THE COLLECTED PAPERS OF ALBERT EINSTEIN

Volume 3

The Swiss Years:

Writings, 1909–1911

Anna Beck, Translator

Don Howard, Consultant

Princeton University Press

Princeton, New Jersey

Published by Princeton University Press, 41 William Street, Princeton, New Jersey 08540
In the United Kingdom: Princeton University Press, Chichester, West Sussex

ISBN: 0-691-10250-3

Publication of this translation has been aided by a grant
from the National Science Foundation

Princeton University Press books are printed on acid-free paper
and meet the guidelines for permanence and durability of the
Committee on Production Guidelines for Book Longevity of the
Council on Library Resources

Printed in the United States of America

10 9 8 7 6 5 4 3 2

Contents

List of Texts

Publisher's Foreword

We are pleased to be publishing this translation of Volume III of *The Collected Papers of Albert Einstein*. While every effort has been made to ensure the scientific accuracy of this translation, it is not a literary translation and is not intended for use without the documentary edition. The documentary edition provides the extensive editorial commentary necessary for a full understanding of the source documents.

Recent advances in desktop typesetting are responsible for the enhanced quality of the typography in this and future translation volumes.

We are grateful to Dr. Anna Beck and Professor Don Howard for their efforts and dedication to this project. All translations in this volume were prepared by Dr. Beck in consultation with Professor Howard; all translations in the documentary edition were prepared separately by the editors of that volume.

We wish to acknowledge gratefully the continuing grant from the National Science Foundation which has made this publication possible.

Princeton University Press
August 1993

Preface

This volume presents new English translations of all of the documents in Volume 3 of *The Collected Papers of Albert Einstein,* with the exception of the "Scratch Notebook, 1909-1914," which is published as Appendix A in the documentary edition. The documentary edition presents twenty-four of these documents in German versions, the remaining three appearing in their published French versions. The translation volume does not reproduce the annotations or editorial apparatus of the documentary edition, which the reader should consult. We have, however, included in this volume the editorial footnote numbers that correspond to the footnotes in the documentary edition; they are placed within square brackets. Bracketed numbers in the margins that are preceded by a "p." refer to pages in Einstein's notebooks. Angle brackets indicate crossed-out material. For the most part, misprints and errors in the original documents have not been corrected, except for the occasional correction of misspelled names.

The purpose of the translation project, in accordance with the agreement between Princeton University Press and the National Science Foundation, is to provide "a careful, accurate translation that is as close to the German original as possible while still producing readable English." Therefore, our aim has not been to produce a "literary translation," so style has been sacrificed to literalness in some places to enable readers who are not fluent in German to make a scholarly evaluation of the content of the documents. We hope, nevertheless, that the quality of the original German prose shines through.

Some of the technical vocabulary found in the original documents is peculiar to the time and place of their composition. We have tried, whenever possible, to provide not modern translations but English equivalents commonly employed in the contemporary physics literature; otherwise we supply literal translations. Perhaps the most significant exception to this rule is our translation of "Spannung" as "voltage" or as "potential difference" (depending on the context), there being no one standard English equivalent in common use in the first two decades of this century. Similarly, we have reproduced all notations and equations in a form as close as possible to the original.

Three documents in this volume presented a special challenge—the three sets of lecture notes (Docs. 1, 4, and 11). As might be expected, the style of these notes is often fragmentary and telegraphic. To the greatest extent possible, we have sought to reproduce Einstein's abbreviations, repetitions, and errors of grammar and spelling, so as to preserve the feel of the original notes, except in the few cases where such literalness would have produced an impossibly unclear translation.

We would like to thank the staff at the Einstein offices in Boston for their help at various stages in the preparation of this translation. We owe a debt to Walter Lippincott, director of Princeton University Press, for his support and encouragement of this project. Alice Calaprice, senior editor at Princeton University Press, has provided invaluable assistance, for which we thank her. Our thanks go as well to Michael Perlman for his technical assistance in the preparation of the final camera-ready copy of the volume, and to Charles Creesy and Michael Volk for seeing to our computer needs.

ANNA BECK, TRANSLATOR
DON HOWARD, CONSULTANT

Texts

Doc. 1

Lecture Notes for Introductory Course on Mechanics at the University of Zurich, Winter Semester 1909/1910

[18 October 1909–5 March 1910][1]

Mechanics is the science of motion of ponderable matter. It establishes the conditions [p. 1] under which the motion of matter ceases (statics).[2] It seeks to reduce the manifold phenomena of motion to the smallest possible number of elementary laws of the simplest possible form, from which it seeks to reconstruct the more complicated phenomena.

I. Mechanics of the Material Point

We shall first discuss the motion of a body whose dimensions are of no importance in the motions we will discuss, that is, can be regarded as ∞ small. While in motion, such a body will, in general, carry out rotations and change its shape. But we disregard these circumstances, that is, treat it as if it were pointlike; we designate it as a "material point."

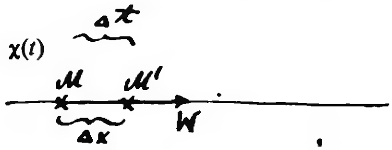
Before we investigate the motion of a m. p.[3] as a function of the motive causes, we must discuss the means and the auxiliary quantities that we use in order to *describe* the motion of an m. point.

A. Kinematics of the M. P.

One cannot speak of the motion of a body (and hence also of a m. p.) *in and for itself,* [p. 2] but only of a relative motion of bodies with respect to each other. If we wish to describe the motion of an m. p., we must describe its motion with respect to a second body. For the latter we choose a system of 3 mutually perpendicular rigid rods. (Coordinate system). We conceive of times as being measured by an arbitrary clock, in that we assume that means are available for ascertaining the readings of the clock that are simultaneous with particular individual positions that the m. p. assumes during its motion.

Obviously, the motion of the <body> m. p. is given if the coordinates x, y, z are given with respect to the c.s.[4] as a function of time. Equations of the following type obtain here:

$$x = \varphi(t) \quad y = \psi(t) \quad z = \chi(t)$$



Rectilinear motion

(a) uniform

$$x = a + bt$$

$$x + \Delta x = a + b(t + \Delta t)$$

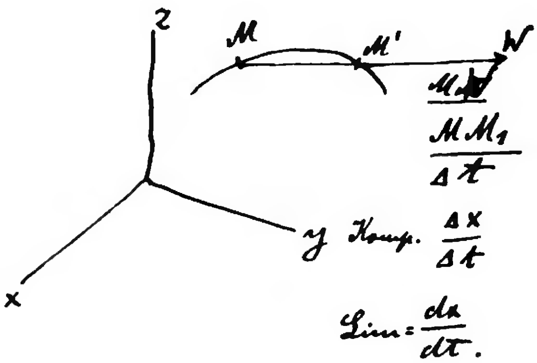
$$\Delta x = b\Delta t \qquad \frac{\Delta x}{\Delta t} = b$$

(b) non-uniform

[p. 3] *Arbitrary Curvilinear Motion*

(Definition of mean and instantaneous velocities)

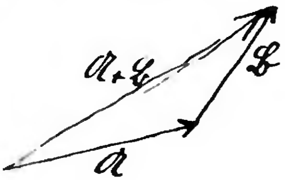
[5]



Velocity is a vector (structure defined by magnitude, direction, and orientation). Graphically represented by an arrow of a given direction & magnitude. Usually denoted by German letters (e.g., $\mathfrak{A}$). Components $\mathfrak{A}_x$, $\mathfrak{A}_y$, $\mathfrak{A}_z$.

Two vectors $\mathfrak{A} = (\mathfrak{A}_x, \mathfrak{A}_y, \mathfrak{A}_z)$
and $\mathfrak{B} = (\mathfrak{B}_x, \mathfrak{B}_y, \mathfrak{B}_z)$

One speaks of the sum $\mathfrak{A} + \mathfrak{B}$ of these vectors. By this one means the vector
$(\mathfrak{A}_x + \mathfrak{B}_x, \mathfrak{A} + \mathfrak{B}_y, \mathfrak{A}_z + \mathfrak{B}_z)$
Geometrically
The commutative law applies.
In the case of several vectors,
the associative and distributive

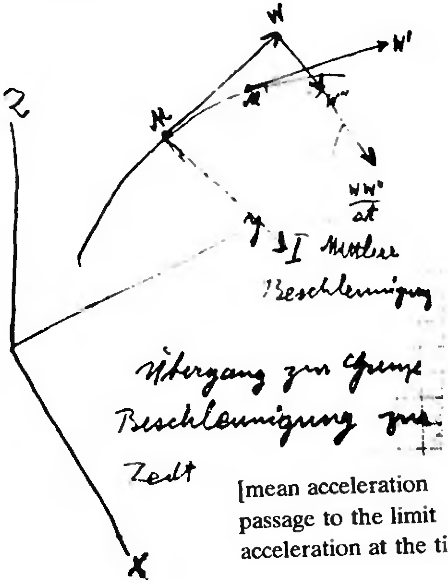


laws apply.[6]

Acceleration [p. 4]

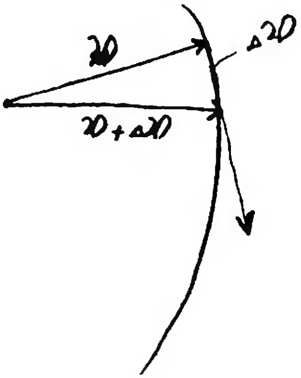
If one divides these by Δt and passes to the limit, one obtains

$$\frac{d^2x}{dt^2} \quad \frac{d^2y}{dt^2} \quad \frac{d^2z}{dt^2}$$



[mean acceleration
passage to the limit
acceleration at the time]

Hodograph[7]



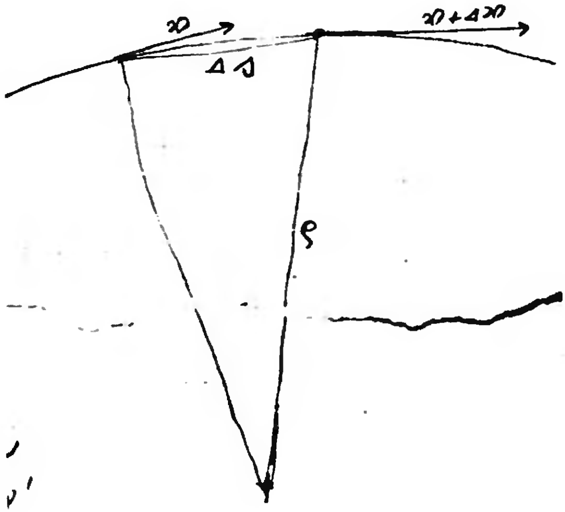
[p. 5]

Tangential and Normal Acceleration

There exists a particularly noteworthy way of resolving the acceleration vector into components, namely, the resolution into a tangential and into a normal component.

$$N = \frac{dv}{dt}$$

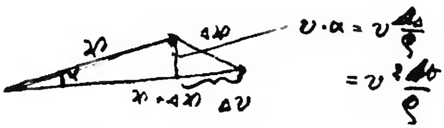
$$T = \frac{v^2}{\rho}$$



Direction cos of N: α β γ
Direction cos of tangt α' β' γ'

[8]

$$\frac{d^2x}{dt^2} = \alpha\frac{dv}{dt} + \alpha'\frac{v^2}{\rho}$$



[p. 6]

B. Dynamics of the Material P.[9]

1. <Galileo's> principle of inertia a m.p. that is <present alone in space> not acted upon by other bodies[10] moves without acceleration.

(a) In a certain sense, this law is an empirical law; (billiard ball, railway car). Strictly speaking, however, it has the character of a definition. For we say that other bodies do not act on a body if this body moves uniformly in a straight line. But for all practical purposes the law can be designated as empirical, because experience happens to be such that the law can be carried out or maintained without any artificial-looking assumptions.[11]

(b) But this law does not hold for an arbitrary state of motion of the c.s. But it holds to a certain degree of approximation for systems at rest relative to the earth, and to an even closer approximation for a system whose origin is at rest relative to the center of gravity of the solar system, and whose axes are directed permanently toward 3 fixed stars.

~~2. If other~~ bodies act upon a material point, the acceleration vector $\left(\frac{d^2x}{dt^2}, \frac{d^2y}{dt^2}, \frac{d^2z}{dt^2}\right)$ is generally different from zero. The ~~cause of the~~ acceleration of a m.p. is called a force. We have a certain direct representation of force, and ~~this is the feeling~~ of exertion or pressure that we experience when, for example, we use our hand to set into ~~motion a~~ [p.7] body that was originally at rest.

2. The accelerations imparted by A to B, & and by B to A are directed along the connecting line and are oppositely oriented.

3. The ratio of accelerations of two mat. p. defines the ratio of masses. Explanation of the empirical laws involved.

$$\frac{B_{\lambda\mu}}{B_{\mu\lambda}} = \frac{m_\mu}{m_\lambda}$$

One mass can be chosen arbitrarily. The rest of the masses can be derived from it by experiment.

4. The addition theorem for accelerations.

If one introduces a vector (x, y, z) that is equal to the acceleration of the point multiplied by m, that is, if one sets

$$m\frac{d^2x}{dt^2} = X \; - \; -,$$

then, for two masses interacting with each other, this vector has the property of being equal for the two, and of opposite direction. We call this vector the force acting on the mass point. Thus, this force always fulfills the condition of equivalence of action and [p. 8] reaction.

The equations of motion given above have the character of definitional equations for the force, thus they can be neither confirmed nor refuted by experience. Nonetheless, we could find ourselves compelled by experience to abandon them; this would happen if the description of facts by means of the equations $m\frac{d^2x}{dt^2} = X \cdot\cdot$ would lead to our having to assume expressions for the force components $X \cdot\cdot\cdot$ in a very complicated manner. One would then reject the equations of motion as unsuitable.

Example: identical springs, stretched in the same way, act in the same direction upon a free body. If the acceleration were not proportional to the number of springs acting, then it would follow from the equations that the force would also not be proportional to the number of springs. This does not represent a logical contradiction, but it would
[p. 9] result in our presuming that we could arrive at a simpler, i.e., preferable theory of motion, if we based ourselves on other equations of motion.

General Remarks on the Motion of the Material Point

For our equations of motion to be useful, the expressions for the force components X etc. may not contain higher than first-order time derivatives of the coordinates. Because the second derivative can be eliminated by solving the equations. However, the occurrence of higher derivatives would make a solution for the second derivative seem unjustified. Hence, for a general theory we have to consider X etc. as functions of $xyz\frac{dx}{dt}\ldots$ and t alone. We have then 3 simultaneous equations of the second order.

The general integrals of these eq. contain 6 arbitrary constants. For the motion is completely determined only if, for a time t_0, $x \cdot\cdot$ and $\dot{x}\,\dot{y}\,\dot{z}$ are given. If $X \cdot\cdot$ are unique functions, then the solution is thereby uniquely det. For we can write

$$\begin{array}{llll} \frac{d\dot{x}}{dt} = \frac{1}{m}X(x\ldots\dot{x}\ldots t) & \frac{dx}{dt} = \dot{x} \quad d\dot{x} = X(\quad)\,dt & & dx = \dot{x}\,dt \\ \text{- - - - - - -} & \text{oder - - - - -} & & \text{- - -} \\ \text{- - - - - - -} & \quad\ \text{- - - - -} & & \text{- - - - -} \end{array}$$

Thus, if $x \cdot\cdot \dot{x} \cdot\cdot$ are given for a time t, they can be calculated for the time $t + dt$ etc.

[p. 10] ~~—In certain cases it proves possible to integrate the equations of motion once (first integrals), so that one arrives at first-order equations~~

(1) The eq. of mot. can be written

$$\frac{d}{dt}\left(m\frac{dx}{dt}\right) = X \quad \text{etc.}$$

~~If the right-hand side can be directly integrated with respect to time, if X<YZ> vanishes or <are> is at least independent of $x \cdot\cdot \dot{x} \cdot\cdot$ etc.~~
~~Example. The force is everywhere parallel to a given direction. We choose the one parallel to the direction Z. Then~~

$$m\frac{dx}{dt} = a \qquad m\frac{dy}{dt} = b \qquad \begin{array}{l|l} m\frac{dx}{dt} = a & dy \\ m\frac{dy}{dt} = b & -dx \end{array}$$

~~From this,~~ $x = a't + c_1 \quad y = b't + c_2$
~~The motion takes place on a plane, because~~ $a\,dy - bdx = 0$, $ay - bx = \text{const.}$
~~Remark: the above equation contains the vector~~ $(m\dot{x}, m\dot{y}, m\dot{z})$, ~~the velocity vector multiplied by mass. It plays a role in many derivations. We call it~~ $b = (b_x, b_y, b_z)$.
We have

$$db_x = Xdt \qquad b_x = \int Xdt$$

~~The momentum is equal to the time integral of the force acting on the body (material point)~~ [p. 11]
[12]

$$\begin{array}{l|l} m\frac{d^2x}{dt^2} = X & -y \\ m\frac{d^2y}{dt^2} = Y & x \end{array}$$

$$m\left(x\frac{d^2y}{dt^2} - y\frac{d^2x}{dt^2}\right) = xY - yX$$

~~Left-hand side $(\ ) = \frac{d}{dt}\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right)$~~

~~hence $\frac{d}{dt}\left\{m\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right)\right\} = xY - yX$~~

Analogous equations hold for the other axes.

<Free Fall. Force of Gravity>
Practical and CGS-Unit of Mass

We measure time in <average> seconds, $\frac{1}{24 \cdot 60 \cdot 60}$ of the aver. solar day, and lengths in cm. 1 cm is the hundredth part of the distance between two marks of a specific meter-stick kept in Sèvres near Paris.

Besides the quantities that depend only on length and time $\left(\frac{d^2x}{dt^2} \cdot\cdot\right)$, our equation of motion contains two additional quantities, namely m and $X \cdot\cdot$ It suffices to establish a unit for *one* of these quantities, because the equations of motion make it possible then to establish the second one.

[p. 12] For if we have defined a unit for m, then we can define

$$mB = K$$

as the unit of force that force which imparts to the unit of mass the acceleration 1.

Conversely, if we have established a unit for the force, than the unit mass is that mass to which the force 1 imparts the acceleration 1.

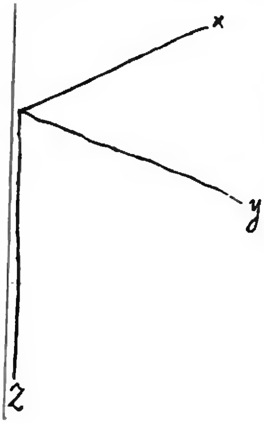
From the theoretical point of view it does not matter for which of the quantities we will establish a unit, but from a pragmatic point of view it does.

<Earlier (before Gauss) one could>[13] A unit for the *force* can be defined in the following way in Paris the earth exerts a quite specific force on 1 cc water at 4°. I call this unit of force 1 gram. (In addition, the kg is also used.)

This definition suffers (for precise investigations) from the following drawback. If people who are not in Paris wish to measure a force accurately, they must compare the force to be measured with the force that terrestrial gravitation exerts on 1 cc water *in Paris*. To this extent the application of the definition is cumbersome.

[p. 13] But the situation is different if the unit of *mass* (also called "gram") as the mass of one cc water at 4°. In this definition, <which now in physical investigations> no particular location on the earth plays a role. The mass 1 gr can be realized at once at any location at which a cm-measuring rod and water are available. For this reason, this system of measures seems now generally to be used in physics.

The advantage over the other system is only a formal one.



Free Fall

At a place close to the earth's surface we imagine a coord. sys. whose Z-axis is directed vertically upwards. We inquire into the motion of a material point with respect to this system. In order to solve this problem, we must know the magnitude of the force exerted by the earth on the <material> body.

One would expect a priori that this force

1) is proportional to the <mass of the> m. p.
2) depends on the physical quality of the point
3) The force could also depend on the velocity.[14]

<For reasons of symmetry> From the choice of the position of the coordinate system it follows that [p. 14]

$$X = 0 \quad Y = 0$$

Further, one arrives at a correct description of the phenomena if one assumes that gravity does not depend either on the quality or the velocity of the m.p. In this way one obtains

$$\cancel{m}\frac{d^2x}{dt^2} = 0$$

$$\cancel{m}\frac{d^2y}{dt^2} = 0$$

$$\cancel{m}\frac{d^2z}{dt^2} = \cancel{m}g$$

From the two first equations one obtains

$$\begin{array}{l|l} \dfrac{dx}{dt} = a & dy \\ \dfrac{dy}{dt} = b & -dx \\ \hline \end{array}$$

$$a\,dy - b\,dx = 0$$

$$ay - bx = c$$

The motion takes place in the vertical plane. We choose this to be the *x*-*Z*-plane. In that case, we have permanently $y = 0$, and we obtain by direct integration of our equations

$$x = c_1 t + c_2 \quad \frac{dx}{dt} = c_1$$

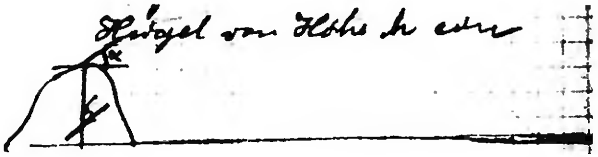
$$z = \frac{g}{2}t^2 + c_3 t + c_4 \quad \frac{dz}{dt} = c_3 + gt$$

[p. 15] We shall now assume that for $t = 0$, $x = z = 0$ and $\frac{dx}{dt} = \frac{dz}{dt} = 0$; in that case all $c = 0$, and we obtain

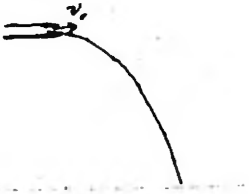
$$x = 0$$

$$z = \frac{g}{2}t^2.$$

The familiar formula for free fall. By calculating the constants c from the conditions of the problem, we can solve each problem concerned with free fall.
Example: On a hill of height h is a cannon whose elev. is α. The initial velocity of the shell is v_0. Where will it strike? The most favorable elevation angle? (Air resistance neglected).



Example 2: Water pipe
Equation of the curve.



<For> The fact that the gravitational force is independent of the material <we have no explanation.> shows a close relationship between inertial mass, on the one hand, and the effect of gravitation, on the other hand.[15]

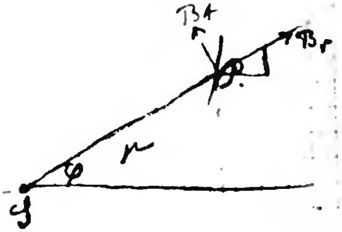
Let us now find the law of interaction between <masses> the sun & the planets through <gravity> gravitation, the way Newton found it from Kepler's laws. [p. 16]

These laws of Kepler are as follows:

(1) The radius vector sun-planet sweeps out equal areas in equal times.

(2) The planet travels in an ellipse, in which the sun occupies one focus.

(3) The squares of the planets' periods of orbit vary as the third power of the major axes of the ellipses.



The nature of the present problem makes it seem expedient to use polar coordinates for the description of planetary motion. In order to apply these, we will seek to express the acceleration vector in polar coordinates.

First, we have

$$x = r\cos\varphi \qquad \frac{dx}{dt} = \cos\varphi\frac{dr}{dt} - r\sin\varphi\frac{d\varphi}{dt}$$

$$y = r\sin\varphi \qquad \frac{dy}{dt} = \sin\varphi\frac{dr}{dt} + r\cos\varphi\frac{d\varphi}{dt}$$

[p. 17]

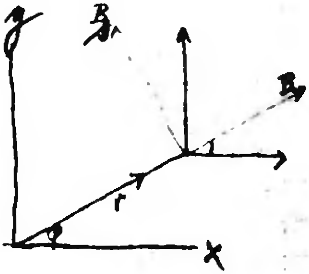
$$\begin{array}{c|c|l}
-\sin\varphi & \cos\varphi & \dfrac{d^2x}{dt^2} = \cos\varphi\dfrac{d^2r}{dt^2} - 2\sin\varphi\dfrac{dr}{dt}\dfrac{d\varphi}{dt} - r\cos\varphi\left(\dfrac{d\varphi}{dt}\right)^2 \\
 & & \qquad - r\sin\varphi\dfrac{d^2\varphi}{dt^2} \\
\cos\varphi & \sin\varphi & \dfrac{d^2y}{dt^2} = \sin\varphi\dfrac{d^2r}{dt^2} + 2\cos\varphi\dfrac{dr}{dt}\dfrac{d\varphi}{dt} - r\sin\varphi\left(\dfrac{d\varphi}{dt}\right)^2
\end{array}$$

$$+ r\cos\varphi\frac{d^2\varphi}{dt^2}$$

$$B_r = \frac{d^2x}{dt^2}\cos\varphi + \frac{d^2y}{dt^2}\sin\varphi$$

$$B_s = -\frac{d^2x}{dt^2}\sin\varphi + \frac{d^2y}{dt^2}\cos\varphi$$

$$\boxed{\begin{aligned} B_r &= \frac{d^2r}{dt^2} - r\left(\frac{d\varphi}{dt}\right)^2 \\ B_s &= 2\frac{dr}{dt}\frac{d\varphi}{dt} + r\frac{d^2\varphi}{dt^2} = \frac{1}{r}\frac{d}{dt}\left(r^2\frac{d\varphi}{dt}\right) \end{aligned}}$$



From (1) it follows that $r \cdot r\,d\varphi = c\,dt$ & $r^2\dfrac{d\varphi}{dt} = c$

From this it follows first that $B_s = 0$. Thus, the acceleration vector of planetary motion lies in the direction sun–planet.

We now calculate B_r by means of (1) and (2)
Because of (2) we have

$$r=\frac{p}{1-e\cos\varphi}\qquad \frac{dr}{dt}=-\frac{p}{(1-e\cos\varphi)^2}\cdot e\sin\varphi\frac{d\varphi}{dt}=-\frac{e}{p}r^2\sin\varphi\frac{d\varphi}{dt}$$

$$\frac{dr}{dt}=-\frac{e}{p}c\sin\varphi$$

$$\frac{d^2r}{dt^2}=-\frac{e}{p}c\cos\varphi\frac{d\varphi}{dt}$$

$$\begin{array}{l|l}
\cos\varphi=\dfrac{r-p}{er} & \dfrac{d^2r}{dt^2}=\dfrac{c^2}{p}\dfrac{r-p}{r^3}\text{[16]} \\[2ex]
\dfrac{d\varphi}{dt}=\dfrac{c}{r^2} & r\left(\dfrac{d\varphi}{dt}\right)^2=\dfrac{c^2}{r^3} \\ \cline{2-2}
\end{array}$$

$$B_r=\frac{c^2}{p}\cdot\frac{1}{r^2}$$

The acceleration imparted to a planet = $\frac{\text{const.}}{r^2}$. There is still the question whether this constant has the same value for all planets. To find this out, we must introduce the period of orbit. We have [p. 18]

The area of the ellipse = $\frac{1}{2}cT = ab\pi \qquad c = \frac{2ab\pi}{T}$

$$\frac{c^2}{p}=\frac{4\pi^2a^2b^2}{T^2p}.\quad \text{But since}\quad \frac{b^2}{a}=p \quad \& \quad \frac{b^2}{p}=a$$

hence $\frac{c^2}{p}=4\pi^2\frac{a^3}{T^2}$

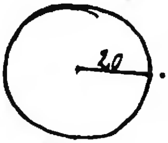
But since according to Kepler's 3d law $\frac{a^3}{T^2}$ has the same value for all planets, we can [p. 19] set the accel[eration] caused by the sun = $\frac{C}{r^2}$.[17] Force with which the sun attracts a planet = mass·accel. = $\frac{\mu \cdot f}{r^2}$, where f a factor that is independent of the planet. For reasons of symmetry, the numerator must depend on the mass of the sun M just as it depends on the mass of the planet, hence $f = M\kappa$, where κ does not depend either on the sun or on the planet. Hence we have

$$\text{force} = \kappa \cdot \frac{Mm}{r^2} \qquad \left(\text{or from accel.} = \kappa \frac{M}{r^2}\right)$$

There is still the question: What is the value of constant κ? In order to find it, we must know for one case both masses, the force, and the distance. In order to determine κ, we must know the magnitude of the acting mass. This is only possible for relatively small masses. (Example of this. Earth as gravitating center mass of the earth)

It has been found that $\kappa = 6.70 \cdot 10^{-8}$.

More about the methods later on.[18]



$$B = 6.7 \cdot 10^{-8} \cdot \frac{4}{3}\pi \cdot \frac{20^3}{20^2} 13.6$$

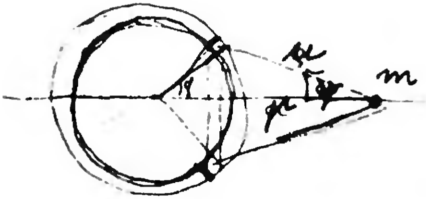
$$= \frac{\overline{6.7} \cdot 10^{-8} \cdot \bar{8} \cdot 10^1 \cdot 13.6 \cdot 3.14}{3}$$

$$= 7{,}6 \cdot 10^{-5} \text{ cm weniger als } \frac{1}{1000} \text{ mm.}$$

10^{-7} der Schwere

$$\frac{6 \cdot \cancel{10^8}}{\cancel{2} \cdot 20} \cdot \frac{2{,}5 \cdot 20}{6 \cdot 10^8}$$

[p. 20] $$\frac{R dR\, d\varphi \cdot R \sin\varphi \langle d\psi \rangle 2\pi}{\rho^2} \cos\psi$$ [19]



Anz[iehung] $\dfrac{\kappa m\rho 2\pi R^2 \sin\varphi\, d\varphi}{u^2}\cos\psi$

$$\int_0^\pi = \kappa m\rho 2\pi R^2 \int \frac{\sin\varphi\, d\varphi}{u^2}\cos\psi$$

$$u^2 = R^2 + r^2 + 2Rr\cos\varphi^{[20]}$$

$$u\,du = Rr\sin\varphi\, d\varphi$$

$$\sin\varphi\, d\varphi = \frac{u\,du}{Rr}$$

$$\cos\psi = \frac{r^2 + u^2 - R^2}{2ur}$$

$$\frac{1}{2Rr^2}\int \frac{du}{u^2}(r^2 - R^2 + u^2)$$

$$= \frac{1}{2Rr^2}\left\{(r^2 - R^2)\int \frac{du}{u^2} + du\right\} = \frac{1}{2Rr^2}\left|\left\{-\frac{(r^2 - R^2)}{u} + u + \text{konst}\right\}\right|_{r-R}^{r+R}$$

$$= -(r^2 - R^2)\left\{\frac{1}{r+R} - \frac{1}{r-R}\right\} + (r - R) - (r + R)^{[21]}$$

$$= -(2R) + 2R$$

[p. 21]

General Remarks on the Motion of the M. P.

1)
$$m\frac{d^2x}{dt^2} = X$$

$$m\frac{d^2y}{dt^2} = Y$$

$$m\frac{d^2z}{dt^2} = Z$$

Integrating the equations of motion between two time limits t_1 and t_2, one obtains

$$\left| m\frac{dx}{dt} \right|_{t_1}^{t_2} = \int X\,dt$$

etc. Thus, the time integral of the force acting on a m. p. causes a completely determined change of $m\frac{dx}{dt^2}$ ·· of the material point. These quantities are called the components of the momentum *(mv)* of the m. p. Conversely, we see that the total effect that a force lasting for a certain time has on the state of motion of the m. p. is determined only by $\int X dt$ etc. (impulse.

2) The momentum <motion in a plane>

$$\begin{array}{l|l} m\frac{d^2x}{dt^2} = X & -y \\ & \\ m\frac{d^2y}{dt^2} = Y & +x \end{array} \qquad m\left(x\frac{d^2y}{dt^2} - y\frac{d^2x}{dt^2}\right) = xY - yX\ .$$

But since $x\frac{d^2y}{dt^2} - y\frac{d^2x}{dt^2} = \frac{d}{dt}\left\{x\frac{dy}{dt} - y\frac{dx}{dt}\right\}$, one obtains

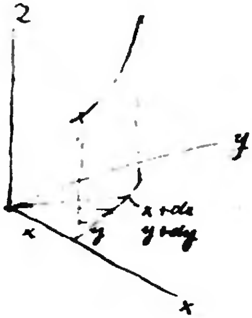
$$\frac{d}{dt}\left\{m\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right)\right\} = xY - yX$$ [p. 22]

$$\frac{d}{dt}\left\{m\left(y\frac{dz}{dt} - z\frac{dy}{dt}\right)\right\} = yZ - zY$$

$$\frac{d}{dt}\left\{m\left(z\frac{dx}{dt} - x\frac{dz}{dt}\right)\right\} = zX - xZ$$

$$\frac{d}{dt}\left\{m\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right)\right\} = xY - yX$$

If the right-hand side of one of these equations vanishes, i.e., if $\frac{X}{Y} = \frac{x}{y}$, i.e., if the force intersects the Z-axis, one obtains an integral.



We have then $x\frac{dy}{dt} - y\frac{dx}{dt} = \text{const.}$

$$ds = \frac{1}{2}\begin{vmatrix} x & y \\ x+dx & y+dy \end{vmatrix} = \frac{1}{2}(x\,dy - y\,dx)$$

$$\frac{ds}{dt} = \frac{1}{2}\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right) = \text{konst.}$$

The areal velocity of the rad.
vect. of the x-y projection is const.
The reverse also holds.
If central force, then 3 integrals, because all three right sides are then = 0. Then

$$y\frac{dz}{dt} - z\frac{dy}{dt} = A$$

$$z\frac{dx}{dt} - x\frac{dz}{dt} = B$$

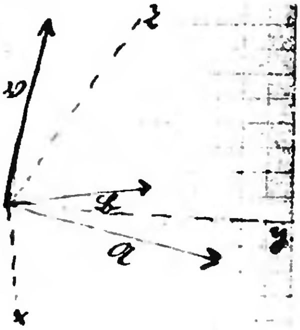
Mult by x, y z and add. $Ax + By + Cz = 0$. Equation of a plane passing through [*sic*] through the O-point.

[p. 23]

Geometrical Interpretation of the Law of Areas

Given are two vectors $\mathfrak{A}$ and $\mathfrak{B}$, with the components $\mathfrak{A}_x$, $\mathfrak{A}_y$, $\mathfrak{A}_z$ and $\mathfrak{B}_x$, $\mathfrak{B}_y$, $\mathfrak{B}_z$. In addition to these two vectors we can construct a third vector $\mathfrak{v}$ in the following way:

1. $\mathfrak{v}$ is perpendicular to the plane laid through $\mathfrak{A}$ and $\mathfrak{B}$
2. The quantity, or (as it is called) the tensor[22] of the vector is twice the area of the triangle that is to be constructed from A & B
3. The direction of $\mathfrak{v}$ is such that a rotary motion of $\mathfrak{A}$ toward $\mathfrak{B}$, together with a translational motion in the direction of the arrow of $\mathfrak{v}$ leads to a right-hand coil



This vector $\mathfrak{v}$ is called the vector product of $\mathfrak{A}$ and $\mathfrak{B}$.
Components of the vector product. The plane ($\mathfrak{A}\mathfrak{B}$) is $\perp$ to $\mathfrak{v}$. Hence the angle between $\mathfrak{v}$ and the Z-axis is equal to the angle between the plane $\mathfrak{A}\mathfrak{B}$ and the plane XY

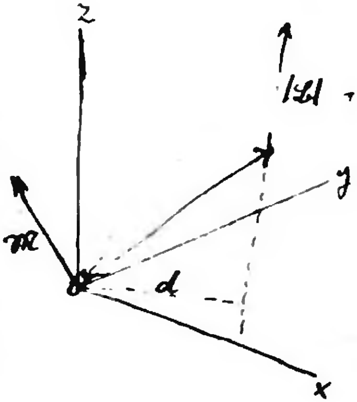
$$|\mathfrak{v}| = \Delta$$

$\mathfrak{v}_z = \Delta \cos \mathfrak{v}_z = \Delta_{xy}$, where Δ_{xy} denotes the area of the proj[ection] of Δ on the xy plane.

$$\mathfrak{v}_z = \mathfrak{A}_x\mathfrak{B}_y - \mathfrak{A}_y\mathfrak{B}_x$$

Analogously for the other two components of $\mathfrak{v}$.

[p. 24] We consider the special case where one of the vectors is the radius vector drawn from O to the point of application of the other vector. The vector product of the radius vector and the given vector $\mathfrak{B}$ is called the moment of the vector $\mathfrak{B}$ with respect to the point O.



It is perpendicular to the plane $O\mathfrak{B}$, and its tensor is equal to twice Δ, and thus to the product of the magnitude $|\mathfrak{B}|$ of the vector and the latter's distance d from O.

According to the aforesaid, the components are

$$y\mathfrak{B}_z - z\mathfrak{B}_y$$

$$- - - - -$$

$$- - - - - -$$

In the equation developed above, components of moments appear on both sides. On the left, the moment of the momentum, on the right, the moment of the forces acting on the m p. In a central motion, the moment of momentum, and thus also that of the velocity, is a spatially and temporally constant vector.

The Law of Kinetic Energy

[p. 25]

$$\begin{array}{l|l} m\dfrac{d^2x}{dt^2} = X & \dfrac{dx}{dt}dt = dx, \\[2ex] m\dfrac{d^2y}{dt^2} = Y & \dfrac{dy}{dt}dt = dy \\[2ex] m\dfrac{d^2z}{dt^2} = Z & \dfrac{dz}{dt}dt = dz \end{array}$$

where dx is the projection of the element of the trajectory.

$$m\left(\frac{dx}{dt}\frac{d^2x}{dt^2} + \frac{dy}{dt}\frac{d^2y}{dt^2} + \frac{dz}{dt}\frac{d^2z}{dt^2}\right)dt = X\,dx + Y\,dy + Z\,dz$$

$$v^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2$$

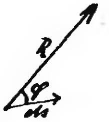
$$d(v^2) = 2\left\{\frac{dx}{dt}\frac{d^2x}{dt^2} + \cdot + \cdot\right\}$$

From this $d\left(m\dfrac{v^2}{2}\right) = Xdx + Ydy + Zdz$

$m\dfrac{v^2}{2}$ is called kinetic energy. The right-hand side is the product, resultant force · path element · cos of the angle between the two. Because one can write it as

$$R \cdot ds \cdot \left\{\frac{X}{R}\frac{dx}{ds} + \frac{y}{R}\frac{dy}{ds} + \frac{Z}{R}\cdot\frac{dz}{ds}\right\} = R\ ds\ \cos\varphi\ ,$$

where the bracketed fractions are direction cos. of R & ds.



This is the work that the force $X\ Y\ Z$ transmits to the mat. p. during time dt.

The law can also be derived directly, by resolving the acceleration into a tangential & ⊥ component, B_r and B_s.

[p. 26]

$$B_t = \frac{dv}{dt}$$

$$F_t = m\frac{dv}{dt}$$

$$F_t\ ds = mv\frac{dv}{dt}dt = d\left(\frac{mv^2}{2}\right)$$

Integrating the equation obtained, one gets

$$\frac{mv^2}{2} - \frac{mv_0^2}{2} = \int_{t_0}^{t_1} Xdx + Ydy + Zdz\ .$$

In the special case where $X\ Y\ Z$ depend only on $x\ y\ z$, it is possible to calculate the integral on the right-hand side if the trajectory is given.

But in an even more special case, namely if $X\ Y\ Z$ are of the form

$$X = +\frac{\partial U}{\partial x}$$ [23]

$$Y = +\frac{\partial U}{\partial y}$$

$$Z = +\frac{\partial U}{\partial z},$$

it is not even necessary to know the trajectory in order to carry out that integration. <Such forces we call forces derivable from a potential>
For in this case we have

$$\int X dx + Y dy + Z dz = +\int\left(\frac{\partial U}{\partial x}dx + \frac{\partial U}{\partial y}dy + \frac{\partial U}{\partial z}dz\right) = +\int dU$$
$$= +(U - U_0)$$

In this case $\left(\frac{mv^2}{2} + U\right) = \left(\frac{mv_0^2}{2} + U_0\right)$ = const., where U is, thus, a function of the coordinates alone. If the m. p. turns up twice at the same spatial point, then P, and thus also v, will have the same value if P is single-valued.[24] The temporally constant [p. 27] quantity that we found here for the case when a m. p. is under the influence of temp. const. forces that are derivable from a <potential> force function, we will call the "energy" of the system considered. —P is called the potential energy. The law we found can then be formulated thus:

"The sum . . . . remains constant."

As the condition for the work $\int$ to be independent of the integration path, we found the equations

$$X = +\frac{\partial P}{\partial x}, \quad Y = +\frac{\partial P}{\partial y}, \quad Z = +\frac{\partial P}{\partial z}$$

This condition can also be expressed in another form. If one differentiates the third d. eq. with respect to y, and the second with respect to z, one obtains

$$\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z} = 0$$

analog. $$\frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x} = 0$$

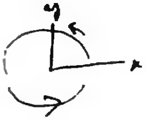
$$\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y} = 0$$

An example of multivalued U $\quad U = \operatorname{arctg} \frac{y}{x}$

$$dU = \frac{1}{1 + \frac{y^2}{x^2}} \left(\frac{dy}{x} - \frac{y}{x^2} dx \right)$$
$$= \frac{xdy - ydx}{x^2 + y^2}$$

$$\left. \begin{aligned} X &= -\frac{y}{r^2} \\ Y &= \frac{x}{r^2} \end{aligned} \right\}$$

$xX + yY = 0$, thus the force is perpendicular to the radius. Its magnitude is $\frac{1}{r}$.



[p. 28]

Equilibrium of the Material Point

If force function present

$$X = Y = Z = \frac{\partial U}{\partial x} = \frac{\partial U}{\partial y} = \frac{\partial U}{\partial z} = 0$$

The law of kinetic energy reveals a case in which it is certain that the equilibrium is stable.

$$\frac{m}{2} v^2 + P = \frac{m}{2} v_1^2 + P_1 \qquad \text{Equilibrium at } P_0(x_0\, y_0\, z_0)$$

$$= P_0 + \varepsilon$$

$$\frac{m}{2}v^2 + P - P_0 = \varepsilon$$

$$P - P_0 \leqq \varepsilon$$

Central Forces That Depend Only on the Distance

We have seen that motion takes place in a plane.

1. Law of areas

$$x\frac{dy}{dt} - y\frac{dx}{dt} = c$$

or also $r^2\dfrac{d\varphi}{dt} = c$ (1)

2. The energy law.

$$d\left(\frac{mv^2}{2}\right) = F\left(\frac{x}{r}dx + \frac{y}{r}dy + \frac{z}{r}dz\right)$$

$$= Fdr \qquad (2)$$

These two equations determine completely r & φ as funct of time. [p. 29]

We have

$$v^2 = \frac{dr^2 + r^2 d\varphi^2}{dt^2}$$

From this we obtain by means of the law of areas

$$v^2 = \left(\frac{dr}{dt}\right)^2 + \frac{C^2}{r^2} \qquad (3)$$

&

$$v^2 = C^2\frac{1}{r^4}\left(\frac{dr}{d\varphi}\right)^2 + \frac{C^2}{r^2}$$

&
$$v^2 = C^2 \left\{ \left(\frac{\partial \frac{1}{r}}{\partial \varphi} \right)^2 + \frac{1}{r^2} \right\} \cdots \qquad (4)$$

By substituting $\frac{2}{m}\int F dr$ for v^2, one obtains dt and $d\varphi$ as funct of r.

Now we write (2) in the form

$$\frac{1}{2}\frac{dmv^2}{dt} = F\frac{dr}{dt}$$

$$\frac{d}{dt}\left(\frac{m}{2}\left[\left(\frac{dr}{dt}\right)^2 + \frac{C^2}{r^2}\right]\right) = F\frac{dr}{dt}$$

Differentiating,

$$\frac{m}{2}\left[2\cancel{\frac{dr}{dt}}\frac{d^2r}{dt^2} - 2\frac{C^2}{r^3}\cancel{\frac{dr}{dt}}\right] = F\cancel{\frac{dr}{dt}}$$

& $$m\left[\frac{d^2r}{dt^2} - \frac{C^2}{r^3}\right] = F$$

& $$m\frac{d^2r}{dt^2} = \Gamma + m\frac{C^2}{r^3} \qquad \ldots\ldots (5)^{[25]}$$

[p. 30] From $\frac{1}{2}\frac{dmv^2}{d\varphi} = \Gamma\frac{dr}{d\varphi}$

one obtains, by inserting v^2 from 4

$$\frac{d}{d\varphi}\left\{\frac{mC^2}{2}\left[\left(\frac{d\frac{1}{r}}{d\varphi}\right)^2+\left(\frac{1}{r}\right)^2\right]\right\}=F\frac{dr}{d\varphi}$$

$$\frac{mC^2}{2}\left\{-\frac{2}{r^2}\cancel{\frac{\partial r}{\partial\varphi}}\frac{\partial^2\frac{1}{r}}{\partial\varphi^2}-\frac{2}{r^3}\cancel{\frac{\partial r}{\partial\varphi}}\right\}=F\cancel{\frac{dr}{d\varphi}}$$

$$\Gamma=-\frac{mC^2}{r^2}\left\{\frac{\partial^2\frac{1}{r}}{\partial\varphi^2}+\frac{1}{r}\right\}\quad\ldots\ldots(6)$$

We now determine t & φ as funct. of r. We set $\frac{2}{m}\int F\,dr=\varphi(r)+h=v^2$ according to eq. 2

Then eq. 3 becomes

$$\psi(r)=\varphi(r)-\frac{c^2}{r^2}+h=\frac{dr^2}{dt^2}$$

$$dt=\frac{dr}{\pm\sqrt{\psi(r)}}$$

Suppose we know the sign of $\frac{dr}{dt}$ for $t=t_0$

The sign of the square root is thereby determined up to the moment when $\frac{dr}{dt}$ becomes 0 again. Then $\frac{dr}{dt}$ usu. changes sign (at r_0). This is generally easy to detect in the special case under discussion. It can be determined unambiguously from the sign of $m\frac{d^2r}{dt^2}=\Gamma+\frac{mc^2}{3}$.

From this it is easy to derive $d\varphi$, since, according to the law of areas, [p. 31]

$$r^2\frac{d\varphi}{dt} = c, \quad \text{hence}$$

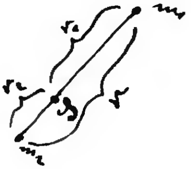
$$d\varphi = \frac{c}{r^2}dt = c\frac{dr}{r^2\sqrt{\psi(r)}}$$

Insert (should come before "Central Forces.")[26]

Let us discuss the case of two mat. points acting on each other by central forces that depend only on the distance.

