Frequenz, die unser System bei freien Schwingungen haben würde, wenn die Reibung $r$ gleich Null wäre, kurz gesagt, die *„Eigenschwingungszahl des ungedämpften Systemes“* (vgl. S. 393). Die wirkliche Schwingungszahl des freien Systemes ist wegen des Vorhandenseins der Reibung $r$ nicht $\omega_0$, sondern

$$\nu = \sqrt{\frac{k}{m} - \frac{r^2}{4m^2}},$$

wobei wir voraussetzen, daß $4km - r^2 > 0$ ist — andernfalls hat das freie System keine Schwingungszahl, sondern ist aperiodisch.

Die Funktion $\varphi(\omega)$ wird sich asymptotisch dem Werte 0 nähern, wenn die erregende Frequenz $\omega$ gegen Unendlich strebt, und zwar wird sie von der Größenordnung $\frac{1}{\omega^2}$ verschwinden. Es ist ferner $\varphi(0) = \frac{1}{k}$; mit anderen Worten: bei einer erregenden Kraft von der Frequenz Null und der Amplitude 1, d. h. einer konstanten Kraft der Größe 1 wird ein bestimmter Ausschlag $\alpha$ des schwingenden Systems von der Größe $\frac{1}{k}$ erzeugt werden. Die Ableitung $\varphi'(\omega)$ kann im Bereich positiver $\omega$ nur dort verschwinden, wo die Ableitung des Ausdruckes $(k - m\omega^2)^2 + r^2\omega^2$ verschwindet, d. h. an einer Stelle $\omega = \omega_1 > 0$, für welche $-4m\omega(k - m\omega^2) + 2r^2\omega = 0$ gilt. Damit eine solche Stelle vorhanden ist, muß offenbar $2km - r^2 > 0$ sein; alsdann ist

$$\omega_1 = \sqrt{\frac{k}{m} - \frac{r^2}{2m^2}} = \sqrt{\omega_0^2 - \frac{r^2}{2m^2}}.$$

Diese Stelle wird, da die Funktion $\varphi(\omega)$ für $\omega > 0$ überall positiv ist, für kleine positive Werte von $\omega$ monoton wächst und im Unendlichen verschwindet, ein Maximum der Funktion darstellen. Wir nennen die Frequenz $\omega_1$ die *„Resonanzfrequenz“* unseres Systems.

Der Wert des Maximums ist, wie wir durch Einsetzen des Ausdruckes von $\omega_1$ finden,

$$\varphi(\omega_1) = \frac{1}{r\sqrt{\frac{k}{m} - \frac{r^2}{4m^2}}}.$$

Dieser Wert wächst bei $r \to 0$ über alle Grenzen. Für $r = 0$, also für ein ungedämpft schwingendes System würde die Funktion $\varphi(\omega)$ an der Stelle $\omega = \omega_1$ einen Unendlichkeitspunkt besitzen, ein Grenzfall, den wir nachher noch gesondert zu betrachten haben.

Die Kurve für die Funktion $\varphi(\omega)$ heißt die *Resonanzkurve* des schwingungsfähigen Systemes. Die Tatsache, daß für $\omega = \omega_1$ — also bei kleinem $r$ in der Nähe der Eigenschwingung $\omega_0$ des Systemes — die Amplitudenverzerrung $\alpha = \varphi(\omega)$ besonders groß ist, stellt mathematisch das

*„Resonanzphänomen"* dar, welches bei gleichbleibenden Werten von $m$ und $k$ um so ausgeprägter in die Erscheinung tritt, je kleiner die Reibungskonstante $r$ ist.

In Figur 126 ist unter der Annahme $m = 1$ und $k = 1$, also auch $\omega_0 = 1$ eine Schar von Resonanzkurven gezeichnet, die sich alle nur durch den verschiedenen Wert von $D = \frac{r}{2}$ unterscheiden. Man sieht, daß bei kleinem $D$ sehr stark ausgeprägte Resonanz nahe bei $\omega = 1$ herrscht — im Grenzfall

$$D = r = 0$$

würde für $\omega = 1$ eine Unendlichkeitsstelle statt eines Maximums von $\varphi(\omega)$ auftreten —, daß bei wachsendem $D$ die Maxima nach links rücken und daß für