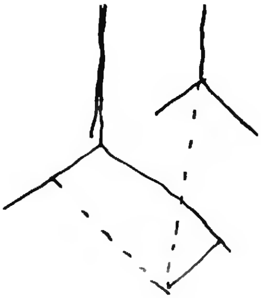
$$m_1\frac{d^2x_1}{dt^2} = F(r)\frac{x_1 - x_2}{r} \qquad m_2\frac{d^2x_2}{dt^2} = F(r)\frac{x_2 - x_1}{r}$$

- - -



- - - $r = \sqrt{(x_2 - x_1)^2 + \cdot + \cdot}$

These equations do not change their form with the introduction of a uniformly moving coordinate system



$$x_1 = x_1' + \alpha_x t + \beta_x \qquad x_2 = x_2' + \alpha_x t + \beta_x$$

- - - - - - - -   - - - - - - - -

Second derivat. & differ[ence] $x_2 - x_1$ do not change under transformation.

$$m_1\frac{d^2x_1}{dt^2} + m_2\frac{d^2x_2}{dt^2} = 0$$

$$m_1x_1 + m_2x_2 = \alpha t + \beta$$

In order to interpret, we define

$$\xi = \frac{m_1x_1 + m_2x_2}{m_1 + m_2}$$

$\eta =$

$\zeta =$

the so-called center of gravity of both masses

$$\xi = (m_1 + m_2)\,(\alpha t + \beta) \quad \text{moves uniformly}$$

New coordin. syst. that mov. unif., hence is at rest relative to the cen. of grav. *O* placed at the center of gravity. Then

$$m_1 x_1 + m_2 x_2 = 0$$

or

$$m_1 x_1 = -m_2 x_2 \qquad x_1 : y_1 : z_2 = x_2 : y_2 : z_2$$

$$m_1 y_1 = -m_2 y_2 \qquad \text{sq. \& add}$$

$$\text{- - - - - - -} \qquad m_1 r_1 = m_2 r_2$$

$$r_2 = \frac{m_1}{m_2} r_1$$

$$r = r_1 + r_2 = \frac{m_1 + m_2}{m_2} r_1$$ [p. 32]

$$m_1 \frac{d^2 x_1}{dt^2} = F\left(\frac{m_1 m_2}{m_1} r_1\right) \cdot \frac{x_1}{r_1} = F_1(r_1) \frac{x_1}{r_1}$$

This is the same equation as the equation of motion with a fixed center of force.

Application to the Solar System. Sun & Planet

Force $F(r) = \kappa \dfrac{m_1 m_2}{r^2} = \kappa \dfrac{mM}{r^2}$

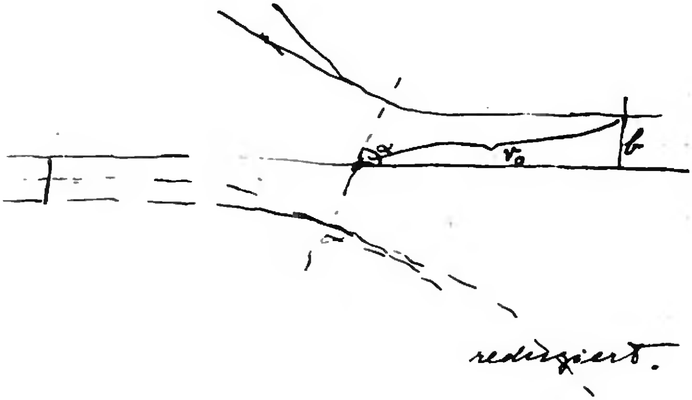
$$F\left(\frac{m_1+m_2}{m}r_1\right)=\kappa\frac{mM}{r_1^2\left(\frac{m+M}{M}\right)^2}=\kappa\frac{mM}{r_1^2}\left(\frac{M}{m+M}\right)^2$$

$$=\frac{c^2}{p}\cdot\frac{1}{r_1^2}=4\pi^2\,\frac{a^3}{T^2}\cdot\frac{1}{r_1^2}$$

3rd Kepler's law not strictly valid. Neither is the second law.

Example with regard to central forces Force law between two identical gas molecules

$F=\frac{\alpha}{r^5}$ Collision law[27]



$$\Gamma_1(r_1)=\frac{\alpha}{(2r_1)^5}=\overset{\beta}{\boxed{\frac{\alpha}{32}}}\cdot\frac{1}{r_1^5}$$

Problem reduced to a central-force problem.

$$\varphi(r)+h=\frac{2}{m}\int\frac{\beta}{r^5}dr=-\frac{2\beta}{m5}\cdot\frac{1}{r^4}+h=v^2$$

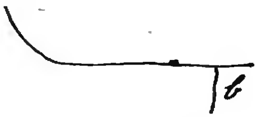
If v_0 denotes the velocity at ∞ distance, then $h = v_0^2$

$$\psi(r) = \Phi(r) + h - \frac{c^2}{r^2} = -\frac{2\beta}{m\cdot 5}\cdot\frac{1}{r^4} - \frac{c^2}{r^2} + v_0^2$$ [p. 33]

$$\left.\begin{aligned} dt &= \int \frac{dr}{\pm\sqrt{\psi r}} \\ d\varphi &= c^2 \int \frac{dr}{r^2\sqrt{\psi(r)}} \end{aligned}\right|\quad \text{Here } c = +bv_0$$

$$\varphi = 2\alpha = 2c^2 \int_{\infty}^{r_0} \frac{dr}{r^2\sqrt{\psi(r)}}$$

Example. Mass penetrates into the solar system, branch of hyperbola. What direction does it have afterwards?[28]



From the last equation (6) it is very easy to derive the force law from Kepler's 2nd law.

$$\Gamma = -\frac{mc^2}{2}\left\{\frac{\partial^2 \frac{1}{r}}{\partial\varphi^2} + \frac{1}{r}\right\}$$

$$\left.\begin{aligned} \frac{1}{r} &= \frac{1 - e\cos\varphi}{p} \\ \frac{\partial^2 \frac{1}{r}}{\partial\varphi^2} &= +\frac{e\cos\varphi}{p} \end{aligned}\right\} = \frac{1}{p}.$$

[p. 34] *The Motion of a Point That Must Stay on a <Plane> Surface*

$$m\frac{d^2x}{dt^2} = X_a + X_f$$

$$m\frac{d^2y}{dt^2} = Y_a + Y_f$$

$$m\frac{d^2z}{dt^2} = Z_a + Z_f$$

X_f, Y_f, Z_f is here the force that the surface exerts on the point. We assume that the surface exerts on the point a counterpressure that is normal to the former. In that case the components $(X_f : Y_f : Z_f)$ vary as the direction cosines of the normal to the surface. Let $\varphi(x,y,z,t)$ be the equation of the surface, then these direction cosines vary as $\frac{\partial\varphi}{\partial x} : \frac{\partial\varphi}{\partial y} : \frac{\partial\varphi}{\partial z}$, hence $X_f = \lambda\frac{\partial\varphi}{\partial x}$ $Y_f = \lambda\frac{\partial\varphi}{\partial y}$.

$$m\frac{d^2x}{dt^2} = X + \lambda\frac{\partial\varphi}{\partial x}$$

$$\varphi\,(x,\, y,\, z,\, t) = 0$$

$$m\frac{d^2y}{dt^2} = Y + \lambda\frac{\partial\varphi}{\partial y}$$

- - - - -

Together, these equations determine the four variables x, y, z and λ.
If surface at rest & forces derivable from a potential, then the conservation of energy holds.
Examples:

Simple Pendulum

Z-axis downwards
Equation of the surface

$$x^2 + y^2 + z^2 - l^2 = 0$$

[p. 35]

$$\frac{\partial\varphi}{\partial x} : \frac{\partial\varphi}{\partial y} : \frac{\partial\varphi}{\partial z} = x : y : z \;.$$

Hence equations

$$m\frac{d^2x}{dt^2} = 0 + 2\lambda' x$$

$$m\frac{d^2y}{dt^2} = 0 + 2\lambda' y \quad \text{\& substituting } \frac{2\lambda}{m} \text{ instead of } \lambda$$

$$m\frac{d^2z}{dt^2} = mg + 2\lambda' z \quad \frac{2\lambda'}{m} = \lambda \quad \text{string tension } 2\lambda' l = m\lambda l$$

$$\left.\begin{array}{l} \dfrac{d^2x}{dt^2} = \lambda x \\[2ex] \dfrac{d^2y}{dt^2} = \lambda y \end{array}\right| \begin{array}{l} -y \\ \quad x^2 + y^2 + z^2 = l^2 \\ x \end{array}$$

$$\frac{d^2z}{dt^2} = g + \lambda z \;.$$

We need two relations for a complete solution.
1) Energy principle. Because $xdx + ydy + zdz = 0$

$$\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2 = 2gz + h$$

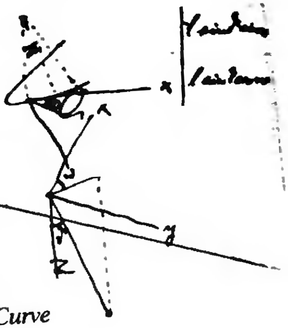
$$\&\quad \left.\begin{array}{r} dx^2 + dy^2 + dz^2 = (2gz + h)dt^2 \\ xdy - ydx = cdt \end{array}\right\}$$

Because the distance from the coordinate origin is const., it is advantageous to introduce polar coordinates.

$x = l \sin\vartheta \cos\omega \quad dx = l\{\cos\vartheta \cos\omega\, d\vartheta - \sin\vartheta \sin\omega\, d\omega\}$
$y = l \sin\vartheta \sin\omega \quad dy = l\{\cos\vartheta \sin\omega\, d\vartheta + \sin\vartheta \cos\omega\, d\omega\}$
$z = l \cos\vartheta \quad dz = l\{-\sin\vartheta \cdot d\vartheta\}$

From this

$$dx^2 + dy^2 + dz^2 = l^2\{d\vartheta^2 + \sin^2\vartheta \; d\omega^2\}$$
$$xdy - ydx = l^2\sin^2\vartheta \; d\omega$$



Inserting this into our equation, we obtain

$$l^2\{d\vartheta^2 + \sin^2\vartheta d\omega^2\} = (2gl \cos\vartheta + h)dt^2$$
$$l^2\sin^2\vartheta d\omega^{<2>} = cdt.$$

[p. 36]

Motion of a Point Along a Given Fixed Curve

§ 1.

Thus far we have been addressing the problem of finding the force when the motion was given, or finding the motion when the force was given. But there are problems in which conditions for the motion are given. Imagine, for example, a small perforated body pulled along a rigid wire and acted upon by given external forces. Besides the given external force, a reactive force of the wire, to be viewed as unknown for the time being, also acts upon the point. All that we assume for the time being about this reactive force $\mathfrak{R}$ is that it is perpendicular to the tangent on the wire, so that we will have

$$\mathfrak{R}_x dx + \mathfrak{R}_y dy + \mathfrak{R}_z dz = 0.$$

This implies that the reactive force does not perform any work.

To find the mot. of the p., we can replace the wire by the reactive force exerted by it. Formally, this reduces the case of the point pulled along the wire to the case of the freely moving point. We can therefore set

$$m\frac{d^2x}{dt^2} = X + \mathfrak{R}_x$$

$$m\frac{d^2y}{dt^2} = Y + \mathfrak{R}_y$$

$$m\frac{d^2z}{dt^2} = Z + \mathfrak{R}_z$$

We can evaluate the condition assumed for 𝕽 by multiplying these equations by $dx\,dy\,dz$ [p. 37] & summing, & we obtain

$$d\left(\frac{mv^2}{2}\right) = Xdx + Ydy + Zdz \qquad (1)$$

Thus, the equation for the kinetic energy is here valid and is sufficient for the solution of any problem of motion, as can be seen from the following. A single variable (q)[29] suffices for the descrip. of the motion of the point. x, y, & z are to be considered the given functions of this single variable s. First, we have $v^2 = \left(\frac{ds}{dt}\right)^2$

$$\frac{dx}{dt} = \frac{dx}{ds}\frac{ds}{dt}\;\ldots \qquad v^2 = \left(\frac{dx^2}{ds} + \frac{dy^2}{ds} + \frac{dz^2}{ds}\right)\frac{ds^2}{dt^2}$$

$$= (x'^2 + y'^2 + z'^2)\,\frac{ds^2}{dt^2}$$

Further, we have

$X\left(x, y, z, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}, t\right)$, thus also X(~~$s, x'\; y'\; z'$~~ $\left(s, \frac{ds}{dt}, t\right)$) .

Further, we have to set (as known,

$$x = \varphi(s) \qquad y = \chi(s) \qquad z = \psi(s)$$

Further,

$$\frac{dx}{dt} = \varphi'\cdot\frac{ds}{dt}$$

$$v^2 = (\varphi'^2 + \chi'^2 + \psi'^2)\frac{ds^2}{dt}$$

X dep. on $x\;y\;z$, $\frac{dx}{dt}$ . . and t, Hence, since one has to set $x = \varphi(q)$ . . (known func. of s), $\frac{dx}{dt} = \varphi'(q)\frac{dq}{dt}$. Thus, X a known func. of q and $\frac{dq}{dt}$. Thus, the above equation yield diff. eq. for s. We can write above equation as

$$d\left\{\frac{m}{2}(\varphi'^2 + \chi'^2 + \psi'^2)\left(\frac{dq}{dt}\right)^2\right\} = \{X\varphi' + Y\chi' + Z\psi'\}ds = Qds. \qquad (1')$$

where X, Y, Z are to be thought of as expressed in the new variables. If X, Y, Z depend only on the coordinates, then { } of the right-hand side (Q) depends only on s. The equation can then be integrated right away. <If we set $\int Q\,ds = f$> one obtains

$$\frac{m}{2}v^2\ \frac{m}{2}v_0^2 = \int_{s_0}^{s} QdQ \text{ or } \left\langle \text{solved for } v^2 = \frac{dq^2}{dt},\ \left(\frac{dq}{dt}\right)^2 = f(q)\ (1'') \right\rangle$$

$$<\text{From this,}\quad t - t_0 = \int \frac{dq}{\pm\int[\sqrt{fq}?]}\quad>$$

[p. 38]

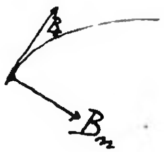
§ 2
Geometric Derivation of the Fundamental Equation

We resolved the acceleration into normal and a tang. components.[30] We resolve the total force R acting on the point in an analogous way.

$$B_t = \frac{dv}{dt} \qquad R_t = K_t^{<(a)>}$$

analogously, total force

$$B_n = \frac{v^2}{\rho} \qquad R_{<s>n} = K_n^{<(a)>} + N$$



$$R_s = 0 = K_s + N'_{<s>}\,.$$ [31]

From the tangential components

$$mB_t = K_t \ \ldots\ldots\ (2a)$$

$$\&\qquad m\frac{dv}{dt} = K_t$$

K_t is gen. known as a func of $s\frac{ds}{dt}$ & t. Mult both sides by $ds = vdt$, we obtain

$$mv\frac{dv}{dt} = K_t ds$$

& $d\left(\frac{m}{2}v^2\right) = K_t ds$ integrable if K_t dep[ends] only on s.

The total force K is composed of the external force $K^{(a)}$ and the reactive force of the curve of magnitude N this is perpendicular to the curve &
just like $K_n^{(a)}$ is taken as positive with respect to the center of the curvature.
After the solution of the motion problem, the normal components yield the reaction of the curve. We have

$$m\frac{v^2}{\rho} = K_n^{<(a)>} + N \qquad (2b)$$

& $N = \dfrac{mv^2}{\rho} - K_n^{(a)}$ (2b) If v^2 is found as a funct of s, then this eq. yields norm. reaction.

§ 3

[p. 39]

There Exists a Single-valued Force Function. Physical Meaning.

We return to equation (1)

$$d\left(\frac{m}{2}v^2\right) = Xdx + Ydy + Zdz$$

We have already seen that this eq. is integrable if X, Y, Z depend only on s. We now further assume that there exists a force function for X, Y, Z that only depends on $x\ y\ z$, so that

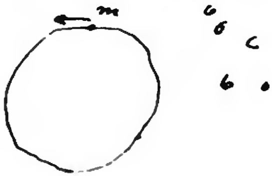
$$X = \frac{\partial U}{\partial x}, \quad Y = \frac{\partial U}{\partial y}, \quad Z = \frac{\partial U}{\partial z}$$

The right-hand side is then equal to dU, so that equation is integrable.

$$\frac{mv^2}{2} = U + h \ldots\ldots (3)$$

<Since $v = \dfrac{ds}{dt}$, therefore> The solution of the problem is obtained from eq. (3) by a single integration.

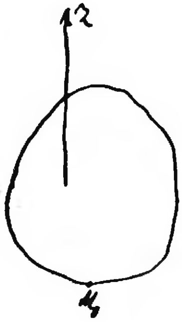
Let us add here a general remark. Suppose the force $X\ Y\ Z$ derives from a system of bodies that does not experience any spatial or other kind of change during the motion of the m p. <If <force U> dep. only on the *position* of the m. p., then> What does the existence of a single-valued force function mean in this case? Suppose the point moves, perchance, along an endless wire without changing the sign of its vel. Then v^2 is always of the same magnitude at the same location otherwise mechanism for the construction of perpetuum mobile Forces exerted by unchang. sys., which must depend only *on the position*



[p. 40]

<5> *Gravitation*

z-axis upwards



Then $Xdx + Ydy + Zdz = -mg\,dz \quad U = -mgz$

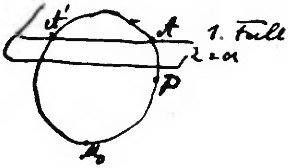
$$m\frac{v^2}{2} = -mg + h' \qquad \frac{v^2}{2} = gz + h'$$

If we write $\frac{h}{g} = a$

$$\underline{v^2 = 2g\,(a - z)}$$

$v_0^2 = 2g\,(a - z_0)$ thus a can be made arbitrarily large.

Suppose we lay on a plane $z = a$, then this plane either cuts the curve or lies above it.



1) $a - z$ cannot become negative, hence, in the first case, the curve cannot cross the plane $z = a$. The material point turns around at $z = a$, but at no other point, because at no other point can we have $v = 0$.

2) If $z = a$ lies above the curve, then $a - z$ is positive for all points of the curve. Then the point travels without turning around.

1st Case thus shuttling to & fro between A and A', with the velocity being the same at each point of the curve. We calculate the time the mobile needs from M_0 to P.

[p. 41]

$$v^2 = \frac{ds}{dt^2} = 2g(a - z)$$ [32]

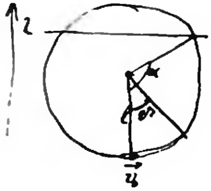
$$dt = \frac{1}{2g}\int_{M_0}^{P} \frac{ds}{\pm\sqrt{a - z}}$$

In the integral z & s are related by the equation of the curve.

§ (5). Example. Circle in a Vertical Plane (Simple Pendulum)[33]

$$v^2 = 2g\,(a - z)$$

Constant from velocity at the lowest point



$$v_0{}^2 = 2g(a + l) \qquad a = -l + \frac{v_0^2}{2g}$$

1st Case. $z = a$ intersects the circle

$$-l + \frac{v_0^2}{2g} < l \qquad v < 2\sqrt{lg}$$

We set $\quad z = -\,l\cos\vartheta \quad a = -\,l\cos\alpha$

$$v = \frac{ds}{dt} = l\frac{d\vartheta}{\delta t}$$

$$(1 - \cos\alpha) - (1 - \cos\vartheta)$$

$$l^2\left(\frac{d\vartheta}{dt}\right)^2 = 2gl(\cos\vartheta - \cos\alpha)$$

or $$l\left(\frac{d\vartheta}{dt}\right)^2 = 4g\left\{\sin^2\frac{\alpha}{2} - \sin^2\frac{\vartheta}{2}\right\}$$

$$\sqrt{\frac{g}{l}}\,dt = \frac{d\left(\frac{\vartheta}{2}\right)}{\sqrt{\sin^2\frac{\alpha}{2} - \sin^2\frac{\vartheta}{2}}} \qquad \sqrt{\frac{g}{l}}\,t = \int_0^{\vartheta}\frac{d\frac{\vartheta}{2}}{\sqrt{\sin^2\frac{\alpha}{2} - \sin^2\frac{\vartheta}{2}}}$$

$$\sin\frac{\vartheta}{2} = u\sin\frac{\alpha}{2} \qquad \sqrt{} = \sin\frac{\alpha}{2}\sqrt{1 - u^2}$$

$$\cos\frac{\vartheta}{2}\frac{d\vartheta}{2} = du\sin\frac{\alpha}{2}$$

$$d\frac{\vartheta}{2} = \frac{\sin\frac{\alpha}{2}}{\sqrt{1-\sin^2\frac{\vartheta}{2}}}du = \frac{\sin\frac{\alpha}{2}du}{\sqrt{1-\sin^2\frac{\alpha}{2}u^2}} = \frac{\sin\frac{\alpha}{2}du}{\sqrt{1-\kappa^2u^2}}$$

where $\kappa = \sin\frac{\alpha}{2}$

$$\sqrt{\frac{g}{l}}t = \int_0^u \frac{du}{\sqrt{(1-u^2)(1-\kappa^2u^2)}}$$

if K infinitely small

$$\sqrt{\frac{g}{l}}t = \arcsin u$$

$$u = \frac{\sin\frac{\vartheta}{2}}{\sin\frac{\alpha}{2}} = \sin\sqrt{\frac{g}{l}}t$$

$$\vartheta = \alpha\sin\sqrt{\frac{g}{l}}t.$$

[p. 42]

$$u = \frac{\sin\frac{\vartheta}{2}}{\sin\frac{\alpha}{2}} = \text{s[i]n}\left(\sqrt{\frac{g}{l}}t\right)$$

$$\sin\frac{\vartheta}{2} = \sin\frac{\alpha}{2}\,\text{s[i]n}\left(\sqrt{\frac{g}{l}}t\right)$$

$$\cos\frac{\vartheta}{2} = \sqrt{1-\kappa^2\text{s[i]n}(\;)} = du\left(\sqrt{\frac{g}{l}}t\right)$$

The duration of a simple oscillation

$$\sqrt{\frac{g}{l}}\,\frac{T}{4} = \int_0^1 \frac{du}{\sqrt{(1-u^2)(1-\kappa^2u^2)}} = K$$

We develop T as a funct of κ.

$$\frac{1}{\sqrt{1-\kappa^2u^2}} = (1-\kappa^2u^2)^{-1/2} = 1 + \frac{1}{2}\kappa^2u^2$$

$$+\frac{1\cdot 3}{2\cdot 4}\kappa^4u^4 + \cdots\cdots \frac{1\cdot 3\cdots\cdots 2n-1}{2\cdot 4\cdots\cdots 2n}\kappa^{2n}n^{2n}$$

$$\int_0^1 \frac{u^{2n}\,du}{\sqrt{1-u^2}} = \frac{\pi}{2}\,\frac{1\cdot 3\cdot 5\cdots\cdots(2n-1)}{2\cdot 4\cdot 6\cdots\cdots 2n}$$

$$\kappa = \frac{\pi}{2}\left[1+\left(\frac{1}{2}\right)^2\kappa^2+\left(\frac{1\cdot 3}{2\cdot 4}\right)^2\kappa^4\cdots\cdots\right]$$[34]

$$T_g = 2\pi\sqrt{\frac{l}{g}}\underbrace{\left[1+\left(\frac{1}{2}\right)^2\sin^2\frac{\alpha}{2}+\left(\frac{1\cdot 3}{2\cdot 4}\right)^2\sin^4\frac{\alpha}{2}\cdots\cdots\right]}_{\text{2nd approx[imation]}\ 1+\frac{\alpha^2}{16}.}$$

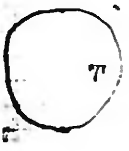
2nd Case

$$-l+\frac{v_0^2}{2g} > l$$

$$l^2\left(\frac{d\vartheta}{dt}\right)^2 = 2g(a+l\cos\vartheta) = 2g\left(a+l-2l\sin^2\frac{\vartheta}{2}\right)$$

$$= 2g(a+l)\left(1-\underbrace{\frac{2l}{a+l}}_{\kappa^2}\sin^2\frac{\vartheta}{2}\right)$$

$$\underbrace{\frac{1}{2}\frac{\sqrt{2g(a+l)}}{l}}_{\lambda}dt = \frac{d\frac{\vartheta}{2}}{\sqrt{1-\kappa^2\frac{\sin^2\vartheta}{2}}} \qquad \frac{\sin\vartheta}{2} = u$$ [35]



$$\lambda t = \int_0^u \frac{du}{\sqrt{(1-u^2)(1-\kappa^2u^2)}}$$

$$\lambda T = 2\int_0^1 = \pi\left\{1+\left(\frac{1}{2}\right)^2\kappa^2+\left(\frac{1\cdot 3}{2\cdot 4}\right)^2\kappa^4\ldots\right)$$

[p. 43] *3rd Case.* Limiting case

$$l^2\left(\frac{d\vartheta}{dt}\right)^2 = 2g(l+l\cos\vartheta) = 4gl\cos^2\frac{\vartheta}{2}$$

$$\left\langle\sqrt{\frac{g}{l}}\frac{d\frac{\vartheta}{2}}{\cos\vartheta_2} = \right\rangle\sqrt{\frac{g}{l}}dt = \frac{\langle\sin\rangle d\frac{\vartheta}{2}}{\cos\frac{\vartheta}{2}}$$

$$\sqrt{\frac{g}{l}}t = \log tg\left(\frac{\Theta}{4}+\frac{\pi}{4}\right)$$

String tension $R_n = K_n + N$

$$R_n = +\frac{mv^2}{\rho} \qquad K_n = -mg\cos\vartheta$$

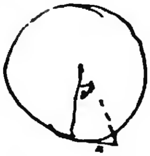
$$N = \frac{\underbrace{mv^2}_{2g(a-z)}}{l} + \text{mg}\underbrace{\cos\vartheta}_{-\frac{z}{l}} = \frac{m}{l}g\{2a-3z\}$$

to discuss!

§ 3. Approximative Treatment of the Pendulum Problem

$$m\frac{d^2s}{dt^2} = K_t = -mg\sin\vartheta \approx -mg\,\vartheta$$

$$l\frac{d^2\vartheta}{dt^2} = -g\vartheta$$



or, if we introduce the abscissa x, $\vartheta l = x$

$$\frac{d^2x}{dt^2} = -\frac{g}{l}x$$

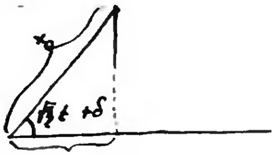
Solution $A\sin\left(\sqrt{\frac{g}{l}}t\right) + B\cos\sqrt{\frac{g}{l}}t$, as differentiation shows at once This can be reformulated $\sqrt{A^2+B^2}\left\{\underbrace{\frac{A}{\sqrt{A^2+B^2}}}_{\cos\delta}\sin(\ ) + \underbrace{\frac{B}{\sqrt{A^2+B^2}}}_{\sin\delta}\cos(\ )\right\} = W\sin\left(\sqrt{\frac{g}{l}}t + \delta\right)$

$$= x_m\sin\left(\sqrt{\frac{g}{l}}t + \delta\right)$$

$$x_m\sin\left(\frac{2\pi}{T}t + \delta\right)$$

$$T = 2\pi\sqrt{\frac{l}{g}}.$$

This is also the form of the general solution.
Graphic representation



Rotating vector. X_0 Amplitude. $\sqrt{\frac{g}{l}}t + \delta$

Phase angle. δ = phase ∡ for $t = 0$.
This graphic representation corresponds to calculation with complex numbers.[36] [p. 44]

If $G(x) = 0$ is a homogeneous linear diff. eq. in the variable t with real, constant coefficients, which is solved by means of the complex numbers $\alpha(t) + j\beta(t)$, where α & β are real, then we can write symbolically

$$G(\alpha + j\beta) = 0$$

Now, the real funct. remains real when it is diff, and the im function remains likewise im when diff. It is therefore easy to prove that

$$G(\alpha + j\beta) = G(\alpha) + jG(\beta)$$

Thus, the equation $G(\alpha + j\beta) = 0$ is equivalent to

$G(\alpha) + jG(\beta) = 0$ & to the two equations

$$G(\alpha) = 0 \text{ and } G(\beta) = 0$$

Thus, if we found the complex function $\alpha + j\beta$ that satisfies the eq. $G = 0$, then its real component also satisfies the equation.

Application to the prev examp

$$\frac{d^2x}{dt^2} + \frac{g}{l}x = 0 \text{ is linear eq.}$$

We seek the solution of the form $e^{\alpha t}$ Inserted

$$\alpha^2\cancel{e}\text{---} + \frac{g}{l}\cancel{e}\text{---} = 0 \qquad \alpha = \sqrt{-\frac{g}{l}} = j\sqrt{\frac{g}{l}}$$

Solution $e^{j\frac{g}{l}t}$ Real part $\cos\sqrt{\frac{g}{l}}t$

Since the starting point of t is arbitrary, we thus arrive at the previous solution.

[p. 45] 2nd Example. Infinitely small pendular oscillation with friction

$$m\frac{d^2x}{dt^2} = -\frac{mg}{l}x - R\frac{dx}{dt}$$

$$\frac{d^2x}{dt^2} + \frac{R}{m}\frac{dx}{dt} + \frac{g}{l}x = 0$$

$$e^{\alpha t} \text{ solution} \qquad \alpha^2 + \frac{R}{m}\alpha + \frac{g}{l} = 0$$

$$\alpha = -\frac{R}{2m} \pm \sqrt{-\frac{g}{l} + \left(\frac{R}{2m}\right)^2}$$

We now want the friction to be small. Then with root negative.

We write $\alpha = -\frac{R}{2m} \pm j\sqrt{\frac{g}{l} - \left(\frac{R}{2m}\right)^2}$

$$\Re\left(e^{\alpha t}\right) = e^{-\frac{R}{2m}t}\cos\sqrt{}\,t = e^{-\alpha t}\cos\frac{2\pi}{T}t$$

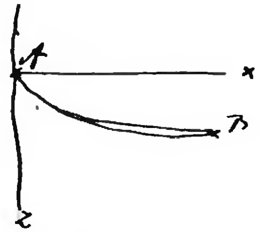
α determines damping

$$\frac{2\pi}{\mathrm{T}} = \sqrt{\frac{g}{l} - \left(\frac{R}{2m}\right)^2} = \sqrt{\frac{g}{l}}\sqrt{1 - \frac{l}{g}\left(\frac{R}{2m}\right)^2} \approx \sqrt{\frac{g}{l}}\left(1 - \frac{l}{8g}\left(\frac{R}{m}\right)^2\right)$$

$$\mathrm{T} = 2\pi\sqrt{\frac{l}{g}}\left(1 + \frac{l}{8g}\left(\frac{R}{m}\right)^2\right)$$

The effect of friction on oscillation is of the second order.

Brachystochrone



$$\left(\frac{ds}{dt}\right)^2 = 2gz + (h) \qquad h = 0$$

$$\sqrt{2g}\,dt = \int_A^B \frac{ds}{\sqrt{z}} \qquad \text{Minimum.}$$

Motion of a M P. Relative to a Fixed or Movable Plane [p. 46]

$$f(x, y, z, t) = 0$$

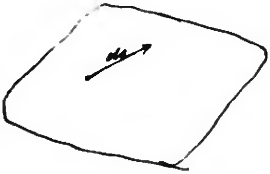
Reaction ⊥ to the plane, hence proportion to

$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$$

$$m\frac{d^2x}{dt^2} = X + \lambda\frac{\partial f}{\partial x}$$

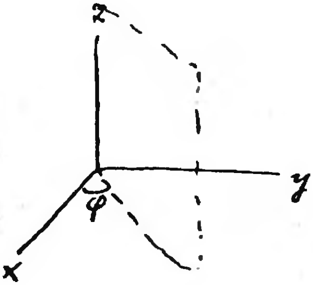
$$m\frac{d^2y}{dt^2} = Y + \lambda\frac{\partial f}{\partial y}$$

- - - - - - - - -



These 4 eq. determine *x*, *y*, *z* & λ completely.

Example Plane rotates with const. angular velocity ω about the *z*-axis. How does a point move on it?



$$\frac{y}{x} = tg\,\omega t = \frac{\sin\omega t}{\cos\omega t}$$

$$f = x\sin\omega t - y\cos\omega t = 0$$

$$\left.\begin{array}{ll} \cos & m\dfrac{d^2x}{dt^2} = \lambda\sin\omega t \\ \sin & m\dfrac{d^2y}{dt^2} = -\lambda\cos\omega t \\ & m\dfrac{d^2z}{dt^2} = 0 \end{array}\right\}$$

z linear function of time. Assume special case where $z = 0$.

$$\frac{d^2x}{dt^2}\cos\omega t + \frac{d^2y}{dt^2}\sin\omega t = 0$$

$$x = \rho\cos\varphi = \rho\cos\omega t \qquad \frac{dx}{dt} = \frac{d\rho}{dt}\cos - \omega\rho\sin$$

$$y = \rho\sin\varphi = \rho\sin\omega t \qquad \frac{dy}{dt} = \frac{d\rho}{dt}\sin + \omega\rho\cos$$

$$
\begin{array}{l|c|c}
\dfrac{d^2x}{dt^2} = \dfrac{d^2\rho}{dt^2}\cos - 2\omega\dfrac{d\rho}{dt}\sin - \omega^2\rho\cos & \cos & \sin \\
\dfrac{d^2y}{dt^2} = \dfrac{d^2\rho}{dt^2}\sin + 2\omega\dfrac{d\rho}{dt}\cos - \omega^2\rho\sin & \sin & -\cos \\
\hline
\end{array}
$$

$$\frac{d^2\rho}{dt^2} - \omega^2\rho = 0$$

$\rho = e^{\alpha t}$, Then $\alpha^2 = \omega^2 \quad \alpha = \pm\,\omega$ [p. 47]
$\rho = Ae^{\omega t} + Be^{-\omega t} \qquad \varphi = \omega t.$

It remains for us to find λ.

From the two equations

$$\lambda = m\left\{\sin\,\omega t\frac{d^2x}{dt^2} - \cos\,\omega t\frac{d^2y}{dt^2}\right\} = -2\omega\frac{d\rho}{dt} = -2\omega^2\{Ae^{\omega t} - Be^{-\omega t}\}$$

If $B = 0$ logarithmic spiral.

Special case: curve at rest. $dx\ dy\ dz$ is then a line element on the plane, so that also

$$\frac{\partial f}{\partial x}dx + . + . = 0$$

Multiplying the eq. by $dx = \dfrac{dx}{dt}dt$ etc., & summing, one gets

$$d\left(m\frac{v^2}{2}\right) = Xdx + Ydy + Zdz$$

Further, in the special case when X Y Z are derivable from a potential, then integrable.

$$m\frac{v^2}{2} = U + h$$

i.e., <there exists> one can give an integral equation (law of the kinetic energy.)

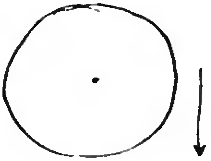
Example. Spherical Pendulum

Point remains on the sphere $f = l^2 - x^2 + y^2 + z^2 = 0$

$$m\frac{d^2x}{dt^2} = -2\lambda x$$

$$m\frac{d^2y}{dt^2} = -2\lambda y$$

$$m\frac{d^2z}{dt^2} = -2\lambda z + mg$$



or, if we introduce $\frac{2\lambda}{m} = \mu$,

$$\frac{d^2x}{dt^2} = -\mu x$$

$$\frac{d^2y}{dt^2} = -\mu y$$

$$\frac{d^2z}{dt^2} = -\mu z + g$$

[p. 48] ∞ small oscillations $z = l$ $\frac{d^2z}{dt^2} = 0$

Thus, from the last equation $\mu = \frac{g}{l}$

The other two equations become:

$$\frac{d^2x}{dt^2} = -\frac{g}{l}x$$

$$\frac{d^2y}{dt^2} = -\frac{g}{l}y$$

x & y components behave totally independently of each other. We obtain

$$x = A\cos\sqrt{\frac{g}{l}}t + B\sin\sqrt{\frac{g}{l}}t \qquad \frac{dx}{dt} = -A\sqrt{}\sin + B\sqrt{}\cos$$

$$y = A'\cos \qquad + B'\sin \qquad \frac{dy}{dt} = -A'\sqrt{}\sin + B'\sqrt{}\cos$$

There will exist points of time at which velocity is perpendicular to radius vector. We count the time from such a point on, and choose the X-axis such that the ZX plane passes through this point.

Then for $t = 0$ $y = 0$ and $\frac{dx}{dt} = 0$

$A' = 0$ $B = 0$. thus

$$x = A\cos\sqrt{\frac{g}{l}}t$$

$$y = B'\sin\sqrt{\frac{g}{l}}t$$

Thus, $\left(\frac{x}{A}\right)^2 + \left(\frac{y}{B}\right)^2 = 1$ Ellipse

$$\text{Oscillation period} = 2\pi\sqrt{\frac{l}{g}}.$$

Oscillations of Arbitrary Amplitudes [p. 49]

We have $<U> = mgz \mid m\frac{v^2}{2} = mgz + \text{const.}$, hence

$$v^2 = 2gz + h \quad \cdots\cdot \quad (1)$$

$$\begin{array}{l:l} \dfrac{d^2x}{dt^2} = -\mu x & -y \\[2ex] \dfrac{d^2y}{dt^2} = -\mu y & x \end{array}$$

Since neither the external force nor the reaction have a momentum with respect to the Z-axis, the law of areas is valid with respect to the xy-plane. In fact; if one mult. the second eq by x and the first by $-y$, & adds them up, one obtains

$$x\frac{d^2y}{dt^2} - y\frac{d^2x}{dt^2} = 0$$

$$x\frac{dy}{dt} - y\frac{dx}{dt} = c \quad \text{(law of areas)}$$

or also

$r^2\frac{d\vartheta}{dt} = c$; if one sets $x^2 + y^2 = r^2$ & ϑ is the ∡ between the x axis & r

We choose r, ϑ, and z as coordinates. v is to be expressed in them. We have

$$v^2 = \frac{ds^2}{dt^2} = \frac{dr^2 + r^2d\vartheta^2 + dz^2}{dt^2},$$

so that equation (1) becomes

$$\frac{dr^2 + r^2d\vartheta + dz^2}{dt^2} = 2gz + h \quad \cdots\cdot \quad (1')$$

We have further $r^2 + z^2 = \ell^2 \cdot\cdot\cdot\cdot (3)$

By eliminating ϑ and r with the aid of (2) and (3), we obtain an equation between z and t.

$$rdr + zdz = 0; \quad dr = \frac{-zdz}{\underline{\sqrt{l^2 - z^2}}}$$

$$\frac{d\vartheta}{} = \frac{cdt}{r^2} = \frac{cdt}{\underline{l^2 - z^2}}$$

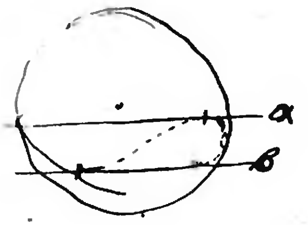
[p. 50] Inserting this in (1′), we obtain:

$$l^2\,dz^2 = \underbrace{[(2gz + h)(l^2 - z^2) - C^2]\,dt^2}_{\psi(z)}$$

$$dt = \frac{ldz}{\pm\sqrt{\psi(z)}} \quad \cdots\cdot \quad (1'')$$

Because of (2), $\frac{d\vartheta}{dt}$ never changes its sign & never becomes zero. In contrast, $\frac{dz}{dt}$ becomes zero if $\psi(z) = 0$. Only values between $-l$ and $+l$ come under consideration.

Positive values of $\psi(z)$ are certain to exist between them, because otherwise (1″) could not be satisfied. Thus, there are at least *two* zero-points between them, and no more than two because ψ is of the third order. We call them α and β. Thus, the m. p. moves always to and fro between two <points> planes $z = \alpha$ & $z = \beta$. To traverse the space between two specific horizontal planes, it always requires the same time.



From 2) $$d\vartheta = \frac{cdt}{r^2} = \frac{cldz}{l^2 - z^2\sqrt{\psi(z)}}$$

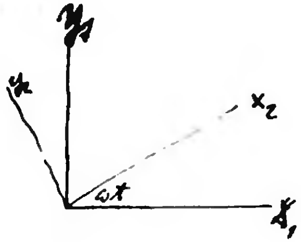
The ∡ descr[ibed] between two planes is thus also always the same.

$$\vartheta_2 - \vartheta_1 = \int_\alpha^\beta \frac{cldz}{(l^2 - z^2)\sqrt{\psi(z)}}$$

(∡ $\frac{\pi}{2}$ for the case where α & β are positive

Laws of Motion Relative to the Earth[37]

[p. 51]



$$\begin{array}{c|c|l}
-\sin & \cos & m\dfrac{d^2x_1}{dt^2} = X_1 \\
\cos & \sin & m\dfrac{d^2y_1}{dt^2} = Y_1 \\
 & & m\dfrac{d^2z}{dt^2} = Z_1
\end{array}$$

X_1 etc. shall be indep. of earth's rotation

transferred to the comoving system We have here the equations

$x_2 = x_1 \cos \omega t + y_1 \sin \omega t$ $\quad X_2 = X_1 \cos \omega t + Y_1 \sin \omega t$

$y_2 = -x_1 \sin \omega t + y_1 \cos \omega t$ $\quad$ - - -

$z_2 = z_1$

This shows us the factors with which one must mult. the original eq. in order to obtain new ones.

$$x_1 = x_2\cos\omega t - y_2\sin\omega t \qquad\qquad y_1 = x_2\sin\omega t + y_2\cos\omega t$$

$$\frac{dx_1}{dt} = \frac{dx_2}{dt}\cos - \frac{dy_2}{dt}\sin + \omega(-x_2\sin - y_2\cos) \qquad \frac{dy_1}{dt} = \frac{dx_2}{dt}\cos + \frac{dy_2}{dt}\sin$$

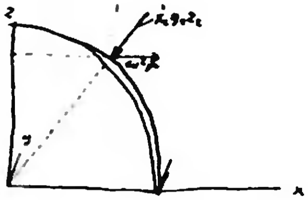
$$+ \omega(x_2\cos - y_2\sin)$$

$$\begin{array}{llllllll|l|l}
\frac{d^2x_1}{dt^2} = & \frac{d^2x_2}{dt^2}\cos + & \frac{d^2y_2}{dt^2}\sin + & 2\omega\left(-\frac{dx_2}{dt}\sin\right. & \left.- \frac{dy_2}{dt}\cos\right) & + \omega^2(-x_2\cos & + y_2\sin) & & \cos & -\sin \\
\frac{d^2y_1}{dt^2} = & \sin & \cos & \cos & -\sin & -\sin & -\cos & & \sin & \cos.
\end{array}$$

$$m\left\{\frac{d^2x_2}{dt^2} - 2\omega\frac{dy_2}{dt} - \omega^2 x_2\right\} = X_2$$

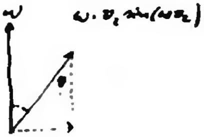
$$m\left\{\frac{d^2y_2}{dt^2} + 2\omega\frac{dx_2}{dt} - \omega^2 y_2\right\} = Y_2$$

$$m\left\{\frac{d^2z_2}{dt^2}\right\} = Z_2$$



If we set $X_2 + \omega^2 x_2 = X'_2$ etc., we get

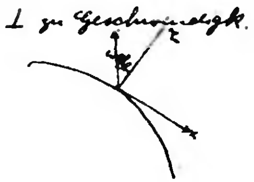
$$\begin{array}{l}
m\frac{d^2x_2}{dt^2} = X'_2 + 2m\omega\frac{dy_2}{dt} \\[2ex]
m\frac{d^2y_2}{dt^2} = Y'_2 - 2m\omega\frac{dx_2}{dt} \\[2ex]
m\frac{d^2z_2}{dt^2} = Z'_2 + 0
\end{array}
\;\vdots\;
\begin{array}{l}
\frac{dx}{dt} \\[2ex]
\frac{dy}{dt} \\[2ex]
\frac{dz}{dt}
\end{array}$$



Interpret supplementary forces <‖> In the *xY* plane, thus ⊥ to ω. ⊥ to velocit.

[p. 52]

$$-2m \left| \begin{array}{ll} \omega_y \dfrac{dz}{dt} - \omega_z \dfrac{dy}{dt} & \omega_x = -\omega \cos\varphi \\ \omega_z \dfrac{dz}{dt} - \omega_x \dfrac{dy}{dt} & \omega_y = 0 \\ \omega_x \dfrac{dz}{dt} - \omega_y \dfrac{dy}{dt} & \omega_z = +\omega \sin\varphi \end{array} \right.$$



$$\cancel{m}\frac{d^2x}{dt^2} = 0 + 2\cancel{m}\omega \sin\varphi \frac{dy}{dt} \qquad X$$

$$\cancel{m}\frac{d^2y}{dt^2} = 0 - 2\cancel{m}\omega\left(\sin\varphi \frac{dx}{dt} + \cos\varphi \frac{dz}{dt}\right) \quad +\text{ev.} \qquad Y$$

$$\cancel{m}\frac{d^2z}{dt^2} = -\cancel{m}g + 2\cancel{m}\omega \cos\varphi \frac{dy}{dt} \qquad Z$$

Foucault's pendulum. If motion takes place in the x-y plane, then $\frac{dz}{dt} = 0$. Then the first two equat., in conjunction with the eq., set up earlier for x_2 etc., show that the system behaves as a system rotating with velocity ω sin φ. Thus, apparent rotation of the plane of the pendulum. The 3rd equation shows that, due to the rotation of the earth, reaction force = $-2m\omega\cos\varphi\frac{dy}{dt}$.

We consider a free-falling m p.

$$\left. \begin{array}{l} \dfrac{dx}{dt} = a + 2\omega \sin\varphi\, y \\ \\ \dfrac{dz}{dt} = b - gt + 2\omega \cos\varphi y \end{array} \right| \begin{array}{l} = a + 2\omega \sin\varphi(d + ct) \\ \\ = b - gt + 2\omega \cos\varphi(d + ct) \end{array}$$

$$\frac{d^2y}{dt^2} - 2\omega\{a\sin\varphi + (b - gt)\cos\varphi\}$$

$$\frac{dy}{dt} = c - 2\omega\{a\sin\varphi + b\cos\varphi\}t + \omega g\cos\varphi t^2$$

$$y = ct - \omega(a\sin\varphi + b\cos\varphi)t^2 + \frac{1}{3}\omega g\cos\varphi t^3$$

$$x = (a + 2\omega\sin\varphi d)t + \omega c\sin\varphi\frac{t^2}{2}$$

$$z = (b + 2\omega\cos\varphi d)t - (g + 2\omega c\cos\varphi)\frac{t^2}{2}$$

If for $t = 0$ $\dot{x}\,\dot{y}\,\dot{z} = 0$, $a = 0$, $b = 0$ $c = 0$, and $d = 0$

Then Reich's experiments in Freiburg[38]

$$x = 0$$

$$y = \frac{1}{3}\omega g\cos\varphi t^3$$

158,5 m $\varphi = \text{ca.}51°/$

$y = 27.5$ mm Reich fand 28.4.[40]

$$z = \frac{g}{2}t^2.$$

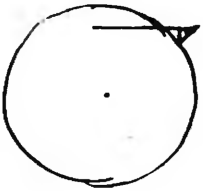
$$t = \sqrt{\frac{2}{g}z} \quad t^3 = \frac{2}{g}\sqrt{\frac{2}{g}}z^3$$

$$y = \frac{2}{3}\omega\sqrt{\frac{2z^3}{g}}$$ [39]

[39] [40]

Elementary Consideration

[p. 53]



$$\omega(\rho - \rho')\tau = y$$

$$\rho - \rho' = z\cos\varphi$$

$$z = \frac{g\tau^2}{2} \qquad \sqrt{\frac{2}{g}z} = \tau$$

$$y = \omega z\cos\varphi\sqrt{\frac{2}{g}z} = \omega\cos\varphi\sqrt{\frac{2}{g}z^3}$$

Factor $\frac{2}{3}$ missing.

Foucault's Pendulum

$$\begin{array}{l|c|c} \frac{d^2x}{dt^2} = -\frac{g}{l}x + 2\omega\sin\varphi\frac{dy}{dt} & dx & -y \\ \frac{d^2y}{dt^2} = -\frac{g}{l}y - 2\omega\sin\varphi\frac{dx}{dt} & dy & x \end{array}$$

$$d\left(\frac{v^2}{2}\right) = -\frac{g}{2l}r^2$$

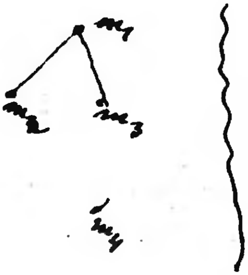
$$\frac{d}{dt}\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right) = \omega'\langle\sin\varphi\rangle\frac{d}{dt}(r^2)$$

$$\left(r^2\frac{d\vartheta}{dt}\right) = -\omega' r^2 + C.$$

Set $\vartheta + \omega' t = \vartheta'$. Then the ordinary pendulum law.

Dynamics of the Systems

[p. 54]



For any of these points

$$m\frac{d^2x}{dt^2} = \Sigma X_i + \Sigma X_a$$

Law of the momentum.

Law of the Center of Gravity

$$\Sigma m \frac{d^2x}{dt^2} = \Sigma\Sigma X_i + \Sigma\Sigma X_a$$

$$\Sigma\Sigma X_i = 0$$

for $\Sigma m \frac{d^2x}{dt^2}$ we can then set

$$\frac{d}{dt}\left\{\Sigma m \frac{dx}{dt}\right\} = \Sigma\Sigma X_a$$

- - - - - - - - - -

- - - - - - - - - -

{ } = ΣX components of the momenta of all points of the system = momentum of the system.

A different mode of expression We define the center of gravity $\xi\eta\zeta$ of a system of masses

$$M\xi = \Sigma mx$$

etc.

[p. 55] thus $\Sigma m \frac{dx}{dt} = M \frac{d\xi}{dt}$

Thus, $M \frac{d^2\xi}{dt^2} = \Sigma\Sigma X_e$

etc.

The center of gravity of a syst moves like a mat p. of mass M that is acted upon by the resultant of all external forces of the system.

Such a law is necessary because the dynamics of the mat. point can otherwise not be maintained.

Examp. Heavy body falling freely in empty space.

Examp. Masses attracted by the center proportionally to their distance & mass.

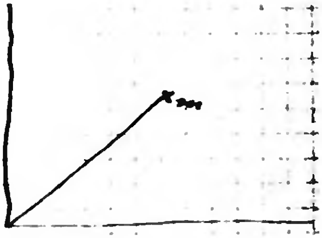
Force$_x$ $= -\kappa mx$

Resultant $= -\kappa\,\Sigma mx = -M\xi\kappa$

$$\cancel{M}\frac{d^2\xi}{dt^2} = -\kappa\cancel{M}\xi$$

- - - - - - - -

- - - - - - - -



Identical with the equation of motion of a single mat. point. Central force, thus in the plane, motion in ellipse as with spherical pendulum with small amplitude.
Examp. Recoil of firearms. $MV + mv = 0$
Examp. Vibration of the ground caused by the operation of a machine with reciprocating parts.
Examp. Light pressure. Mass of energy.[41]

Law of Moments of Momenta

[p. 56]

Repetition of the law of areas for the mat point. Here

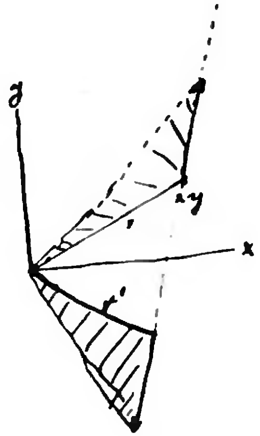
$$m\left(x\frac{d^2y}{dt^2} - y\frac{d^2x}{dt^2}\right) = \Sigma(xY_i - yX_i) + \Sigma(xY_a - yX_a)$$

$$\frac{d}{dt}\left\{m\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right)\right\} = \cdots\cdots$$

$$\frac{d}{dt}\,\Sigma m\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right) = \Sigma\Sigma(xY_i - yX_i) + \Sigma\Sigma(xY_a - yX_a)$$

$$xY_i - yX_i$$

Can also be conceived vectorially



The sum of the moments of internal forces vanishes. What remains is.

$$\frac{d}{dt}\sum m\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right) = \sum\sum xY_a - yX_a.$$

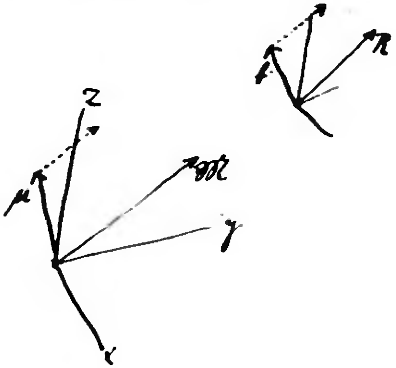
If system is closed, then sum of areal velocities const with respect to every plane.

$\sum m\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right)$ is then const. As regards rotation, analogous to the law of mot. of the mom of momentum. Difference.

[p. 57]

Law of Areas for Isolated Systems

<Graphic> Geometr. depiction.



If $\mathfrak{M} = 0$, then μ is a constant.[42] If the $X - Y$ plane is chosen ⊥ to μ, then only one component of the resultant momentum, and this permanently.

The moment of momentum of a solid body rotating about an axis.[43]

$$\frac{d}{dt}\sum m\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right) = \sum(xY_a - yX_a) = 0$$

$$I\omega + M\omega' = 0$$

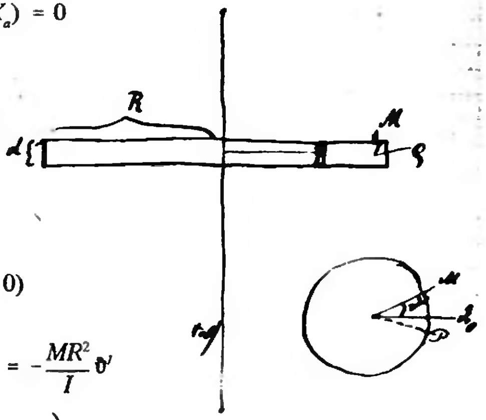
$$I\frac{d\vartheta}{dt} + MR^2\frac{d\vartheta'}{dt} = 0$$

Angle person traverses relative to the body is $\vartheta - \vartheta'$

At the start $\vartheta = 0$ & $\vartheta' = 0$

At the end $\vartheta' - \vartheta = 2\pi$ Now,

$I\vartheta + M\vartheta' = \text{const} = 0$ (because for $t = 0$)



$$\vartheta = -\frac{MR^2}{I}\vartheta'$$

$$\left(1 - \frac{MR^2}{I}\right)\vartheta' = 2\pi$$

$$r^2 \cdot \underline{2\pi} r \cdot dr \cdot \underline{d} \cdot \underline{\rho} = 2\pi d\rho \cdot \frac{r^4}{4} = \frac{1}{2\pi} dr^4 \rho.$$

The story of the cat.[44]

). 58] Complete the discussion about reaction of machines on foundation.

The moment of the momentum must also be constant over time

$$\sum m \left(y\frac{dz}{dt} - z\frac{dy}{dt} \right) = \text{const.}$$

extended over all moving masses of the machine. Otherwise torques & thus vibration of the foundation.

The law of <areas> moments of mom. with respect to the center of gravity.

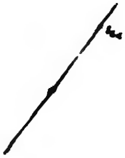
Derivation.

If the moments vanish.

Heavy rigid body in homog. gravitational field.

Moments vanish (mov), thus the law of areas.

Special case of rod.



$$x' = ra \qquad \frac{dx'}{dt} = r\dot{a}$$

$$y' = rb \qquad \frac{dy'}{dt} = r\dot{b}$$

$$z' = rc \qquad \frac{dz'}{dt} = r\dot{c}$$

Thus, the law of areas $(a\dot{b} - b\dot{a})\sum(mr^2) = c$ etc.

Deformable bodies in gravitational field. The law of areas holds as regards the center of gravity. The cat again.

The Law of Kinetic Energy[45]

$$\varphi(r) \cdot \frac{x}{r} \cdots$$

). 59]

$$X_i = \varphi(r)\frac{(x - x')}{r} \qquad X_i \; Y_i \; Z_i$$

$$X_i' \; Y_i' \; Z_i'$$

Derivation ....

Work by internal forces

Sum over all combinations

$$\varphi(r)\left\{\frac{x}{r}dx + \cdot + \cdot - \frac{x'}{r}dx' + \cdot + \cdot\right\}$$

$$r = \sqrt{(x - x')^2 +}$$

of
$$\cancel{F_{jk}\cdot\left(\frac{x-x'}{r}dx+\cdot+\cdot\right) - F_{jk}\left(\frac{x-x'}{r}\,dx'+\cdot+\cdot\right)}$$
$$= F_{jk}\left\{\frac{(x-x')d(x-x')+.+}{r}\right\} = F_{jk}\cdot dr_{jk}$$

$$\cancel{dr = \frac{(x-x')(dx-dx')+\cdot+}{r}}$$
$$\cancel{\varphi(r) = \frac{\partial\Phi(r)}{\partial r} \qquad \frac{x-x'}{r} = \frac{\partial r}{\partial x}}$$

hence $X_i = \frac{\langle\partial\rangle\, d\Phi}{\partial r}$

We assume that the F depend on the distance alone.

$$F = -\frac{d\varphi_{jk}}{dr} \quad \text{Then work} = \Gamma dr = -d(\varphi_{jk})$$

$$\text{Total work} = -d(\Sigma\varphi_{jk}) = -d\Pi$$

From this the energy principle holds in pure mechanics.

A part of the external & internal forces of the system may consist in binding forces (threads stationary surfaces etc.), especially important the special force where binding forces do not perform work. In that case the energy principle holds without these bind. forces appearing in it.

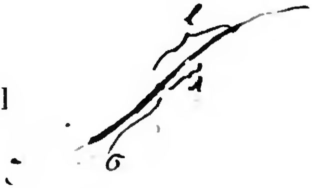
Systems with *one* degree of freedom can be solved *compl.* with the aid of the law of kinetic energy.

Example. Gliding chain.

$z = \varphi$ (bend)

Work = $-\rho\, d\lambda\, dz = -\rho g d\lambda\ \varphi'(\sigma+\lambda)d\sigma$

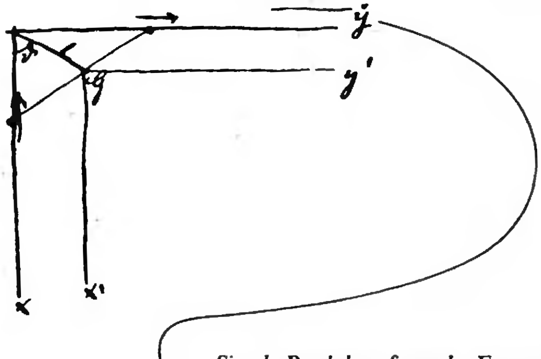
Integrated over λ: $-\rho g d\sigma[\varphi(\sigma+l) - \varphi(\sigma-l)]$



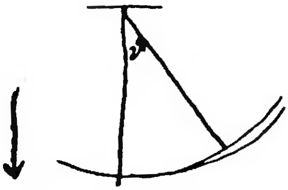
kin. energy $\rho l\left(\frac{d\sigma}{dt}\right)^2$

d(kin. energy) = work. When is motion independent of l?

[p. 60] König's theorem.[46] The law of kin energy with regard to the center of gravity.



Simple Pendulum from the Energy Principle



$z = l \cos \vartheta$

Work $mg\,dz = -\,mgl \sin \vartheta \, d\vartheta$

$$d\left(\frac{\cancel{m}l^{\cancel{2}}}{2}\left(\frac{d\vartheta}{dt}\right)^2\right)$$

$$-\frac{2g}{l}\sin\vartheta = \frac{d^2\vartheta}{dt^2}$$ [47]

Pot Energ $-\,mgz$

$$-\cancel{m}g\cancel{l}\cos\vartheta + \frac{\cancel{m}l^{\cancel{2}}}{2}\left(\frac{d\vartheta}{dt}\right)^2 = \text{konst.}$$

$$-\frac{\langle 2\rangle g}{l}\cos\vartheta + \left(\frac{d\vartheta}{dt}\right)^2 = \text{konst.}$$

Moments of Inertia[47]

1) Moment of inertia w. resp. to plane $\Sigma m\delta^2$	Σmx^2
2) Moment of inertia w. resp. to axis Σmr^2	$\Sigma m(x^2+y^2)$
3) Moment of inertia w. resp. to point	$\Sigma m(x^2+y^2+z^2)$

[p. 61] Example: moment of inertia of a sphere w. resp. to the center.

$\Sigma\, mr^2$ first over shell of the sphere $= r^2\,\Sigma\, m$ for the shell

$$4\pi r^2 dr\cdot\rho = \sum m \quad \text{over the shell}$$

$$r^2\sum m = 4\pi\rho\, r^4 dr$$

Integrated. over sphere yields $\frac{4}{5}\pi\rho R^5$

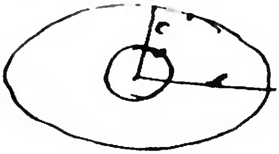
With respect to a plane $\frac{4}{15}\pi\cdot\rho\cdot R^5$

With respect " " axis $\frac{8}{15}\pi\rho R^5 = \frac{2}{5}R^2 M \quad k = R\sqrt{\frac{2}{5}}$

Example: Homog ellipsoid $\sum mz^2 = \sum \rho d\sigma z^2 = I_{xy}$

$$x' = \frac{x}{a} \quad y' = \frac{y}{b} \quad z' = \frac{z}{c}$$

Then boundaries unit sphere

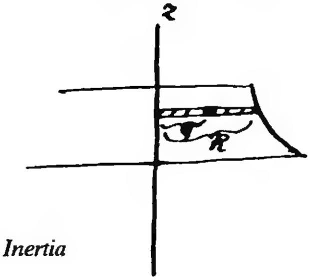


$$d\sigma' = \frac{1}{abc}d\sigma \quad z^2 = c^2 z'^2$$

hence $I_{xy} = \rho\ abc^3 \int d\sigma' z'^2 = \frac{4}{15}\pi\rho abc^3 = \frac{1}{5}Mc^2$

Example. Body of rotation bounded by two planes.

$$\rho dz \int_0^R 2\pi r dr\cdot r^2 = \frac{\pi}{2}R^4 dz\cdot\rho$$



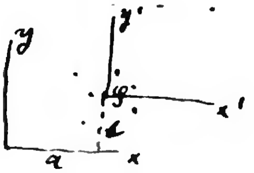
For the whole body of rotation

$$I_z = \frac{\pi\rho}{2}\int R^4 dz \quad R = \varphi(z)$$

Examp cylinder. $\frac{\overline{\pi\rho}}{2}R^{\overset{2}{4}}\overline{h} = M\frac{R^2}{2} \quad k = \frac{R}{\sqrt{2}}$

General Theorems on Moments of Inertia

1) Known I for axis through center of gravity.
Sought for arbitrary axis.



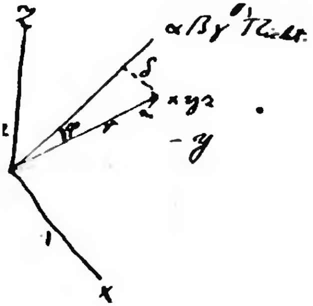
$$I = \sum m(x^2 + y^2) = \sum m\{(x' + a)^2 + (y' + b)^2\}$$

$$= \sum m(x'^2 + y'^2) + 2\overbrace{\sum max'}^{0} + 2\overbrace{\sum mby'}^{0} + (a^2 + b^2)\sum m$$

From that the theorem. [p. 62]

analogous theorems for the other two kinds of moments of inertia.

2) Moment of inertia dependent on direction.



$$I = \sum m\delta^2 = \sum mr^2 \sin^2\varphi$$

$$= \sum m(r^2 - (r\cos\varphi)^2)$$

$$r\cos\varphi = \alpha x + \beta y + \gamma z$$

$$r^2 - r^2\cos^2\varphi$$

$$= (x^2 + y^2 + z^2)(\alpha^2 + \beta^2 + \gamma^2) - (\alpha x + \beta y + \gamma z)^2$$

$$m|x^2(\beta^2 + \gamma^2) + \cdot + \cdot - 2\beta\gamma yz - \cdot - \cdot$$

In sum

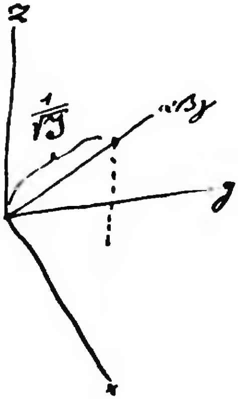
$$\alpha^2\{\sum m(y^2+z^2)\}+\beta^2\{\sum m(z^2+x^2)\}+\gamma^2\sum m(x^2+y^2)-2\beta\gamma\sum yz+\cdot+\cdot$$

$$I = A\alpha^2 + B\beta^2 + C\gamma^2 - 2D\beta\gamma - 2E\gamma\alpha - 2F\alpha\beta.$$

$$X = \frac{\alpha}{\sqrt{I}} \qquad Y = \frac{\beta}{\sqrt{I}} \qquad Z = \frac{\gamma}{\sqrt{I}}$$

$$\alpha = X\sqrt{I} \quad \text{etc.}$$

$$1 = AX^2 + BY^2 \ldots. - 2FXY$$



If the axes of this ellipsoid are chosen as coordinate axes, then D, E, F vanish. The axes are called major axes of inertia with respect to the point. Condition for the major axis of inertia z to be symmetrical to the xy plane, thus, for it not to change when the sign of γ changes. So, $D = E = 0$.

Each of the major axes of inertia with respect to the center of gravity is also the major axis of inertia with respect to any other of its points.

$$D = \sum yz = 0 \quad E = \sum zx = 0 \quad z = \alpha + z'$$

$$D = \sum y(\alpha + z') = 0$$

$$\sum \alpha y = 0$$

hence $D = D' = 0$.

Systems of Forces Acting on a Rigid Body

[p. 63]

The position of a rigid body is determined by 6 variables (6 degrees of freedom). We need therefore 6 equations for the complete determination of its motions. These 6 necessary and also sufficient equations are furnished by the momentum law and the law of moments.

$$\frac{d}{dt}\left\{\sum\left(m\frac{dx}{dt}\right)\right\} = \sum\sum X_a \qquad \frac{d}{dt}\left\{\sum m\left(y\frac{dz}{dt} - z\frac{dy}{dt}\right)\right\} = \sum\left(yZ_a - zY_a\right)$$

— — — — — — — — — — — — —

— — — — — — — — — — — — —

Since these equations must completely suffice for the calculation of the motion of the motion of the rigid body, two systems of forces acting on a rigid body are equivalent if they have the same geometric sum of forces and moments. Such systems of forces can be substituted for one another.

From this follows the elementary law of displaceability of the point of application of a force vector along the straight line in which the force vector lies.

$$X = X'' = K\alpha$$

$$\left.\begin{aligned} yZ - zY &= 2\Delta \\ y'Z' - z'Y' &= 2\Delta' \end{aligned}\right\}\Delta = \Delta'$$

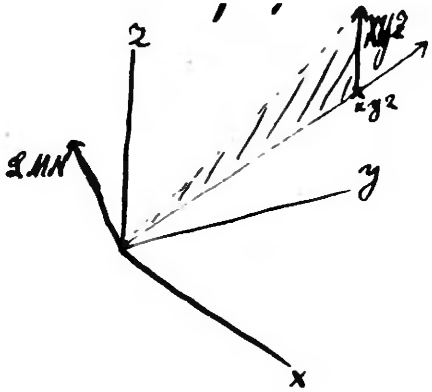
Proof also vectorially

[p. 64]

2 vectors are called equal if they have equal components

$(XYZ) = (X'Y'Z')$

$$\left.\begin{aligned} yZ - zY &= L \\ zX - xZ &= M \\ xY - yX &= N \end{aligned}\right\} \quad (L\,M\,N) \text{ can be represented by a vector that is constructed as follows}$$



This construction shows immediately that the force can be displaced along its straight line without changing the moment.

From this the law that two forces acting along the same line in opposite directions cancel each other.

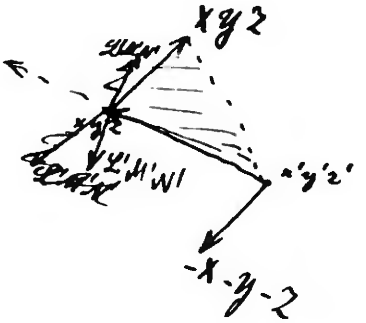
We put $\sum X_a = X^x \cdot \cdot \cdot \cdot \sum yZ_a - zY_a = \sum N = N^x$.

The effect of the system of forces on rigid body is completely determined by $X^x\,Y^x\,Z^x \cdot \cdot \cdot \cdot$ N^x. $X^x\,Y^x\,Z^x$ resultant force. $L^x\,M^x\,N^x$ resultant moment. In general, it is impossible to indicate *one* force that is equivalent to a system of forces. We seek simplest possible [p. 65] representation of a system of forces. To this end pairs of forces introduced. Two forces of equal magnitude and opposite direction.
We investigate this system of forces
and obtain

$$X^x = X - X = 0$$

$$\text{----------}$$

$$\text{----------}$$



$$N^{x} = (xY - yX) + (x'\ (-y) - y'(-X))$$

$$= (x - x')Y - (y - y')X$$

- - - - - - - - - - - -

- - - - - - - - - - - - - -

The force couple has no resultant force vector but only a moment of rotation. This is a vector product of the point-connecting vector and the force vector.

1) <Direction of ve> absolute position of the system without effect on the moment
2) Magnitude = 2 Δ = force · distance
3) Direction and sense of the vector ⊥ to the plane of force couple ([*r* ?], force, mom.) = right-hand system Force couple determined by vector with a totally arbitrary point of application. Each moment vector can be replaced by a force couple.

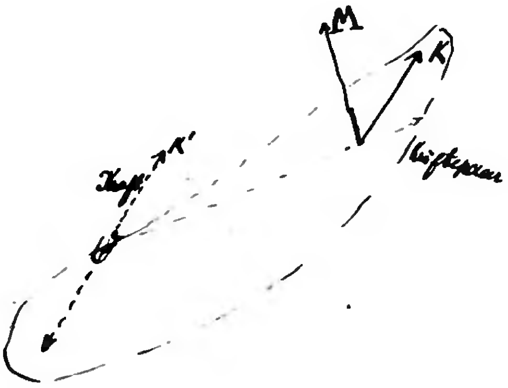
From what has been said, it follows that

1) arbitrary system of forces acting on a rigid body can be replaced by
the system $X^{x} Y^{x} Z^{x}$ $L^{x} M^{x} N^{x}$
thus also by $x^{x} y^{x} z^{x}$ 0 0 0
and 0 0 0 $L^{x} M^{x} N^{x}$

That is, by a force through the coordinate origin and force couple. [p. 66]

Geometric derivation.

K′ and M separately added geometrically and each combined to a resultant.



Equilibrium Condition for Rigid Body

$$X^x = Y^x = \ldots\ldots\ldots N^x = 0$$

Special case. All forces lie in a plane.

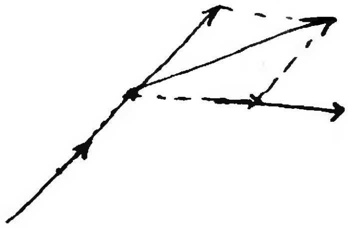
Analytically

$$X^x = \Sigma X = 0$$

$$Y^x = \Sigma Y = 0$$

$$N^x = \Sigma xY - yX = 0$$

Graphically



Here replaceability by resultant force.

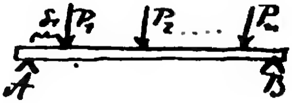
[p. 67] Even more special, all forces ∏ X-axis

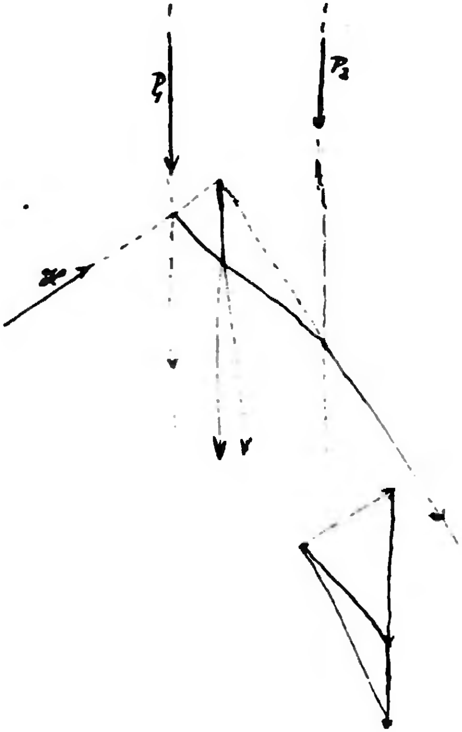
Then $\Sigma X = 0 \qquad -\Sigma yX = 0$

Example. Reaction of beam resting at 2 places[48]

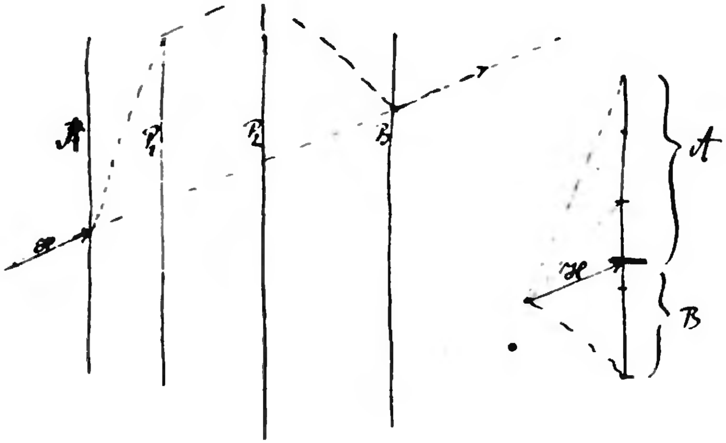
$A + B = \Sigma P \quad Bl = \Sigma P\delta$

If forces ∥, then treated graphically as follows





[p. 68]



2. Special case, all points have the same direction, but points of application

$$X_\nu = P_\nu \alpha; \qquad Y_\nu P_\nu \beta; \qquad Z_\nu = P_\nu \gamma^{[49]}$$

Seeking resultant system

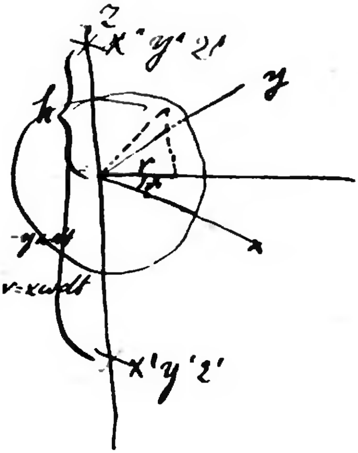
$$X^x = \alpha \sum P_\nu \qquad L^x = \sum (y_\nu Z_\nu - z_\nu Y_\nu) = \gamma \sum y_\nu P_\nu - \beta \sum z_\nu P_\nu$$

$$Y^x = \beta \sum P_\nu \qquad M^x = \quad \text{— — — — — — — — —}$$

$$Z^x = \gamma \sum P_\nu \qquad N^x = \quad \text{—————————}$$

Now we choose the coordinate origin such that $\Sigma x_\nu P_\nu = \Sigma y_\nu P_\nu = \Sigma z_\nu P_\nu = 0$. Is always possible unless $\Sigma P_\nu = 0$. Then moment vanishes for all $\alpha \beta \gamma$. Thus, if one also changes directions, one can always replace the forces by resultants passing through the coordinate origin. Center of gravity of the force system Case of gravity special case. Here the magnitude of the force acting on the individual point is $P_\nu = m_\nu g$. Thus $\Sigma P_\nu x_\nu$ [p. 69] $= g \Sigma m_\nu x_\nu = 0$. I.e., resultant passes through the point that we have designated in the general dynamics of systems as the center of gravity of the system.

Motion of a Rigid Body about an Axis



$$\sum m\frac{v^2}{2}=\frac{I}{2}\omega^2$$

$$d\left(\frac{I}{2}\omega^2\right)=\sum(X\,dx+Y\,dy+Z\,dz)$$

$$x=r\cos\vartheta \qquad dx=-r\sin\vartheta\,d\vartheta+\overset{0}{\cos}\,\vartheta\,dr=-y\omega\,dt$$

$$y=r\sin\vartheta \qquad dy=r\cos\vartheta\,d\vartheta+\overset{0}{\sin}\,\vartheta\,dr=x\omega\,dt$$

$$z=z \qquad dz=0$$

$$d\left(\frac{I}{2}\omega^2\right)=\omega\,dt\sum(xY-yX)$$

$$I\frac{d\omega}{dt}=\sum xY-yX$$

We now insert the reaction of the axis (X',Y',Z) & $(X'',Y'',Z''$ & treat the body as free. We obtain[50]

$$\sum m\frac{d^2x}{dt^2}=X'+X''+\sum X \qquad \sum m\left(y\frac{d^2z}{dx^2}-z\frac{d^2x}{dt^2}\right)=\sum(yZ-zY)-hY''$$

$$\sum m\frac{d^2y}{dt^2}=Y'+Y''+\sum Y \qquad \sum m\left(z\frac{d^2x}{dt^2}-x\frac{d^2z}{dt^2}\right)=\sum(zX-xZ)+hX''$$

$$\sum m\frac{d^2z}{dt^2}=Z'+Z''+\sum Z$$

$$\begin{array}{ll|c|c}
\frac{dx}{dt}=-\omega y & \frac{d^2x}{dt^2}=-\omega\frac{dy}{dt}-y\frac{d\omega}{dt}=-\omega^2x-y\frac{d\omega}{dt} & z & \\
\frac{dy}{dt}=\omega x & \frac{d^2y}{dt^2}=+\omega\frac{dx}{dt}+x\frac{d\omega}{dt}=-\omega^2y+x\frac{d\omega}{dt} & & -z \\
\frac{dz}{dt}=0 & \frac{d^2z}{dt^2}=0 & -x & y
\end{array}$$

[p. 70]

$$-\omega^2 \sum mx - \frac{d\omega}{dt} \sum my = X' + X'' + \sum X$$

$$-\omega^2 \sum my + \frac{d\omega}{dt} \sum mx = Y' + Y'' + \sum Y$$

$$0 = Z' + Z'' + \sum Z$$

$$2) \quad -\omega^2 \sum mxz - \frac{d\omega}{dt} \sum myz = \underline{hX''} + \sum (zX - xZ)$$

$$1) \quad +\omega^2 \sum myz - \frac{d\omega}{dt} \sum mxz = -\underline{hY''} + \sum (yZ - zX)$$

$$I \frac{d\omega}{dt} = \sum (xY - yX)$$

We seek to reformulate equations in such a way that ω shall be the only variable occurring. To this end we introduce a co-rotating system

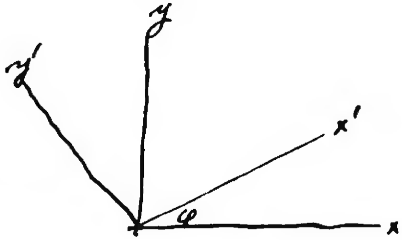
$$x = x' \cos \varphi - y' \sin \varphi$$

$$y = x' \sin \varphi + y' \cos \varphi$$

$$z = z'$$

$$\sum myz = \sin \varphi \sum x'z' + \cos \varphi \sum y'z'$$

$$\sum mxz = \cos \varphi \sum x'z' - \sin \varphi \sum y'z'$$



We consider the special case where the given external forces are not present. [?] In that case the sum on the right-hand side vanishes.

When does P'' experience no reaction? or when does one need no X'', Y'' to maintain the rotation about the axis? We must have $\Sigma mxz = \Sigma myz = 0$. The Z-axis, i.e., the axis of rotation, must be a major axis of inertia with [p. 71] respect to P''.

In addition to that, when does P' experience no reaction? We must have $\Sigma mx = \Sigma my = 0$. Thus, the axis of rotation must pass through the center of gravity, and one of the major axes of inertia must be through the center of gravity. So that none of the two bearings would experience a reaction. No reactions will then take place even if the body is accelerated by a torque N.

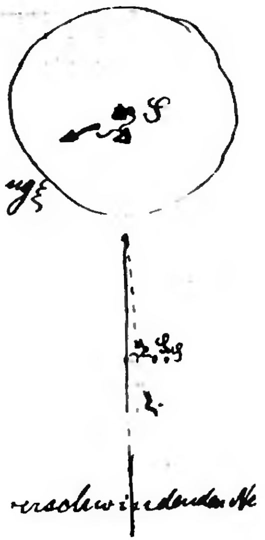
Example of rigid bodies on elastic axis

1) Disc perpendicular to the Z-axis. Symmetric with respect to the x - Y plane. $\Sigma xz = \Sigma yz = 0$. $X'' = Y'' = 0$ The first two equations yield[51]

$$-\omega^2\sum mx = (X' + X'') \qquad -\omega^2 M\xi = X' + X''$$

&

$$-\omega^2\sum my = (Y' + Y'') \qquad -\omega^2 M\eta = Y' + Y''$$



$$\omega^2 M\Delta = K_a$$

Let us have an elastic axis that is mounted somewhat eccentrically. Rotation is accompanied by a sag ξ[52]

$$\Delta = \Delta_0 + \xi$$

On the other hand, we will have $K_a = E\xi$ The above reads accordingly

$$\omega^2 M(\Delta_0 + \xi) = E\xi$$

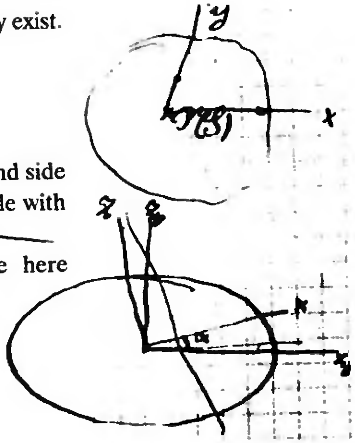
$$\xi = \frac{\omega^2 M\Delta_0}{e - \omega^2 M}$$

Critical Angular Velocity for Vanishing Denominator

[p. 72]

We have here considered the case in which the axis does not pass through the center of gravity. Now let the axis pass through the center of gravity but not be a major axis of inertia. Let the center of gravity coincide with P'. No external forces may exist. From equations 4 & 5 we then obtain

$$-\omega^2\,\underline{\sum mxz} = h\underline{X}''$$
$$-\omega^2\,\underline{\sum myz} = h\underline{Y}''$$

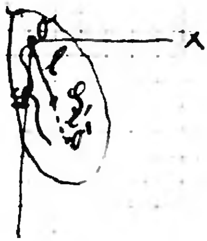


If the Z-axis is the major axis of inertia, then the right-hand side vanishes. But if the major axis of inertia does not coincide with the axis of rotation, but about Y-axis

The sign and the magnitude of the reaction are here determined by the centrifugal moments.

Physical Pendulum

$$I\frac{d^2\varphi}{dt^2} = \sum xY - yX = -M\,gl\,\sin\vartheta$$ [53]



& if one sets $I = Mk^2$

$$\frac{d^2\vartheta}{dt^2} = -\frac{gl}{k^2}\sin\vartheta$$

[p. 73] This determines the motion. Synchronous with simple pendulum of length L,

$$\frac{d^2\vartheta}{dt^2} = -\frac{g}{L}\sin\vartheta \text{ if } L = \frac{k^2}{l}.$$

We now introduce the radius of inertia (radius of gyration) for the center of gravity. We have

$$I = I_s + Ml^2$$

or $$Mk^2 = Mk_s^2 + Ml^2$$

& $$k^2 = k_s^2 + l^2$$

Substituted in the above relation

$$L = l + \underbrace{\left(\frac{k_s^2}{l}\right)}_{l'}$$

$$\frac{\partial L}{\partial l} = 0 \qquad 1 - \frac{k_s^2}{l^2} = 0$$

$$\underline{l = k_s}$$

$$L_{\min} = 2\,k_s$$

We now imagine that the pendulum is suspended at O′ and that the L^x of the accompanying seconds pendulum has been determined. For this we can apply the relation just obtained and insert the quantity l' instead of l:

$$L^x = l' + \frac{k_s^2}{l'} \text{ or, because } l' = \frac{k_s^2}{l}$$

$$L^x = \frac{k_s^2}{l} + l, \text{ hence } = L\ .$$

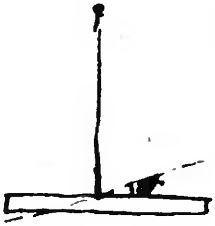
The relation between the points 0 and 0′ is thus reciprocal. Therefore, by observing the oscillation of any rigid pendulum one can find out the length of a mathematical pendulum of the same oscillation period.

The above formula makes it possible to determine the minimum oscillation period on the axis.

Further discussion of above formula. Minimal oscillation period that can be attained with the body about the axis of the minor moment of inertia.

Experimental Determination of Moments of Inertia & Torsional Forces by Means of Oscillations of a Suspended Rigid Body

[p. 74]



$$I\frac{d^2\vartheta}{dt^2} = -\Theta\vartheta$$

$$\vartheta = A\sin\sqrt{\frac{\Theta}{I}}t \qquad \frac{2\pi}{T} = \sqrt{\frac{\Theta}{I}}$$

$$T = 2\pi\sqrt{\frac{I}{\Theta}}$$

It is not possible to determine both quantites Θ & I from such measurements, but this can be done using a second experiment in which the moment of inertia is increased by adding two cylinders.[54] For each such cylinder

$$I_z = m\{k^2 + \delta^2\} \quad \text{where } k = \frac{R}{\sqrt{2}}$$

$$I_z = m\left\{\frac{1}{2}R^2 + \delta^2\right\}$$

$$I_{tot} = m\{R^2 + 2\delta^2\}$$

$$T' = 2\pi\sqrt{\frac{I + m(R^2 + 2\delta^2)}{\Theta}}$$

The two equations yield R and Θ. Modification of the method, in case torsion not independent of the added weights.

[p. 75]

General Principles of Mechanics

Principle of virtual moments (statics).
Equilibrium of the point
Equilibrium condition of a point
$X = Y = Z = 0$

We think of the point as infinit. displaced δx δy δz
Work of the force $\delta A = X\delta x + Y\delta y + Z_{\delta}z \underline{\underline{= 0}}$.
No joke. Happens only when a part of the forces is not given but determined by conditions (connections). These forces have the characteristic property that their work vanishes in its entirety. Let this always be assumed.

Example Point is forced to stay on a stationary surface $(f(x,y,z,\langle t\rangle) = 0$

Force of the surface on the point $\lambda\frac{\partial f}{\partial x}, \lambda\frac{\partial f}{\partial y}, \lambda\frac{\partial f}{\partial z}$.

Total force when, in addition, another force X, Y, Z acts on the point.

$$X + \lambda\frac{\partial f}{\partial x}$$

$$Y + \lambda\frac{\partial f}{\partial y}$$

$$Z + \lambda\frac{\partial f}{\partial z}$$

Equilibrium condition. Vanishing of those I components.
Can again be replaced by

$$\left(X + \lambda\frac{\partial f}{\partial x}\right)\delta x + \left(Y + \lambda\frac{\partial f}{\partial y}\right)\delta y + \left(Z + \lambda\frac{\partial f}{\partial z}\right)\delta z = 0.$$

This holds for every arbitrary system δx, δy, δz. But for those displacements in which [p. 76] the point does not leave the surface, we have the relation

$$\lambda\frac{\partial f}{\partial x}\delta x + \cdot + \cdot = 0.$$

(Special case of the law that the connective forces do not perform work.) If we confine ourselves to the consideration of *such* displacements, then the connective forces make no contribution to the virtual work. Thus, for such displacements, which do not violate the conditions, we have the equation

$$X\,\delta x + Y\,\delta y + Z\,\delta z = 0$$

if δx δy & δz are connected by the relation

$$\frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial z}\delta z = 0$$

These equations are really sufficient for the calculation of the coordinates of the equilibrium position. Because if one eliminates δx from the first equation by means of the second one, one obtains an equation of the form $B\,\delta y + C\,\delta z = 0$. This is satisfied for an arbitrary choice of δy & δz only if one chooses $B = 0$ and $C = 0$. To these two equations is added as the third eq. $f = 0$.

Generalization. Let there be a system of n material points $P_1\,P_2 \cdots P_n$ We seek the [p. 77] general condition for the equilibrium of this system of points. Let each point be acted upon by connective forces $X_\nu\,Y_\nu\,Z_\nu$ and explicitly considered forces $X\,Y\,Z \cdots$ Then we have for each point

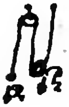
$X + X_\nu = 0$
- - - - - hence also $(X + X_\nu)\delta x + \cdot + \cdot = 0$
- - - - -

thus also $\sum_n (X + X_\nu)\delta x + \cdot + \cdot = 0$

This equation holds for any arbitrary displacement of the points (even for one incompatible with the given conditions). But if the displacements are chosen such that the conditions are not violated by them, then the connective forces do not perform any work during displacement, i.e., we have $\Sigma(X_0\delta x + \cdot + \cdot) = 0$. Thus, we also have $\Sigma X\delta x + \cdot + \cdot = 0$. The sum of the virtual works vanishes for every virtual displacement compatible with the conditions of the system. Proof that sufficient restrictive eq. for the solution of the problem.

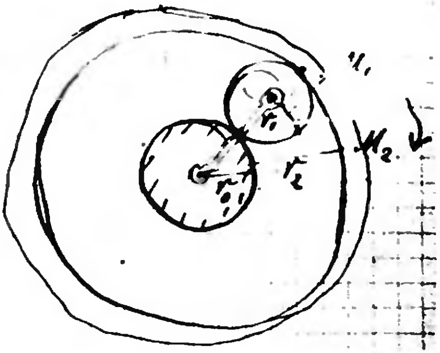
The advantage of this principle consists in the fact that the connective forces do not [p. 78] have to be investigated & that the virtual work can often be calculated without using a Cartesian coordinate system.

$P_2 = 2P_1$



* * *

Example. Epicyclic wheel. Consider infinitely small rotation of the outer wheel. Infinitely small rotation of the outer wheel α_2, the epicentric wheel α_1, the epicentric arm α_{01}



There are two conditions between these displacements because 0-1 & 1-2 do not slide, i.e., they experience identical displacements with their places of contact. Hence, we must have

$$\left.\begin{aligned}(r_0+r_1)\alpha_{01}-r_1\alpha_1&=0\\(r_0+r_1)\alpha_{01}-r_1\alpha_1&=r_2\alpha_2\end{aligned}\right\}\text{ eliminate } \alpha_1$$

$$2(r_0+r_1)\alpha_{01}=r_2\alpha_2$$

According to the principle of virtual moments

$$M_{01}\alpha_{01}+M_2\alpha_2=0$$

$$M_{01}+M_2\frac{2(r_0+r_1)}{r_0}=0$$

Special case of the principle when the forces are derivable from a potential
If Φ is the potential energy, then we have for each point

$$X_\nu=\frac{\partial\Phi}{\partial x_\nu}\qquad Y_\nu=\frac{\partial\Phi}{\partial y_\nu}\qquad \text{——}$$

Principle then takes the form $\sum\left(\frac{\partial\Phi}{\partial x_\nu}\delta x_\nu+\cdot+\cdot\right)=0$

or $\delta(\Phi)=0$ for every displacement compatible with the conditions.
Further, if only a part of the explicitly considered forces (e.g., all apart from the external [p. 79] forces) are derivable from a potential, and we call them residual (e.g., external) forces $X\ Y\ Z$, then we can write

$$\sum X\delta x+Y\delta y+Z\delta z-\delta\Phi=0$$

D'Alembert's Principle

Considerations that are analogous to those regarding the eqilibrium of the material point.
Equations of motion of the mat point (freely movable)

$$\begin{array}{r|l} m\frac{d^2x}{dt} - X = 0 & \delta x \\[2em] m\frac{d^2y}{dt^2} - Y = 0 & \delta y \qquad \left(m\frac{d^2x}{dt^2} - X\right)\delta x + \cdot + \cdot = 0 \\[2em] - - - - - - & \delta z \end{array}$$

Triviality

Secondly, we assume that the point is subjected to two kinds of forces, namely explicitly considered forces and connective forces. Purely formally, as above, the following eq. will then be valid:

$$\left(m\frac{d^2x}{dt^2} - X - X_v\right)\delta x + \left(m\frac{d^2y}{dt^2} - Y - Y_v\right)\delta y + \cdot = 0$$

The quantities $\delta x\ \delta y\ \delta z$ determine for each moment a position infinitely close to the real position of the material point. We will now choose these infinitely close positions in such a way that at any moment the point could be displaced from $x\ y\ z$ to $x + \delta x,\ y + \delta y,\ z + \delta z$ without violating the conditions of the system.