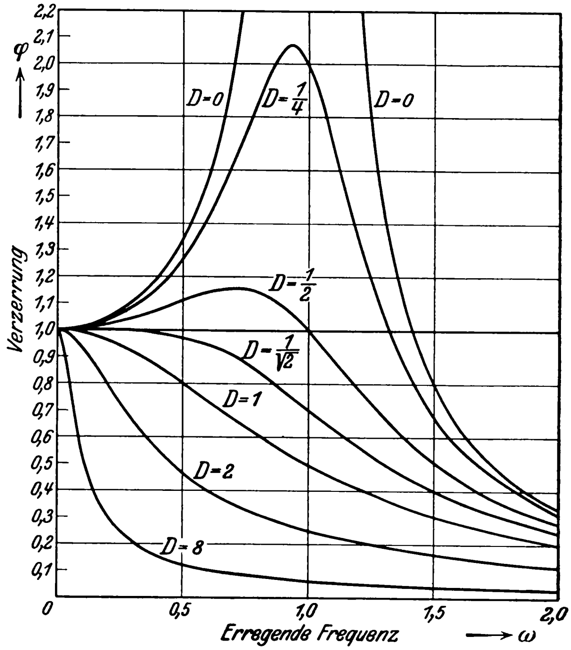


Fig. 126. Resonanzkurven.

den Wert $D = \frac{1}{\sqrt{2}}$ gerade $\omega_1 = 0$ wird. Die Stelle mit horizontaler Tangente ist im letzteren Falle in den Nullpunkt gerückt, das Maximum ist ganz verschwunden. Für $D > \frac{1}{\sqrt{2}}$ gibt es keine Nullstelle von $\varphi'(\omega)$, die Resonanzkurve besitzt kein Maximum mehr, es tritt keine Resonanz ein.

Allgemein hört jede Resonanzerscheinung auf, sobald die Bedingung

$$2km - r^2 \leqq 0$$

besteht. Gilt das Gleichheitszeichen, so setzt die Resonanzkurve für $\omega_1 = 0$ in der Höhe $\varphi(0) = \frac{1}{k}$ mit horizontaler Tangente ein und fällt nach anfänglichem annähernd horizontalem Verlauf nach Null ab.

## 4. Nähere Diskussion des Schwingungsablaufes.

Mit der gegebenen Diskussion können wir uns noch nicht begnügen. Es muß für das wirkliche Verständnis des Vorganges der erzwungenen Bewegung noch ein Punkt hervorgehoben werden. Die spezielle Lösung $c\alpha e^{i\omega(t-\delta)}$ ist als Grenzzustand anzusehen, der sich aus der allgemeinen Lösung $x(t) = c\alpha e^{i\omega(t-\delta)} + c_1 u_1 + c_2 u_2$ im Laufe der Zeit immer genauer ergibt, indem die der speziellen erzwungenen Bewegung sich überlagernde freie Bewegung $c_1 u_1 + c_2 u_2$ mit der Zeit abklingt. Dieser Abklingungsvorgang wird um so langsamer vor sich gehen, je kleiner $r$ ist, um so schneller, je größer $r$ ist.

Nehmen wir z. B. an, es sei zu Beginn der Bewegung, d. h. für $t = 0$ das System in Ruhe, d. h. es sei $x(0) = 0$ und $\dot{x}(0) = 0$. Hieraus bestimmen sich die beiden Konstanten $c_1$ und $c_2$. Auch wenn die erregende Frequenz $\omega$ annähernd oder genau gleich $\omega_1$ ist, wenn wir also Resonanz haben, wird nicht etwa von Anfang an die verhältnismäßig große Amplitude $\alpha = \varphi(\omega_1)$ zum Vorschein kommen, sie wird vielmehr durch die Funktion $c_1 u_1 + c_2 u_2$ verdeckt werden und erst mit dem Abklingen dieser Funktion in Erscheinung treten, also um so langsamer, je kleiner $r$ ist.

Im Falle des ungedämpften Systemes, also für $r = 0$, versagt unsere Lösung überhaupt, falls die erregende Frequenz $\omega$ gleich der Eigenfrequenz $\omega_0 = \sqrt{\frac{k}{m}}$ ist, weil dann $\varphi(\omega_0)$ unendlich wird. Wir erhalten also keine Lösung der Gleichung $m\ddot{x} + kx = e^{i\omega t}$ in der Form $\sigma e^{i\omega t}$. Trotzdem können wir sofort eine Lösung unserer Gleichung in der Form $x = \sigma t e^{i\omega t}$ angeben. Setzen wir diesen Ausdruck in die Differentialgleichung ein, so finden wir wegen

$$\dot{x} = \sigma e^{i\omega t}(1 + i\omega t), \qquad \ddot{x} = \sigma e^{i\omega t}(2i\omega - t\omega^2)$$

sogleich

$$\sigma(2im\omega - m\omega^2 t + kt) = 1,$$

also wegen $m\omega^2 = k$

$$\sigma = \frac{1}{2im\omega}.$$

Wir gewinnen so im Falle der Resonanz bei dem ungedämpften System die Lösung

$$x = \frac{t}{2im\omega} e^{i\omega t} = \frac{t}{2i\sqrt{km}} e^{i\omega t},$$

oder in reeller Schreibweise: für $f(t) = \cos\omega t$ ergibt sich $x = \frac{t}{2\sqrt{km}} \sin\omega t$,

für $f(t) = \sin\omega t$ ergibt sich $x = -\frac{t}{2\sqrt{km}} \cos\omega t$.