[p. 80] If we are dealing, for example, with the motion of a m. p. on an arbitrarily moving surface, then let $\delta x\ \delta y\ \delta z$ for time t be chosen such that $f(x + \delta x, :, t) = 0$ The law to the effect that the work of the connective forces vanishes holds for such displacements. Thus, for example, for the mat. p. on a surface, because connective force ⊥ to the plane, but displacement *in* the surface. We will have, therefore, for such a displacement $X_v\delta x + Y_v\delta y + Z_v\delta z = 0$ Since the above equation is valid for every virtual displacement, and thus also for such ones that do not violate the conditions of the system, we also have for virtual displacements of the last-mentioned kind the equations

$$\left(m\frac{d^2x}{dt^2} - X\right)\delta x + \cdot + \cdot = 0$$

where $\delta x\ \delta y\ \delta z$ are connected by those relations to which the connective forces are to be traced.

Analogous argument for *systems* of material p. If we again introduce connective forces and other forces, we get

$$\sum_{1-n}\left(m\frac{d^2x}{dt^2} - X - X_\nu\right)\delta x + \cdot + \cdot = 0$$

for every system of δx δy δz. If $\delta x \cdot \cdot$ are specially chosen such that the conditions of the system are not violated, then

$$\sum_{1-n}(X_\nu \delta x + \cdot + \cdot) = 0 \ ,$$

so that subtraction yields [p. 81]

$$\delta x_\nu = \sum_q \frac{\partial x_\nu}{\partial q_k}\delta q_k$$

$$\sum_\nu \sum_q \left\{\left(m_\nu \frac{d^2x_\nu}{dt^2} - X_\nu\right)\frac{\partial x_\nu}{\partial q_k}\delta q_k + \cdot + \cdot\right\} = 0$$

$$\sum_{1 \text{ to } n}\left\{\left(m\frac{d^2x_\nu}{dt^2} - X_\nu\right)\delta x_\nu + \cdot + \cdot\right\} = 0$$

This is D'Alembert's principle. Next we have to show that this equation gives the solution of every problem of motion.

Let a virtual displacement that does not violate the conditions of the system be determined by k mutually independent quantities $\delta q_1 \cdot \cdot \cdot \cdot \delta q_k$ (k degrees of freedom)

$$\delta x_1 = \frac{\partial x_1}{\partial q_1}\delta q_1 + \frac{\partial x_1}{\partial q_2}\delta q_2 \cdots \frac{\partial x_1}{\partial q_k}\delta q_k \qquad \delta y_1 \cdots \delta z_1 \cdots$$

$$- - - - - - - - -$$

$$\delta x_n = \frac{\partial x_n}{\partial q_1}\delta q_1 - - - - - - \frac{\partial x_n}{\partial q_k}\delta q_k \qquad \cdots\cdots \qquad \cdots\cdots$$

If one sets
$$\sum_\nu m \frac{d^2x_\nu}{dt^2}\frac{\partial x_\nu}{\partial q_1} = Q_1 \qquad \sum_\nu X_\nu \frac{\partial x_\nu}{\partial q_1} = R_1$$

$$\sum_\nu m \frac{d^2x_\nu}{dt^2}\frac{\partial x_\nu}{\partial q_2} = Q_2 \qquad \sum_\nu Y_\nu \frac{\partial x_\nu}{\partial q_2} = R_2$$

$$\text{— — — — — — —} \qquad \text{————————}$$

then the above system of equations assumes the form

$$(Q_1 - R_1)\delta q_1 + (Q_2 - R_1)\delta q_2 \cdot\cdot\cdot\cdot\cdot (Q_k - R_k)\delta q_k = 0$$

Since all q_k are independent of each other, we have

$$Q_1 = R_1 \quad Q_2 = R_2 \cdot\cdot\cdot\cdot\cdot Q_k = R_k$$

These k equations are just sufficient for the solution of the problem.—

If the conditions between the $\delta x \cdot\cdot\cdot\cdot$ can be represented as equations between the $x_1 \cdot\cdot\cdot\cdot z_n$ & t, that is, in the form $f(x_1 \cdot\cdot\cdot\cdot z_n t) = 0$, then the system is said to be *holonomic*.

[p. 82] The equations of motion for such a system can be found in the following fashion, which was first presented by Lagrange. We have

$$\sum_\nu \left\{ X_\nu - m\frac{d^2x_\nu}{dt^2}\delta x_\nu + \cdot + \cdot \right\} = 0 \quad \text{and} \quad f_m(x_1 \cdots\cdot z_n, t) = 0$$

mc from 1 to *h*

For such displacements, which are compatible with these conditions, we have

$$\begin{array}{c|l}
\lambda_1 & \sum_\nu \left(\frac{\partial f_1}{\partial x_\nu}\delta x_\nu + \cdot + \right) = 0 \\
\lambda_2 & \sum_\nu \frac{\partial f_2}{\partial x_\nu}\delta x_\nu + \cdot + \cdot = 0 \\
 & \text{- - - -} \quad \text{———} \quad \text{———} \\
\lambda_h & \sum_\nu \frac{\partial f_k}{\partial x_\nu}\delta x_\nu + \cdot + \cdot = 0 \\
\hline
\end{array}$$

h variations can be expressed from the rest
$3n - h = k$ equations from the first

Restrictive equations mult. by factors λ & add.

$$\sum_{\nu}\left\{X_{\nu} - m\frac{d^2x_{\nu}}{dt^2} + \lambda_1\frac{\partial f_1}{\partial x_{\nu}} + \lambda_2\frac{\partial f}{\partial x_{\nu}} \cdots\cdot + \lambda_h\frac{\partial f}{\partial x_{\nu}}\right\}\delta x_{\nu} + \cdot + \cdot = 0$$

3*n* such {} are present. We can choose *h* of those = 0 by choosing λ in the appropriate way. These terms of the sum will then vanish. The δx δy δz of the rest are then arbitrary, since of the δ $3n - h = k$ can be chosen arbitrarily. From this it follows that the remaining {} must also vanish. Thus, one also obtains the equations of motion of a system of points in the form

[p. 83]

$$\left.\begin{array}{l} m_{\nu}\dfrac{d^2x_{\nu}}{dt^2} = X_{\nu} + \lambda_1\dfrac{\partial f_1}{\partial x_{\nu}} + \lambda_2\dfrac{\partial f_2}{\partial x_{\nu}} + \cdots \lambda_h\dfrac{\partial f_h}{\partial x_{\nu}} \\[2ex] m_{\nu}\dfrac{d^2y_{\nu}}{dt^2} = Y_{\nu} + \lambda_1\dfrac{\partial f_1}{\partial x_{\nu}} + \cdots\cdots\cdots\cdot \\[2ex] \text{— — — — — — — — —} \end{array}\right\} \quad \nu \text{ from } 1 \text{ to } n$$

Equations of Motion of Lagrange

Of historical interest only.[55]

<Principle of Least Action>
Hamilton's Principle. Lagrange's Eq. of Motion

We start out from d'Alembert's principle

$$\sum_{\nu}\left\{\left(X_{\nu} - m_{\nu}\frac{d^2x_{\nu}}{dt^2}\right)\delta x_{\nu} + \cdot + \cdot\right\} = 0$$

for all virtual displacements compatible with the cond. of the syst. We can put

$$-\delta\left(\frac{1}{2}\frac{dx_\nu^2}{dt}\right)$$

$$\frac{d^2x_\nu}{dt^2}\delta x_\nu = \frac{d}{dt}\left\{\frac{dx_\nu}{dt}\delta x_\nu\right\} - \frac{dx_\nu}{dt}\frac{d(\delta x_\nu)}{dt} = \frac{d}{dt}\left\{\frac{dx_\nu}{dt}dx_\nu\right\} - \frac{dx_\nu}{dt}\delta\left(\frac{dx_\nu}{dt}\right) =$$

The second term can be written in another form. We shall show that it is equal to the variation of velocity

$$\frac{d\,\delta x}{dt} = \frac{d(x+\delta x)}{dt} - \frac{dx}{dt} = \delta\left(\frac{dx}{dt}\right)$$

$$\frac{d^2x_\nu}{dt^2}\delta x = \frac{d}{dt}\left\{\frac{dx_\nu}{dt}\delta x\right\} - \delta\left\{\frac{1}{2}\frac{dx_\nu^2}{dt}\right\}$$

$$\underbrace{\sum X_\nu \delta x_\nu}_{\delta A} - \underbrace{\sum \frac{d}{dt} m_\nu \frac{dx_\nu}{dt}\delta x_\nu}_{0} + \underbrace{\sum \delta \frac{m_\nu}{2}\left(\frac{dx_\nu}{dt}\right)^2}_{\delta L} = 0$$

Integr. over time limits. All δ shall vanish at time limits.

$$\int_{t_0}^{t_1} (A_\delta + \delta L)dt = 0$$

[p. 84] We shall rewrite the work A_δ according to the follow. principle we choose indep. variables $p_1 \cdots\cdot p_n$, whose number is equal to the number of degrees of freedom. A_δ will then have the form $P_1\delta p_1 + P_2\delta p_2 \cdots\cdots$

We now specialize the problem a little. Let the forces be derivable in part from a force function. Let Π be the potential energy. The part of the virtual work A_0 derived from them is $-\sum_\nu \left(\frac{\partial\Pi}{\partial x_\nu}\delta x_\nu + \cdot + \cdot\right) = -\delta\Pi$

In addition, forces might be present that are either given or sought as functions of time. If X'_ν Y'_ν Z'_ν are components of these forces for 1 point, then the corresponding term of A' is

$$\sum (X'_\nu \delta x_\nu + \cdot + \cdot)$$

If we set $\delta x_\nu = \sum \frac{\partial x_\nu}{\partial p_\mu} \delta p_\mu$ [56]

then we recognize that the part of work considered can be expressed in the form $\sum\limits_\mu P_\mu \delta p_\mu$. If the P_μ are to be viewed as functions of time only, then we can set

$$\sum P_\mu \delta p_\mu = \delta \cdot \sum P_\mu \, p_\mu .$$

Substituting both terms for A_8 into the above formula, one obtains [p. 85]

$$\delta \left\{ \int\limits_{t_0}^{t_1} (\Pi - L - \sum (P_\mu \, p_\mu)) \delta t \right\} = 0$$

This is <the least action principle> Hamilton's principle. If all forces can be derived from a potential (Π), the latter assumes the simple & familiar form

$$\delta \left\{ \int\limits_{t_0}^{t_1} (\Pi - L) dt \right\} = 0.$$

In this principle the Cartesian coordinates of mass points no longer occur. It is valid no matter what coordinates we choose to determine the position of the points of the system.

Now we start out from the general form of Hamilton's principle

$$\int\limits_{t_0}^{t_1} (\delta L + A_8) dt = 0$$

We use general coordinates, which determine completely the state of the system $(p_1 \cdots p_n)$. Then we can set

$$A_\delta = \sum P_\nu \delta p_\nu$$

Further, we have to examine how L depends on the p_ν.

We have $L = \sum \frac{m_\nu}{2}\left(\frac{dx_\nu^2}{dt} + \cdot + \cdot\right)$ $\qquad x_\nu = \varphi_{\nu x}(p_1 \cdots p_n)$

$$\frac{dx_\nu}{dt} = \frac{\partial \varphi_{\nu x}}{\partial p_1}\frac{dp_1}{dt} + \cdot + \cdot$$

Thus, L_ν is a function of p_ν & $\frac{dp_\nu}{dt} = p_\nu'$

[p. 86] For that reason we have to set

$$\delta L = \sum \frac{\partial L}{\partial p_\nu}\delta p_\nu + \sum \frac{\partial L}{\partial p_\nu'}\delta p_\nu'.$$

Replacing A_δ and δL with their values, one obtains

$$\int \left\{\sum \frac{\partial L}{\partial p_\nu}\delta p_\nu + \sum \frac{\partial L}{\partial p_\nu'}\delta p_\nu' + \sum P_\nu \delta p_\nu\right\} dt = 0$$

The factors of δp_ν & $\delta p'_\nu$ do not have to vanish individually!
But we have

$$\frac{\partial L}{\partial p_\nu'}\delta p_\nu' = \frac{\partial L}{\partial p_\nu'}\frac{d}{dt}\delta p_\nu = \frac{d}{dt}\left(\frac{\partial L}{\partial p_\nu'}\delta p_\nu\right) - \delta p_\nu \frac{d}{dt}\left(\frac{\partial L}{\partial p_\nu'}\right)$$

But since the δp_ν must vanish at the limits, it follows that

$$\int_0^t \sum \left(\frac{\partial L}{\partial p_\nu} \delta p'_\nu \right) = \underset{\overset{\|}{0}}{\left| \frac{\partial L}{\partial p'_\nu} \delta p_\nu \right|_0^t} - \int_0^t \delta p_\nu \frac{d}{dt} \left(\frac{\partial L}{\partial p'_\nu} \right) dt$$

The above equation changes therefore into

$$\int_0^t \sum \left\{ \frac{\partial L}{\partial p_\nu} - \frac{d}{dt} \left(\frac{\partial L}{\partial p'_\nu} \right) + P_\nu \right\} \delta p_\nu dt = 0$$

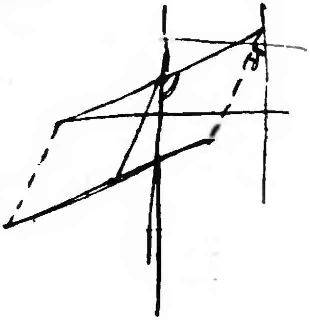
Since the δp_ν can be chosen totally arbitrarily, as long as they are continuously variable, one obtains

$$\frac{\partial L}{\partial p_\nu} - \frac{d}{dt} \left(\frac{\partial L}{\partial p'_\nu} \right) + P_\nu = 0. \qquad \Bigg| \qquad \underbrace{\frac{d}{dt} \left(\frac{\partial L}{\partial p'_\nu} \right)} + \frac{\partial \Pi - L}{\partial p_\nu} - P_\nu = 0$$

These are the important equat. of motion of Lagrange.
In the special case in which part of the forces can be derived from a potential, part of the forces has the form $-\frac{\partial \Pi}{\partial p_\nu}$

Example. [p. 87]

Two identical rods are linked to each other at their ends by means of threads. One of the rods is situated so as to pivot around its middle.[57]



$$L = \frac{2}{2} M k^2 \varphi'^2 + \frac{M}{2} l^2 \vartheta'^2$$

$$= \frac{M}{2} \{ 2k^2 \varphi'^2 + l^2 \vartheta'^2 \}$$

$$\Pi = - Mg\, l \cos \vartheta$$

From this at once equations of motion

$$\frac{d}{dt}\frac{\partial L}{\partial \varphi'} + \frac{\partial(\Pi - L)}{\partial \varphi} = 0$$

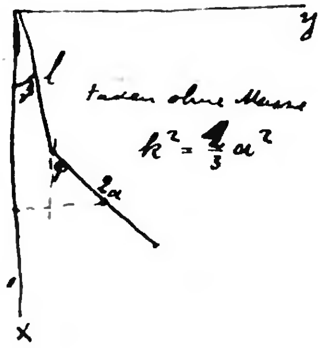
$$\frac{d}{dt}\frac{\partial L}{\partial \vartheta'} + \frac{\partial(\Pi - L)}{\partial \vartheta} = 0$$

The equations yield:

$$\frac{d}{dt}(2mk^2\varphi') = 0 \qquad \varphi' = \text{const.}$$

$$\frac{d}{dt}(Ml^2\vartheta') + Mgl\sin\vartheta = 0 \qquad \vartheta'' = \frac{g}{l}\sin\vartheta\ .$$

2nd Example.



$$\xi = l\cos\vartheta + a\cos\varphi \qquad \xi' = -l\sin\vartheta\vartheta' - a\sin\varphi\varphi'$$

$$\eta = l\sin\vartheta + a\sin\varphi \qquad \eta' = l\cos\vartheta\vartheta' + a\cos\varphi\varphi'$$

$$\Pi = +Mg\{l(1-\cos\vartheta) + a(1-\cos\varphi)$$

Calculation of the kin. energy

1) If the mass were concentrated at the center of gravity

[58]

$$\frac{M}{2}(\xi'^2 + \eta'^2) = \frac{M}{2}(l^2\vartheta'^2 + a^2\varphi'^2 + 2al\cos(\vartheta - \varphi)$$

2) Kin. energ. with respect to the center of gravity $\frac{1}{6}Ma^2\varphi'^2$

[p. 88]

$$L = \frac{M}{2}\left[l\vartheta'^2 + \frac{4}{3}a^2\varphi'^2 + 2al\vartheta'\varphi'\cos(\vartheta - \varphi)\right]$$

For infinitely small oscillations only the smallest terms are retained

$$\Pi = \frac{mg}{2}(l\vartheta^2 + a\varphi^2)$$

$$L = \frac{M}{2}(l\vartheta'^2 + \frac{4}{3}a^2\varphi'^2 + 2al\vartheta'\varphi')$$

The Lagrange eq. without P_v $\quad \frac{d}{dt}\left(\frac{\partial L}{\partial p_v'}\right) + \frac{\partial}{\partial p_v}(\Pi - L) = 0$

$$(Ml^2\vartheta'' + Mal\varphi'') + Mgl\vartheta = 0 \qquad \Big| \; l^2\vartheta'' + al\varphi'' = -gl\vartheta$$

$$\left(\frac{4}{3}Ma^2\varphi'' + Mal\vartheta''\right) + mga\varphi = 0 \qquad \Big| \; al\vartheta'' + \frac{4}{3}a^2\varphi'' = -ga\varphi$$

Linear homogeneous equations that can be solved trigonometrically. Set

$$\vartheta = \lambda_1 \cos(\omega t + \delta) \qquad (-\omega^2 l^2 + gl)\lambda_1 - al\omega^2\lambda_2 = 0$$

$$\varphi = \lambda_2 \cos(\omega t + \delta) \qquad -al\omega^2 \lambda_1 - \frac{4}{3}a^2\omega^2 + ga\lambda^2 = 0$$ [59]

$$(\omega^2 l - g)\left(\frac{4}{3}a\omega^2 - g\right) - al\omega^4 = 0$$

Biquadratic equation for frequency (ω) From this ω_1 & ω_2

The equations also yield $\frac{\lambda_2}{\lambda_1}$

As the general solution, one finds

$$\vartheta = a\omega_1^2\mu_1 \cos(\omega_1 t + \delta_1) + a\omega_2^2\mu_2(\cos \omega_2 t + \delta_2)$$

$$\varphi = (g - l\omega_1^2)\mu_1 \cos(\omega_1 t + \delta_1) + (g - l\omega_2^2)\mu_2(\cos \omega_2 t + \delta_2)$$

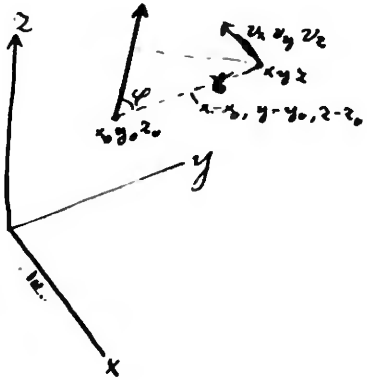
Superposition of two mutually completely independent oscillations of different periods.

[p. 89]

Rigid Body
Kinematics

Representation of rotation by vector ω velocity of a mat. point at distance 1 from the axis of rotation. Can be represented by vector whose point of application is of no importance. Length ω sense that rot with vector right-hand screw.

We denote the projections of the rotation vector on the coordinate axes by p, q, r. We consider an arbitrary point of the body. Seeking vector V. It is perpendicular to vector (ω), perpendicular to vector $(x - x_0, y - y_0, z - z_0) = (r)$. Is equal to the product of the magnitudes of the two vectors mult by the sine of the enclosed angle. Sequence . . $v_x\ v_y\ v_z$ is the vector product of the vectors (ω) and (r), hence



$$\begin{array}{llll} v_x = q(z - z_0) - r(y - y_0) & & & \\ v_y = r(x - x_0) - p(z - z_0) & p & q & r \\ v_z = p(y - y_0) - q(x - x_0) & x - x_0 & y - y_0 & z - z_0 \end{array}$$

Important formulas.

[p. 90]

Composition of Angular Velocities

Body with rotation $(p\ q\ r)$ about a given point (coordinate origin) describes with its point $x\ y\ z$ the path

$$v_x dt = q dt z - r dt y$$

$$- - - = - - - - - - - -$$

$$- - - = - - - - - - - -$$

the second vector $(p^x_{<1>}\ q^x_1\ r^x_{<1>})$ considered rot vector $(\omega^x_{<1>})$ imparts during time *dt*

$$v'_x dt = q^x dt z - r^x dt y$$

$$- - - - - - - - - - - -$$

$$- - - - - - - - - - - -$$

Adding the two motions Sum of the displacements

$$(v + v^x_x)dt = (q + q^x)dt z - (r + r^x)dt y$$

$$- -$$

$$- -$$

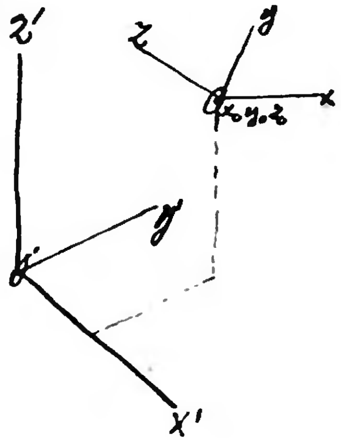
The result is, thus, rotation with the vector $(p + p^x, q + q^x, r + r^x)$ Thus, the rotational vectors are to be added according to the law of parallelograms if the two axes of rotation intersect one another.

Description of the Most General Motion of a Rigid Body

The motion is referred to the coordinate system at rest $X'Y'Z'$. In addition, we introduce a second coordinate system that is rigidly connected with the body (X,Y,Z). Determined by coordinates of 0 with resp. to 0′, x_0, y_0, z_0, and by the direction cosines [p. 91] of the coordinate axes.

	x	y	z
x'	α_1	α_2	α_3
y'	β_1	β_2	β_3
z'	γ_1	γ_2	γ_3

Relation between the direc. cos. How does point of the body move in space?



$$x' = x_0 + \alpha_1 x + \alpha_2 y + \alpha_3 z$$

$$y' = y_0 + \beta_1 x + \beta_2 y + \beta_3 z$$

$$z' = z_0 + \gamma_1 x + \gamma_2 y + \gamma_3 z$$

$$\left(\alpha_1 \frac{d\alpha_3}{dt} + \beta_1 \frac{d\beta_3}{dt} + \gamma_1 \frac{d\gamma_3}{dt}\right) = -\left(\alpha_3 \frac{d\alpha_1}{dt} + \beta_3 \frac{d\beta_1}{dt} + \gamma_3 \frac{d\gamma_1}{dt}\right) = q$$

etz.

$$\begin{array}{l|l}
V'_x = \dfrac{dx'}{dt} = \dfrac{dx_0}{dt} + \dfrac{d\alpha_1}{dt} x + \dfrac{d\alpha_2}{dt} y + \dfrac{d\alpha_3}{dt} z & \alpha_1 \\
V'_y = \dfrac{dy'}{dt} = \dfrac{dy_0}{dt} + \dfrac{d\beta_1}{dt} x + \dfrac{d\beta_2}{dt} y + \dfrac{d\beta_3}{dt} z & \beta_1 \\
V'_z = \dfrac{dz'}{dt} = \dfrac{dz_0}{dt} + \dfrac{d\gamma_1}{dt} x + \dfrac{d\gamma_2}{dt} y + \dfrac{d\gamma_3}{dt} z & \gamma_1
\end{array}$$

From this we calculate the velocity with respect to *X Y Z* Now we seek the same velocity with respect to a system that coincides with *X Y Z*. We

$$V_x = V'_x \alpha_1 + V'_y \beta_1 + V'_z \gamma_1 = V^0_x + qz - ry$$

[p. 92] $$= \left(\alpha_1 \frac{dx_0}{dt} + \beta_1 \frac{dy_0}{dt} + \gamma_1 \frac{dz_0}{dt}\right) + \cdot$$

$$\left.\begin{aligned} V_x &= V_x^0 + qz - ry \\ V_y &= V_y^0 + rx - pz \\ V_z &= V_z^0 + py - qx \end{aligned}\right\}$$

Velocity of a point with respect to a coord. syst. momentarily coinciding with $x\ y\ z$, but

Superposition of a translation and a rotation.
The derived formula is fundamental.

The Kinetic Energy of a Rigid Body

<1) Kinetic Energy>

$$v^2 = v_x^2 + v_y^2 + v_z^2 = v_x^{02} + v_y^{02} + v_z^{02} + p^2(y^2 + z^2) + q^2(z^2 + x^2)$$
$$+ r^2(x^2 + y^2) + 2yzqr + 2zxrp + 2xypq$$

$$\sum \frac{m}{2} v^2 = \sum \frac{mv^{02}}{2} + \frac{1}{2} p^2 \overbrace{\sum (y^2 + z^2)}^{A} + \frac{1}{2} q^2 \overbrace{\sum (z^2 + x^2)}^{B} + \frac{1}{2} r^2 \overbrace{\sum (x^2 + y^2)}^{C}$$
$$- qr \underbrace{\sum yz}_{D} - rp \underbrace{\sum zx}_{E} - pq \underset{F}{\sum xy}$$

$$2L = Mv_0^2 + Ap^2 + Bq^2 + Cr^2 - 2Dqr - 2Erp - 2Fpq$$

In the special case where the principal axes of inertia of the body with respect to 0′ the axes are chosen as $x\ y\ z$, we have $2L = Mv^{02} + Ap^2 + Bq^2 + Cr^2$.

<2)> Moment of Momentum of a Rigid Body
Moving About a Point.

Moment of the momenta $\sigma_x = \sum m \left(y \frac{dz}{dt} - z \frac{dy}{dt} \right) = \sum m (yv_z - zv_y)$

$$\sigma_x = \sum m \{ y(py - qx) - z(rx - pz) \} = p \sum m (y^2 + z^2) - q \sum mxy - r \sum xz$$

$$\sigma_x = Ap - Fq - Er = \frac{\partial L}{\partial p}$$

$$\sigma_y = Bq - Dr - Fp = \frac{\partial L}{\partial q}$$

$$\sigma_z = Cr - Ep - Dq = \frac{\partial L}{\partial r}$$

Special case where $x\ y\ z$ are the principal axes of inertia

$$\sigma_x = Ap$$

$$\sigma_y = Bq$$

$$\sigma_z = Cr.$$

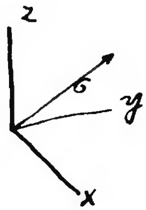
note that the moment of momentum does not coincide with the axis of rotation.

[p. 93] We obtain the equations of motion by applying the law of moments with respect to the axes $x\ y\ z$, or, to be more exact, with respect to a system that coincides with $x\ y\ z$ at the moment in question, but does not participate in the motion of the body. We denote the time derivative with respect to this system by $\left(\frac{d}{dt}\right)'$. We obtain then as the expression of the law of the conservation of the momentum the

$$\left(\frac{d}{dt}\right)' \sigma_x = L$$ [60]

$$\left(\frac{d}{dt}\right)' \sigma_y = M$$

$$\left(\frac{d}{dt}\right)' \sigma_z = N$$



We compose the change experienced by σ from two parts.

1) from the change experienced by σ due to the fact that this vector also changes relative to the moving system. This yields for the X component in dt the change

$$\frac{d\sigma_x}{dt}$$

2) the change experienced by the vector, even if it does not change relative to the moving system $X\ Y\ Z$, *because this system is moving.*[61] It changes absolutely in space as the components $\sigma_x\ \sigma_y\ \sigma_z$ of a material point of a rigid body connected with $x\,y\,z$. This

[p. 94] yields the contribution $q\sigma_z - r\sigma_y$ Thus, one obtains the equation

$$\frac{d\sigma_x}{dt} + q\sigma_z - r\sigma_y = L \quad \text{etc.}$$

We now replace σ_x, σ_y, σ_z by Ap, Bq, Cr in that we assume that the axes X, Y, Z coincide with the principal axes of inertia, and then we obtain

$$A\frac{dp}{dt} + (C - B)qr = L$$

$$B\frac{dq}{dt} + (A - C)rp = M$$

$$C\frac{dr}{dt} + (B - A)pq = N$$

These are Euler's equations of motion.[62]
Euler's equations:

[p. 95]

$$A\alpha^2p^2 + B + C\alpha^2qr = 0$$

$$B\frac{dq}{dt} + (A - C)rp = M$$

$$C\frac{dr}{dt} + (B - A)pq = N$$

$p(t)$, $q(t)$, $r(t)$
$\alpha(p\ (\alpha t)\ .\ .$
also solution
$A\alpha^2p^2 + B - C\alpha^2qr = 0$
thus, α times faster

Motion of the rigid body not acted upon by any forces.

Leads to elliptical, that is, periodical functions for p q r as funct of t. Simplest special case rotation about principal axis of inertia $q = 0$, $r = 0$. The second & third equation are then identically satisfied, while the first one yields p = const. Rotation is not *stable* about *every* principal axis of inertia. As a proof, consider motion deviating slightly from rotation about the principal axis of inertia q and r ∞ small to first order. ∞ small second-order terms neglected. The first equation yields then $A\frac{dp}{dt} = 0$ p = const. the second and third yield

$$\frac{dq}{dt} + \overset{\alpha}{\boxed{p\frac{C - A}{B}}}\, r = 0$$

$$\frac{dr}{dt} + \underbrace{\boxed{p\frac{A-B}{C}}}_{\beta} q = 0$$

One obtains for q the eq. $\frac{d^2q}{dt^2} - \alpha\beta q = 0$.

$\alpha\beta > 0$

Solutions $e^{\sqrt{\alpha\beta}\,t}$ $e^{-\sqrt{\alpha\beta}\,t}$

Leads to exponential function.

$$Ae^{\sqrt{\alpha\beta}\,t} + Be^{-\sqrt{\alpha\beta}\,t}$$

For positive & negative ∞ large t,
q (& r?) becomes finite;[63]
the larger the p, the faster this process

$(C-A)(A-B)>0$ $(A-C)(A-B)<0$

Axis mean principal moment of inertia

$\alpha\beta<0$ $(C-A)(A-B)<0$ $(A-C)(A-B)>0$

A largest or smallest mom. of inertia

$$q = A \sin\sqrt{-\alpha\beta}\,t + B\cos\sqrt{-\alpha\beta}\,t$$

$$r = -\frac{1}{\alpha}\frac{dq}{dt} = -A\sqrt{-\frac{\beta}{\alpha}}\cos + B\sqrt{\frac{\beta}{\alpha}}\sin()$$

Rotation stable

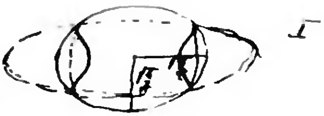
Axis of rotation describes ellipse relative to body

[p. 96] second analysis

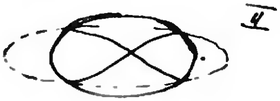
$$Ap^2 + Bq^2 + Cr^2 = h$$

$$A^2p^2 + B^2q^2 + C^2r^2 = \sigma^2$$

$p\ q\ r$ considered as rectangular coordinates. Sum of the possible $p\ q\ r$ line of intersection of two ellipsoids, with main axes $\left(\sqrt{\frac{h}{A}}, \sqrt{\frac{h}{B}}, \sqrt{\frac{h}{C}}\right)$ and $\frac{\sigma}{A}, \frac{\sigma}{B}, \frac{\sigma}{C}$



If h is kept const & σ is allowed to grow, one obtains all types of curves of intersection.:



One sees in this way as well the types of possible motion, especially the stability of the extreme principal axes of inertia, and the lability of the middle ones.

Relation to System at Rest. Euler Angle.

[p. 97]

With regard to the system ($x_1\, y_1\, z_1$), α_1 .... γ_1 can be determined as functions of time by means of additional integrations if p, q, r have already been determined as funct. of time by means of Euler's equations. Because we have

$$\left(\alpha_3\frac{d\alpha_2}{dt} + \beta_3\frac{d\beta_2}{dt} + \gamma_3\frac{d\gamma_2}{dt}\right) = -\left(\alpha_2\frac{d\alpha_3}{dt} + \beta_2\frac{d\beta_3}{dt} + \cdot\right) = p$$

— — — — — — — — — — — — — — — — — — $= q$

— — — — — — — — — — — — — — — — — — $= r,$

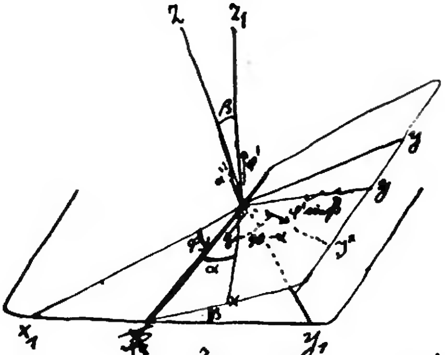
which equations together with the 6 independent relations between the α to γ determine all α to γ as functions of time. But generally one makes use of the so-called Euler angle to determine the position of the body.

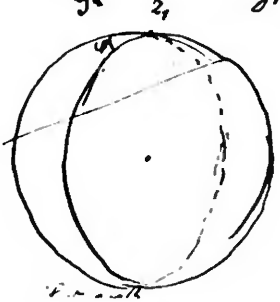
p	q	r
$\varphi' \sin\beta \sin\alpha$	$\varphi' \sin\beta \cos\alpha$	α'
$\beta' \cos\alpha$	$-\beta' \sin\alpha$	$\varphi' \cos\beta$

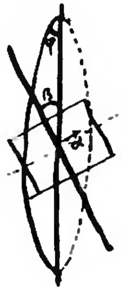
$$p = \varphi' \sin\beta \sin\alpha + \beta' \cos\alpha$$

$$q = \varphi' \sin\beta \cos\alpha - \beta' \sin\alpha$$

$$r = \alpha' + \varphi' \cos\beta.$$







three differential equations in order to determine α, β, & φ if p, q & r are given. Special case β is infinitely small, then $r = \alpha' =$ const. $\alpha = \omega t$

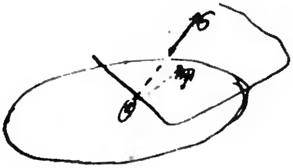
~~second~~ derivation of Euler's equations

$$\frac{d}{dt}\frac{\partial L}{\partial \varphi'} = 0 \qquad \frac{1}{2}\frac{\partial L}{\partial \varphi'} = Ap\frac{\partial p}{\partial \varphi'} + Bq\frac{\partial q}{\partial \varphi'} + Cr\frac{\partial r}{\partial \varphi'}$$

$$= Ap\sin\beta \, \sin\alpha + Bq\sin\beta \, \cos\alpha + Cr\cos\beta$$

Geometrical Solution of the Problem of a Body Moving in the Absence of Forces

[p. 98]



1) $2L = \dfrac{\omega^2}{\overline{OP}^2}$

If $\overline{OP}$ is the ellipsoid of inertia.

2) Ellipsoid of inertia $Ax^2 + By^2 + Cz^2 = 1$

Tangent plane in $x\,y\,z$

$$Axx' + Byy' + Czz' - 1 = 0$$

Direction cos. as $Ax : By : Cz$

thus also as $\underline{Ap} : \underline{Bq} : \underline{Cr}$

Thus, normal parallel to the moment vector σ.

3) Distance of the tangent plane from 0. One brings the tangent plane into the normal form

$$\delta = \frac{1}{\sqrt{(Ax)^2 + (By)^2 + (Cz)^2}} = \frac{\sqrt{2L}}{\sqrt{A^2p^2 + B^2q^2 + C^2r^2}} = \sqrt{\frac{2L}{\sigma^2}}$$

$$x = \overline{OP}\frac{p}{\omega} = \frac{\omega}{\sqrt{2L}} \cdot \frac{p}{\omega} = \frac{1}{\sqrt{2L}}p$$

$$y = \qquad = \frac{1}{\sqrt{2L}}q$$

Neither direction cos nor distance of the tangent plane changes. Thus, the plane stays fixed in space.

The ellipsoid of inertia rolls on this plane because P, as the point of the momentary [p. 99] axis of rotation, does not move. The velocity of rotation is found from 1) to be

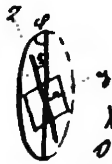
$$\omega = \sqrt{2L} \cdot \overline{OP}.$$

Motion of a Body with Two Identical Principal Moments of Inertia

~~$2L = A^2 \cdot \beta'^2 + B^2\{\varphi'^2 \sin^2\beta + (\alpha' + \varphi'\cos\beta)^2\}$~~

~~$2L = A^2\beta'^2 + B^2\{\varphi'^2 + \alpha'^2 + 2\alpha'\varphi'\cos\beta\}$~~

~~p set, by turns, equal α, β & φ~~



X	Y	Z
0	~~φ sin~~	$\varphi'\cos\beta$
β'	0	0
0	0	α'

~~$2L = A(\alpha' + \varphi\cos\beta)^2 + B(\beta'^2 + \varphi'^2 \sin^2\beta)$~~

~~$0 = \frac{d}{dt}\left(\frac{\partial L}{\partial \alpha'}\right) = \frac{d}{dt} 2A(\alpha' + \varphi'\cos\beta)$~~

~~$0 = \frac{d}{dt}\left(\frac{\partial L}{\partial \beta'}\right) = \frac{d}{dt}\{2B\beta'\}$~~

~~$0 = \frac{d}{dt}\left(\frac{\partial L}{\partial \varphi'}\right) = \frac{d}{dt}\{2A(\alpha' + \varphi'\cos\beta)\cos\beta + 2B\sin^2\beta\varphi'\}$~~

~~$\frac{\partial L}{\partial \alpha} = 0 \qquad \frac{\partial L}{\partial \beta} = 2B\varphi'^2 \sin\beta\cos\beta \qquad \frac{\partial L}{\partial \varphi} = 0$~~

~~Thus, Lagrange's equations become~~

~~$\frac{d}{dt}\{\alpha' + \varphi'\cos\beta\} = 0$~~

~~$\frac{d}{dt}\{\beta'\} - \varphi'^2 \sin\beta\cos\beta = 0$~~

~~$\frac{d}{dt}\{A(\alpha' + \varphi'\cos\beta)\cos\beta + B\sin^2\beta\varphi'\} = 0$~~

$$2L = A(\alpha' + \varphi'\cos\beta)^2 + B(\beta'^2 + \varphi'^2\sin^2\beta)$$

[p. 100]

$$\frac{\partial L}{\partial \alpha'} = A(\alpha' + \varphi'\cos\beta) \qquad \frac{\partial L}{\partial \alpha} = 0$$

$$\frac{\partial L}{\partial \beta'} = B\beta' \qquad \frac{\partial L}{\partial \beta} = -A(\alpha' + \varphi'\cos\beta)\varphi'\sin\beta + B\varphi'^2\sin\beta\cos\beta$$

$$\frac{\partial L}{\partial \varphi'} = A(\alpha' + \varphi'\cos\beta)\cos\beta + B\varphi'\sin^2\beta \qquad \frac{\partial L}{\partial \varphi} = 0$$

$$\begin{array}{r|l}
 & \text{when external forces} \\
 & P_\alpha\alpha' + P_\beta\beta' + P_\varphi\varphi' \\
\frac{d}{dt}\{A(\alpha' + \varphi'\cos\beta) = 0 & = +P_\alpha \qquad 0 \\
\frac{d}{dt}\{A(\alpha' + \varphi'\cos\beta)\cos\beta + B\varphi'\sin^2\beta\} = 0 & = +P_\varphi \qquad 0 \\
A(\alpha' + \varphi'\cos\beta)\varphi'\sin\beta & \\
- B\varphi'^2\sin\beta\cos\beta + \frac{d}{dt}(B\beta') = 0 & = \quad P_\beta + \quad C\sin\beta
\end{array}$$

Consider the case when external forces do not act.

$$A(\alpha' + \varphi'\cos\beta) = R_1$$

$$A(\alpha' + \varphi'\cos\beta)\cos\beta + B\varphi'\sin^2\beta = R_2 \quad \& \quad R_1\cos\beta + B\varphi'\sin^2\beta = R_2$$

The general case can be treated by calculating φ' & α' from these two equations and substituting them in the third equation. But we confine ourselves to the case of constant β The last term of the third equation will then vanish; the equation assumes the form

$$A\alpha'\varphi'\sin\beta + (A - B)\varphi'^2\sin\beta\cos\beta = 0$$

We divide by $(A - B)$ $\sin\beta$ $\cos\beta$ under the assumption that none of these quantities vanishes. Then it turns into

$$\varphi' \left\{\varphi' + \frac{A\alpha'}{(A - B)\cos\beta}\right\} = 0$$

[p. 101] We exclude the trivial case $\varphi' = 0$ $\quad \frac{d}{dt}\frac{\partial L}{\partial p_\nu'} + \frac{\partial(\Phi - L)}{\partial p_\nu} = P_\nu$

$$\varphi' = -\frac{A\alpha'}{(A - B)\cos\beta}$$

φ changes in the opposite direction from α, if A is the axis of the greatest moment of inertia, otherwise the opposite. In the case of nearly spherical shape or small B,[64] φ rotates <always> faster than α, since $\frac{A}{(A - B)\sin\beta\cos\beta}$ [65] is an improper fraction if B is not very large. In that case the two rotations in the same direction.

Special case when β is very small. Then the rotational velocity of the spinning top is

$$D = \alpha' + \varphi' \neq = \varphi'\left(1 - \frac{A - B}{A}\right) = \varphi'\frac{B}{A}$$

$$\varphi' = D \cdot \frac{A}{B}$$

Axis describes conical shell with angular velocity φ'.

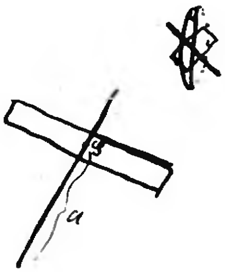
How must α' β' & φ' be chosen for such a motion to take place? Answer

$\beta' = 0$ & $\varphi' = -\frac{A\alpha'}{(A - B)\cos\beta}$, because then we have in the first moment $\beta'' = 0$

Argument then repeated.

Motion of the spinning top.

$$P_\beta = + \overset{C}{\boxed{Mg\alpha}} \sin\beta$$



The differential equations are [p. 102]

$$\frac{d}{dt}\{A(\alpha' + \varphi'\cos\beta)\} = 0$$

$$\frac{d}{dt}\{A(\alpha' + \varphi'\cos\beta)\cos\beta + B\varphi'\sin^2\beta\} = 0$$

$$-C\sin\beta + A(\alpha' + \varphi'\cos\beta)\varphi'\sin\beta - B\varphi'^2\sin\beta\cos\beta + B\beta'' = 0$$

We examine again the case of constant β. The third equation yields

$$-C\sin\beta + A\alpha'\varphi'\sin\beta + (A - B)\varphi'^2\sin\beta\cos\beta = 0$$

We divide by $(A - B)\sin\beta\cos\beta$, assuming again that none of these factors vanishes. From this results the eq.

$$\varphi'^2 + \frac{A\alpha'}{(A-B)\cos\beta}\varphi' - \frac{C}{(A-B)\cos\beta} = 0$$

$$\varphi' = -\frac{1}{2}\frac{A\alpha'}{(A-B)\cos\beta} \pm \sqrt{\frac{1}{4}\frac{A^2\alpha'^2}{(A-B)^2\cos^2\beta} + \frac{C}{(A-B)\cos\beta}}$$
$$= -\frac{1}{2}\frac{A\alpha'}{(A-B)\cos\beta}\left\{1 \mp \sqrt{1 + \frac{4(A-B)\cos\beta C}{A^2\alpha'^2}}\right\}$$

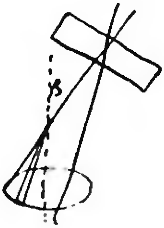
If α' is sufficiently large, then $\{\} = 1 \mp \left(1 + \frac{2C(A-B)\cos\beta}{A^2\alpha'^2}\right)$

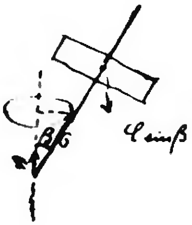
As above, we obtain two solutions for the body moving in the absence of force

$$\varphi_1' = \frac{C}{A\alpha'} \qquad \varphi_2' = -\frac{A\alpha'}{(A - B)\cos\beta} - \frac{C}{A\alpha'}$$

The second one corresponds to the fast rotation of the axis of the top about the vertical, [p. 103] also if $C = 0$. But the first goes over into invariant position of axis A for $c = 0$ yields slower rotation about the vertical the greater the moment of rotation. Is, further, independent of β, because C is independent of β.

We shall derive this solution, in which we are specially interested, in yet another way. σ nearly coincides with the axis of rotation



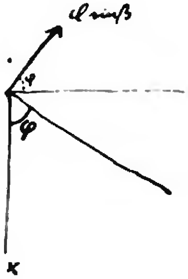


$$\frac{d\sigma_x}{dt} = L \quad \frac{d\sigma_y}{dt} = M \quad \frac{d\sigma_z}{dt} = N$$

$$\sigma_x = \sigma\sin\beta\cos\varphi \qquad L = -C\sin\beta\sin\varphi$$

$$\sigma_y = \sigma\sin\beta\sin\varphi \qquad M = C\sin\beta\cos\varphi$$

$$\sigma_z = \sigma\cos\beta \qquad N = 0$$



$$\cancel{+}\sigma\cancel{\sin\beta}\cancel{\sin\varphi}\frac{d\varphi}{dt} = \cancel{+}C\cancel{\sin\beta}\cancel{\sin\varphi}$$

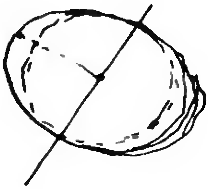
$$\sigma\cancel{\sin\beta}\cancel{\cos\varphi}\frac{d\varphi}{dt} = C\cancel{\sin\beta}\cancel{\cos\varphi}$$

$$\frac{d\varphi}{dt} = \frac{C}{\sigma} = \frac{C}{A\alpha'}$$

$$M = C\sin\beta$$

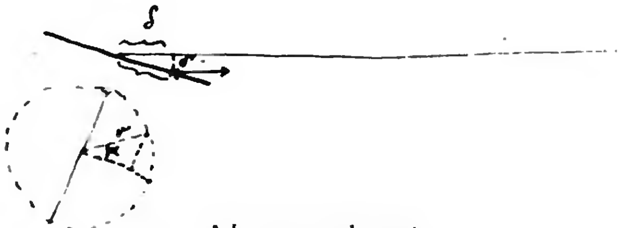
$$\sigma\cancel{\sin\beta}\varphi' = C\cancel{\sin\beta}$$

$$\varphi' = \frac{C}{\sigma} = \frac{M}{\sigma\sin\beta} \qquad C = Mg\delta$$



Precession of the earth.
→S

Bulge on the equator terrestrial radius *r*. Sun—earth *R* [p. 104]



$r\lambda d\alpha$ = mass element

$R^x = R - \delta$ = distance from the sun

$= R - r\cos\alpha\cos\beta$

$\delta = r\,\underline{\cos\alpha}\,\cos\beta \qquad \delta' = r\,\underline{\cos\alpha}\,\sin\beta$

<Potential with respect to the sun> Force = $\dfrac{k^2Mr\lambda\,d\alpha}{(R - r\cos\alpha\cos\beta)^2}$

Moment of this force with respect to the axis of the vernal equinox

$$k^2Mr\lambda\frac{d\alpha}{(R-r\cos\alpha\cos\beta)^2}r\cos\alpha\sin\beta = k^2Mr\lambda\,d\alpha\frac{\partial}{\partial\beta}\left(\frac{1}{R-r\cos\alpha\cos\beta}\right)$$

$$\left\langle\frac{1}{(\quad)^2}=\frac{1}{R^2}\left(1+\right\rangle\right. = \frac{\partial}{\partial\beta}\left\{k^2Mr\lambda\,d\alpha\left(\frac{1}{R}+\frac{r\cos\alpha\cos\beta}{R^2}+\frac{r^2\cos^2\alpha\cos^2\beta}{R^3}\right)\right.$$

$$\int\cos^2\alpha\,d\alpha = \frac{1}{2}\int(1+\cos 2\alpha)\,d\alpha$$

$$= \frac{1}{2}$$

$$\int_0^{2\pi} - d\alpha = \frac{\partial}{\partial\beta}\left(k^2Mr\lambda\left\{\frac{2\pi}{R}+0\ \frac{r^2\cos^2\beta}{2R^3}\right\}\right)$$

$$= k^2Mr\lambda\frac{r^2\sin\beta\cos\beta}{R^3} = k^2M\frac{m}{2\pi}\frac{r^2\sin\beta\cos\beta}{R^3}$$

We can estimate m to be $\dfrac{4}{3}(-r^3 + rr'^2)\ \pi\cdot\bar{\rho} = \dfrac{8}{3}r^3\varepsilon\pi\cdot\bar{\rho}$

$r' = r(1 + \varepsilon)$

Here we assume; that the mass acts as if it were concentrated on the bulge, $\bar{\rho}$ is mean density. The mean annual torque approximately one-half of the one calculated in this way. One thus arrives at a value of φ' that is of the same order of magnitude as the one observed.

$$M = \frac{2}{3}k^2 M_g \varepsilon \bar{\rho} \frac{r^5 \sin\beta \cos\beta}{R^3} \qquad \text{hence} \quad \varphi' = \frac{2\pi}{T} = \frac{M}{I\alpha' \sin\beta}$$ [66]

T can be calculated. $M_e r^2 \sqrt{\frac{2}{5}}$

Foucault's Gyroscope [p. 105]

We investigate general motion of a rigid body at a fixed point with respect to a coordinate system which itself rotates but is not rigidly connected with the body under consideration. The analysis is wholly analogous to that applied in establishing Euler's equations. The system $X^x\ Y^x\ Z^x$ is introduced, which momentarily coincides with the system $X\ Y\ Z$ but does not rotate together with it. Then the law of moments will apply

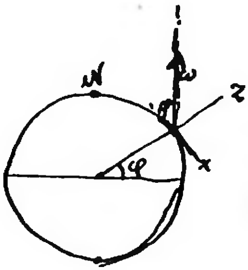
$$\frac{d^x}{dt}(\sigma_x) = L$$

If we denote by $p\ q\ r$ the instantaneous rotation of $X\ Y\ Z$, then we have $\left(\frac{d}{dt}\right)^x(\sigma_x) = \frac{d\sigma_x}{dt} + q\sigma_z - r\sigma_y$, so that the law of moments reads

$$\frac{d\sigma_x}{dt} + q\sigma_z - r\sigma_y = L$$

$$\frac{d\sigma_y}{dt} + r\sigma_x - p\sigma_z = M$$

$$\frac{d\sigma_z}{dt} + p\sigma_y - q\sigma_x = N$$



$p = -\omega\cos\varphi$

$q = 0$

$r = \omega\sin\varphi$

Specifically for a point on the surface of the earth[67]

$$\frac{d\sigma_x}{dt} - \omega\sin\varphi\,\sigma_y = L$$

$$\frac{d\sigma_y}{dt} + \omega\sin\varphi\sigma_x + \omega\cos\varphi\sigma_z = M$$

$$\frac{d\sigma_z}{dt} - \omega\cos\varphi\sigma_y = N$$ [67]

[p. 106] $N = 0$ because reaction of the guide normal

$\sigma_z \infty$ klein $\quad L = -M\sin\alpha$

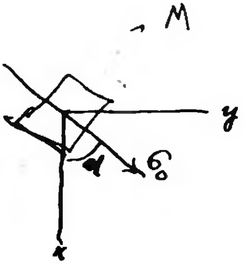
$\sigma_x = \sigma_0\cos\alpha \quad M = +M\cos\alpha$

$\sigma_y = \sigma_0\sin\alpha \quad N = 0$

$$\frac{d\sigma_z}{dt} = \omega\cos\varphi\sigma_0\sin\alpha$$

$$\sigma_z = I_z\frac{d\alpha}{dt}$$

$$I_z\frac{d^2\alpha}{dt^2} = \omega\cos\varphi\sigma_0\sin\alpha$$



We also introduce $\beta = \pi - \alpha$

$$\frac{d^2\beta}{dt^2} = -\frac{\omega\cos\varphi\;\sigma_0}{I_z}\sin\beta \qquad \sigma_0 = I_r\alpha'$$

This is the pendulum equation. If the initial rotation about the Z-axis is small, then oscillation of the pendulum about the northerly direction with oscillation period

$$T = 2\pi\sqrt{\frac{I_z}{I_r}\frac{1}{\cos\varphi\;\omega\alpha'}}$$

Example $I_z = I_r \quad \varphi = 0 \quad \alpha' = 2\pi\cdot 100 \quad T = 29.4.$

For the calculation of the moment of reaction we use the first or the second equation of motion.

$$\frac{d\sigma_x}{dt} < + q\sigma_z - r\sigma_y > - \omega \sin\varphi \; \sigma_y = L$$

$$-\sigma_0 \sin\alpha \; \frac{d\alpha}{dt} - \omega \sin\varphi \; \sin\alpha \; \sigma_0 = -M \sin\alpha$$

ω is small compared with $\frac{d\alpha}{dt}$. Hence $M = \sigma_0 \frac{d\alpha}{dt}$ can be considerable.

Introduction of the Kinetic Potential[68]

[p. 107]

$$P_\nu = \frac{d}{dt}\left(\frac{\partial L}{\partial p_\nu'}\right) + \frac{\partial(\Pi - L)}{\partial p_\nu}$$

We set $\Pi - L = H$. Because Π is independent of p'_ν, we can then write the equation in the following way

$$P_\nu + \frac{d}{dt}\left(\frac{\partial H}{\partial p_\nu'}\right) - \frac{\partial H}{\partial p_\nu} = 0$$

Thus, the knowledge of a single function is all that is needed to determine the motion of a system. One calls H the kinetic potential. If the function H is introduced into Hamilton's principle, the latter assumes the form

$$\int_{t_0}^{t_1} (\delta H - P_\nu \delta p_\nu) dt = 0$$

This equation is a direct consequence of the one just given. If we denote by P_ν the force applied by the system to the environment rather than the force applied by the environment to the system, then P_ν has to be replaced by $-P_\nu^-$, so that one has to set

$$P^-_\nu - \frac{d}{dt}\left(\frac{\partial H}{\partial p_\nu'}\right) + \frac{\partial H}{\partial p_\nu} = 0$$

Helmholtz found that these general equations are suitable for representing the dynamical properties of physical systems far beyond the domain of mechanics.[69] However, it can happen that we know not to conceive of H as $\Pi - L$, & also do not want to be so constrained.

[p. 108] We ask therefore whether the energy principle is maintained if H takes an arbitrary form.

To this end we multiply the generalized Lagrange equation by $dp = p'dt$ and sum over all coordinates

$$\sum P_\nu^+ \, dp_\nu + \sum \underbrace{\frac{d}{dt}\frac{\partial H}{\partial p'_\nu}}_{\frac{d}{dt}\left(\frac{\partial H}{\partial p'_\nu}p'_\nu\right) - \frac{\partial H}{\partial p'_\nu}p''_\nu} p'_\nu \, dt - \sum \frac{\partial H}{\partial p_\nu} dp_\nu = 0$$

$$\sum P_\nu^- \, dp_\nu + d\sum\left(\frac{\partial H}{\partial p'_\nu}p'_\nu\right) - \sum\underbrace{\left(\frac{\partial H}{\partial p'_\nu}dp'_\nu + \frac{\partial H}{\partial p_\nu}dp_\nu\right)}_{dH} = 0$$

thus

$$\sum P_\nu^+ \, dp_\nu = d\underbrace{\left(H - \sum\left(\frac{\partial H}{\partial p'_\nu}p'_\nu\right)\right)}_{E}$$

From this we see that the generalized Lagrange equations (Hamilton's principle) involve the energy principle. We also show that in the special case of ordinary mechanics one arrives back at the customary expression for the energy. Here $H = \phi - L$

$$2L = A_{11}p_1'^2 + 2A_{12}p_1'p_2' + 2A_{13}p_1'p_3' \cdots\cdots = \sum_\mu \sum_\nu A_{\mu\nu}p'_\nu p'_\mu$$

$$+ A_{22}p_2'^2 + 2A_{23}p_2'p_3' \cdots\cdots$$

where $A_{\mu\nu} = A_{\nu\mu}$. We obtain

$$\cancel{2}\frac{\partial L}{\partial p'_1} = \cancel{2}A_{11}p'_1 + \cancel{2}A_{12}p'_2 + \cdots\cdots \qquad p'_1 \qquad \delta L = 2\cancel{\delta}L = \sum\frac{\partial L}{\partial p'_\nu}\cdot\cancel{\delta}p'_\nu$$

[p. 109]

$$\frac{\partial L}{\partial p'_2} = A_{21}p'_1 + A_{22}p'_2 + \cdots\cdots \qquad p'_2$$

From this $\sum\frac{\partial L}{\partial p'_\nu}p'_i = A_{11}p'^2_1 + 2A_{12}p'^2_2\cdots\cdots = 2L$

Euler's law

Hence $$E = H - \sum\frac{\partial H}{\partial p'_\nu}p'_\nu = \Phi - L + 2L = \Phi + L.$$

Application to Electrodynamics

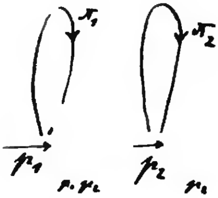
Two circuits

vanish

$$P_\nu^- = \frac{d}{dt}\left(\frac{\partial H}{\partial p'_\nu}\right) - \frac{\partial H}{\partial p_\nu}$$

$$\Pi_\nu = \frac{d}{dt}\left(\frac{\partial H}{\partial \pi'_\nu}\right) - \frac{\partial H}{\partial \pi_\nu}$$

vanish



p_1 $\qquad$ $p_1 p_2$ $\qquad$ p_2

$$H = -\frac{1}{2}(L_1\pi'^2_1 + 2M\pi'_1\pi'_2 + L_2\pi'^2_2)^{[70]}$$

We assume that the L M are independent of the p'

$$P_1^- = +\frac{1}{2}\frac{\partial L_1}{\partial p_1}\pi_1^2 + \frac{\partial M}{\partial p_1}\pi_1\pi_2 \qquad \Pi_1 = -\frac{d}{dt}(L_1\pi_1) - \frac{d}{dt}(M\pi_2)$$

$$P_2 = \frac{\partial M}{\partial p_2}\pi_1\pi_2 + \frac{1}{2}\frac{\partial L_2}{\partial p_2}\pi_2^2 \qquad \Pi_2 = -\frac{d}{dt}(M\pi_1) - \frac{d}{dt}(L_2\pi_2)$$

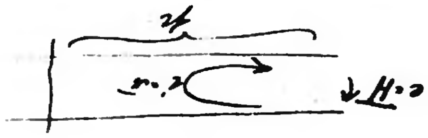
[p. 110]

$$P_1^- = \frac{d}{dt}\left(\frac{\partial H}{\partial p_1'}\right) - \frac{\partial H}{\partial p_1}$$

$$\Pi_1^- = \frac{d}{dt}\left(\frac{\partial H}{\partial \pi_1'}\right) - \frac{\partial H}{\partial \pi_1}$$

1) A circuit

We apply this to a circuit. Let its configuration be determined by *one* coordinate p



and one coordinate $\pi = i$. We assume that H is homogeneous to second order in π'.

$H = -\frac{1}{2}L\pi'^2$ L is here a function of p. The above equations then yield

$$P_1^- = -\frac{1}{2}\pi^2\frac{\partial L}{\partial p} = +\frac{1}{2}i^2\frac{\partial L}{\partial p}$$

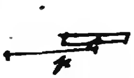
$$\Pi_1^- = -\frac{d(L\pi')}{dt} = -\frac{d}{dt}(Li)$$

In agreement with the familiar laws of electrodynamics.

2) moving magnet & solid conductor

We set $H = -W \cdot \pi'$.

$P^- = \pi'\frac{\partial W}{\partial p} = i\frac{\partial W}{\partial p}$ $P^- dp = i dW = i\Pi\, dt$



$\Pi^- = -\frac{dW}{dt}$ work delivered = work of the electromotive force.

The previous page follows [p. 111]

Limit of the applicability of the principle of least action

$$\cancel{P_\nu^- = \frac{d}{dt}\left(\frac{\partial H}{\partial p'_\nu}\right) - \frac{\partial H}{\partial p_\nu}}$$

canonical equations

$$P_\nu = +\frac{d}{dt}\underbrace{\left(\frac{\partial L}{\partial p'_\nu}\right)}_{q_\nu} + \frac{\partial(\Pi - L)}{\partial p_\nu} = 0$$

$$\delta L = \sum \frac{\partial L}{\partial p_\nu}\delta p_\nu + \sum \frac{\partial L}{\partial p'_\nu}\delta p'_\nu$$

$$= \sum \frac{\partial L}{\partial p_\nu}\delta p_\nu + \sum q_\nu \delta p'_\nu$$

$$\underbrace{\delta(L - \sum q_\nu p'_\nu)}_{T}$$

$$= \sum \frac{\partial L}{\partial p_\nu}\delta p_\nu - \sum p'_\nu \delta q_\nu$$

$$\cancel{P_\nu = -\frac{d}{dt}\underbrace{\left(\frac{\partial H}{\partial p'_\nu}\right)}_{-q_\nu} + \frac{\partial H}{\partial p_\nu}}$$

$$\cancel{\delta H = \sum \frac{\partial H}{\partial p_\nu}\delta p_\nu - \sum q_\nu \delta p_\nu}$$

$$\cancel{\underbrace{\delta(H + p_\nu q_\nu)}_{T} = \sum \frac{\partial H}{\partial p_\nu}\delta p_\nu + \frac{\partial \cancel{H}}{\partial q_\nu}\sum p'_\nu \delta q_\nu}$$

$$\cancel{\frac{\partial T}{\partial p_\nu} = \frac{\partial H}{\partial p_\nu}\,\frac{\partial T}{\partial q_\nu} = \cancel{q}_\nu p'_\nu}$$

$$\frac{dq_\nu}{dt} = P_\nu - \frac{\partial T}{\partial p_\nu} \qquad T = \Phi + L$$

$$\frac{dp_\nu}{dt} = \frac{\partial T}{\partial q_\nu}$$

in case of

ponderable mechanics = E.

If T is viewed, secondly, as function of p_ν & q_ν, then

$$\delta T = \sum \frac{\partial T}{\partial p_\nu}\delta p_\nu + \sum \frac{\partial T}{\partial q_\nu}\delta q_\nu$$

Hence, we have

$$\frac{\partial L}{\partial p_\nu} = \frac{\partial T}{\partial p_\nu} \qquad -p'_\nu = \frac{\partial T}{\partial q_\nu}$$

Therefore,

$$\frac{dp_\nu}{dt} = -\frac{\partial T}{\partial q_\nu}$$

$$\frac{dq_\nu}{dt} = +P_\nu - \frac{\partial \Pi}{\partial p_\nu} - \partial$$

[p. 112]

Energy Principle

$$\left.\begin{array}{l} \dfrac{dp_\nu}{dt} = \dfrac{\partial E}{\partial q_\nu} \\[2ex] \dfrac{dq_\nu}{dt} = -\dfrac{\partial E}{\partial p_\nu} + P_\nu \end{array}\right| \quad \begin{array}{c} \dfrac{dq_\nu}{dt} \\[2ex] -\dfrac{dp_\nu}{dt} \end{array} \quad \sum\left(\frac{\partial E}{\partial q_\nu}\frac{dq_\nu}{dt} + \frac{\partial E}{\partial p_\nu}\frac{\partial p_\nu}{dt}\right) - \sum P_\nu \frac{dp_\nu}{dt} = 0$$

The chief value of these equations for physics consists in the fact that they are the most convenient basis for the equations of statistical mechanics.[71]

More on the physical applications of the principle of least action and the Lagrange equations.

$$P_\nu^- = -\frac{\partial H}{\partial p_\nu}$$

$$\Pi_\mu^- = \frac{d}{dt}\left(\frac{\partial H}{\partial \pi'_\mu}\right)$$

We apply the equations to reversible changes of state of the unit mass of a substance $p = \text{v}$ P = pressure. Heat is a cyclic process, characterized by a velocity π', which we identify with temperature.[72] Then we have

$$P = -\frac{\partial H}{\partial \text{v}} \qquad \Pi d\pi = \Pi\pi' \, dt = -q$$

$$-\frac{q}{T} = d\left(\frac{\partial H}{\partial T}\right)$$

[p. 113]

From the second eq. $-\int\frac{q}{T} = \frac{\partial H}{\partial T} = -S$]

$$P = -\frac{\partial H}{\partial v} \quad \vdots \quad -dv$$

$$-P\,dV - S\,dT = dH$$

$$S = -\frac{\partial H}{\partial T} \quad \vdots \quad -dT$$

$$\underbrace{-P\,dV + T\,dS}_{dE} = dH + d(TS) = dE \qquad dH = d(E - TS)$$

General conclusions from the equations for cyclic motion.

1) If one considers only processes in which the <coordinates> cyclic velocities π' are constant (currents, temperatures), then the forces can be derived from a potential.

2) The same is also true if forces do not act upon the cyclic coordinates. In that case we have m equations $\frac{\partial H}{\partial \pi'_\mu} = \text{const.}$, by means of which one can eliminate the π' from H. (Interaction between magnets & resistance-free short closed circuits. Adiabatic processes.

~~3)~~ We write down two cyclical eq.

$$\Pi_\rho' = \frac{d}{dt}\left(\frac{\partial H}{\partial \pi_\rho'}\right) = \sum \frac{\partial^2 H}{\partial \pi_\rho' \partial p_a} p_a' + \sum \frac{\partial^2 H}{\partial \pi_\rho' \partial \pi_a'} \pi_a'' \qquad \frac{\partial \Pi_\rho'}{\partial \pi_\sigma''} = \frac{\partial^2 H}{\partial \pi_\rho' \partial \pi_\sigma'}$$

$$\Pi_\sigma' = \frac{d}{dt}\left(\frac{\partial H}{\partial \pi_\sigma'}\right) = \sum \frac{\partial^2 H}{\partial \pi_\sigma' \partial p_b} p_b' + \sum \frac{\partial^2 H}{\partial \pi_\sigma' \partial \pi_b'} \pi_b'' \qquad \frac{\partial \Pi_\sigma'}{\partial \pi_\rho''} = \frac{\partial^2 H}{\partial \pi_\sigma' \partial \pi_\rho'}$$

equal to each other

[p. 114]

4) $P_\nu = -\frac{\partial H}{\partial p_\nu}$

$$\Pi_\mu = \frac{d}{dt}\left(\frac{\partial H}{\partial \pi_\mu'}\right) = \sum \frac{\partial^2 H}{\partial \pi_\mu \partial p_a} p_a' + \sum \frac{\partial^2 H}{\partial \pi_\mu' \partial p_a'} \pi_a''$$

So far as one can disregard an independence of the terms $\frac{\partial^2 H}{\partial \pi'_\mu \partial \pi'_a}$ from the quantities

The Reciprocity Laws

$$P_a^- = -\frac{\partial H}{\partial p_a} + \frac{d}{dt}\left(\frac{\partial H}{\partial p_a'}\right)$$

$$= -\frac{\partial H}{\partial p_a} + \sum \frac{\partial^2 H}{\partial p_v \partial p_a'} p_v' + \sum \frac{\partial^2 H}{\partial p_a' \partial p_v'} p_v''$$

a)

$$\frac{\partial P_a}{\partial p_b''} = \frac{\partial^2 H}{\partial p_a' \partial p_b'} = \frac{\partial P_b}{\partial p_a''}$$

Examples

1)

$$P_a = e_1 \qquad P_b = e_2 \qquad \frac{\partial e_1}{\partial i_2'} = \frac{\partial e_2}{\partial i_1'}$$

$$p_a'' = i_1' \qquad p_b'' = i_2''$$

Equality of mutual induction of two circuits also with algebraic sign

2) Conductor in magnetic field

$$P_a = P \qquad P_b = e$$

$$p''_a = p'' \qquad p''_b = i'$$

$\frac{\partial P}{\partial i'} = \frac{\partial e}{\partial p''} = 0$ That one equals zero has the consequence that the other one also equals zero.

[p. 115] b) Reciprocity laws that concern velocities.

$$\frac{\partial P_a}{\partial p_b'} = -\frac{\partial^2 H}{\partial p_a \partial p_b'} + \frac{\partial^2 H}{\partial p_b \partial q_a'} + \underbrace{\sum_v \frac{\partial}{\partial p_v}\left(\frac{\partial^2 H}{\partial p_a' \partial p_b'}\right) p_v' + \sum \frac{\partial}{\partial p_v'}\left(\frac{\partial^2 H}{\partial p_a' \partial p_b'}\right) p_v''}_{\frac{d}{dt}\left(\frac{\partial^2 H}{\partial p_a' \partial p_b'}\right)}$$

If one constructs the corresponding equation and adds, one obtains

$$\frac{\partial P_a}{\partial p'_b} + \frac{\partial P_b}{\partial p'_a} = \frac{d}{dt}\left(\frac{\partial^2 H}{\partial'_a \partial p'_b}\right)$$

If the system is cyclical, only states with constant p are considered, and H is a second-order function of the p' (= π'), the right hand side vanishes. The same is true if derived for a cyclical & a noncycl.

Examples.

$$1)\ \frac{\partial P}{\partial i} = -\frac{\partial e}{\partial p'}$$

Lenz's law

$$2)\ \frac{dp}{dT} = \frac{d\left(\frac{dS}{dT}\right)}{d\left(\frac{dv}{dt}\right)} = \frac{d\left(\frac{q}{T}\right)}{d\frac{dv}{dt}} \qquad P_a = p \qquad P_b = -\frac{dS}{dt}$$

not correct $\quad p_a' = \dfrac{dv}{dt} \qquad p_b' = T$

$$\text{Peltier}\ \frac{\partial e}{\partial T} = \frac{\partial q_1}{\partial i}\cdot\frac{1}{T} \qquad P_a = e \qquad P_b = -\frac{dS}{dt} = -\frac{q}{dt}\cdot\frac{1}{T}$$

$$\frac{dp_a}{dt} = i \qquad \frac{dp_b}{dt} = T$$

Another Derivation of the Fundamental Equations of the Material Point[73]

[p. 116]

At some place on the surface of the earth

let there be a number of identical weights as well as a spring. We suspend from the spring 0 1 2 . . . of the weights in succession, & thus obtain the lengths $l_0, l_1, l_2 \ldots$ of the spring.

We set the force exerted by the weights upon the spring equal to the number of the suspended weights and thus obtain a relative[74] measure of the force exerted upon the spring. We postulate that the force exerted by the spring upon the weights is of the same magnitude. Now we can use the spring to apply forces of specified magnitude to a given mass.

We know that a mat. p. that is not acted upon by external causes moves without acceleration. For it, $\frac{d^2x}{dt^2}$ etc. equal zero. Imagine that the question as to how the acceleration is related to the force for a freely suspended body is investigated with the help of our spring. If we assume that, in the case of some arbitrary, already present motion of the point and an arbitrarily large force,[75] the acceleration is always proportional to the acting force and is directed in the same way, then we obtain

[p. 117]

$$m\frac{d^2x}{dt^2} = X \quad m\frac{d^2y}{dt^2} = Y \quad m\frac{d^2z}{dt^2} = Z$$

<if we assume that the force acts in the direction *X*>
Because these equations state that
1) acceleration and force have the same direction

$$X : Y : Z = \frac{d^2x}{dt^2} : \frac{d^2y}{dt^2} : \frac{d^2z}{dt^2}$$

2) If *m* is taken as constant, then the magnitude of the acceleration

$$\sqrt{\left(\frac{d^2x}{dt^2}\right)^2 + \left(\frac{d^2y}{dt^2}\right)^2 + \left(\frac{d^2z}{dt^2}\right)^2}$$

is proportional to the magnitude of the force $\sqrt{X^2 + Y^2 + Z^2}$

If the acting force is not that of our gauge-spring but some other force, then it will be replaced by that force of the gauge-spring that produces the same motion. Then what was said above will hold for arbitrary forces.

Doc. 2

The Principle of Relativity and Its Consequences in Modern Physics

by A. Einstein

[*Archives des sciences physiques et naturelles*
29 (1910): 5–28; 125–144]

§1. The Ether

When it was realized that a profound analogy exists between the elastic vibrations of ponderable matter and the phenomena of interference and diffraction of light, it could not be doubted that light must be considered as a vibratory state of a special kind of matter. Since, moreover, light can propagate in places devoid of ponderable matter, one was forced to assume for the propagation of light a special kind of matter that is different from ponderable matter, and that was given the name "ether." Since the velocity of light propagation in bodies of low density, such as gases, is more or less the same as in a vacuum, it had to be assumed that the ether is the principal carrier of light phenomena in these bodies as well. Finally, the hypothesis of the presence of ether in the interior of liquids and solids was also necessary in order to make it possible to understand the propagation of light in these bodies, since it was impossible to explain the great velocity of propagation by the elastic properties of ponderable matter alone. For all these reasons, the existence of a special medium permeating all matter seemed beyond dispute, and the ether hypothesis formed an essential part of the picture of the universe which presented itself to the physicists of the last century.

The introduction of the electromagnetic theory of light brought about a certain modification of the ether hypothesis. At first the physicists did not doubt that the electromagnetic phenomena must be reduced to the modes of motion of this medium. But as they gradually became convinced that none of the mechanical theories of ether provided a particularly impressive picture of electromagnetic phenomena, they got accustomed to considering the electric and magnetic fields as entities whose mechanical interpretation is superfluous. Thus, they have come to view these fields in the vacuum as special states of the ether that do not require an analysis in greater depth.

What the mechanical interpretation of optical and electromagnetic phenomena has in common with the purely electromagnetic interpretation is the fact that both view the electromagnetic field as a special state of a hypothetical medium filling the whole of space. This is where these two interpretations differ fundamentally from the emission theory proposed by Newton, according to which light consists of particles in motion. According to the latter theory, a space containing neither ponderable matter nor light rays should be considered totally empty, while according to the mechanical and electromagnetic theories such a space should be considered as filled by ether. [1]

[2] §2. *The Optics of Moving Bodies and the Ether*

Once one accepts the ether hypothesis, one faces the question as to the kind of mechanical bonds that link ether to matter. When matter is in motion, does the ether participate in this motion completely, or is it only partly carried along, or else, is the ether completely stationary? These questions are fundamental for the optics and electrodynamics of moving bodies.

The simplest hypothesis is to assume that moving bodies carry along completely the ether they contain. It is on the basis of this hypothesis that Hertz developed an [3] electrodynamics of moving bodies that is free of contradictions. However, it follows from [4] a famous experiment by Fizeau that this hypothesis is not acceptable. This experiment, which can be considered an *experimentum crucis*, is based on the following considerations: Let u' be the velocity of propagation of light in a transparent and immobile medium. Suppose we impart to this medium a uniform translational motion of velocity v. If the medium completely carries along the ether it contains, then the light will propagate *with respect to the medium* in the same way as if the medium were at rest. In other words, u' will also be the velocity of propagation of the light with respect to the moving medium. To find the velocity with respect to an observer not taking part in the motion of the medium, it suffices, according to the rule of addition of velocities, to add vectorially the velocity v to the velocity u'. In the special case when u' and v have the same direction, one obtains either $u' + v$ or $u' - v$ for the desired sum, depending on whether u' and v are in the same or the opposite direction. But even the greatest velocities that could be imparted to a body are very small compared with the velocity of light; a very sensitive method is therefore needed in order to demonstrate the effect of the motion of the medium on this velocity. Fizeau devised the following experiment: We consider two light rays capable of interfering with each other, and two tubes filled with the same liquid. We pass one of the rays axially through each tube in such a way that each ray will interfere with the other after both exit from the tubes: the position of the fringes will be changed if the liquid moves axially in the tube.

From the different positions of the fringes when the velocity of the flow is varied, one can determine the propagation velocity of the light[1] in the moving liquid, i.e., in the medium, with respect to the walls of the tube. Proceeding in this way, Fizeau did not obtain the value $u' \pm v$, as had to be expected from what we have said above, but the value $u' \pm \alpha v$, where α is a number between 0 and 1 that depends on the refractive index n:

---

[1] More exactly, the propagation velocity of the planes of equal phase of the light beam.

$$\alpha = 1 - \frac{1}{n^2}.$$[2]

Thus, the light is indeed carried along by the moving liquid, but only partially.

This experiment showed the hypothesis of the complete carrying along of the ether to be unacceptable, so that only two possibilities remained:

1. The ether is completely immobile, i.e., it does not take part in the motion of the matter at all.
2. The ether inside the moving matter is movable, but it moves with a velocity different from that of the matter.

One cannot go very far in developing the second hypothesis without introducing arbitrary assumptions about the relationship between the ether and matter in motion. In contrast, the first hypothesis is perfectly simple, and its development with the aid of Maxwell's theory does not necessitate any arbitrary assumption that might complicate the foundations of the theory.

Assuming that ether is completely immobile, H. A. Lorentz conceived in 1895 a very satisfactory theory of electromagnetic phenomena,[3] a theory which not only permitted a quantitative prediction of Fizeau's experiment, but also provided a simple explanation of almost all the experiments that one can imagine in this sphere.

According to Lorentz, matter consists of elementary particles at least part of which are provided with electrical charges. A charged particle moving with respect to the ether may be compared to an element of current. The actions of the electromagnetic field on the particle, and the reactions of the particle on the field, are the only bonds that bind matter to the ether. In the regions of ether where the space is not already occupied by a particle, the magnetic and electric field strengths are expressed by Maxwell's equations for the free ether, if one assumes that the equations refer to a coordinate system that is at rest with respect to the ether. The fecundity of Lorentz's theory is due to the fact that the states of matter that play a role in optics and electromagnetism are explained unambiguously by the relative positions and the motions of the charged particles.

§3. Experiments and Consequences Not Reconcilable with the Theory

From Fizeau's experiment one had to conclude that the ether is not carried along completely by matter in motion but that, instead, there occurs a relative displacement of the one with respect to the other. The earth being a body that rotates around its axis and revolves around the sun with velocities that change their directions in the course of

---

[2] This expression neglects the effects of dispersion.

[3] H.A. Lorentz, *Versuch einer Theorie der elektrischen und optischen Erscheinungen in bewegten Körpern.* Leiden, 1895. New edition, Leipzig, 1906. [5]

the year, one was bound to believe that, in our laboratories, the ether would take a slight part in the motion of the earth the way it did in the motion of the liquid in Fizeau's investigations. From this it would seem to follow that the relative velocity of the ether with respect to our equipment must vary with time, and that one therefore should expect that an apparent spatial anisotropy be observed in optical phenomena, i.e., that these phenomena should depend on the orientation of the equipment. Thus, in vacuum or in the atmosphere, light should propagate faster in the direction of the earth's motion than in the opposite direction. Experimental verification of this consequence of the theory was unthinkable, because the order of magnitude of the term considered is that of the ratio of the velocity of the earth to the velocity of light, i.e., of the order of 10^{-4}, and one could not hope to attain such precision in the direct determination of the velocity of light. Also—and this is a most important point—all terrestrial methods for measuring the velocity of light employ light rays that travel along a closed (back and forth) rather than a simple path, this due to the fact that the times of departure and arrival of the rays must be determined with the help of one and the same device, as for example a notched wheel.

Many optical phenomena are known in which variations in the velocity of light of the order of 10^{-4} become capable of being detected, and when observing these phenomena one should have expected to find, according to the theory, that the results vary with the orientation of the apparatus with respect to the motion of the earth. Without dwelling on these experiments, let me only say that all of them gave negative results. Thus, [6] Fizeau's experiment led to the hypothesis of the relative motion of the ether with respect to moving bodies. None of the other experiments confirmed this hypothesis. The theory of Lorentz[4] provided, at least partly, a key to the solution of this puzzle. A uniform translation of velocity v of the apparatus with respect to the ether does have an influence on the phenomena, but this influence on the distribution of visible light intensities can only be detected starting with the terms of the order of $\left(\frac{v}{c}\right)^2$ in the Lorentz equations that give this distribution, c being the velocity of light in a vacuum. This seemed, therefore, to explain the negative results of the experiments aimed at showing the relative motion of the earth with respect to the ether. Still, the negative result of one of these experiments presented a real headache for the theoreticians: I am speaking of the famous investigations by Michelson and Morley.[5] These physicists based themselves on the following observation: Let M and N be two points of a solid body; suppose a ray begins at M and travels toward N, where it is reflected, and then returns to M. In this case, if the body undergoes a uniform translation with respect to the ether,

---

[4] It should be added, for the sake of completeness, that Mr. Lorentz did not consider bodies capable of rotating the plane of polarization when they are not in a magnetic field (naturally active bodies).

[7] [5] A. A. Michelson and E. W. Morley, *Amer. Jour. of Science* 34 (1887): 333–345.

the theory predicts that the time t needed for the light to travel the closed path MNM will vary depending on whether the body is moving in the direction MN or perpendicular to it. True, the difference is very small, being of the order of $\left(\frac{v}{c}\right)^2$, that is to say, of the order of magnitude 10^{-8} if v is taken as the velocity of the earth. But Michelson and Morley were able to devise an interference experiment in which this slight difference would have to become detectable. The essential features of their arrangement were as follows: Light rays coming from the source S (Fig. 1) are split into two beams by means of a transparent mirror at A. One of the beams is then reflected at B and returns to A, where it splits and yields a ray traveling to I. The other crosses the mirror and travels to B', where it is reflected toward A; there it splits, yielding, too, a ray that goes to I, where the two rays interfere. The position of the fringes depends on the difference between the routes ABA and $AB'A$ taken by the two rays during their travels. This difference between the routes should have depended on the orientation of the equipment; one should have observed a displacement of the fringes the moment AB', instead of AB, coincided with the direction of the earth's motion. However, nothing of the kind was observed, and as a result the foundation of Lorentz's theory seemed extremely shaky. To save the theory, Lorentz and FitzGerald resorted to a strange hypothesis: they assumed that *each body* in motion with respect to the ether contracts in the direction of motion [8] by a fraction equal to $\frac{1}{2}\left(\frac{v}{c}\right)^2$, or—which amounts to the same if only terms of second

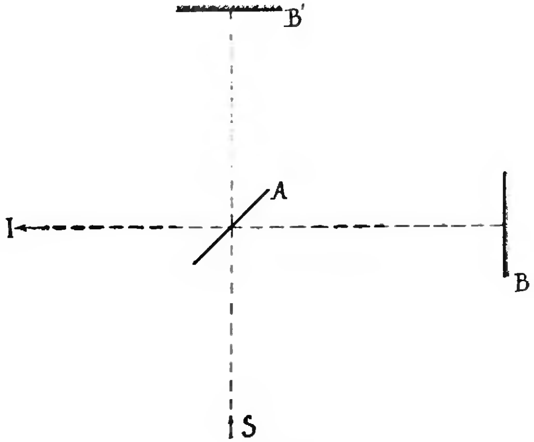


Fig. 1.

order are considered—that the length of the body is diminished in that direction in the ratio $1 : \sqrt{1 - \frac{v^2}{c^2}}$.

This hypothesis succeeded, in effect, in eliminating the disagreement between theory and experiment. But the theory did not offer an intellectually very satisfying whole. It was based on the existence of an ether that one had to conceive as being in motion with respect to the earth, with the consequences of this motion forever unverifiable by experiment; this peculiarity could only be explained by introducing *a priori* implausible hypotheses into the theory. Could one really believe that, by a curious accident, the laws of nature present themselves to us in such a highly unusual way that none of them allows us to know about the fast motion of our planet through the ether? Is it not more plausible to assume that some wrong or defective argument had led us to this impasse?

Before explaining how these difficulties were overcome, we will show that even in particular cases, the theory based on the existence of the ether does not always offer an intellectually satisfying explanation of the phenomena, even though that representation is not in direct conflict with experiment.

[9] Let us consider, for example, a magnetic pole moving with respect to a closed circuit. If the number of lines of force that cross the surface enclosed by the circuit changes with time, a current will be generated in the conductor. As we know, the generated current depends only on the rate of change of the flux that crosses the circuit. This rate depends only on the *relative motion* of the pole with respect to the circuit; in other words, from the point of view of the result produced, it does not make any difference whether it is the circuit that moves, while the pole is at rest, or whether the opposite is the case. But to understand this phenomenon from the point of view of the ether theory, one must attribute fundamentally differing states to the ether, depending on whether it is the pole or the circuit that is moving with respect to the ether. In the first case one must consider that the motion of the pole has the effect of changing the strength of the magnetic field at the various points of the ether at each instant of time. The change thus engendered produces an electric field with closed lines of force, the existence of which is independent of the presence of the circuit. This field, like every field of electrical force, possesses a certain energy; it is this field that produces the electric current in the circuit. If, in contrast, it is the circuit that is in motion while the pole remains at rest, then no electric field will be generated. In that case, the electrons present in the conductor are subjected to ponderomotive forces arising from the motion of the electrons in the magnetic field, forces that make the electrons travel, thus producing the induced electric current.

Thus, in order for them to be understood with the aid of the ether theory, two experiments that are not essentially different in themselves require that substantively different states be attributed to the ether. Moreover, such a split, alien to the nature of facts, is introduced every time one resorts to the presence of the ether to explain phenomena caused by relative motions of two bodies.

§4. The Principle of Relativity and the Ether

What is the source of the difficulties we have just seen?

Lorentz's theory contradicts the purely mechanical models to which physicists hoped to reduce all the phenomena of the universe. For while mechanics in effect admits of no absolute motion, but only the motions of bodies relative to each other, there is a particular state in Lorentz's theory that corresponds physically to the state of *absolute rest*: that is the state of a body which is not in motion with respect to the ether.

If the fundamental equations of Newtonian mechanics, referred to a coordinate system that is not undergoing accelerated motion, are referred by means of the relations

$$(1) \qquad \begin{cases} t' = t \\ x' = x - vt \\ y' = y \\ z' = z \end{cases}$$

to a new coordinate system which is in uniform translational motion with respect to the first, one obtains equations in t', x', y', z' that are identical to the original equations in t, x, y, z. In other words, the Newtonian laws of motion transform to laws of the same form when one passes from one coordinate system to another one that is in uniform translational motion with respect to the first. This is the property we express when we say that the principle of relativity is satisfied in classical mechanics.

More generally, we will state the *principle of relativity* in the following way:

The laws governing natural phenomena are independent of the state of motion of the coordinate system with respect to which the phenomena are observed, provided that this system is not in accelerated motion.[6]

If one transforms the fundamental equations of Lorentz's theory by means of the transformation equations (1), one obtains equations of another form, in which the quantities x', y', z' no longer occur symmetrically. Thus, the theory of Lorentz, based on the ether hypothesis, does not admit of the principle of relativity. The difficulties

---

[6] In all this we assume that the notion of acceleration has an objective meaning, or in other words, that an observer attached to a coordinate system is able to determine by experiment whether the system is or is not in accelerated motion. *From now on we will consider only coordinate systems in nonaccelerated motion.* [10]

encountered so far are mainly due to this fact; the deeper reasons will become apparent later on. Be that it as it may, it is so much the more improper to accept a theory that is not consistent with the principle of relativity, the more so because this principle has not been contradicted by a single experimental fact.

§5. On Two Arbitrary Hypotheses Contained Implicitly in the Customary Notions of Time and Space

We have seen that, assuming the existence of the ether, we were led by experiment to view this medium as stationary. Further, we have seen that the theory so founded permits us to predict the main experimental facts but leaves something to be desired in one respect: It does not admit of the principle of relativity, contrary to everything that experimental investigations have taught us. The question arises: *Is it really impossible to reconcile the essential foundations of Lorentz's theory with the principle of relativity?*

If we wish to attempt such a reconciliation, the first step we must take is to *give up the ether.* For, on the one hand, we have been obliged to admit that the ether is stationary, whereas, on the other hand, the principle of relativity demands that the laws of natural phenomena referred to a uniformly moving coordinate system S' be identical with the laws of these same phenomena referred to a system S at rest with respect to the ether. But there is no reason to assume the immobility of ether, which is demanded by theory and experiment, any more with respect to the system S' than with respect to the system S; these two systems cannot be distinguished from each other, and it is therefore improper to make one of them play a special role by saying that it is at rest with respect to the ether. From this it follows that the only way to arrive at a satisfactory theory is to give up the notion of a medium filling all of space.

This is the first step to be taken.

To go a step further, we must reconcile the principle of relativity with an essential consequence of Lorentz's theory, because giving up this consequence would amount to giving up the most fundamental formal properties of the theory. And here is the consequence in question:

A ray of light in vacuum always propagates with the same velocity c, *which velocity is independent of the motion of the body that emits the ray.*

We will see in section 6 that we are raising this consequence to the the status of a principle. For the sake of brevity, we will from now on call it *the principle of the constancy of the velocity of light.*

In Lorentz's theory this principle holds only for a system in a special state of motion: In effect, the system must be at rest relative to the ether. If we want to preserve the principle of relativity, we must assume that the principle of the constancy of the velocity of light holds for any arbitrary system not in accelerated motion. At first glance this seems impossible. For let us consider a light ray that propagates with velocity c with respect to the system S, and suppose that we seek to determine the velocity of propagation relative to a system S′ that is in uniform translational motion with respect to the first system. Applying the rule of addition of velocities (the rule of the parallelogram of velocities), we will generally find a velocity different from c; in other words, the principle of the constancy of the velocity of light that is valid with respect to S is not valid with respect to S'.

So that the theory based on these two principles should not lead to contradictory results, one must renounce the customary rule of addition of velocities or, better, replace it with another rule. Well founded as this rule may seem to be at first glance, it conceals no less than two arbitrary hypotheses, which consequently, as we shall see, hold sway over all of kinematics. It is these hypotheses that made us think that, with the aid of the transformation equations (1), the incompatibility of Lorentz's theory with the principle of relativity can be demonstrated.

The first hypothesis we wish to discuss concerns the physical notion of time measurement. To measure time, we use clocks. What is a clock? By a clock we understand any thing characterized by a phenomenon passing periodically through identical phases so that we must assume, by virtue of the principle of sufficient reason, that all that happens in a given period is identical with all that happens in any arbitrary period.[7] If the clock comes in the form of a mechanism that is provided with clock [11] hands, then to mark the positions of the clock's hands is tantamount to counting the number of moments elapsed. By definition, to measure the time interval during which an event takes place means to count the number of time periods indicated by the clock from the beginning till the end of the event in question.

The meaning of this definition is perfectly clear as long as the clock is sufficiently close to the place at which the event occurs, so that the clock and the event can be observed simultaneously. If, on the contrary, the event is taking place in some corner far away from the clock, then it will no longer be possible to establish immediately a correspondence between the different phases of the event, and the different positions of the clock's hands. The definition is therefore deficient and needs to be completed. Until now one has completed it unawares.

To determine the time at each point in space, we can imagine it populated with a very great number of clocks of *identical construction*. Let us consider the points A, B, C, …, each of which is furnished with a clock and is referred to a system in nonaccelerated motion with the aid of time-independent coordinates. We will now be able to know the time at any of the locations at which we choose to put a clock. If we choose a sufficiently large number of clocks, so that we can ascribe to each of them a sufficiently small domain, we will be able to fix any instant whatsoever, at any location in space, to any

---

[7] Thus, we postulate that two identical phenomena are of the same duration. The perfect clock thus defined plays a role in the measurement of time that is analogous to the role played by the perfect solid body in the measurement of lengths.

degree of accuracy desired. But we cannot obtain in this manner a definition of time useful to a physicist, because we did not say what the position of the clock hands should be at a given instant of time at different spatial points. We forgot to synchronize our clocks, and it is clear that the intervals of time elapsed during some event have a certain extension that will vary considerably as the event occupies this or that point in space. Suppose, for example, that we are studying the motion of a material point whose trajectory passes through the points $A, B, C \ldots.$ At the moment when the point passes through A, we will note the instant t_A, indicated by the clock located at this point. In the same way we will register the instants t_B, $t_C, \ldots$ of the passage through the points $B, C \ldots.$ Since the coordinates of the points $A, B, C, \ldots$ on the axes of the system S can be obtained directly—by performing measurements with a graduated measuring rod, for example—then, by relating the coordinates x_A, y_A, $z_A \ldots$ of the points A, B, C, $\ldots$ to the instants t_A, t_B, $t_C \ldots$, we can obtain the coordinates x, y, $z \ldots$ of the moving material point as a function of a variable t, which we will call time. It is clear that the form of this function will depend essentially on the way the clocks had been set before they were placed at their respective locations.