Wir sehen, daß wir es hier mit einer Funktion zu tun haben, die ebenfalls als Schwingung bezeichnet werden kann, deren Amplitude aber proportional mit der Zeit wächst. Die überlagerte freie Schwingung klingt zwar jetzt überhaupt nicht mehr ab, da sie ungedämpft ist; aber sie behält ihre anfängliche Amplitude stets bei und kann das Anwachsen der Amplitude der speziellen erzwungenen Bewegung nicht hemmen. Die Tatsache also, daß die Lösung in diesem Falle bei wachsendem $t$ zwischen absolut genommen immer größer werdenden positiven und negativen Grenzen hin- und herpendelt, ist der eigentliche Sinn des Unendlichwerdens unserer Resonanzfunktion im Falle eines ungedämpften Systems.

### 5. Bemerkungen über den Bau von Registrierinstrumenten.

Die Bedeutung der Diskussion aus Nr. 4 für die verschiedenartigsten Anwendungen in Physik und Technik ist außerordentlich groß. Bei vielen Instrumenten, wie Galvanometern, Seismographen, elektrischen Schwingungskreisen in Radioempfangsapparaten, Mikrophonmembranen usw., kommt es darauf an, einen Schwingungsausschlag $x$ oder analoge sonstige Größen als die Wirkung äußerer periodischer Kräfte zu registrieren. Die Größe $x$ genügt dabei, wenigstens in erster Annäherung, unserer Schwingungsdifferentialgleichung. Ist $T$ die Schwingungsdauer der äußeren periodischen Kraft, so können wir uns diese nach dem vorigen Kapitel in eine Fouriersche Reihe der Form

$$f(t) = \sum_{l=-\infty}^{\infty} \gamma_l e^{il\frac{2\pi}{T}t}$$

entwickelt oder besser durch einen nur aus endlich vielen Gliedern bestehenden trigonometrischen Ausdruck $\sum_{l=-N}^{N} \gamma_l e^{il\frac{2\pi}{T}t}$ mit genügender Annäherung dargestellt denken. Nach unserem Superpositionsprinzip von S. 395 werden wir nun die Lösung $x$ unserer Differentialgleichung, abgesehen von der überlagerten freien Bewegung, durch eine unendliche Reihe[1]) der Form

$$x(t) = \sum_{l=-\infty}^{\infty} \sigma_l e^{il\frac{2\pi}{T}t}$$

oder angenähert durch einen endlichen Ausdruck der Form

$$x(t) = \sum_{l=-N}^{N} \sigma_l e^{il\frac{2\pi}{T}t}$$

darstellen können. Hierbei ist auf Grund unseres allgemeinen Ergebnisses

$$\sigma_l = \gamma_l \alpha_l e^{-i\delta_l \frac{2\pi l}{T}}$$

und

$$\alpha_l^2 = \frac{1}{\left(k - m\,l^2 \frac{4\pi^2}{T^2}\right)^2 + r^2 l^2 \frac{4\pi^2}{T^2}}, \qquad \operatorname{tg} \frac{2\pi l}{T}\delta_l = \frac{2\pi l r}{T\left(k - m\frac{4\pi^2 l^2}{T^2}\right)}.$$

[1]) Von Konvergenzbetrachtungen sehe ich hier ab.

Wir können also die Wirkung einer beliebigen periodischen äußeren Erregung auf unser System so beschreiben: Zerlegt man die Erregung in rein periodische Teilerregungen, die einzelnen Glieder unserer Fourierschen Reihe, so erfährt jede Teilerregung eine besondere Amplitudenverzerrung und Phasenverschiebung, und die einzelnen Wirkungen setzen sich dann wieder additiv zusammen. Wenn es uns hauptsächlich auf die Amplitudenverzerrungen ankommt — die Phasenverschiebungen spielen in den Anwendungen keine so wichtige Rolle[1]) und können im übrigen ganz ähnlich wie die Amplitudenverzerrungen diskutiert werden —, so erhalten wir aus dem Studium unserer Resonanzfunktion oder Resonanzkurve ein vollständiges Bild davon, in welcher Art der Registrierapparat durch seine Ausschläge $x$ die äußere Erregung wiedergibt. Für sehr große Werte von $l$ oder $\omega = \frac{2\pi}{T} l$ wird von den erregenden Frequenzen in der Elongation $x$ kaum noch etwas zu bemerken sein; dagegen werden alle in der Nähe der Resonanzfrequenz $\omega_1$ liegenden erregenden Frequenzen sehr stark in der Größe $x$ zum Vorschein kommen.