To get a complete physical definition of time, we have to take an additional step: We have to say in what manner all of the clocks have been set at the start of the experiment. We will proceed as follows: First, we furnish ourselves with a means of sending signals, be it from A to B, or from B to A. This means should be such that we have no reason whatsoever to believe that the phenomena of signal transmission in the direction AB will differ in any way whatsoever from the phenomena of signal transmission in the direction BA. In that case there is, obviously, only one way of regulating the clock at B against the clock at A in such a manner that the signal traveling from A to B would take the same amount of time—measured with the clocks described above—as the signal traveling from B to A. If we denote by

t_A	the reading of the clock at	A	at the moment signal	AB	leaves A
t_B	"	B	"	AB	arr. at B
t_B'	"	B	"	BA	leaves B
t_A'	"	A	"	AB	arr. at A

then we have to set the clock at B against that at A in such a way that

[12]
$$t_B - t_A = t_A' - t^{B'} .$$

For these signals we can use, for example, sound waves that propagate between A and B through a medium that is at rest with respect to these points.[8] We can just as

---

[8] The medium must be at rest—or at the very least must not have any velocity component in the direction AB—so that the paths AB and BA can be equivalent.

well use light rays propagating through the vacuum or through a homogeneous medium at rest with respect to A and B. It does not make any difference whether we choose this or that kind of signals. If two kinds of signals were to produce discrepant results, we would have to conclude that, for at least one of the two kinds of signals, the condition of equivalence of the paths AB and BA was not satisfied.

Still, of all the signals that can be used, we are going to prefer those that make use of light rays propagating in the vacuum, because the synchronization requires that the path out and the path back be equivalent, and in our case this equivalence is satisfied by definition, since, by virtue of the principle of the constancy of the velocity of light, light in the vacuum always propagates with the velocity c.

Hence we will have to synchronize our clocks in such a way that the time spent by a signal traveling from A to B be equal to the time spent by an identical signal traveling from B to A.

Now we possess a well-defined method by which to synchronize two clocks with respect to each other. Once the synchronization has been done, we will say that the two clocks are *in phase*. If, step by step, we regulate clock B against clock A, clock C against clock $B \ldots$, we obtain a series of clocks such that any of them is in phase with the preceding one. Moreover, any two nonconsecutive clocks in the series must also be in phase by virtue of the principle of the constancy of the velocity of light.

The totality of the readings of all of these clocks in phase with one another is what we will call the *physical time*.

By an *elementary event* we will understand an event that is supposed to be concentrated in one point and is of infinitely short duration. By the *time coordinate* of an elementary event we will understand the indication, at the instant of the event's occurrence, of a clock that is situated infinitely close to the point at which the event takes place. An elementary event is thus defined by four coordinates, namely the time coordinate and the three coordinates that define the spatial position of the point in which the event is supposed to be concentrated.

Thanks to our physical definition of time, we can give a perfectly defined meaning to the concepts of simultaneity and nonsimultaneity of two events occurring at locations removed from one another. In the same way, the introduction of the coordinates x, y, z of a spatial point gives a completely defined meaning to the concept of position. Thus, for example, to say that the abscissa of a point P situated on the axis is x, is the same as saying that we must hit upon the point P if, starting from the origin, we apply, with a ruler, a unit length x times along the axis. We proceed in the same way to fix the position of a point if all three coordinates x, y, z are different from zero, except that the operations are a little more complicated. Be it as it may, the specification of the particular coordinates always involves the idea of a well-defined experiment concerned with the position of solid bodies.[9]

Let us now make an important remark: In order to define the physical time with respect to a coordinate system, we used *a group of clocks in a state of rest relative to that system.* According to this definition, the time readings or the establishment of the simultaneity of two events have meaning only if the motion of the group of clocks or that of the coordinate system is known.

Consider two nonaccelerated coordinate systems S and S' in uniform translational motion with respect to one another. Suppose that each of these systems is provided with a group of clocks invariably attached to it, and that all clocks belonging to the same system are in phase. Under these conditions, the readings of the group attached to S will define the physical time with respect to S; analogously, the readings of the group attached to S' define the physical time with respect to S'. Each elementary event will have a time coordinate t with respect to S, and a time coordinate t with respect to S'. *But, we have no right to assume* a priori *that the clocks of the two groups can be set in such a manner that the two time coordinates of the elementary event would be the same, or in other words, in such a way that* t *would be equal to* t'. To assume this would mean to introduce an arbitrary hypothesis. This hypothesis has been introduced into kinematics up to the present time.

The second arbitrary hypothesis introduced in kinematics concerns the configuration of a body in motion. Consider a bar AB moving in the direction of its axis with velocity v with respect to a coordinate system S not in accelerated motion. What should we understand by the "length of the bar"? One is at first inclined to believe that this concept does not require any special definition. However, we will immediately see that nothing of the sort is true if we consider the following two methods of determining the length of the rod:

1. One accelerates the motion of an observer furnished with a measuring rod until he attains the velocity v, i.e., until he is at relative rest with respect to the bar. The observer then measures the length AB by successively applying the measuring rod along the bar.

2. Using a group of clocks in phase with each other and at rest with respect to the system S, one determines the two points P_1 and P_2 of S where one finds the two ends of the bar at the instant t; after that, one determines the length of the straight line

---

[9] We do not claim that the time and space coordinates must necessarily be defined in such a way that their definitions could serve as the basis of measurement methods that permit the experimental determination of these coordinates—the way it has been done above. But whenever the quantities t, x, y, z are introduced in the capacity of purely mathematical variables, equations in physics will have meaning only if they allow the elimination of these quantities.

connecting the two points P_1 and P_2 by successively applying the measuring rod along the line P_1P_2, which is assumed to be a material line.

As one can see, it is with some justification that the results obtained in the first and in the second case are designated as the "length of the bar." But in no way does this mean *a priori* that these two operations must necessarily lead to the same *numerical value* for the length of the bar. All that one can deduce from the principle of relativity—and this is easy to demonstrate—is that the two methods lead to the same numerical value for the length only when the bar AB is at rest relative to the system S. But in no way is it possible to assert that the second method yields a numerical value for the length independently of the velocity v of the bar.

More generally, if the configuration of a body in uniform translational motion with respect to S is determined by ordinary geometric methods, by means of measuring rods or other solid bodies moving in exactly the same way the results of measurement turn out to be independent of the velocity v of the translation: these results give us what we will call the *geometric configuration* of the body. By contrast, if one marks in the system S the positions of various points of the body at a given instant, and determines the configuration formed by these points by geometric measurements using measuring rods at rest with respect to S, one obtains as a result what we will call the *kinematic configuration* of the body with respect to S. [13]

The second hypothesis used unconsciously in kinematics can thus be expressed as follows: The kinematic configuration and the geometric configuration are identical.

[Continued in the 15 February issue of *Archives*, pp. 125–144]

§6. The New Transformation Equations (the Lorentz Transformation) and Their Physical Meaning [14]

To emphasize the considerations discussed in the preceding section, it is easy to see that the rule of the parallelogram of velocities, which made one think that Lorentz's theory cannot be reconciled with the theory of relativity, is based on unacceptable arbitrary hypotheses. In fact, this rule leads to the following transformation equations,

$$t' = t, \quad x' = x - vt, \quad y' = y, \quad z' = z,$$

or more generally,

$$t' = t, \ x' = x - v_x t, \ y' = y - v_y t, \ z' = z - v_z t.$$

The first of these equations expresses, as we have seen, an ill-founded hypothesis about the time coordinates of an elementary event taken with respect to two systems S and S' that are in uniform translational motion with respect to each other. The other three equations express the hypothesis that the kinematic configuration of the system S'

with respect to the system S is identical with the geometric configuration of the system S'.

If one abandons the ordinary kinematics and builds a new kinematics based on the new foundations, one arrives at transformation equations different from those given above. And now, we are going to show[10] that based on

1. *The principle of relativity* and
2. *The principle of the constancy of the velocity of light*, we arrive at transformation equations that allow us to see that Lorentz's theory is compatible with the principle of relativity.

The theory based on these principles we shall call *the theory of relativity.*

Let S and S' be two equivalent coordinate systems, i.e., systems in which lengths are measured in the same unit, and each of which possesses a group of clocks that run in synchrony when the two systems are at relative rest with respect to each other.[11] According to the principle of relativity, physical laws must be identical for the two systems regardless of whether the systems are at relative rest or in uniform translational motion with respect to each other. Thus, in particular, the velocity of light in a vacuum must be expressed by the same number in the two systems. Let t, x, y, z be the coordinates of an elementary event with respect to S, and t', x', y', z' the coordinates of the same event with respect to S'. We seek to find the relations that link these two groups of coordinates. It can be shown that these relations must be linear because of the homogeneity of time and space,[12] and time t is therefore linked with time t' by a formula of the form

$$t' = At + Bx + Cy + Dz. \tag{2}$$

Furthermore, for an observer linked to S it follows from this, in particular, that the three coordinate planes of S' are planes in uniform motion; but, in general, these three planes will not form a rectangular triad even though we assume that the system S' is rectangular for an observer connected with this system. However, if, referring to the system S, we have chosen the position of the x'-axis parallel to the direction of the motion of S', it will follow for reasons of symmetry that the system S' will appear as rectangular. In particular, we may choose the relative position of the two coordinate systems in such a way that the x-axis will permanently coincide with the x'-axis, and the

---

[10] A. Einstein, *Ann. der Phys.* 17 (1905): 891–921, and *Jahrbuch der Radioaktivität und Elektronik* 4 (1907): 411–462.

[11] It should be noted that we will always implicitly assume that the fact of a measuring rod or a clock being set in motion or brought to rest does not change the length of the rod or the rate of the clock. [15]

[12] Cf. footnote 15.

y'-axis will remain parallel to the y-axis, and that, in addition, the same-named axes have the same orientation for the observer connected with S. We will count the time from the instant when the origins of the two systems coincide. Under these conditions, the relations sought are homogeneous, and the following equations

$$\begin{aligned} x' &= 0 \text{ and} & x - vt &= 0 \\ y' &= 0 \text{ and} & y &= 0 \\ z' &= 0 \text{ and} & z &= 0 \end{aligned}$$

are equivalent, or, in other words, the coordinates x, y, z, x', y', z' are linked by relations of the form

$$(3) \qquad \begin{cases} x' = \mathrm{E}(x - vt) \\ y' = \mathrm{F}y \\ z' = \mathrm{G}z \end{cases}$$

To determine the constants A, B, C, D, E, F, G entering equations (2) and (3), we assert that, according to the principle of the constancy of the velocity of light, the velocity of propagation has the same value c with respect to the two systems, or, in other words, that the two equations

$$(4) \qquad \begin{cases} x^2 + y^2 + z^2 = c^2t^2 \\ x'^2 + y'^2 + z'^2 = c^2t^2 \end{cases}$$

are equivalent. Replacing in the second of these equations t', x', y', z' by their values [16] obtained from (2) and (3), and equating it with the first equation, one can easily find that the transformation equations sought are of the form

$$(5) \qquad \begin{cases} t' = \varphi(v) . \beta . \left(t - \dfrac{v}{c^2}x\right) \\ x' = \varphi(v) . \beta . (x - vt), , \\ y' = \varphi(v) . y \\ z' = \varphi(v) . z \end{cases}$$

where

$$\beta = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}},$$

and $\varphi(v)$ is a function of v that is to be determined. We can easily find $\varphi(v)$ by introducing a third coordinate system S'', which is equivalent to the first two systems, is moving relatively to S' with a uniform velocity $-v$, and is oriented with respect to S' as S' is oriented with respect to S. Then, applying equation (5) twice, we obtain

$$\begin{aligned} t'' &= \varphi(v)\,.\,\varphi(-v)\,.\,t \\ x'' &= \varphi(v)\,.\,\varphi(-v)\,.\,x \\ y'' &= \varphi(v)\,.\,\varphi(-v)\,.\,y \\ z'' &= \varphi(v)\,.\,\varphi(-v)\,.\,z \end{aligned}$$

Since the origins of S and S'' are permanently coincident, the axes have the same orientation, and the systems are equivalent, we must necessarily have

$$\varphi(v)\,.\,\varphi(-v) = 1.$$

Since, moreover, the relation between y and y' (as also that between z and z') does not depend on the sign of v, we have

$$\varphi(v) = \varphi(-v).$$

From this it follows that

$$\varphi(v) = 1$$

($\varphi(v) = -1$ is here inappropriate), and that the transformation equations are

$$\text{(I)} \qquad \left\{ \begin{aligned} t' &= \beta\left(t - \frac{v}{c^2}x\right) \\ x' &= \beta(x - vt)\,, \\ y' &= y \\ z' &= z \end{aligned} \right.$$

where

$$\beta = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

These transformation equations have been introduced into electrodynamics in a very felicitous manner by Mr. Lorentz. We will call them *Lorentz transformations.* [17]

If we solve these equations with respect to t, x, y, z we obtain equations of the same form, except that the primed letters are replaced by nonprimed ones, and v is replaced by $-v$. Moreover, this result is an obvious consequence of the principle of relativity: relative to S', S is moving with velocity $-v$ parallel to the x and x' axes.

By combining the transformation equations with the equations expressing the rotation of one system with respect to another one, we can obtain the most general transformations of coordinates.

§7. *Physical Interpretations of the Transformation Equations*

1. Consider a body attached to S'. Let x'_1, y'_1, z'_1 and x'_2, y'_2, z'_2 be coordinates of two points of the body. At any instant t of the system S we will have the following relations between these coordinates:

$$(6)\qquad \begin{cases} x_2 - x_1 = \sqrt{1 - \dfrac{v^2}{c^2}}\,(x'_2 - x'_1) \\ y_2 - y_1 = y'_2 - y'_1 \\ z_2 - z_1 = z'_2 - z'_1 \end{cases}$$

This shows us that the kinematic configuration of a body in uniform translational motion with respect to a coordinate system depends on the velocity v of the translation. Furthermore, the kinematic configuration differs from the geometric configuration solely by a contraction in the direction of the motion, a contraction which is in the ratio $1 : \sqrt{1 - \dfrac{v^2}{c^2}}$. A relative motion of two reference systems with a velocity v that exceeds the velocity of light in vacuum is incompatible with the principle here assumed.

One recognizes at once in these equations the hypothesis of Messrs. Lorentz and FitzGerald (§3). This is the hypothesis that looked so strange to us and that had to be introduced to explain the negative results of the experiment of Michelson and Morley. Here this hypothesis appears naturally as an immediate consequence of the principles assumed.

2. Let us consider a clock H' which is at rest at the origin of S', and which runs p_0 times faster than one of the clocks used for the determination of physical time in the systems S or S'. In other words, when the two clocks are compared while at relative rest, clock H' will indicate p_0 periods during the unit time indicated by the other clock.

How many periods will clock H' indicate during unit time if observed from the system S?

Clock H' will indicate the end of a period at the times

$$t_1' = \frac{1}{p_0}, \quad t_2' = \frac{2}{p_0}, \quad t_3' = \frac{3}{p_0}, \ldots\ldots t_n' = \frac{n}{p_0}.$$

Since we seek the time with respect to S, the first of the transformation equations (I) will have to be written

$$t = \beta\left(t' - \frac{v}{c^2}x'\right)$$

and since clock H' is at rest at the origin of S', we must always have

$$x' = 0,$$

which yields

$$t_n = \beta t_n' = \frac{\beta}{p_0}n.$$

Observed from S, clock H' thus indicates

$$p = \frac{p_0}{\beta} = p_0\sqrt{1 - \frac{v^2}{c^2}}$$

periods in a unit time. In other words, a clock moving uniformly with velocity v with respect to a reference system runs, as observed from this system, $1 : \sqrt{1 - \frac{v^2}{c^2}}$ times slower than an identical clock that is at rest with respect to this system.

And here is an interesting application of the preceding formula. In 1907, Mr. J. Stark[13] remarked that canal ray ions emit spectral lines that give rise to a sort of Doppler phenomenon, namely, a displacement of spectral lines proceeding from the motion of the source. Since the oscillatory phenomena that produce a spectral line must be viewed as intra-atomic phenomena whose frequencies are uniquely determined by the nature of the ions, we can use these ions as clocks. The frequency p_0 of the oscillatory motion of the ions provides us with a means to measure time. This frequency is found by observing the spectrum produced by ions of the same kind but at rest with respect to

---

[18] [13] J. Stark, *Ann. der Phys.* 21 (1906): 401–456.

the observer. The preceding formula thus shows that besides the phenomenon known as the Doppler phenomenon, there exists an effect of motion on the source that diminishes the apparent frequency of the ion. [19]

3. Let us consider the equations of motion of a point moving in uniform translation with velocity u' with respect to S':

$$\begin{aligned} x' &= u_x' t' \\ y' &= u_y' t' . \\ z' &= u_z' t' \end{aligned}$$

If one replaces x', y', z', t' by their values as functions of x, y, z, t by means of equation (I), one obtains x, y, z as functions of t and, hence, the components u_x, u_y, u_z of the velocity u of the point with respect to the system S. In this way it is possible to obtain the formula that expresses the theorem of the addition of velocities in its general form, and one can immediately see that the law of the parallelogram of velocities is valid only in first approximation. In the special case when the velocity u' has the same direction as the velocity v of the translation of S' with respect to S, one easily obtains

$$u = \frac{v + u'}{1 + \frac{vu'}{c^2}}. \tag{7}$$

This equation shows that if one adds two velocities, each smaller than the velocity of light in a vacuum, one always obtains a resultant velocity that is smaller than the velocity of light. For if one sets $v = c - \lambda$, $u' = c - \mu$, where λ and μ are positive and smaller than c, one gets

$$u = c\frac{2c - \lambda - \mu}{2c - \lambda - \mu + \frac{\lambda\mu}{c}} < c.$$

From this it also follows that when one adds the velocity of light c and a velocity smaller than c, one always obtains the velocity of light. Now we can understand why Fizeau could not find $u + v$ for the sum of the velocity u' of the light in a liquid and the velocity v of the liquid in the tube (§2). For if higher than first-order terms are neglected, and the ratio $\frac{c}{u'}$ is replaced by the refractive index n of the liquid,[14] equation (7) can be written:

$$u = u' + v\left(1 - \frac{1}{n^2}\right).$$

This equation is identical to the one found experimentally by Fizeau.

Yet another consequence, as strange as it is interesting, follows immediately from the addition theorem. It can be shown that no means exist by which to send signals that [20] would travel faster than light in a vacuum. Let us consider a bar that is moving uniformly along the x-axis of S with velocity $-v(|v|<c)$, and by means of which one can send signals propagated with velocity u' with respect to the bar. Suppose that an observer A is located at the point $x = 0$ of the x-axis and an observer B at the point $x = x$ of the same axis, and that both are at rest in S. If the observer A sends a signal to B by means of the bar, the signal will be transmitted with velocity $\dfrac{v - u'}{1 - \dfrac{vu'}{c^2}}$ with respect to these observers. The time needed for this transmission will thus be

$$T = x_1 \frac{1 - \dfrac{vu'}{c^2}}{v - u'},$$

where v can take any value smaller than c. Now, if we assume that u' is larger than c, then one can always choose v in such a way that T would be negative. There would have to exist a transmission phenomenon such that the signal would arrive at its goal before having been emitted: The effect would precede the cause. Even though such a result is not inadmissible from the logical point of view, it so contradicts all of our empirical knowledge that we can consider that the impossibility of having $u' > c$ has been demonstrated.

4. The theory of relativity based on the principles assumed here permits us also to obtain the formulas expressing the Doppler and aberration phenomena in their general form. All we have to do is to compare the vector proportional to

---

[14] Strictly speaking, the index n does not correspond to the refractive index of the liquid for the frequency of the source used in the experiment, but to the index of the liquid for the frequency an observer moving with the liquid would ascertain.

$$\sin \omega\left(t - \frac{lx + my + nz}{c}\right),$$

which is the vector of a planar light wave propagated in a vacuum with respect to S, with the vector proportional to

$$\sin \omega'\left(t' - \frac{l'x' + m'y' + n'z'}{c}\right),$$

which is the vector of the same wave with respect to S'. By replacing t', x', y', z' in the last expression by their values obtained from the transformation equations (I), and equating it with the first expression, we find the relation connecting ω', l', m', n' with ω, l, m, n. By means of these relations it is easy to establish the formulas for aberration and the Doppler phenomenon.

The fundamental importance of the transformation equations (I) is first of all due to the fact that they provide a criterion that permits us to check the correctness of a physical theory. In effect, it is necessary for any equation that expresses a physical law to transform to an equation of identical form if the variables t', x', y', z' are substituted for the variables t, x, y, z by means of the transformation equations. In the second place, the transformation equations provide a means for finding the laws applicable to a body in rapid motion if one already knows the laws applicable to the same body when at rest or in infinitely slow motion.[15]

§8. Remarks about Some Formal Properties of the Transformation Equations

Let us consider two coordinate systems Σ and Σ' the origins of which coincide and which have the same orientation.

There are two kinds of coordinate transformations in Newtonian mechanics that do not alter the laws of motion. These are

1. A change in orientation of the system Σ' with respect to the system Σ about the common origin. This first transformation is characterized by equations linear in x', y', z' and x, y, z, between the coefficients of which there exist relations such that the condition

---

[15] Now it is easy to understand what we meant by the homogeneity of time and space in §6, or, in other words, why we assumed *a priori* that the transformation equations must be linear. For if a rate of a clock at rest with respect to S' is observed from S, this rate does not have to depend on the location of the clock in S' nor on the value of the time of S' in the vicinity of the clock. An analogous remark applies to the orientation and length of a bar linked with S' and observed from S. Only when the transformation equations are linear are these conditions satisfied.

$$x'^2 + y'^2 + z'^2 = x^2 + y^2 + z^2 \tag{1}$$

is identically satisfied.

2. Uniform motion (translation) of the system Σ' with respect to the system Σ. This second transformation is characterized by the equations

$$\begin{aligned} x' &= x + \alpha t \\ y' &= y + \beta t \\ z' &= z + \gamma t, \end{aligned} \tag{2}$$

where α, β, γ are constants.

For these two kinds of transformation, the condition

$$t' = t \tag{3}$$

must be satisfied. In other words, time is an invariant under these two transformations.

Combining the transformations (1) and (2), we obtain the most general transformation by means of which one can transform the equations of mechanics without altering them. This transformation is characterized by the equation (3) and by three equations that express x', y', z' as linear functions of x, y, z, t. The coefficients of these three equations are connected with each other by relations that, for $t = 0$, satisfy condition (1) identically.

Let us now consider the most general coordinate transformation compatible with the theory of relativity. From what we have seen, this transformation is characterized by the fact that x', y', z', t' must be linear functions of x, y, z, t, such that the condition

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = x^2 + y^2 + z^2 - c^2t^2 \tag{a}$$

will be satisfied identically. It should be noted that the transformations compatible with Newtonian mechanics can be obtained at once by setting $c = \infty$ in condition (a). Thus, if we take the same route as before, we arrive at the equations of ordinary kinematics if, instead of the principle of the constancy of the velocity of light, we assume the existence of signals whose propagation does not require any time.

The group characterized by equation (a) contains the transformations that correspond to a change in the orientation of the system. These are the transformations compatible with the condition

$$t = t'.$$

The simplest transformations compatible with condition (a) are those for which two of the four coordinates of an elementary event remain invariant. Let us consider, for example, the transformations under which x and t do not change. Instead of the general condition (a), we will have the special condition

$$\begin{aligned} t' &= t \\ x' &= x \\ y'^2 + z'^2 &= y^2 + z^2. \end{aligned} \tag{a$_1$}$$

To this condition corresponds a rotation of the system about the x-axis.

If, on the other hand, we consider transformations under which two of the spatial coordinates, for example y and z, remain invariant, we will have instead of the general condition (a) the special condition

$$\begin{aligned} y' &= y \\ z' &= z \\ x'^2 - c^2t'^2 &= x^2 - c^2t^2 \end{aligned} \tag{a$_2$}$$

These are the transformations we have encountered in the preceding section while investigating a system in uniform motion parallel to the x-axis of an identically oriented system at rest.

The formal analogy between the transformations (a$_1$) and (a$_2$) is immediately evident. The two systems of equations differ only by a change of sign in the third condition. But even this difference can be made to disappear if one chooses, with Minkowski, to take ict instead of t as a variable, where i is the imaginary unit.[16] In that case this imaginary temporal coordinate plays the same role in the transformation equations as the spatial coordinates. If we set

$$\begin{aligned} x &= x_1 \\ y &= x_2 \\ z &= x_3 \\ ict &= x_4 \end{aligned}$$

and consider x_1, x_2, x_3, x_4 as the coordinates of a point in a four-dimensional space such that to each elementary event there corresponds a point in this space, we reduce everything that happens in the physical world to something static in the four-dimensional space. In that case the condition (a) will be written as

$$x'^2_1 + y'^2_2 + x'^2_3 + x'^2_4 = x_1^2 + x_2^2 + x_3^2 + x_4^2.$$

---

[16] H. Minkowski, *Raum und Zeit*. Leipzig, 1909. [21]

This is the condition that corresponds to a rotation without relative translation of a four-dimensional coordinate system.

The principle of relativity demands that the laws of physics not be altered by the rotation of the four-dimensional coordinate system to which they are referred. The four coordinates x_1, x_2, x_3, x_4 must appear in the laws symmetrically. To express the different physical states, one can use four-dimensional vectors which behave in the calculations in a manner analogous to ordinary vectors in three-dimensional space.

§9. Some Applications of the Theory of Relativity

Let us apply the transformation equations (I) to the Maxwell-Lorentz equations representing the magnetic field. Let E_x, E_y, E_z be the vector components of the electric field , and M_x, M_y, M_z the components of the magnetic field, with respect to the system *S*. Calculation shows that the transformed equations will be of the same form as the original ones if one sets

$$(1)\quad \begin{cases} E'_x = E_x & M'_x = M_x \\ E'_y = \beta(E_y - v/c\ M_z) & M'_y = \beta(M_y + v/c\ E_z) \\ E'_z = \beta(E_z + v/c\ M_y) & M'_z = \beta(M_z - v/c\ E_y) \end{cases}$$

The vectors (E'_x, E'_y, E'_z) and (M'_x, M'_y, M'_z) play the same role in the equations referred to *S′* as the vectors (E_x, E_y, E_z) and (M_x, M_y, M_z) play in the equations referred to *S*. Hence the important result:

The existence of the electric field, as well as that of the magnetic field, depends on the state of motion of the coordinate system.

The transformed equations permit us to know an electromagnetic field with respect to any arbitrary system in nonaccelerated motion *S′* if the field is known relative to another system *S* of the same type.

These transformations would be impossible if the state of motion of the coordinate system played no role in the definition of the vectors. This we will recognize at once if we consider the definition of the electric field strength: the magnitude, direction, and orientation of the field strength at a given point are determined by the ponderomotive force exerted by the field on the unit quantity of electricity, which is assumed to be concentrated in the point considered and *at rest with respect to the system of axes.*

The transformation equations demonstrate that the difficulties we have encountered (§3) regarding the phenomena caused by the relative motions of a closed circuit and a magnetic pole have been completely averted in the new theory.

For let us consider an electric charge moving uniformly with respect to a magnetic pole. We may observe this phenomenon either from a system of axes *S* linked with the magnet, or from a system of axes *S′* linked with the electric charge. With respect to

S there exists only a magnetic field $(\mathrm{M}_x, \mathrm{M}_y, \mathrm{M}_z)$, but not any electric field. In contrast, with respect to S' there exists—as can be seen from the expression for E'_y and E'_z—an electric field that acts on the electric charge at rest relative to S'. Thus, the manner of considering the phenomena varies with the state of motion of the reference system: all depends on the point of view, but in this case these changes in the point of view play no essential role and do not correspond to anything that one could objectify, which was not the case when these changes were being attributed to changes of state of a medium filling all of space.

As we have already noted, we can find at once the laws applying to a body in rapid motion if we know the laws applying to a body at rest. In this way we can obtain, for example, the equations of motion for a material point of mass m carrying a charge e (an electron, for example) and subjected to the action of an electromagnetic field. We know, in fact, the equations of motion of a material point at the instant when its velocity is zero. According to Newton's equations and the definition of the electric field strength, we have

$$m\frac{d^2x}{dt^2} = e\mathrm{E}_x \tag{2}$$

and two other, similar equations with respect to the coordinates y and z. Applying the transformation equations (I) and the equations (1) given above, we find then for a point in any motion whatever

$$\frac{d}{dt}\left\{\frac{m\dfrac{dx}{dt}}{\sqrt{1-\dfrac{u^2}{c^2}}}\right\} = \mathrm{F}_x, \tag{3}$$

where
$$u = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}$$

and
$$\mathrm{F}_x = e\left[\mathrm{E}_x + \frac{1}{c}\left(\frac{dy}{dt}\mathrm{M}_z - \frac{dz}{dt}\mathrm{M}_y\right)\right]$$

and two, other, similar equations for the other two axes. These equations make it possible to follow the path of cathode rays and β-rays in an electromagnetic field. Their accuracy is almost beyond doubt, more so than the experiments of Bucherer and Hupka. [22]

If one wants to retain the relation between the force, mechanical work, and the theorem of the conservation of momentum, then the vectors F_x, F_y, F_z entering these equations have to be viewed as the vector components of the ponderomotive force acting on the material point in motion. Under these conditions, equations (3) have to be considered as *the most general equations of motion of a material point* compatible with the principles here assumed, whatever the nature of the force (F_x, F_y, F_z) might be.

If one expresses mathematically, first with respect to the system S, and then with respect to the system S', the fact that the principles of conservation of energy and momentum retain their validity in the emission and absorption of radiant energy by a body, one is led to the important conclusion that the *mass* of any arbitrary body *depends* [23] *on the quantity of energy* it contains. If m denotes the mass for a certain quantity of energy contained in a body, then the mass of the body will become $m + \frac{W}{c^2}$ if one increases the energy of the body by W (c always represents the velocity of light in vacuum). Thus, the principle of the conservation of mass assumed in Newtonian mechanics is valid only for a system whose energy remains constant. Mass and energy become equivalent quantities like heat and mechanical work, for example, and it takes only one additional step to consider mass as an enormous concentration of energy. Unfortunately, the change of mass $\frac{W}{c^2}$ is so slight that one cannot hope for its detection [24] by experiment for the time being.

Doc. 3

Response to Manuscript of *Planck 1910a*

[Before 18 January 1910]

On page 6 of your manuscript you say: "If, therefore, the oscillations of the emitting particles are subjected to certain fluctuations, these fluctuations will also manifest themselves in the intensity of the emitting light." Here you refer to the *very* point that first and foremost makes it seem to me that an explanation of the fluctuations of the radiation by the quantum character of the emission *alone* is out of the question. For, obviously, there must not exist any dependence of the statistical properties of radiation on the distance of the emitting wall. Let us compare the two cases:

A receives radiation one time from the surface *f* and another time from *F*. Suppose that *f* and *F* are made of the same material and are at the same temperature. If the radiation is produced in quanta of the same finite magnitude at *f* and *F*, but is distributed over the space in spherical waves, then the fluctuations will be smaller in the second case than in the first, because a greater number of quantum-like acts of emission, with a smaller percentage of the energy for each individual act, will work together. One candle produces at a distance of 1 m a strongly flickering light; 100 candles of the same kind produce at a distance of 10 m light of the same intensity but with less flicker. Further, I did not introduce the constant *h* in the dimensional argument that was to yield the fluctuation of the radiation pressure, because this constant just does not belong in the pure undulatory theory of radiation. As far as I see it, even at an arbitrary distance from the emitting wall, the latter theory does not allow for any fluctuations other than the fluctuations due to interference. One of these days I will recommend a more exact investigation of this problem to a doctoral candidate.

Further, you consider it a weakness of the quantum conception that one cannot see how to conceive of static and stationary fields. In this matter I am definitively of the opinion that the development of relativistic electrodynamics will lead to a localization of energy different from the one we are now accustomed to assume without any good reason. Without an ether, energy continuously distributed in space seems to me an absurdity. It can also easily be shown that the localization of energy as the old action-at-a-distance theory had it is compatible with Maxwell's theory; one of these days I will publish this in connection with some other material. Though Faraday's intuitive representation rendered important services in the development of electrodynamics, one cannot conclude from this, in my opinion, that it must be retained in all its details.

Doc. 4

Lecture Notes for Course on the Kinetic Theory of Heat at the University of Zurich, Summer Semester 1910

[19 April–5 August 1910][1]

$u_1 \quad u_2 \quad u_3 \cdots\cdot u'_1 \quad u'_2 \quad u'_3 \cdots\cdot$ [2] [fly-leaf]

$$m\frac{du}{dt} = k \qquad K = \sum - k = p$$

$$x|m\ddot{x}_1 = X_1 \quad , \quad ,$$

$$\sum m x\ddot{x} = \sum X x$$

$$x\ddot{x} = \frac{d}{dt}(x\dot{x}) - \dot{x}^2$$

$$\frac{d}{dt}\underset{\substack{\|\\ 0}}{(\sum m x\dot{x})} - 2L = \sum X x + \cdot + \cdot$$

$$p\,ds\cos\alpha\cdot r = p3\frac{dV}{}$$

$$pV = \frac{2}{3}L.$$

Anders

$$\int_0^1 p\,dt = \int_0^1 \sum - k\,dt$$

$$= -\sum\int_0^1 k\,dt = 2\sum m u_a$$

$$= 2\sum n_1 u_1 m u_1$$

$$\int k\,dt = m(u_e - u_a) = -2mu_a$$

$$= 2\sum n_1 m u_1^2 = 2\left|\frac{L_1}{3}\right|_{v=1} = \frac{2}{3}\frac{L}{V}$$

$$\underline{pV = \frac{2}{3}L}$$

applied to one gram-molecule

$$pV = \frac{2}{3}L$$

$$pV = RT \quad \text{gas eq.}$$

$$\bar{L}_{\text{mol}} = \frac{3}{2}\frac{RT}{N}$$

Test $\frac{dL}{dT} = c_v = \frac{3}{2}R$ $|c_p - <c_v> = C_v + R = \frac{5}{2}R$ $\frac{c_p}{c_v} = \frac{5}{3}$

The law $pV = \frac{2}{3}L$ also holds if several different kinds of molecules

If r_1 gram.-molecules of the first gas, r_2 of the second gas

$pV = \frac{2}{3}\left\{\frac{3}{2}(RTr_1 + RTr_2)\right\} = RT(r_1 + r_2)$ Dalton's law.[3]

Virial theorem

$$x|m\frac{d^2x}{dt^2} = X \qquad \sum(Xx + \cdot + \cdot) = \sum m\left(x\frac{d^2x}{dt^2} + \cdot + \cdot\right)$$

$$= \frac{d}{dt}\sum m\left(x\frac{dx}{dt} + \cdot + \cdot\right) - \sum mv^2$$

$$L = -\frac{1}{2}\sum(Xx + \cdot + \cdot)$$

$$ds\, r\cos nr = \left\langle\frac{2}{3}\right\rangle 3\, d\varphi$$

$$L = \frac{3}{2}pV$$

Pressure of an ideal gas. Analysis as in Boltzmann. Also using the virial theorem, [p. 1]

$$L = \frac{3}{2}pV \tag{1}$$

On the other hand, according to the equation of state

$pV = RT$, if one gram-molecule is present[4]

From this, $RT = \frac{2}{3}L$

or $\frac{\overline{mc^2}}{2} = \frac{L}{N} = \frac{3}{2}\frac{R}{N}T \qquad \sqrt{\overline{c^2}} = \sqrt{\frac{3RT}{M}}$[5]

Thus, according to the equation of state, the mean kinetic energy of a monatomic gas depends on the temperature but not on the constitution (mass) of the molecule and not on the density of the molecules. We shall see later that it is also possible to prove this law on the basis of purely molecular-theoretical considerations without resorting to the equation of state, that is to say, the equation of state can be deduced entirely by means of the molecular theory.[6]

<From the constant of the equation of state> We can use (1) to calculate the mean velocities of the gas molecules,[7] & this calculation obviously applies in the case of polyatomic molecules as well. Here L denotes then the kin. energy of translational motion.

$L = \frac{3}{2}pV$ <applied to a unit volume> $\frac{L}{V} = \frac{3}{2}p = n\frac{\overline{mc^2}}{2} = \frac{\rho}{2}\overline{c^2}$

[p. 2] $\underline{\overline{c^2} = \frac{3p}{\rho}}$ or also $L = \frac{3}{2}pV = n\frac{\overline{mc^2}}{2} =$ $\qquad \frac{3}{2}\frac{8.3\cdot 10^7\cdot 300}{2}$

$\frac{3}{2}RT = M\frac{\overline{c^2}}{2} \qquad \underline{\overline{c^2} = \frac{3\,RT}{2\,M}}$[8] $\qquad 18\cdot 10^9$

$1.8\cdot 10^{10}$

$1.3\cdot 10^5$ cm.

For $T = 273$, one obtains in this way about 1840 m/sec for hydrogen, etc.

<Our analysis leads further to Avogadro's rule.> The rule that at the same temperature and pressure a given space always contains the same number of molecules can be presented as a consequence of the theory only after it has been proved by purely molecular-theoretical considerations that the mean kinetic energy of the translational motion of a molecule depends *only* on the temperature. [9]

The Specific Heat of a Monatomic Molecule

$$L = \frac{3}{2}pV \quad \text{for a gram-molecule}$$

$$L = \frac{3}{2}RT$$

$$\frac{dL}{dT} = c_v = \frac{3}{2}R \qquad c_v^x = \frac{3}{2}\frac{R}{4.2\cdot 10^7} \approx 3$$ [10]

$$c_p = c_v + R = \frac{5}{2}R$$

$$\frac{c_p}{c_v} = \frac{\frac{5}{2}R}{\frac{3}{2}R} = \frac{5}{3} = 1.66\cdots$$

$$\sqrt{\frac{dp}{d\rho}}$$

$$\sqrt{\frac{5}{3}\frac{p}{\rho}} = \sqrt{\frac{5}{3}\frac{RT}{M}}$$

$$pV = RT$$

$$\frac{1}{\rho}M = V$$

Proved to be exact for monatomic gases.[11] For polyatomic gases

$$E = L + E_i = \frac{3}{2}RT + E_i$$

$$\frac{dE}{dT} = c_v = \frac{3}{2}R + \frac{dE_i}{dT} = \frac{3}{2}R + c_i$$

$$c_p = \qquad\qquad\qquad \frac{5}{2}R + c_i$$

$$\frac{c_p}{c_v} = \frac{5 + c_i\frac{2}{R}}{3 + c_i\frac{2}{R}}$$

Since c_i obviously increases with increasing molecular size, $\frac{c_p}{c_v}$ is here $<\frac{5}{3}$ and approaches unity with increasing c_i.

[p. 3]

Approximate Theory of Heat Conduction, Viscosity, and Diffusion

To simplify the calculation, we shall assume that at the same location all molec. have the same velocity $\sqrt{\overline{c^2}}$. We designate by $\bar{\lambda}$ the mean path length of a molecule on which the quantities to be calculated depend.

Transport of Any Molecular Quantity through the Gas[12]

n number of molecules in unit volume.
$n d\kappa$ number " " " " " of spec. κ

During time dt, $\nu = \dfrac{n \underline{d\kappa} c \cos\vartheta dt}{4\pi}$

fly from top to bottom

$$\frac{d\kappa}{4\pi} = \frac{\sin\vartheta d\vartheta d\omega}{4\pi}$$

Each molecule carries along a certain quantity of something, with this amount depending only on where the molecule's last collision took place. traveled freely the distance λ up to the layer comes from the layer $z = z_0 - \lambda \cos\vartheta$ There the value $G(z_0 - \lambda\cos\vartheta) = G_0 - \frac{\partial G_0}{\partial z}\lambda\cos\vartheta$ obtains permanently. Thus, the ν molecules carry along during dt the quantity

$$\nu\left(G_0 - \frac{\partial G_0}{\partial z}\lambda\cos\vartheta\right).$$

In the same way, we find that the amount delivered in the opposite direction, characterized by the opposite cone, is

$$-\nu'\left(G_0 + \frac{\partial G_0}{\partial z}\lambda\cos\vartheta\right).$$

Here $\nu = \nu'$, so that we have in all (for $\kappa + \kappa'$)

$$-2\nu\frac{\partial G}{\partial z}\bar{\lambda}\cos\vartheta = -2n\frac{d\kappa c}{4\pi}dt\frac{\partial G}{\partial z}\lambda\cos\vartheta^2$$

$$dF = -\frac{2}{4\pi}nc\,dt\,\lambda\frac{\partial G}{\partial z}\cos\vartheta^2\sin\vartheta\,d\vartheta\,d\omega$$

This we have to integrate:[13] ω from 0–2π, and ϑ from 0 to π/2.

$$\frac{F}{dt} = -nc\lambda\frac{\partial G}{\partial z}\cdot\left|\frac{\sin^3\vartheta}{3}\right|_0^{\frac{\pi}{2}} = -\frac{1}{3}nc\lambda\frac{\partial G}{\partial z}$$

$$F = -\frac{1}{3}nc\lambda\frac{\partial G}{\partial z}.$$

Free Path

[p. 4]

Assume again that all molecules of a given kind have the same velocity.

Collisions of a moving molecule (radius R_1) with stationary molecules (radius R_2)

Collision as soon as the distance between centers $R_1 + R_2 = \sigma$

Volume covered in unit time

$$\pi c\sigma^2$$

Number of col. of a molecule of kind 1 with molecules of the second kind

$$Z_{12} = n_2\pi c_1\sigma^2$$

Number of collisions of mol. of kind 1 with mol. of kind 1

$$Z_{11} = n_1\pi c_1 s_1^2,$$

where s is the diameter of the molecule of kind 1.

We now calculate more precisely by taking into account the fact that the molecules that are hit are set in motion. First, collisions with molecules of the kind previously emphasized $(d\kappa)$[14]

$$c_r^2 = c_1^2 + c_2^2 - 2c_1c_2\cos\vartheta$$

$$dZ_{12} = n_2\pi\sigma^2 c_r\, \underset{\substack{| \\ \frac{\sin\vartheta\, d\omega\, d\vartheta}{4\pi}}}{d\kappa}$$

$$Z_{12} = \frac{n_2\pi\sigma^2}{4\pi}\int_{\vartheta=0}^{\pi}\int_{\omega=0-2\pi} \times \sqrt{(c_1^2 + c_2^2 - 2c_1c_2\cos\vartheta)}\sin\vartheta\, d\omega\, d\vartheta$$

$$= \frac{n_2\pi\sigma^2}{2}\cdot\frac{1}{2c_1c_2}\cdot\frac{2}{3}\underbrace{\underbrace{\left|(c_1^2 + c_2^2 - 2c_1c_2\cos\vartheta)^{3/2}\right|_0^{\pi}}_{(c_1+c_2)^3-(c_1-c_2)^3}}_{6c_1^2c_2+2c_2^3}$$

$$Z_{12} = n_2\pi\sigma^2\left(c_1 + \frac{1}{3}\frac{c_2^2}{c_1}\right)$$

$$Z_{11} = n_1\pi s_1^2\left\{\frac{4}{3}c_1\right\}$$

[p. 5] We obtain the mean free path by dividing the velocity c_1 of the molecule by the number of collisions per second.

$$\lambda_1 = \frac{c_1}{Z_{12} + Z_{11}} = \frac{1}{\frac{4}{3}n_1\pi s_1^2 + n_2\pi\sigma^2 \underset{\left(1+\frac{1}{3}\frac{m_1}{m_2}\right)}{\left(1 + \frac{1}{3}\frac{c_2^2}{c_1^2}\right)}} \qquad \left| \quad \frac{c_2^2}{c_1^2} = \frac{m_1}{m_2}\right.$$

If only one kind of molecule is present, we get

$$\lambda = \frac{1}{\frac{4}{3}n\pi s^2}$$

Thermal Conductivity

Heat conduction from the fact that, on average, each molecule the amount of heat $\frac{C_v T}{N}$ with it. Hence the heat conduction will be

$$F = -\frac{1}{3}nc\lambda\frac{C_v}{N}\frac{\partial T}{\partial z} = -k\frac{\partial T}{\partial z} \qquad k = \frac{1}{3}nc\lambda\frac{C_v}{N}$$

Since $\frac{n}{N}$ is concentration η, we have

$$\boxed{\begin{aligned} k &= \frac{1}{3}\eta c\lambda C_v \\ &= \frac{1}{3}\rho c\lambda c_v \end{aligned}} \qquad\qquad \begin{aligned} \rho &= \eta M \\ v_v &= \frac{C_v}{M} \end{aligned}$$

For monatomic gas $C_v = \frac{3}{2}R$

Further, if one inserts the value for λ into the first equ. for k, one obtains

$$\boxed{k = \frac{1}{4}\frac{cC_v}{\pi s^2 N}}$$

It turns out that the thermal conductivity is independent of the density of the gas. (Limit of validity of this law.[15]

$$\begin{aligned} &\frac{1}{4}\frac{\eta cC_v}{\pi s^2 n} \\ &= \frac{1}{4}\frac{cC_v}{\pi s^2 N} \end{aligned}$$

[p. 6]

Viscosity

$$K = F_{\text{(momentum}}$$
$$G = mu$$

$$F = K - \frac{1}{3} nc\lambda m \frac{\partial u}{\partial <x> z} = -\frac{1}{3} \rho c \lambda \frac{\partial u}{\partial x} = -\mathfrak{R} \frac{\partial u}{\partial z}$$

$$\boxed{\mathfrak{R} = \frac{1}{3} \rho \bar{c} \lambda}$$

We compare this with our expression

$$\boxed{k = \frac{1}{3} \rho \bar{c} \lambda c_v},$$

so that we get $\frac{k}{\mathfrak{R}} = c_v$

Accurate to an order of magnitude[16]

		$\mathfrak{R}$	κ observed	κ calculated	quotient
H_2	2	$1.850 \cdot 10^{-4}$	$0.35 \cdot 10^{-3}$	$0.21 \cdot 10^{-3}$	1.6
O_2	32	$1.880 \cdot 10^{-4}$	$0.056 \cdot 10^{-3}$	$0.029 \cdot 10^{-3}$	1.9.

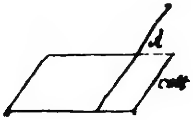
We can set $\rho = nm$

$$k = \frac{1}{4}\frac{m\bar{c}c_v}{\pi s^2} \qquad R = \frac{1}{4}\frac{mc}{\pi s^2}$$

From this one sees that k should be independent of the density, which has been amply proved by experiment.[17] Further, we have $\bar{c} \propto \sqrt{T}$, which should thus hold for k as well. But it does not hold, because s decreases for higher temperatures (the elastic spheres model not completely applicable.[18]

Diffusion of Molecules of Kind 1 in Molecules of Kind 2[19]

[p. 7]



$$c\,dt\cos\varphi\cdot\frac{d\kappa}{4\pi}n_1(z_0-\lambda\cos\varphi)$$

$$c\,dt\cos\varphi\frac{d\kappa}{4\pi}\left(n_1-\frac{\partial n_1}{\partial z}\lambda\cos\varphi\right)$$

$$-c\,dt\cos\varphi\frac{d\kappa}{4\pi}\left(n_1+\frac{\partial n_1}{\partial z}\lambda\cos\varphi\right)$$

$$dZ = -2c\,dt\cos\varphi\frac{\sin\varphi\, d\omega\, d\varphi}{4\pi}\lambda\cos\varphi\frac{\partial n_1}{\partial z}$$

$$Z = -c\lambda\frac{\partial n_1}{\partial z}\underbrace{\int_0^{\pi/2}\cos^2\varphi\sin\varphi\,d\varphi}_{\left|-\frac{\cos^3\varphi}{3}\right|_0^{\pi/2}=\frac{1}{3}} = -\frac{1}{3}c\lambda\frac{\partial n_1}{\partial z}$$

If η denotes the concentration of the gas in $\frac{\text{gr.mol.}}{cm^3}$, one obtains by dividing by N

$$\text{Diffusion gram-molecule} = -\frac{1}{3}\bar{c}\lambda\frac{\partial\eta}{\partial z} = -D\frac{\partial\eta}{\partial z}$$

$$D = \frac{1}{3}\bar{c}\lambda \qquad \boxed{\frac{1}{3}\cdot 5\cdot 10^4\cdot 10^{-5} = 0.16^{[20]}}$$

We found earlier

$$\mathfrak{R} = \frac{1}{3}\bar{\rho} c \bar{\lambda}$$

Comparison yields $\frac{\mathfrak{R}}{D} = \rho$.

This relation also proves to be correct to an order of magnitude[21]

	$\mathfrak{R}$	Mean	D_{observed}	$D_{\text{calculated}}$
O_2	$1.88 \cdot 10^{-4}$	$1.77 \cdot 10^{-4}$	0.17	0.13
N_2	$1.66 \cdot 10^{-4}$			

[p. 8]

Path Length and Actual Size of Molecules[22]

$$\mathfrak{R} = \frac{1}{3}\rho \bar{c} \lambda$$

ρ = density.

Approximately $c = \sqrt{\frac{3RT}{M}}$

λ can be calculated. e.g., oxygen atmospheric pressure 10^{-4} mm. Since now $\lambda = \frac{1}{\frac{4}{3}\pi s^2 n}$,

λ also yields $s^2 n$, and thus $s^2 N = \frac{s^2 n}{\frac{n}{N}} = \frac{s^2 n}{\eta}$

If molecules are packed perfectly tightly, then molecule occupies roughly the space s^3. The value for all molecules of a gram-molecule therefore $s^3 N$. Hence we have

$s^2 N = A$ (from path length)
$s^3 N$ = V (approximate molecular volume in the solid or liquid state.)

One obtains $$s = \frac{V}{A}$$

$$N = \frac{A^3}{V^2}$$

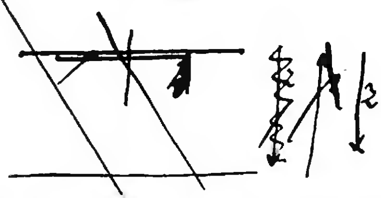
The determination of N very unreliable because of the powers of A (path length) One obtains values between 10^{23} & 10^{24}; $s \approx 10^{-7}$ mm.[23]

Direct Influence of Path Length on the Phenomena

We had considered the flow of the molecular function G in a stationary state. Must be equal to this flow when infinitely close to the wall.

For the molecules flying off of the wall the value of the quantity G is on the average G_0. For those going upwards from below the value is

$$G_{(z_0+\lambda\cos\varphi)} = G_0' + \frac{\partial G}{\partial z}\lambda\cos\varphi$$



When the molecule returns, it has G_0. Hence the molecule transports [p. 9]

$$G_0 - \left(G_0' <+> - \frac{\partial G_0'}{\partial z}\lambda\cos<\varphi>\vartheta\right)$$

There are $n\frac{d\kappa}{4\pi}$ such ($d\kappa$) molecules in a unit volume.

hence $n\frac{d\kappa}{4\pi}c\,dt\cos\varphi$ act in the time element dt

Hence total flow through surface

$$F = \int \left[(G_0 - G_0') <+> - \frac{\partial G_0'}{\partial z} \lambda \cos\vartheta \right] \cdot nc \frac{\sin\vartheta d\vartheta d\omega}{4\pi} \cos <\varphi> \vartheta$$ [24]

$$= \frac{1}{4} nc \langle\lambda\rangle (G_0 - G_0') <+> - \frac{1}{6} \frac{\partial G_0'}{\partial z} nc\lambda$$

$$= -\frac{1}{3} nc\lambda \left(\frac{1}{2} \frac{\partial G'}{\partial z} <+> - \frac{3}{4\lambda} (G_0 - G_0') \right)$$

But F must be of the same magnitude as for the middle cross section, where its value is $-\frac{1}{3} nc\lambda \frac{\partial G}{\partial z}$. Hence,

$$\frac{\partial G}{\partial z} = \frac{1}{2} \frac{\partial G}{\partial z} <+> - \frac{3}{4\lambda} (G_0 - G_0')$$

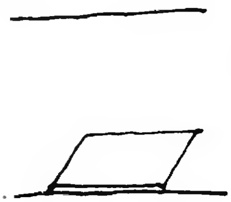
$$\frac{1}{2} \frac{\partial G}{\partial z} = \frac{3}{4\lambda} (G_0 - G_0')$$

$$G_0 - G_0' = \frac{2}{3} \lambda \frac{\partial G}{\partial z}$$ [25]

As we see, there is a jump corresponding to $\frac{2}{3}\lambda$. Has been proved by Kundt and Warburg for viscosity, and by Smoluchowsky for thermal conduction.[26]

[p. 10]

Phenomena at Walls Owing to Finiteness of Path Length



We assume that G changes linearly down to the lower wall. There, but in the gas, G_0' For the molecules flying off of the wall G_0 In the terminal cross-section as well we must have

$$F. = -\frac{1}{3} nc\lambda \frac{\partial G_0}{\partial z}$$

On the other hand, G-flow calculated[27]

$$n\frac{d\kappa}{4\pi}\cdot\left(G_0' - \lambda\cos\varphi\frac{\partial G_0}{\partial z}\right)c\cos\varphi$$

$$d\kappa = \sin\vartheta\, d\vartheta\, d\omega$$

Integrated $F_+ = ncG_0'\int_0^{\pi/2}\frac{\sin\vartheta\, d\vartheta\, d\omega}{4\pi}\cos\varphi - nc\lambda\frac{\partial G_0}{\partial z}\int\frac{\sin\vartheta\cos^2\vartheta\, d\vartheta\, d\omega}{4\pi}$

$$\Gamma_+ = \frac{1}{4}ncG_0' - \frac{1}{6}nc\lambda\frac{\partial G_0}{\partial z}$$

On the other hand, assuming that all molecules coming from the wall have G_0.

$$F_- = \frac{1}{4}ncG_0$$

$$F = F_+ - \Gamma_- = \frac{1}{4}nc(G_0' - G_0) - \frac{1}{6}nc\lambda\frac{\partial G_0}{\partial z} \quad\Bigg|\quad = -\frac{1}{3}nc\lambda\frac{\partial G}{\partial z}$$

Thus $\frac{1}{4}nc(G_0' - G_0) = -\frac{1}{6}nc\lambda\frac{\partial G}{\partial z}$

$$G_0' - G_0 = -\frac{2}{3}\lambda\frac{\partial G}{\partial z}$$

Is a positive quantity. Right side shows how much G changes over the distance $\frac{2}{3}\lambda$ in the gas. The space occupied by the gas would have to be increased downwards by $\frac{2}{3}\lambda$ in order for G to assume the value G_0.

For viscosity, one must set [p. 11]

$$G = mu$$

$$m(u_0' - 0) = -\frac{2}{3}\lambda\frac{\partial mu}{\partial z}$$

or $u_0' = \frac{2}{3}\lambda\left(-\frac{\partial u}{\partial z}\right)$

For heat conduction $G = \frac{C_v T}{N}$

$$\frac{c_v T_0'}{N} - \frac{c_v T_0}{N} = \frac{2}{3}\lambda\left(-\frac{\partial}{\partial z}\left(\frac{c_v T}{N}\right)\right)$$

$$T_0' - T_0 = \frac{2}{3}\lambda\left(-\frac{\partial T}{\partial z}\right)$$

Both consequences have been confirmed to an order-of-magnitude.[28]

Gases in Narrow Channels (Knudsen)[29]

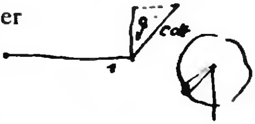
Flow. If tube wide compared with cross section,[30]

$$\begin{array}{l}\text{quantity of flow-through}\\ \text{per unit time}\end{array} = \frac{\pi}{4}\frac{\Delta}{\mathfrak{R}}R^4\rho \ .$$

We now investigate the flow under the assumption that the diameter of the tube is small compared with the path length.—

$$n\frac{d\kappa}{4\pi}c\,\cos\vartheta\,dt = \text{number of incident molecules in time } dt$$

Thus, in unit time from all angles together



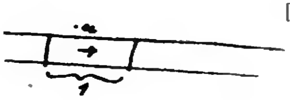
$$nc\int\frac{\sin\vartheta\,d\omega\,d\vartheta}{4\pi}\cos\vartheta = \frac{1}{4}nc$$

The ones that impinge during dt from the cone $d\kappa$ will each bring along the momentum $2mc\cos\vartheta$ each. hence molec $d\kappa$ in unit time

$$n\frac{d\kappa}{4\pi}\,2mc^2\cos^2\vartheta \ .$$

All together $$2nmc^2\int_0^{\frac{\pi}{2}}\frac{\cos^2\vartheta\sin\vartheta\,d\omega\,d\vartheta}{4\pi} = \frac{1}{3}nmc^2 = p.$$

[p. 12]



Now consider narrow tubes. Velocity u in a cross section. Regarded as constant over the cross section. On the average, each impinging molecule brings to the wall the momentum mu[31]; $\frac{1}{4}\,nc$ molecules impinge per unit time. Bring along the momentum $\frac{1}{4}\,nc \cdot mu$ (per unit surface area. hence $\frac{1}{4}\,nmcuP$ per unit length.[32]

This momentum is supplied by pressure forces at the beginning & the end of the layer. If the cross section is q, then the above quantity is equal to

$$qp_1 - qp_2 = q\Delta$$

Hence $q\Delta = \frac{1}{4}nmcuP$

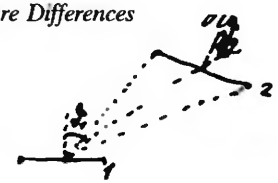
Now we have $\left\langle p = \frac{1}{3}nmc^2\right\rangle\ nm = \rho.c = \sqrt{\frac{3RT}{M}}$

so that $\left\langle u = \left\langle\frac{4}{\sqrt{3}\,\rho\sqrt{3}}\right|\right\rangle \left\langle\frac{q}{P}\Delta\right\rangle \frac{4}{\sqrt{3}}\cdot\frac{1}{\rho}\sqrt{\frac{M}{RT}}\ \Delta\frac{q}{P}\right\rangle.$

Hence quantity per unit time $= \frac{4}{\sqrt{3}}\sqrt{\frac{M}{RT}}\ \frac{q^2}{P}\Delta$

Found & confirmed by Knudsen.[33]

Pressure Differences Produced by Temperature Differences in Capillary Spaces[34]



Consider two surface elements.
First, case of the thermal equilibrium.
<Of the molecules impinging at 1, the fraction>

$n\frac{d\kappa}{4\pi}c\cos\vartheta$ impinge on unit surface area in unit time in the angle $d\kappa$ hence from f_2 on f_1

$$f_1 \cdot n\frac{\kappa_{12}}{4\pi}c\cos\vartheta_1, \text{ or, since } \kappa_{12} = \frac{f_2\cos\vartheta_2}{r^2}:$$

$$\frac{f_1f^2}{4\pi r^2}nc\cos\vartheta_1\cos\vartheta_2$$

Symmetrical with regard to indices 1 & 2, as it must be.

[p. 13] If we assume that f_1 & f_2 are part of the wall of the tube, and that n & c are functions of the abscissa, then the number of molecules sent from f_2 to f_1 will be determined by the values of n & c at f_2. Hence number of molecules sent per sec from f_2 to f_1:

$$\underset{\kappa}{\left(\frac{f_1f_2}{4\pi r^2}\cos\vartheta_1\cos\vartheta_2\right)} \cdot n_2c_2$$

Number of molecules sent per sec from f_1 to f_2:

$$\kappa n_1c_1$$

Let us now assume that no flow takes place. Then the two quantities will be equal on the average. Hence,

$$n_1c_1 = n_2c_2$$

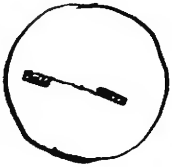
$$\frac{1}{3}n_1m\frac{c_1^2}{c_1} = \frac{1}{3}n_2m\frac{c_2^2}{c_2}$$

$$\text{or}\quad \frac{p_1}{c_1} = \frac{p_2}{c_2}$$

$$\text{or}\quad \frac{p_1}{p_2} = \frac{c_1}{c_2} = \frac{\sqrt{T_1}}{\sqrt{T_2}}$$

In such spaces, the pressures behave like the temperatures.

Brief Discussion of the "Light Mill"[35]



Molecular Processes and Probability[36] [p. 14]

~~Let a point move on a plane according to a law of the form~~

$$\frac{dx}{dt} = \varphi_1(x,y)$$

$$\frac{dy}{dt} = \varphi_2(x,y)$$

~~These equations determine the motion completely if only the initial position of the point is given. If we follow the point, then a variety of cases can occur.[37]~~

1) ~~The point describes a totally closed curve. In that case we have an integral of these equations such that $\psi(xy) = \text{const}$, such that to each x there corresponds a finite number of y values. We shall exclude this case for the time being. One can reduce it to the opposite case by decreasing the number of variables by 1.~~
2) ~~The point describes a curve that does not close.~~

~~Here again we can distinguish two cases.~~

a) ~~The moving point never returns to the immediate vicinity of a point on the plane it has already passed. I.e., x_0y_0 is such a point. We can then mark off a circle with a very small radius R around x_0y_0 such that once the point has left the circle, it never returns into it again.~~
b) ~~The moving point returns to the immediate vicinity of each point of the plane that it has passed before. No matter how small R around x_0y_0 is~~

assumed to be. The moving point enters again into the interior of the small circle.

Example for case 1):

$$\left.\begin{aligned} \frac{dx}{dt} &= -ay \\ \frac{dy}{dt} &= ax \end{aligned}\right| \begin{matrix} x \\ y \end{matrix} \qquad x\frac{dx}{dt} + y\frac{dy}{dt} = \text{const.} \quad r^2 = \text{const.}$$

[p. 15] Point travels along a closed circle.

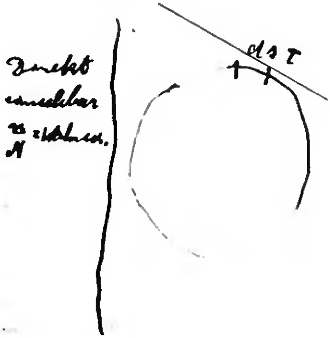
2) a) Motion along a spiral of Archimedes at constant velocity.[38]

$$\frac{dr}{dt} = -\alpha r$$

$$\frac{d\varphi}{dt} = b$$

can also be written in rectangular coordinates. But this is superfluous.

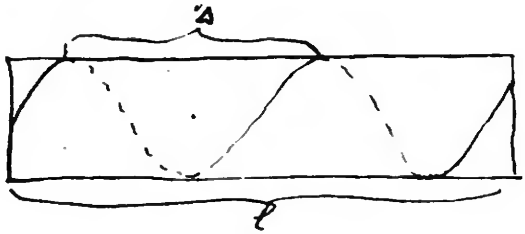
[p. 16] Second proof[39]



To traverse ds, a point needs the time $\frac{ds}{v} = \tau$. During this time, n points <run> are replaced by n other points. But during the time a point traverses the whole trajectory $\frac{T}{\tau}$ times, this process is also repeated $\frac{N}{n}$ times.

Hence, $\boxed{\frac{\tau}{T} = \frac{n}{N}}$ This relation is also valid if ds consists of several parts.

The concept of probability as defined can also be extended to the case where the point moves according to such a law that it does not describe a closed curve.[40] We have in mind a cylinder & a point moving with constant speed on it along a spiral line.



We think of the cylinder as being bent together into a circular ring (torus). If the ratio of *l* to *s* is rational, then the line will close after a finite number of turns, so that the case reduces to the one considered above. But if the ratio is irrational, then a new case. Line does not close. But statistical consideration possible in this case too. We choose surface σ on the torus and observe point for a long time *T*. A fraction τ of this time is distinguished by the fact that point lies within σ. We consider

$$\left(\lim \frac{\tau}{T}\right)_{T=\infty}$$

A limit will exist for this fraction. We have to think of this limit as the probability W_σ [p. 17] of σ, or also as the probability of encountering this point in σ at an arbitrarily chosen instant. Here too the probability can be visualized by means of a stationary point flow as $\frac{n\sigma}{N}$, which can be proved as above. That is to say that, to any approximation desired, one can replace a very large number of non-closed turns with a closed one.

We now generalize the examples considered in that we leave unspecified the law according to which the point moves. We set

$$\frac{dx}{dt} = \varphi_1(xyz)$$

$$\frac{dy}{dt} = \varphi_2(xyz)$$

$$\frac{dz}{dt} = \varphi_3(\qquad)$$

These equations determine the position of the point at time $t + dt$ if it is given at time t. Thus, path completely determined. Obviously, there can be statistical laws for the motion only in the case where the point later on returns arbitrarily close to a point it had already occupied once before. But in correspondence with the previous examples, we must also demand that, based on the assumed law, it be possible to construct a stationary flow from infinitely many *(N)* points.[41] In this flow there will be $nd\tau$ points in the volume element dt, and n <will> should vary continuously with the location.

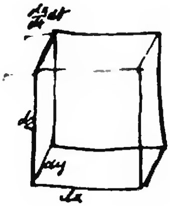
Let the flow be a continuous one. From this we obtain an important formal relation.

$$[nd\tau]_t = [nd\tau]_{t+dt}.$$

Since a change in the number of points present in $d\tau$ can occur only by points entering or leaving the space, we can also say: The sum of the points entering a volume $d\tau$ during time dt is zero.

[p. 18] $n\left(dydz\frac{dx}{dt}dt\right)$ = number of points entering through $dy\ dz$ during dt

$= dy\ dz\ dt \cdot (n\varphi_1)$



In the same way, one finds that the number of points leaving through the opposite surface is

$$dy\ dz\ dt(n'\varphi_1')$$

Since $n\varphi_1 - n'\varphi_1' = -\frac{\partial(n\varphi_1)}{\partial x}dx$, one finds that the excess of entering over exiting points is

$$-\frac{\partial(n\varphi_1)}{\partial x}dxdydzdt$$

The other two pairs of lateral surfaces yield . . so that the total excess of the entering over the exiting points is

$$-\left(\frac{\partial(n\varphi_1)}{\partial x} + \frac{\partial(n\varphi_2)}{\partial y} + \frac{\partial(n\varphi_3)}{\partial z}\right)dxdydzdt.$$

This must vanish for every volume element. Hence we must have in the entire space

$$\frac{\partial n\varphi_1}{\partial x} + \frac{\partial n\varphi_2}{\partial y} + \frac{\partial n\varphi_3}{\partial z} = 0$$

This is the mathematical formulation of our condition. We want to write it in another form. We have

$$n\left(\frac{\partial\varphi_1}{\partial x} + \frac{\partial\varphi_2}{\partial y} + \frac{\partial\varphi_3}{\partial z}\right) + \left(\frac{\partial n}{\partial x}\frac{dx}{dt} + \cdot + \cdot\right) = 0$$

The second term can be written as

$$\frac{\frac{\partial n}{\partial x}dx + \frac{\partial n}{\partial y}dy + \frac{\partial n}{\partial z}dz}{dt}$$

where *dx,dy,dz* are the paths described by a point in the time element *dt*. The numerator is thus the increase in point density in case one passes from a point $x\,y\,z$ to a point that the point present in $x\,y\,z$ reaches in time *dt*. We can always use the symbol "*d*" for these increases & write more concisely

$$n\left(\frac{\partial\varphi_1}{\partial x} + \cdot + \cdot\right) = -\frac{dn}{dt}$$ [p. 19]

or

$$\frac{\partial\varphi_1}{\partial x} + \cdot + \cdot = -\frac{d(\lg n)}{dt.} \qquad (1)$$

Let us now consider a special case to which the general case can be reduced, as will be shown later on. That is, we assume that the functions φ that determine the law of motion of our points satisfy the condition

$$\frac{\partial\varphi_1}{\partial x} + \cdot + \cdot = 0:$$

In this special case our law becomes

$$\frac{d\lg n}{dt} = 0. \cdots \qquad (1')$$

This says: No matter how long we follow a point on its trajectory; the point density n is everywhere the same.[42]

<Two cases are possible> In what follows, we will assume

<1)> In its motion the point covers a three-dimensional space. We will confine our statistical analysis to this space. In all of this space, then, n = const. The point density is constant. Probability of a spatial region $\frac{n d\sigma}{N} = \lim \frac{\tau}{T}$.

2) The law of motion is such that, in its motion, the point remains constantly on a surface. In that case, all that follows from equation (1′) is that n has the same value for all points of the surface. Let this be $\psi(x\, y\, z) = E$, where the value of E can be chosen arbitrarily. E is then determined by the initial conditions. In this case we can conclude that n depends on E alone. $n = \psi(E)$. In this case as well, the statistical properties are thus established to the extent that the conditions of the problem permit it.

In fact, the problem in this case is a two-dimensional one, since the position of the point on the surface E = const. could be completely determined by two coordinates. (Cf. example of the motion on the torus.)[43] Thus, <the statistical laws for an individual system cannot here be deduced immediately> from the statistical laws for a stationary spatial flow.

[p. 20] If we do not assume that $\sum \frac{\partial \varphi_1}{\partial x} = 0$, then equation 1) needs further interpretation. The right-hand side is a time derivative of a spatial function. The left-hand side of 1) is then integrable, & one obtains an equation of the form:

$$dn = N\psi(xyz)dx\, dy\, dz.$$

$$\text{or also } dW = \frac{dn}{N} = \psi(xyz)\, dx\, dy\, dz.$$

Introducing new variables $\xi\eta\zeta$ in the place of xyz, we can obtain

$$dW = \psi'(\xi,\eta,\zeta)D\, d\xi d\eta d\zeta$$

If we choose the substitution such that $\psi' D = 1$, we have again $dW = \text{const.}\, d\xi\, d\eta\, d\zeta$, that is, the same kind of law as above. But in the following we will confine ourselves to the case where $\sum \frac{\partial \varphi_1}{\partial x} = 0$. For this case it has been shown in the foregoing that already for the initial choice of coordinates, the probability of a volume element is equal to a constant multiplied by the size of the element.

Generalization of the theorem to a structure that is completely defined by n quantities $p_1 \cdots p_n$, and the change of which is completely determined by n equations

$$\frac{dp_1}{dt} = \varphi_1(p_1 \cdots p_n)$$

$$\frac{dp_2}{dt} = \varphi_2(p_1 \cdots p_n)$$

$$- - - - - - - - - - -$$

$$\frac{dp_n}{dt} = \varphi_n(p_1 \cdots\cdots p_n)$$

n is an arbtrarily large but finite number.[44] If $\sum \frac{\partial \varphi_\nu}{\partial p_\nu} = 0$, then we have again

$$dn = N \cdot \text{const}\; dp_1 \ldots\ldots dp_n$$

$$dW = \text{const}\; dp_1 \ldots\ldots dp_n$$

Consideration of a Special Case of Particular Importance [p. 21]

We consider again point moving in the space according to an arbitrary law. But we no longer assume that the point sweeps the entire space in the course of its motion, and that after a sufficiently long time it comes arbitrarily close to every point of the space under consideration; instead, the point will always remain on a finite closed surface. One could reduce this case to the more general case considered before insofar as one could completely determine the position of the point on its surface by two coordinates. However, this would lead to great difficulties in applications.

So as to be able to apply three coordinates in this case as well, we consider again not *one* point, but ∞ many points. Each point moves on a finite surface $\varepsilon(xyz) = \text{const.}$ In this case as well we can ask ourselves about the condition that must be satisfied for a stationary point flow. However, we can now no longer infer from the number dn of the points of an elementary volume $d\tau$ the probability of the region for any arbitrary one of the points, precisely because the trajectory of a point does not fill up all surfaces ε = const., but, instead, only one of these surfaces but $\frac{dn}{N}$ is a probability in this case as well, to wit, the probability that a point picked out at random from among all N points will be found in the region $d\tau$.

Treating n again as a continuous spatial function we obtain again as the condition for stationary flow the condition

$$\sum \frac{\partial n\varphi_1}{\partial x} = 0 \quad \text{or} \quad \sum \frac{\partial \varphi_1}{\partial x} = -\frac{d \lg n}{dt}.$$

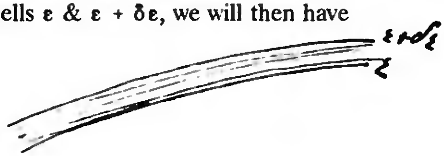
For the individual moving point we will again have n = const. if $\sum = 0$. That is to say, the spatial point density is the same for all points of a surface $\psi = \varepsilon$. But it can be an arbitrary funct. of ε. Hence, we have

$$dn = \text{const} \cdot \psi(\varepsilon) dx\, dy\, dz.$$

Generalization of the theorem to n dimensions.
The simple expedient of considering ∞ many systems instead of a single one puts us in a position to treat systems whose p_ν lie permanently on an n-dimensional surface in a statistical manner, in the same way as those without this property. However, the statistical distribution is here not yet unique by the equations of <motion> change.[45] We can still freely choose $\psi(\varepsilon)$; this is based on the fact that we can still freely choose the number of systems whose ε lies between given limits.

[p. 22] The simplest case that we can choose is that of setting $\psi(\varepsilon)$ constant for values of ε that lie between ε & $\varepsilon + \delta\varepsilon$, but equal to zero for values outside these limits. For a region lying entirely between the energy shells ε & $\varepsilon + \delta\varepsilon$, we will then have

$$dn = \text{const} \cdot dp_1 \ldots\ldots dp_n$$
$$dW = \text{const}\, dp_1 \ldots\ldots dp_n,$$



if, as we will always assume in what follows,

$$\sum \frac{\partial \varphi_\nu}{\partial p_\nu} = 0.$$

Gibbs calls such a phase distribution a microcanonical one.[46] The smaller the $\delta\varepsilon$, the less the paths described by the various n-dimensional points differ from each other, and the more the statistical properties of an individual system are similar to those of the total system.

Canonical Ensemble[47]

Let there again be one, and only one, integral equation $\varepsilon(p_1 \ldots. p_n) = \varepsilon$ for each individual system.[48] We call ε "energy." Let the system stay on energy shell. Then the total system is again governed by the equation

$$dn = \text{const} \cdot \psi(\varepsilon) dp_1 \cdots\cdots dp_n,$$

where ψ can be chosen arbitrarily. We choose $\psi\varepsilon = e^{-\frac{\varepsilon}{\Theta}}$.[49] so that [p. 23]

$$dn = \text{const}\, e^{-\frac{\varepsilon}{\Theta}} dp_1 \ldots.., dp_n,$$

In this case it is not immediately obvious that all systems have practically the same energy. For our exponential factor does become infinitesimally small for large ε but not for small ε. We will demonstrate this right away, however, for the special case where the individual system is an ideal gas.

Ideal Gas

The state variables $p_1 p_2 \cdots p_n$ are given by

$$x_1 y_1 z_1 \quad x_2 y_2 z_2 \cdots x_l y_l z_l$$

$$\xi_1 \eta_1 \zeta_1 \quad \xi_2 \eta_2 \zeta_2 \cdots\cdots \xi_l \eta_l \zeta_l$$

The equations of change become

$$\frac{dx_1}{dt} = \xi_1 \quad \frac{dy_1}{dt} = \eta_1 \quad \frac{dz_1}{dt} = \zeta_1 \quad \ldots \qquad .l..$$

$$\frac{d\xi_1}{dt} = -\frac{1}{m}\frac{\partial \Phi}{\partial x_1} \quad .. \qquad \ldots \qquad .l..$$

$$\varepsilon = E = \Phi + L = \Phi + \sum m_1 \left(\frac{\xi_1^2}{2} + \frac{\eta_1^2}{2} + \frac{\zeta_1^2}{2} \right)$$

The condition $\sum \frac{\partial \varphi_\nu}{\partial p_\nu} = 0$ is here satisfied, since the quantities

$$\frac{\partial \xi_1}{\partial x_1} \cdots\cdots \frac{\partial \zeta_1}{\partial z_1} \quad \& \quad -\frac{1}{m}\frac{\partial}{\partial \xi_1}\left(\frac{\partial \Phi}{\partial x_1}\right) \quad \text{etc.}$$

vanish one by one.[50]

Thus, if we consider a system of ∞ many (n) gases,[51] then the fundamental statistical equation here takes the form

$$dn = \text{const.}\,\psi(E) dx_1 \cdot dx_2 \cdots dz_l d\xi_1 \cdots\cdots d\zeta_l$$

For the canonical composite system we will then have the equation

$$dn = \text{const.}\, e^{-\frac{E}{\Theta}} dx_1 \cdots\cdots dz_l d\xi_1 \cdots d\zeta_l.$$

[p. 24] Now we want to investigate how many systems have an energy lying between E & $E + dE$. Since in the case of an ideal gas the potential energy <for all points in the interior> is very small compared with the kinetic energy in all generally realizable states, we can neglect Φ, and we obtain

$$dn_{dE} = \text{const} \cdot \int\limits_{E \& E+dE} e^{-\frac{E}{\Theta}} dx_1 \cdots dz_l d\xi_1 \cdots\cdots d\zeta_l$$

$e^{-\frac{E}{\Theta}}$ can be taken out. The integration over $dx_1 \cdots dz_l$ yields V^l, (thus a constant). All that remains is to carry out the integration

$$\int\limits_{E/E+dE} d\xi_1 \cdots\cdots d\zeta_l \qquad E = \frac{m}{2}\left(\xi_1^2 \cdots\cdots + \zeta_l^2\right)$$

This integration can be performed by means of the following argument

$\xi_1^2 \cdots\cdots + \zeta_l^2 = \text{const.}$ is the analog of the spherical surface in l dimensions. The radius of the two bounding spheres is

$$\sqrt{\frac{2E}{m}} \qquad \sqrt{\frac{2(E+dE)}{m}}$$

$$R$$

$$\sqrt{\frac{2E}{m}} \quad \text{difference of radii up to constant } \frac{\sqrt{E}}{E}dE = \frac{dE}{\sqrt{E}}$$

$$\sqrt{\frac{2}{m}}\sqrt{E}\left(1 + \frac{1}{2}\frac{dE}{E}\right)$$

l-dimensional spherical surface to be multipled by the difference of the radii.

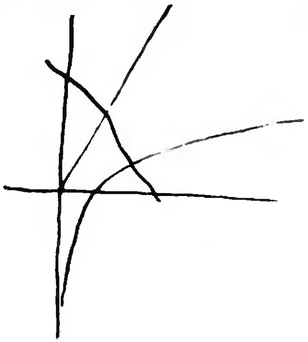
$$\left.\begin{array}{l} \text{spherical surface} = R^{l-1}\text{[52]} \approx E^{\frac{l-1}{2}} \\ \text{difference of radii} \approx dE\, E^{-\frac{1}{2}} \end{array}\right\} \text{product } E^{\frac{l}{2}-1}dE$$

Thus, we obtain

$$dn_{dE} = \text{const. } E^{-\frac{E}{\Theta}}E^{\frac{l}{2}-1}dE$$

We form the logarithm of this function $\left(\frac{l}{2} - 1\right)\lg(E) - \frac{E}{\Theta}$

[p. 25]



$$\langle \lg \rangle \frac{\frac{l}{2} - 1}{E} - \frac{1}{\Theta} = 0 \qquad \Bigg| \qquad \begin{array}{l} E_0(1+\varepsilon)\,\varepsilon \text{ small} \\ -\left(\frac{l}{2} - 1\right)\frac{\varepsilon}{E_0} \end{array}$$

~~There is E_0 where exponent is maximum. If l is very large <& Θ small, however>, this differential quotient changes very rapidly with E. Thus, canonical distribution has in this case as well the required character. Due to the fact that n is very large number. One can say in general. If an <canonical> individual system consists of many elementary structures, then, for a canonical distribution, the larger the n, the more nearly identical will be the energy of all of them. Up to the arbitrarily small, an individual system is thereby reduced for all statistical questions to the corresponding canonical ensemble.~~

<Properties of Canonical Composite Systems>

We set $E = E_0\,(1 + \varepsilon)$

$$\lg E = \lg E_0 + \left(\varepsilon - \frac{\varepsilon^2}{2} + \cdots\right)$$

$$\text{Exponent} = \left(\frac{l}{2} - 1\right)\left[\lg E_0 + \varepsilon - \frac{\varepsilon^2}{2} \cdots\right] - \frac{E_0 + \varepsilon}{\Theta}$$ [53]

Because E_0 is the maximum value, the terms with ε vanish, so that we also have

$$\text{Exponent} = -\left(\frac{l}{2}-1\right)\left(\frac{\varepsilon^2}{2}-\frac{\varepsilon^3}{3}\cdots\right)$$

$$dn_{dE} = \text{konst}\ e^{-((l/2)-1)((\varepsilon^2/2)-(\varepsilon^3/3).....)}dE$$

The quantity in the exponent vanishes for ε that are smaller the greater the l, i.e., the number of molecules. Hence, the greater the number of molecules, the less the energy values of gaseous systems deviate percentagewise from each other. All of the gas systems in our canonical ensemble have approximately the same energy. The same is true for each system in the molecular model that consists of very many molecules.

Practically all of the systems of a canonical ensemble possess the same energy. [p. 26] Instead of investigating the statistical properties of the individual system, we investigate the statistical properties of the canonical ensemble.

Temperature[54]

We consider a microcanonical ensemble whose statistical properties are completely given[55] by the eq.

$$dn = nCe^{-\frac{\varepsilon}{\Theta}}dp_1\cdots dp_l$$

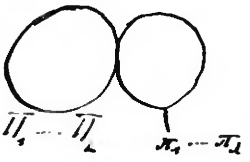
According to the aforesaid, the greater the l, the more the statistical properties of this ensemble resemble those of an individual system of the kind in question. Besides being determined by the <equation of state> functions φ_ν(& ε), the behavior of the ensemble is completely determined by the constants C & Θ. But the latter two constants are not mutually independent, because we must have

$$\int dn = n = \text{total number of systems, hence } 1 = C\int e^{-\frac{\varepsilon}{\Theta}}dp_1\cdots dp_l$$

Thus, the reciprocal value of C is equal to this integral. The statistical behavior of the ensemble (hence also of the individual system) is therefore determined by Θ. For a given composite system, Θ will determine the energy of the system, and vice versa.[56] We want to show that, up to a factor, Θ has the meaning of the temperature.

If we imagine that the system under consideration consists of two parts and that, up to the negligible, $\varepsilon = H + \eta$, where H depends only on the Π, and η only on the π,[57] then

$$dn = \text{const}\, e^{-\frac{(H+\eta)}{\Theta}}\, d\Pi_1 \cdots d\Pi_L \cdot d\pi_1 \cdots d\pi_\lambda$$



[p. 27] We now inquire into the statistical properties of the system of the π, leaving aside the statistical properties of the systems Π. I.e., we ask how many systems $d\nu$ of the π are to be found at an arbitrary instant in a state that is characterized by the region $d\pi_1 \ldots. d\pi_n$?

The number of systems for which not only the π lie in the region $d\pi_1 \cdot\cdot d\pi_n$, but also the Π lie in a specific elementary region, is

$$dn = \text{const} \cdot e^{-\frac{\eta}{\Theta}} d\pi_1 \cdots d\pi_\lambda e^{-\frac{H}{\Theta}} d\Pi_1 \cdots d\Pi_L.$$

If one omits the latter condition, then one has to sum over all regions of the Π while keeping the elementary region of the π fixed. Hence,

$$d\nu = \text{const}\; e^{-\frac{\eta}{\Theta}} d\pi_1 \cdots d\pi_\lambda \int e^{-\frac{H}{\Theta}} d\Pi_1 \cdots d\Pi_L$$

The last integral does not contain the π. Hence the result of the integration is also independent of π, so that we have

$$d\nu = \text{const}\; e^{-\frac{\eta}{\Theta}} d\pi_1 \cdots d\pi_\lambda$$

where "const" denotes a different constant. Thus, the ensemble of the subsystems π again forms a canonical <system> ensemble with the same constant Θ as the ensemble of the original systems.

The same holds of course for the <system> ensemble of the Π, if it is considered by itself. This is also a canonical ensemble with the characteristic constant Θ.

If we ascribe the constant Θ to the individual system rather than to the canonical ensemble, then we can say: Systems that touch each other (for an infinitely long time)[58] have the same Θ. Thus we see that Θ plays the role of the temperature or of a function of the temperature. To be sure, this way of thinking is permissible only if the susbsystems also consist of so many molecules that their canonical distribution is almost without energy fluctuation.

Maxwell's Distribution Law

[p. 28]

The law derived for a subsystem of a canonical ensemble

$$dn = \text{const} \cdot e^{-\frac{\eta}{\Theta}} d\pi_1 \cdots d\pi_\lambda$$

is always correct if the energy of each composite system is composed in the way indicated of the energy of the subsystem and that of the rest of the system, even if the subsystem consists of a single molecule We will consider the latter case, starting with a monatomic ideal gas.

Let the rest of the system be a monatomic molecule of an ideal gas the molecules of which are not acted upon by any external forces.[59] We have then

$$dn = \text{const.}\, e^{-\frac{\frac{m}{2}(\xi^2 + \eta^2 + \zeta^2)}{\Theta}} dx\, dy\, dz\, d\xi\, d\eta\, d\zeta.$$

If we seek the probability for the region $d\xi d\eta d\zeta$, then we have to integrate with respect to x y z, and we obtain

$$dn_1 = \text{const}\, e^{-\frac{(\xi^2 + \eta^2 + \zeta^2)}{\frac{2}{m}\Theta}} d\xi d\eta d\zeta$$

or also

$$dW = \text{const}\, e^{-\frac{(\xi^2 + \eta^2 + \zeta^2)}{\frac{2}{m}\Theta}} d\xi d\eta d\zeta$$

This formula contains the simplest case of Maxwell's distribution law. To be sure, as regards its derivation, it has so far been proved only for a randomly chosen system[60] of a canonical ensemble. But since that probability must be identical, up to the vanishingly small, for the individual systems of the canonical ensemble, because of the nearly identical energy of the composite systems, the formula also holds for the individual system, and the greater the number of the molecules constituting the composite system, the better the approximation to which the formula holds.

[p. 29] The multiplicative constant can easily be determined from the condition $\int dW = 1$

$$1 = \text{konst} \cdot \int e^{\cdots} \, d\xi \, d\eta \, d\zeta = \text{cost} \left[\int_{-\infty}^{+\infty} e^{-\xi^2/(2/m)\Theta} \, d\xi \right]^3 \underset{\sqrt{(2/m)\Theta}}{}$$

$$= \text{const} \left(\frac{2\Theta}{m} \right)^{3/2} \left[\int_{-\infty}^{+\infty} e^{-x^2} \, dx \right]^3$$

$$\int_{-\infty}^{+\infty} = 2 \int_0^{\infty} = 2I$$

$$I^2 = \iint e^{-(x^2+y^2)} \, dx \, dy = \int_0^{\infty} e^{-r^2} 2\pi r \, dr = \pi \int_0^{\infty} e^{-x} 4 \, dx = \pi \cdot 1.^{[61]}$$

$$\int_{-\infty}^{+\infty} = 2\sqrt{\pi}$$

$$1 = \text{konst} \cdot \left(\frac{2\Theta\pi}{m} \right)^{3/2} \cdot (4\pi)^{3/2} \, \text{konst} =$$

We have $dW = \text{const}\, e^{-\frac{\xi^2}{\frac{2}{m}\Theta}} d\xi \cdot e^{-\frac{\eta^2}{\frac{2}{m}\Theta}} d\eta \cdots = dW_\xi dW_\eta dW_\zeta$

The probabilities that ζ lies between specific limits are always the same, no matter what values ξ & η may have. (Probability of mutually independent events.)

$dW_\xi = \text{const}\, e^{-\frac{\xi^2}{\left(\frac{2}{m}\right)\Theta}} d\xi$ We wish to determine the const.

$$\int dW_\xi = 1 \qquad \text{konst} \sqrt{\frac{2}{m}\Theta} \cdot \int e^{-\xi^2/(2/m)\Theta} \frac{d\xi}{\sqrt{\frac{2}{m}\Theta}} = 1$$

$$\text{konst} \cdot \sqrt{\frac{2}{m}\Theta} \underbrace{\int_{-\infty}^{+\infty} e^{-x^2} \, dx}_{\langle 2 \rangle \sqrt{\pi}} = 1 \qquad \underline{\text{konst} = \sqrt{\frac{m}{2\pi\Theta}}}$$

This is the constant of Maxwell's law. If the law is written in three variables, then the constant is to be raised to the third power.

Now we seek to calculate the mean square of the velocity. n systems constitute canonical composite system. 1 specific gas molecule in each system singled out.[62]

In dn_c ξ is between $\xi + d\xi$

We have

$$\overline{\xi^2} = \frac{\Sigma\xi^2 \text{ of the selected molecules of all systems}}{\text{number of systems } (n)} = \frac{\int\xi^2 dn_\xi}{\int dn_\xi}$$

Hence [p. 30]

$$-\frac{1}{2}\int x d(e^{-x^2}) = -\frac{1}{2}\underbrace{\Big|xe^{-x^2}\Big|_{-\infty}^{+\infty}}_{0} + \frac{1}{2}\int_{-\infty}^{+\infty} e^{-x^2}dx$$

$$\overline{\xi^2} = \frac{\displaystyle\int_{-\infty}^{+\infty}\frac{\xi^2}{\frac{2}{m}\Theta}\cdot\cancel{\text{konst}}\, e^{-\xi^2/(2/m)\Theta}\, d\frac{\xi}{\sqrt{\frac{2}{m}\Theta}}}{\displaystyle\int\cancel{\text{konst}}\, e^{-\xi^2/(2/m)\Theta}\, d\frac{\xi}{\sqrt{\frac{2}{m}\Theta}}} = \frac{2}{m}\Theta\frac{\displaystyle\int_{-\infty}^{+\infty}x^2e^{-x^2}dx}{\displaystyle\int_{-\infty}^{+\infty}e^{-x^2}dx} = \frac{\Theta}{m}$$

We obtain

$$\overline{\xi^2} = \frac{\Theta}{m}$$

$$\overline{c^2} = \overline{\xi^2 + \eta^2 + \zeta^2} = \overline{\xi^2} + \overline{\eta^2} + \overline{\zeta^2} = 3\overline{\xi^2} = \frac{3\Theta}{m}$$

$$\boxed{\frac{\overline{mc^2}}{2} = \frac{m}{2}\overline{c^2} = \frac{3}{2}\Theta}$$

We obtain the important theorem that the kinetic energy of a monatomic gas molecule depends only on Θ, and is independent of the mass of the molecule. The mean kinetic energy of a monatomic molecule depends only on the temperature.[63]

We will show that we have thereby proved Avogadro's law for monatomic gases, i.e., that at a given pressure & a given temperature a unit volume always contains the same number of molecules independently of the nature of the gas.

If we have an ideal gas of arbitrary volume, then the virial theorem implies that

$$L = \frac{3}{2}pV$$

If $V = 1$, we have

$$L = Z\frac{\overline{mc^2}}{2} = \frac{3}{2}p$$

If, in accordance with analysis given above, we set $\frac{\overline{mc^2}}{2} = \frac{3}{2}\Theta$, we thus obtain $Z = \frac{p}{\Theta}$, where Θ depends on the temperature alone, which proves Avogadro's law for monatomic gases.

[p. 31] Beyond this, we investigate to what extent the equation of state for ideal gases can be deduced from the investigations carried out thus far. We derived from the virial theorem

$$L = \frac{3}{2}pV$$

If one gram-molecule is present, then V is the molecular volume. We have then, in addition,

$$L = N\frac{\overline{mc^2}}{2} = N\cdot\frac{3}{2}\Theta$$

Hence, by substituting into the above formula,

$$\underline{pV = N\Theta}$$

Thus, pV depends linearly on the temperature function Θ that we introduced, and depends *only* on it.

We have thus shown that according to the kinetics of monatomic gases, pV is a function of temperature alone. One can also deduce from kinetics alone that this function must be equal to universal constant · absolute temperature.

We will delve into this later on, when discussing the second law. At the present we will proceed in the following way.[64] We will conceive of the absolute temperature as defined with the aid of ideal gases by the equation of state

$$pV = RT$$

We have just found from kinetics that

$$pV = N\Theta$$

By comparing the right-hand sides

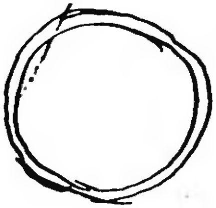
$$\underline{\Theta = \frac{R}{N}T}$$

Thus, Θ is equal to the absolute temperature multiplied by a (numerically very small) constant (about $1.4 \cdot 10^{-16}$).