Bei einem physikalischen Meß- und Registrierapparat hat man die Konstanten $m, r, k$ in weiten Grenzen frei verfügbar. Man wird versuchen, sie so zu wählen, daß die Gestalt der Resonanzkurve den Bedürfnissen der Messung möglichst gut angepaßt ist. Dabei sind gewöhnlich zwei Gesichtspunkte maßgebend. Einmal wird man eine große Empfindlichkeit des Apparates wünschen, d. h., mathematisch gesprochen, einen großen Wert von $\alpha$ für die in Frage kommenden Frequenzen $\omega$ der Anregung. Für kleine Werte von $\omega$ wird angenähert $\alpha$ proportional $\frac{1}{k}$, so daß wir in der Zahl $\frac{1}{k}$ ein Maß für die Empfindlichkeit des Instrumentes bei kleinen Erregungsfrequenzen haben. Steigerung der Empfindlichkeit ist also durch Vergrößerurg von $\frac{1}{k}$, d. h. Verkleinerung der elastischen Bindung möglich.

Zweitens spielt eine wesentliche Rolle die Forderung der relativen *Verzerrungsfreiheit*. Nehmen wir an, wir erhalten für unsere erregende Frequenz durch die Darstellung $f(t) = \sum_{l=-N}^{N} \gamma_l e^{il\frac{2\pi}{T}t}$ eine genügende Annäherung, so werden wir sagen, daß unser Apparat die Erregung $f(t)$ ohne relative Verzerrung registriert, wenn für alle Frequenzen $\omega \leqq N\frac{2\pi}{T}$ der Verzerrungsfaktor $\varphi(\omega)$ annähernd denselben Wert besitzt. Diese Forderung ist ganz entscheidend, wenn wir aus dem Verhalten des Apparates auf den erregenden Vorgang unmittelbare Rückschlüsse ziehen wollen, wenn z. B. im Grammophon oder im Empfangsapparat für drahtlose Telegraphie hohe und tiefe Töne eines Musikstückes mit

[1]) Weil z. B. unser Ohr sie nicht zu erkennen vermag.

verhältnismäßig gleich guter Lautstärke wiedergegeben werden sollen. Der Forderung der relativen Verzerrungsfreiheit kann man natürlich exakt nicht genügen; denn die Resonanzkurven verlaufen niemals genau horizontal. Aber man kann versuchen, die Apparatkonstanten $m$, $r$, $k$ so zu wählen, daß keine ausgeprägte Resonanz mehr auftritt, daß aber zu Beginn die Kurve eine horizontale Tangente hat, so daß $\varphi(\omega) = \alpha$ für kleine Werte von $\omega$ ungefähr konstant bleibt. Nach den obigen Feststellungen erreichen wir dies, indem wir

$$2\,k\,m - r^2 = 0$$

setzen. Wir können bei konstantem $k$ und konstantem $m$ dies durch Wahl einer geeigneten Größe für die Reibung $r$ erreichen (z. B. durch Einschaltung eines elektrischen Widerstandes bei einem Schwingungskreis). Unsere Resonanzkurve zeigt uns, daß dann bis in die Nähe der Eigenschwingung $\omega_0$ des ungedämpften Systemes die Registrierung annähernd Verzerrungsfreiheit liefert und daß oberhalb dieser Frequenz vollständige Abdämpfung herrscht. Man kann also relative Verzerrungsfreiheit in einem gegebenen Intervall erzielen, indem man die Eigenschwingung $\omega_0$ des ungedämpften Systemes durch Vergrößerung von $k$ oberhalb der höchsten in Betracht kommenden erregenden Frequenzen legt und sodann eine Dämpfung $r$ gemäß der Gleichung $2km - r^2 = 0$ wählt.

Die in diesem Kapitel behandelten Differentialgleichungen beziehen sich nur auf Schwingungsvorgänge einfachster Art. Wir werden im zweiten Bande, nachdem wir die Differential- und Integralrechnung der Funktionen mehrerer Veränderlicher entwickelt haben, auf die Differentialgleichungen allgemeinerer Schwingungsvorgänge nochmals zurückkommen.

# Sachverzeichnis.

Die Zahlen geben die Seiten an.