We can also calculate the most common mean values of the velocity of gas molecules [p. 32]

$$dW = \text{const}\; e^{-\frac{(\xi^2+\eta^2+\zeta^2)}{\frac{2}{m}\Theta}}\, d\xi\, d\eta\, d\zeta$$

How large is the probability of c being between c & $c + dc$? Integrate over the spherical shell



$$dW_{dc} = \text{cons}\; e^{-\frac{c^2}{\frac{2}{m}\Theta}}\, c^2 dc$$

we thus obtain

$$\overline{c^2} = \frac{\int e^{-c^2/(2/m)\Theta} c^4\, dc}{\int e^{\cdots} c^2\, dc} = \frac{2}{m}\Theta \frac{\frac{1}{2}\int e^{-x^2} \overset{x^3 \cdot 2x\, dx}{x^4\, dx}}{\int e^{-x^2} x^2\, dx} = \frac{3}{m}\Theta$$

$$\bar{c} = \frac{\int e^{-} c^3\, dc}{\int e^{\cdots} c^2\, dc} = \sqrt{\frac{2}{m}\Theta}\frac{\int e^{-x^2} x^3\, dx}{\int e^{-x^2} x^2\, dx}$$

$$= \sqrt{\frac{2}{m}\Theta}\frac{\frac{1}{2}}{\frac{\sqrt{\pi}}{4}} = \underline{\frac{\sqrt{8}}{\sqrt{\pi}}\sqrt{\frac{\Theta}{m}}}$$

$$\sqrt{\overline{c^2}} = \sqrt{\frac{RT}{M}} \cdot 1.73$$

$$\bar{c} = \sqrt{\frac{RT}{M}} \cdot 1.6$$

$$\frac{1}{2}\int x^2\, d(e^{-x^2})$$

$$\frac{1}{2}\int e^{-x^2}\, d(x^2) = \frac{1}{2}\int e^{-x}\, dx = \frac{1}{2}$$

$$\frac{1}{2}\int x\, d(e^{-x^2}) = \frac{1}{2}\underbrace{\int e^{-x^2}\, dx}_{\frac{\sqrt{\pi}}{2}}$$

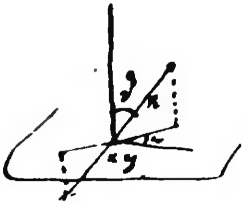
$$\int dW_{dc} = 1 = \left(\frac{2}{m}\Theta\right)^{(3/2)} \int_0^\infty e^{-x^2} x^2\, dx = \frac{1}{2}\left(\frac{\Theta}{2m}\right)^{3/2}$$

$$\frac{1}{2}\int x / d(e^{-x^2}) = \frac{1}{2}\underbrace{\int e^{-x^2}\, dx}_{\frac{\sqrt{\pi}}{2}} = \frac{\sqrt{\pi}}{4}$$

[p. 33] ~~*Diatomic Molecules*~~

~~A molecule determined by~~

~~x y z ϑ ω - - - - - $d\tau$~~
~~ξ η ζ τ o derivatives. $d\tau'$~~



Equations

molecule 1 molecule 2

$\frac{dx}{dt} = \xi$ $\quad \frac{d\xi}{dt} = \varphi_1(x_1 y_1 z_1 \vartheta_1 \omega_1) x_2 \cdots\cdots)$

2. Mol.

$\frac{dy}{dt} = \eta$ $\quad - - - - - - -$

$\overline{}$ $\quad - - - - - - -$ $\quad \cdots\cdots$

$\frac{d\vartheta}{dt} = \tau$ $\quad \frac{d\tau}{dt} = \varphi_4($

$\frac{d\omega}{dt} = o$ $\quad \frac{d\omega}{dt} = \varphi_5($

It follows at once that $\sum \frac{\partial \varphi_\nu}{\partial p_\nu} = 0.$ So our statistical laws are also valid here.

$dW = \text{const}e^{-\frac{E}{\Theta}} dx_x \cdots\cdots do$ (10 variables). E depends only on the derivatives

$$E = \frac{m}{2}(\xi^2 + \eta^2 + \zeta^2) + \frac{m}{2}[R^2 \sin^2 \vartheta o^2 + R^2 \tau^2]$$ [65]

Probability

$$dW = \text{const. } dx\, dy\, dz\, d\vartheta\, d\omega\, e^{-\frac{1}{\Theta}\frac{m}{2}(\xi^2 + \eta^2 + \zeta^2) - \frac{1}{\Theta}\frac{m}{2} R^2(\sin^2\vartheta \omega^2 + \tau^2)} d\xi\, d\eta\, d\zeta\, d\tau\, do.$$

<$dW_{\xi\eta\zeta}$ as if the gas were monatomic.>
Mean kinetic energy of all such atoms

$$\frac{\int E dW}{\int dW} = \frac{\int \left[\frac{m}{2}(\xi^2 + \eta^2 + \zeta^2) + \frac{mR^2}{2}(O^2 \sin^2\vartheta + \tau^2)\right] e^{-\frac{1}{\Theta}[\quad]} d\zeta \cdots do}{\int e^{-\frac{1}{\Theta}[\quad]} d\xi \cdot do}$$

If we introduce new variables so that $\begin{matrix} A_1 \; A_2 \cdots A_5 \\ L = (r_1^2 + r_2^2 \cdots + r_5^2), \end{matrix}$ we obtain

$$\frac{\int_n \sum r^2 e^{-\frac{1}{\Theta}\Sigma r^2} dr_1 \cdots dr_n}{\int_n e^{-\frac{1}{\Theta}\Sigma r^2} dr_1 \cdots dr_n} = \frac{\Theta}{2} \cdot n \quad \text{In our case} \quad \frac{5}{2}\Theta$$

[p. 34] The <specific> heat content is thus $\frac{5}{2}\Theta \cdot N$ per gram-molecule

$$c_v = \frac{5}{2}R$$

But we always have for a gram-mol[ecule]

$c_p = c_v + R$, hence in our case $c_p = \frac{7}{2}R$

$$\frac{c_p}{c_v} = \frac{7}{5} = 1.4$$

Is correct for gases that the chemists regard as diatomic[66]
Generalization of the analysis.

Let $p_1 \cdots p_l$ be coordinates, and $q_1 \ldots\ldots q_l$ velocities[67]
Energy can be written in the form (for an appropriate choice of the q)

$$E = \Pi + \Sigma A_\nu q_l^2$$

P & the A_ν depend <for an appropriate choice of the variables> only on the p.
Canonical distribution.

$$dW = \text{const}\, e^{-(\Pi + \Sigma(A_\nu q_\nu^2))} dp_1 \cdots dp_l dq_1 \cdots dq_l$$

Suppose region $dp_1 \cdots dp_l$ given once and for all. Let us compare the velocity distributions that pertain to this configuration of the p_ν. If we take out everything that does not depend on q_ν, we arrive at

$$dW_p = \text{const}\, e^{-\frac{1}{\Theta}\Sigma A_\nu^2 q_\nu^2} dq_1 \cdots\cdots dq_n$$

Mean value of one of the terms (e.g., $A_1 q_1^2$)

$$\frac{\int \text{const}\, A_1 q_1^2 e^{\cdots} dq_1 \cdots\cdots dq_n}{\int \text{const}\, e^{\cdots} dq_1 \cdots dq_n} = \frac{\int A_1 q_1^2 e^{-\frac{A_1 q_1^2}{\Theta}} \sqrt{A_1}\, dq_1}{\int e^{-} \sqrt{A_1}\, dq_1}$$

$$= \Theta \frac{\int x^2 e^{\cdots} dx}{\int e^{\cdots} dx} = \frac{\Theta}{2}$$ [p. 35]

Thus, mean value of the kinetic energy associated with this configuration $m\frac{\Theta}{2}$. Is independent of the specific configuration. Mean kinetic energy of the system depends in a simple fashion on the number of molecules.

The simplest case for the representation of a solid body

$$E = \underbrace{\sum A_\nu p_\nu^2}_{\Pi} + \underbrace{\sum B_\nu q_\nu^2}_{D}$$

In this case $\overline{E} = n\Theta$.

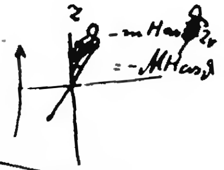
Magnetic Molecule in a Magnetic Field. (Langevin, Weiss)[68]

We think of a molecule that is rigidly connected with an elementary magnet. For the sake of simplicity, we will consider perchance the molecule discussed above (diatomic)

$$dW = \text{konst}\, e^{-(1/\Theta)[(m/2)(\xi^2+\eta^2+\zeta^2)+(mR^2/2)(\sin^2 \vartheta o^2+\tau^2)]|+(1/\Theta)MH\cos\vartheta} dx \cdots\cdots d\omega\, d\xi \cdots\cdots do$$[69]

This formula is valid if no magnetic field is present If magnetic field present, additional term.

~~We ask for probability that molecule is contained in $d\sigma$ (direction cone). Integrated over variables $x\, y\, z\, \xi\, \eta\, \zeta$ gives constant~~



$$dW = \text{const}\, d\vartheta\, d\omega \int e^{-\frac{1}{\Theta}\frac{mR^2}{2}(\sin^2\vartheta o^2 + \tau^2) + \frac{1}{\Theta}MH\cos\vartheta}\, do \quad dt$$

$$= \text{const.}\, d\vartheta\, d\omega\, e^{\frac{1}{\Theta}MH\cos\vartheta} \int e^{-\frac{1}{\Theta}\frac{mR^2}{2}(\sin^2\vartheta o^2)}\, do \int e^{-\frac{1}{\Theta}\frac{MR^2}{2}\tau^2}\, d\tau$$

[p. 36] For a molecule of a monatomic solid substance simplest assumption[70] attraction $ar \quad -ax,\ ay,\ az$

$$E = \frac{m}{2}(\xi^2 + \eta^2 + \zeta^2) + \frac{a}{2}(x^2 + y^2 + z^2)$$

$$dW = ke^{-\frac{1}{\Theta}[\quad]}\, dx\, dy\, dz\, d\xi\, d\eta\, d\zeta$$ [71]

Mean kinetic energy $\frac{m\xi^2}{2}$

$$\overline{\frac{m\xi^2}{2}} = \frac{\int \frac{m}{2}\xi^2 e^{-(1/\Theta)(E)}\, dx \cdots d\zeta}{\int e^{\cdots} \qquad dx \cdots d\zeta}$$

$$= \frac{\int \frac{\frac{m}{2}\xi^2}{\frac{m}{2\Theta}} e^{-(1/\Theta)(m/2)\xi^2} \frac{d\xi}{\sqrt{\frac{m}{2\Theta}}}}{\int e^{\cdots} \qquad \frac{d\xi}{\sqrt{\quad}}} = \Theta \frac{\int x^2 e^{-x^2}\, dx}{\int e^{-x^2}\, dx} = \frac{\Theta}{2}$$

$$\bar{E} = 6\frac{\Theta}{2} = 3\Theta$$

for gram-molecule $\bar{E} = 3RT$ in calories $3 \cdot 1.97 \cdot T = 5.9\, T$

Dulong & Petit law.

Effect of external influences on the relative probability of state regions.

$$dW = \text{const}\, e^{-\frac{E}{\Theta}} dp_1 \cdots\cdot dp_n$$

Let us consider two equally probable regions g_1 & g_2 of the state variables. We have then

$$\frac{dW_2}{dW_1} = \frac{\int\limits_{g_2} e^{-\frac{E}{\Theta}} dp_1 \cdots\cdot dp_n}{\int\limits_{g_1} e^{-\frac{E}{\Theta}} dp_1 \cdots\cdot dp_n} = 1$$

If we now consider the same system, except that an additional term, which we shall denote by ψ,[72] is added to E, then we will have

$$dW = \text{const}\, e^{-\frac{(E+\psi)}{\Theta}} dp_1 \cdots\cdot dp_n$$

[p. 37]

Then

$$\frac{dW_2}{dW_1} = \frac{\int\limits_{g_2} e^{-\frac{(E+\psi)}{\Theta}} dp_1 \cdots\cdot dp_n}{\int\limits_{g_1} e^{-\frac{(E+\psi)}{\Theta}} dp_1 \cdots dp_n} = e^{\frac{(\psi_1-\psi_2)}{\Theta}} \qquad dW = \sim$$

Examples later.[73]

Canonical distribution yields characteristic properties also in the case when not simple mass points but complicated systems—compound molecules in particular—are involved. It is best to apply canonical equations.

L function of p_ν & $\dot{p}_\nu$, in the latter quadratic

We set $\dfrac{\partial L}{\partial \dot{p}_\nu} = q_\nu$

L also funct. of the p_ν & q_ν, in the latter quadratic

$$\left.\begin{aligned} \frac{dp_\nu}{dt} &= \frac{\partial E}{\partial q_\nu} \\ \frac{dq_\nu}{dt} &= -\frac{\partial E}{\partial p_\nu} \end{aligned}\right\} \text{ for closed system}$$

$$\sum \frac{\partial \varphi_\nu}{\partial p_\nu} = \sum \left(\frac{\partial^2 E}{\partial q_\nu \partial p_\nu} - \frac{\partial^2 E}{\partial p_\nu \partial q_\nu} \right) = 0$$

Thus, for this choice of variables the canonical distribution is universally valid

$$dW = \text{const}\, e^{-\frac{\left(\Phi + \left(\frac{1}{2}\right) \sum\sum A_{\mu\nu} q_\mu q_\nu\right)}{\Theta}} dp_1 \cdot \cdots \cdot dp_n dq_1 \cdots dq_n$$

Kinetic energy essentially positive. Hence expression replaceable by $\sum B_\nu r_\nu^2$, where the r are linear functions of the q. We have then also

$$dW = ke^{-\frac{\left(\Phi + \left(\frac{1}{2}\right) \sum (B_\nu r_\nu^2)\right)}{\Theta}} dp_1 \cdots dp_n dr_1 \cdot \cdots \cdot dr_n ,$$

where the constant can certainly now depend on the p_ν.
For specific p_ν we have

$$dW_{p_\nu} = \text{const}\, e^{\frac{-\frac{1}{2}\sum (B_\nu r_\nu^2)}{\Theta}} dr_1 \cdot \cdots \cdot dr_n$$

From this we obtain mean value of $\frac{1}{2} B_\nu r_\nu^2 = \frac{\Theta}{2}$

" " " $\frac{1}{2} \sum B_\nu r_\nu^2 = n \frac{\Theta}{2}$

[p. 38] Thus, the mean kinetic energy is equal to: $\frac{\Theta}{2} \cdot$ number of degrees of freedom. This holds for every configuration of the p, and thus, generally, as long as L is homogeneously quadratic in the q.
diatomic gas $n = 5$

Heat content $\frac{5}{2} \Theta \cdot N$ per gram-molec.

$$\frac{RT}{N}$$

or equal to $\frac{5}{2}RT$

$$\left.\begin{aligned} c_v &= \frac{5}{2}R \\ c_p &= \frac{5}{2}R + R \end{aligned}\right\} \frac{c_p}{c_v} = \frac{7}{5} = 1.4$$

Correct for hydrogen, oxygen, nitrogen, etc.[74]
For molecules in which the atoms form a rigid body <without symmetry> with at least 3 atoms, $n = 6$

$$\frac{c_p}{c_v} = \frac{8}{6} = 1.33.$$

Application to Magnet. Molecule

Magnetic molecule without forces All positions of the magnetic axis equally probable

$$\psi = -MH\cos\vartheta$$

While without field all positions of the axis would be equally probable

$$dW = \text{const}\cdot d\kappa$$

we obtain with field.

$$dW = \text{const.}\, e^{-\frac{MH\cos\theta}{\Theta}} \underset{\substack{| \\ \sin\theta\, d\omega\, d\Theta}}{d\kappa}$$

Mean momentum: [p. 39]

$$\frac{\Theta}{H}\,\frac{\displaystyle\iint \frac{MH\cos\vartheta}{\Theta} e^{-(MH\cos\vartheta)/\Theta}\sin\vartheta\,\cancel{d\omega}\,d\vartheta\cdot\frac{MH}{\Theta}}{\displaystyle\iint e^{-(MH\cos\vartheta)/\Theta}\sin\vartheta\,\cancel{d\omega}\,d\vartheta\,\frac{MH}{\Theta}} = \frac{\displaystyle\int_{-MH/\Theta}^{\overbrace{+MH/\Theta}^{a}} + x\,d(e^{-x})}{\displaystyle\int + d(e^{-x})}$$

[We set; mol. weight; Curie for small; Thus; Molec[ular] field (Weiss); derivable as funct of the temperature]

Setzen wir $\frac{MH\cos\vartheta}{\Theta} = x$

$$= \frac{\left| xe^{-x} \right|_{-a}^{+a} - \int_{-a}^{+a} e^{-x}\,dx}{\int e^{-x}\,dx}$$

$$\frac{NM}{\text{Mol. gew}} = I_0$$

$$\frac{(ae^{-a} + ae^{a}) - (e^{a} - e^{-a})}{e^{a} - e^{-a}}$$

$$\boxed{\begin{array}{l} \text{Curie}^{[75]} \text{ für kl[eine] } a \left\langle \dfrac{1 + a(1-a)}{(1+a)-(1-a)} - \dfrac{1}{a} \right\rangle \\ \dfrac{I}{I_s} = \dfrac{1}{3}\dfrac{M}{\Theta}H \;\Big|\; I = \dfrac{1}{3}\dfrac{I_0^2}{RT}H \end{array}}$$

$$= \frac{\Theta}{H}\left\{ a\frac{\cosh a}{\sinh a} - 1 \right\}$$

$$= M\left(\frac{\cosh a}{\sinh a} - \frac{1}{a} \right)$$

Langevin.[76]

$$\text{Also } \frac{I}{I_0} = \frac{\cosh a}{\sinh a} - \frac{1}{a} \qquad a = \frac{NMH}{RT}$$

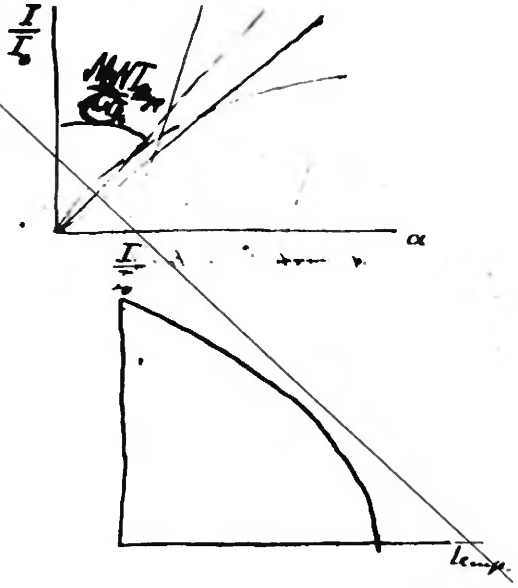
Molek[ular] Feld (Weiss)[77]

$$H_m = N \cdot I$$

$$\frac{MH_m}{\Theta} = \frac{MN}{\Theta}I = \frac{MN}{\Theta}I_0\frac{I}{I_0}$$

$$a_m = \frac{MN}{\Theta}I_0\frac{I}{I_0}$$

$\frac{I}{I_0}$ als Funkt der Temp[eratur] ableitbar



Dependence of saturation magnetization on the temperature. [p. 40]
Variability of the specific heat.
jumpwise change of magnetization in crystals. (Hysteresis).[78]

Application to Suspensions (Brownian Motion)[79]

$$E_{\text{additional}} = V(\rho - \rho_0)g$$

$$dW = \text{const}\, dz e^{-\frac{V(\rho - \rho_0)gz}{\Theta}} = \text{const}\, e^{-\frac{N}{RT}V(\rho - \rho_0)gz} dz$$

Perrin determined with suspensions of resin granules in water. N obtained.[80] One can conclude that Brownian motion takes place. Simpler in the following way

$$dW = \text{const}\, e^{-\frac{\frac{a}{2}x^2}{\Theta}} dx$$

Mean value for $\frac{a}{2}x^2$

$$\frac{a}{2}\overline{x^2} = \frac{1}{2}\Theta = \frac{1}{2}\frac{R}{N}T$$

∞ many such systems treated.

x instantaneous value. Observed after short time τ |path due to force

$$= \frac{\text{force } \tau}{6\pi\eta P}\text{[81]}$$

$$= \frac{-ax\tau}{\underbrace{6\pi\eta P}_{\kappa}}$$

|path due to thermal motion Δ

$$x_{t+\tau} = x_t - \kappa\tau x_t + \Delta = (1 - \kappa\tau)x_t + \Delta$$

$$\overline{x^2_{t+\tau}} = (1 - 2\kappa\tau)\overline{x^2_\tau} + 2(1 - x\tau)\underset{\underset{0}{\parallel}}{\overline{x_\tau\Delta}} + \overline{\Delta^2}$$

$$\overline{x^2_{t+\tau}} = \overline{x^2_\tau}$$

$$2\kappa\tau\overline{x^2_\tau} = \overline{\Delta^2}$$

[p. 41] From this

$$\frac{a}{3\pi\eta P}\cdot\frac{RT}{Na}=\overline{\Delta^2}$$

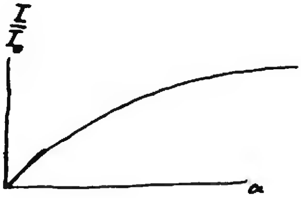
$$\boxed{\overline{\Delta^2}=\frac{RT}{N}\cdot\frac{1}{3\pi\eta P}}$$

This formula determines Brownian motion. Rotational motion can be obtained in an analogous way.[82]

Magnetism

1. Gases

$$\frac{I}{I_0}=\frac{\cosh a}{\sinh a}-\frac{1}{a}=F(a) \qquad a=\frac{\mu H}{\Theta}$$



For all attainable fields a is so small that one can set $F(a)=\frac{1}{3}a$.[83] Thus, practically

$$I=\frac{1}{3}I_0a \qquad =\frac{1}{3}\left(\frac{M}{RT}\right)^2\frac{\langle I\rangle H}{T}$$

$$I_0=\frac{M}{V}=\frac{M}{RT}p \qquad a=\frac{N\mu H}{RT}=\frac{MH}{RT}$$

2. Solids. Ferromagnetism.

To explain the magnetic properties of solids, Weiss assumes that magnetization acts like a magnetizing force parallel to them.[84] (We shall confine ourselves to *one* dimension.) $H_m = NI$.

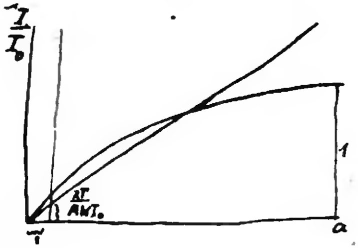
We then have two equations

$$\frac{I}{I_0}=F(a) \qquad a=\frac{M}{RT}(H+H_m)=\frac{M}{RT}(H+WI)$$

First we consider the case where no external force is present; we have then

$$\frac{I}{I_0} = F(a)$$

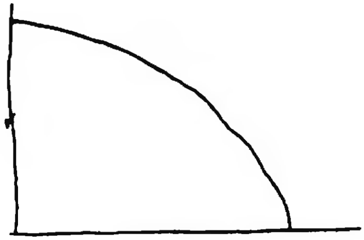
$$\frac{I}{I_0} = \frac{RT}{MWI_0}a$$



In this way we obtain relation between $\frac{I}{I_0}$ and T.

[p. 42]

We seek temperature at which ferromagnetism vanishes. We obtain



$$\frac{I_0 M}{\rho}$$

$$\frac{RT_m}{MWI_0} = \frac{1}{3} \qquad T_m = \frac{1}{3}\frac{MWI_0}{R} = \frac{1}{3}\frac{I_0^2 MW}{\rho R}$$

Hence, the above equations for the determination of the curve can also be written as

$$\frac{I}{I_0} = F(a) \quad \text{(remains)}$$

$$\frac{I}{I_0} = \frac{1}{3}\frac{T}{T_m}\cdot a$$

By eliminating a, one obtains a relation between $\frac{T}{T_m}$ and $\frac{I}{I_0}$ that is independent of the nature of the substance. Shows satisfactory agreement with experiment.[85]

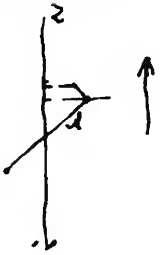
Contribution to Specific Heat

If a molecular magnet changes its position under the influence of the field H, then the work is

$$H\ 2m dz = H\ 2m d(\lambda \cos\vartheta) = H\ d(\mu \cos\vartheta)$$

If many elementary magnets are present, then the work

$H \cdot d\sum \mu\cos\vartheta$, or $H dI$ per unit volume will be performed.



Thus, when the magnetization changes by dI, the energy content of the system increases by

$$-H\,dI.$$

This formula is also applicable when H is the molecular field. In that case $H = WI$, so that

$$E_{\text{magn}} = -WIdI = -\frac{W}{2}I^2$$

Contribution to the specific heat $-\frac{W}{2}I\frac{dI}{dT}$

[p. 43] Jump at $T = T_m$ confirmed experimentally.[86]

Proof that in seemingly not unmagnetized ferromagnetic substances saturation magnetization is really present.[87]

Hysteresis qualitatively explained.

Paramagnetism, with Weiss's Force Taken into Account

$$\frac{I}{I_0} = F(a) \qquad a = \frac{\mu H_{\text{tot}}}{\Theta} = \frac{I_0 M}{\rho R T}(H + WI)$$

$$\frac{R}{N}T \qquad V_\rho = M$$

$$\frac{M}{N} = \frac{I_0 V}{N} = \frac{I_0 M}{\rho N}$$

If only infinitely small fields are admitted, then $a = 3\frac{I}{I_0}$. If this is inserted into the second equation, one obtains

$$I = \enclose{circle}{\left(\frac{1}{3}\frac{I_0^2 M}{\rho R T}\right)}^{\alpha} (H + WI)$$

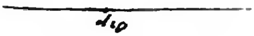
$$I(1 - \alpha W) = \alpha H$$

$$I = \frac{\alpha}{1 - \alpha W} H$$

α and W can be determined from the paramagnetic state. Everything in α is known except for the molecular weight M, hence possible to determine this. Magnetic molecules of iron diatomic for ordinary iron.[88]

Brownian Motion of Rotation[89]

[p. 44]



λ parameter., in the absence of an effective cause, let all λ be equally probable. If force $-a\lambda$, then energy $-a\frac{\lambda^2}{2}$.

Then

$$dW = \text{const} \cdot e^{-\frac{a}{2}\frac{\lambda^2}{\Theta}} d\lambda$$

From this $\overline{\lambda^2} = \frac{\Theta}{a}$

We imagine system, value λ, 2 changes during time τ

1) $-a\lambda \cdot B\tau$ (B = motility)

2) Change $\pm\Delta$ due to irregularity of thermal processes.

$$\lambda_{t+\tau} = \lambda_\tau - aB\lambda_\tau\tau + \Delta = \lambda_\tau(1 - aB\tau) + \Delta$$

$$\overline{\lambda^2_{t+\tau}} = \overline{\lambda^2_\tau} = \underbrace{\overline{\lambda^2_\tau}(1}_{} - 2aB\tau) + \underbrace{2\overline{\lambda_\tau\Delta}}_{\text{vanishes}} - \underbrace{2aB\tau\overline{\lambda_\tau\Delta}}_{\text{vanishes}} + \overline{\Delta^2}$$

$$\overline{\Delta^2} = 2aB\tau\overline{\lambda^2_\tau} = 2B\Theta\tau$$

If we are dealing with the rotation of a sphere, then we have to set λ = angle of rotation

$$B = \frac{1}{8\pi\eta P^3}$$ [90]

$$\overline{\Delta^2} = \frac{1}{4\pi\eta P^3}\frac{RT}{N}\tau$$

[p. 45]

Canonical Distribution & Entropy

$$dW = Ce^{-\frac{E}{\Theta}}dp_1 \cdots dp_n = e^{c-\frac{E}{\Theta}}dp_1 \cdots\cdot dp_n$$

E dependent on $p_1 \cdot\cdot p_n$, and in addition on $\lambda_1 \cdots \lambda_n$.[91]
After ∞ small change of state, canonical distribution obtains again. E has now parameter $\lambda + d\lambda$ We call this function $E + dE. = E + \sum \frac{\partial E}{\partial \lambda} d\lambda$
The temperature is $\Theta + d\Theta$. c then has the value $c + dc$.
Before and after the change we have $\int dW = 1$.

$$\int e^{c+dc-\frac{\left(E+\sum\frac{\partial E}{\partial\lambda}d\lambda\right)}{\Theta+d\Theta}}dp_1 \cdots\cdot dp_n - \int e^{c-\frac{E}{\Theta}}dp_1 \cdots\cdots dp_n = 0$$

The first exponent can be developed

$$\left(c - \frac{E}{\Theta}\right) + \left(dc + \frac{E}{\Theta^2}d\Theta - \frac{\sum\frac{\partial E}{\partial\lambda}d\lambda}{\Theta}\right)$$

The exponential function is

$$e^{c-\frac{E}{\Theta}}\left(1 + dc + \frac{E}{\Theta^2}d\Theta - \sum\frac{1}{\Theta}\frac{\partial E}{\partial\lambda}d\lambda\right)$$

Thus, the above equation becomes

$$\int\left(dc + \frac{E}{\Theta^2}d\Theta - \sum\frac{1}{\Theta}\frac{\partial E}{\partial\lambda}d\lambda\right)e^{c-\frac{E}{\Theta}}dp_1 \cdots dp_n = \int(\ )dW = 0$$

This can also be written as

$$\Theta dc + \frac{\overline{E}}{\Theta}d\Theta - \sum\frac{\overline{\partial E}}{\partial\lambda}d\lambda = 0 \cdots\cdots (1) \text{ or } \left\langle dc + d\left(\frac{\overline{E}}{\Theta}\right)\right\rangle$$

This equation holds for every ∞ small change of state
We omit bars & prove that $\sum\frac{\partial E}{\partial\lambda} = A$ (the work added to the system)

<Now we have>

$$\cancel{dE = \sum \frac{\partial E}{\partial p_\nu} dp_\nu} + \sum \frac{\partial E}{\partial \lambda} d\lambda$$

$$= \sum \frac{\partial E}{\partial \lambda} d\lambda + \cancel{\sum \frac{\partial E}{\partial p_\nu} \frac{dp_\nu}{dt} dt}$$

$$dc + \frac{E}{\Theta^2} d\Theta - \frac{dE - Q}{\Theta} = 0 \qquad \frac{Q}{\Theta} = -dc + d\left(\frac{E}{\Theta}\right) = d\left(\frac{E}{\Theta} - c\right)$$

Thus, $\frac{Q}{\Theta}$ is a total differential

$$\sum \frac{Q}{\Theta} = \frac{E}{\Theta} - c \qquad S = \frac{E}{T} - \frac{R}{N} c$$ [p. 46]

Thus, quantity on right is equal to the entropy. We can rewrite this

$$\int e^{c - \frac{E}{\Theta}} dp_1 \cdots dp_n = 1$$

$$\lg \int = 0 \qquad c + \lg \int e^{-\frac{E}{\Theta}} dp_1 \cdots dp_n = 0$$

Since all systems practically the same energy, E can be replaced in this equation by the mean value E_0

$$c - \frac{E}{\Theta} + \int dp_1 \cdots dp_n = 0$$

Thus also

$$S = \sum \frac{Q}{T} = \frac{R}{N} \lg \int_{E_{[0]}}^{E + \delta E} dp_1 \cdots\cdots dp_{n\,(\text{for given } \lambda)} = \frac{R}{N} \lg(G_\lambda)$$

Entropy = const · magnitude of the region of the elementary variables that belongs to E & $\lambda_1 \cdots\cdots \lambda_n$.

If one considers once and for all a system with a given energy value, but assumes that the λ can assume all possible values (piston does not exist), then

$$dW = \text{const}\, dp_1 \cdots\cdots dp_n$$

If one designates W as the probability of an arbitrary region G_λ, characterized by specific values of λ, we will have

$$W = \text{const} \cdot G_\lambda.$$

Probability of a λ-state at given energy equals the magnitude of the region in question. Hence, up to an inconsequential constant, we can set

$$S = \frac{R}{N} \lg W_\lambda$$

[p. 47] This is Boltzmann's principle.
<Proof that an ∞ small>

Microcanonical and Canonical Ensemble

$$dN = A d\pi_1 \cdots\cdot d\pi_l$$

$$= A d\Pi_1 \cdots\cdot d\Pi_L d\pi_1 \cdots\cdot d\pi_\lambda$$

Energy of the whole between E and $E + \Delta$

$$dN' = A\ d\pi_1 \cdots\cdot dp_\lambda \int\limits_{E-\eta}^{E-\eta+\Delta} \underset{T}{d\Pi_1 \cdots\cdot d\Pi_L}$$

We set $\int\limits_{x}^{x+\Delta} dT = \Delta \cdot \psi(x)$ $\int\limits_{x_0}^{x} dT = \Psi(x) = \int\limits_{x_0}^{x} \psi(x) dx$

$$dN' = A\ d\pi_1 \cdots\cdot d\pi_\lambda \quad \Delta \cdot \psi(E - \eta)$$

Function ψ is sought.

1. Reservoir is ideal gas.

[const const const]

$$dT = \Pi(dx) \qquad \Pi\, d\xi \cdot m^L$$

$$\int dT = V^{L/3} m^L \int_0^{\sum (m/2)\xi^2} \Pi(d\xi) = V^{L/3} m^L \cdot \left(\frac{2}{m}\right)^{L/2} \underbrace{\int_0^{\sum \xi^{x^2}} \pi\, d\xi^x}_{f \cdot \xi_{\max}^{xL}} = \text{konst}\, H^{L/2}$$

$$\psi(x) = \text{konst} \cdot H^{L/2-1}$$

$$f H^{L/2}$$

$$\psi(E - \eta)?$$

$$\lg \psi = \text{konst} + \left(\frac{L}{2} - 1\right) \lg H$$

$$\lg \psi(E - \eta) = \frac{L}{2} - 1 \lg(E - \eta) = \frac{L}{2} - 1 \left[\lg(E) + \lg\left(1 - \frac{\eta}{E}\right)\right]$$

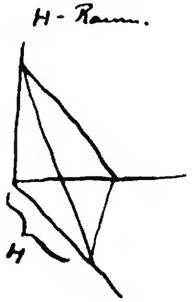
$$= -\frac{\eta}{E} - \frac{\eta^2}{2E^2} \cdots \quad [92]$$

$$\psi(E - \eta) = e^{-(\eta/E)(1+(\eta/2E)+\cdots)((L/2)-1)} = {}^{-\eta/\Theta(1+(\eta/2E)\cdots)}$$

Thus, the larger the E, the more closely [p. 48]

$$dN' = \text{const} e^{-\frac{E}{\Theta}} d\pi_1 \cdots d\pi_l \text{ (Canonical distribution)}$$

2. Reservoir consists of a large number of identical things



$$\Phi = \int_0^H dT = \int_{\text{konst.}}^{H = H_1 + H_2 \cdots H_n} dT_1\, dT_2 \cdots dT_n \langle \omega_1 \ldots \omega_n \rangle$$

$$= \int_0^H \omega_1\, dH_1\, \omega_2\, dH_2 \cdots \omega_n\, dH_n$$

We calculate this integral for the argument $H + \varepsilon$.

In place of every dH_1 we have $dH_1\left(1 + \dfrac{\varepsilon}{H_1}\right)$

" " " " H_1 " " $H_1\left(1 + \dfrac{\varepsilon}{H_1}\right)$

$\Phi' = \displaystyle\int \Pi\omega_1(H_1 + \varepsilon)\Pi d\,H_1\left(1 + \frac{\varepsilon}{H_1}\right)$, where ε is ∞ smaller

$$= \int \Pi\{\omega(H)_1 + \omega'(H_1)\cdot\varepsilon\} \quad \Pi\left(1 + \frac{\varepsilon}{H_1}\right)\Pi(dH_1)$$

$$\Pi\omega(H_1)\left(1 + \varepsilon\sum\frac{\omega' H_1}{\omega H_1}\right) \quad \left(1 + \varepsilon\sum\frac{1}{H_1}\right)\Pi\, dH_1$$

$$\int \Pi\omega(H_1)\Pi(dH_1)\left\{1 + \varepsilon\sum\left(\frac{\omega'(H_1)}{\omega(H_1)} + \frac{1}{H_1}\right)\right\}$$

[p. 49]

The Electron Theory of M[etals][93]

$$F = -\frac{1}{3}nc\lambda\frac{\partial G}{\partial z} \qquad = -\frac{R\overset{\frac{1}{2}}{\langle T\rangle}}{\mu N}nc\lambda\frac{\partial T}{\partial z}$$

$$\frac{1}{2}\,\frac{\partial\mu\frac{c^2}{2}}{\partial T}\,\frac{\partial T}{\partial z} \qquad \langle\mu\rangle c^2 = \frac{3RT}{\mu N} \qquad \kappa = \frac{1}{2}\frac{R}{\langle\mu\rangle N}n\langle c\rangle\lambda c$$

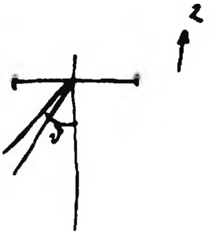
$$\frac{\lambda}{c} = \tau$$

$$-\mathfrak{E}\frac{\varepsilon}{\mu}\frac{\tau^2}{2}\cdot\frac{1}{\tau} = \text{ mean veloc. } = -\frac{1}{2}\mathfrak{E}\frac{\varepsilon}{\mu}\frac{\lambda}{c} = C$$

$$-nC\varepsilon = +\sigma\mathfrak{E} = +\boxed{\frac{1}{2}\,\mathfrak{E}\,\frac{\varepsilon^2}{\mu}\frac{n\lambda}{c}}$$

$$\left.\begin{aligned}\kappa &= \frac{1}{2}\frac{R}{\langle\mu\rangle N}n\lambda c\\ \sigma &= \frac{1}{2}\frac{\varepsilon^2}{\mu}\frac{n\lambda}{c}\end{aligned}\right\}\frac{\kappa}{\sigma} = \cancel{2}\frac{R\mu}{N\varepsilon^2}\underset{\frac{3RT}{\langle\mu\rangle N}}{c^2} = \cancel{6}3\frac{R^2}{N^2\varepsilon^2}T = \cancel{6}3\frac{R^2}{E^2}T.$$

Thermoelectr. Forces[94]



$$c\cos\vartheta\cdot n\frac{d\kappa}{4\pi}$$

$$(cn)_{z-\lambda\cos\vartheta}\cos\vartheta\frac{d\kappa}{4\pi}$$

$$\left((cn)-\frac{\partial(cn)}{\partial z}\lambda\cos\vartheta\right)^{2}\langle\vartheta\rangle\cos\vartheta\frac{d\kappa}{4\pi}$$

$$\frac{1}{4\pi}\int\cos\vartheta\sin\vartheta\,d\vartheta\,d\langle\varphi\rangle\omega=0$$

$$\frac{1}{4\pi}\int\cos^{2}\vartheta\sin\vartheta\,d\vartheta\,d\omega=\frac{1}{2}\left|-\frac{\cos^{3}\vartheta}{3}\right|=\frac{1}{3}$$

$$\text{Electron flow} = -\frac{1}{3}\frac{\partial nc}{\partial z}\lambda$$

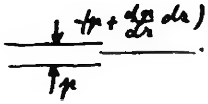
Electron flow under the influence of a field $\mathfrak{E}$: $Cn = -\frac{1}{2}\frac{\varepsilon}{\mu}\frac{\lambda n}{c}\mathfrak{E}$

$$\frac{\partial\varphi}{\partial z}=-\frac{2}{3}\frac{m}{\varepsilon}\frac{c}{n}\frac{\partial(nc)}{\partial z}$$ [p. 50]

Peltier force & Thomson force separately.

Suspended Particles

$$-\left(p+\frac{dp}{dz}dz\right)$$



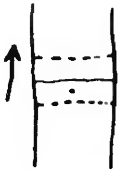
$$V(\rho-\rho_0)gN\eta$$

$$RT\frac{d\eta}{dz}=-V(\rho-\rho_0)gN\eta$$

$$\eta=\eta_0 e^{\cdots}$$

$$n=n_0e^{-V(\rho-\rho_0)gNz/RT}=f(z)$$

Kinetic analysis for equilibrium



B = motility

$u = V(\rho - \rho_0)g \cdot B$

$F'_{\text{[syst?]}} = - un\tau$

Vertical displacement due to irregular motion Δ

$$\left\langle \frac{1}{2}(fn)_{z-\Delta/2} \right\rangle \quad \frac{1}{2}\Delta \cdot n_{z-\Delta/2} \quad -\frac{1}{2}\Delta n_{z+\Delta/2}$$

$$\frac{1}{2}\Delta\left\{n - \frac{dn}{dz}\frac{\Delta}{2} - \left(n + \frac{dn}{dz}\frac{\Delta}{2}\right\}\right.$$

$$F''_{\langle\text{[irr?]}\rangle z} - \frac{1}{2}\Delta^2\frac{dn}{dz}$$

$$un\tau + \frac{1}{2}\Delta^2\frac{dn}{dz} = 0$$

$$\alpha B\tau\langle + \rangle - \frac{1}{2}\Delta^2\alpha\frac{N}{RT}$$

$$\Bigg|\sum \Delta_\nu^2 \frac{dn_\nu}{dz}$$

$$\|$$

$$\sum \frac{\Delta_\nu^2 n_\nu^2}{n}\frac{dn}{dz}$$

$$= \overline{\Delta^2}\frac{dz}{dz}.$$

$$n_\nu = \alpha_\nu n$$

$$\frac{dn_\nu}{dz} = \alpha_\nu \frac{dn}{dz} = \frac{n_\nu}{n}\frac{dn}{dz}$$

$$\sum \Delta$$

$$\Delta^2 = 2\frac{RT}{N}B\tau$$

$$/$$

$$\frac{1}{6\pi\sigma P}$$

$$/$$

Viscosity coefficient.

[p. 51] Holds also in absence of gravitational force. Law of Brownian motion.

$$\sqrt{\overline{\Delta^2}} = \alpha\sqrt{\tau}$$

Explanation $\Delta_{\text{tot}} = \Delta_1 + \Delta_2 \cdots + \Delta_l$ paths independent.

$$\overline{\Delta_{\text{tot}}^2} = l\overline{\Delta^2}$$

$$\sqrt{\overline{\Delta_{\text{tot}}^2}} = \overline{\Delta^2}\sqrt{l}$$

$\frac{f(z)}{n_0}$ probability of the position of the particle as funct. of z.

After this irreversible processes explained.[95]

[p. 52]

f-Correction[96][97]

$$x \Big| \; m\frac{d^2x}{dt^2} = X$$

$$- \Big| \; ------ \cos\vartheta\cos\psi - \cos\vartheta\sin\psi\sin\varphi$$

$$\Big| \; ------$$

$$-\sum m\left[\left(\frac{dx}{dt}\right)^2 + \cdot + \cdot\right] = \sum Xx + \cdot + \cdot$$ [98]

$$L = -\frac{1}{2}\sum rf(r) + \frac{3}{2}pV$$

For collision

$$m_1\frac{d^2x_1}{dt^2} = X \qquad m_2\frac{d^2x_2}{dt^2} = -X$$

$$---- \qquad ----$$

$$----- \qquad -----$$

$$\left\langle \left| m\left(\frac{dx_1}{dt} - \frac{dx_2}{dt}\right)\right|_a^e = 2\int X\,dt \right\rangle$$

r	$f(r)$
$x_1 - x_2$	X
$----$	Y
$----$	Z

$$rf(r) = (x_1 - x_2)X + \cdot + \cdot$$

$$= m_1\left(x_1\frac{d^2x_1}{dt^2} + \cdot + \cdot\right) + m_2(\qquad)$$

$$\frac{d}{dt}(m_1(x_1\xi_1 + \cdot + \cdot) + m_2(x_2\xi_2 —) - m_1v_1^2 + m_2v_2^2$$

$$\begin{array}{c|c} m\dfrac{d^2(x_1-x_2)}{dt^2} = 2X & x_1 - x_2 \\ \text{------} & \text{---} \\ \text{---------} & \text{-----} \end{array}$$

$$\int f(r)r\,dt = 2m\left|\left([x_1-x_2]\frac{d(x_1-x_2)}{dt} + \cdot + \cdot\right)\right|^{\text{[99]}}$$

$$= 2m(\mathfrak{r}v_r)$$

[p. 53]

$$\sum T_{\nu\nu} = \alpha^{\text{[100]}}$$

$$\frac{\partial\varphi}{\partial x_\mu}\sum\frac{\partial^2}{\partial x_\nu^2}(\lg\varphi)$$

$$\frac{\text{\_\_\_\_}G_0'}{} \qquad \frac{G_0^x}{G_0}$$

$$nc\cos\vartheta\frac{d\kappa}{4\pi}G_0'(z-\lambda\cos\vartheta)$$

$$G_0' - \frac{\partial G}{\partial z}\lambda\cos\vartheta$$

$$\frac{nc}{4\pi}\left[G_0'\int_0^{\pi/2}\cos\vartheta\sin\vartheta\,d\vartheta\,d\omega - \lambda\frac{\partial G}{\partial z}\int_0^{\pi/2}\cos^2\vartheta\sin\vartheta\,d\vartheta\,d\omega\right]$$

$$n\frac{c}{2}\left\{G_0'\left|\frac{-\cos^2\vartheta}{2}\right|_0^{\frac{\pi}{2}} - \lambda\frac{\partial G}{\partial 2}\left(\frac{-\cos^3\vartheta}{3}\right)\right\}$$

$$nc\left\{+\frac{1}{4}G_0'\langle+\rangle - \frac{1}{6}\lambda\frac{\partial G}{\partial z}\right\}$$

$$\frac{\langle+\rangle - nc\frac{1}{4}G_0^x}{}$$

$$+\left\langle\frac{1}{4}\right\rangle nc\left[\frac{1}{4}\{G_0' - G_0^x\}\langle+\rangle - \frac{1}{6}\lambda\frac{\partial G}{\partial z}\right] = F = -\frac{1}{3}nc\lambda\frac{\partial G}{\partial z}$$

$$\frac{1}{4}(G_0' - G_0^x) = -\frac{1}{6}\lambda\frac{\partial G}{\partial z}$$

$$G_0' - G_0^x = -\frac{2}{3}\lambda\frac{\partial G}{\partial z}$$

Simplest assumption $G_0^x = G_0$
More complicated assumption

$$\Big\langle G_0^x \quad \frac{G_0 =}{G_0'} \Big\rangle \underline{G_0'(z-\alpha\lambda\cos\vartheta)-G_0}$$

$$G_0^x - G_0 = p\left(G_0' - \frac{\partial G}{\partial z}\alpha\lambda\cos\vartheta - G_0\right)$$

$$\langle p\rangle G_0' = \frac{G_0^x - (1-p)G_0 + p\dfrac{\partial G}{\partial z}\alpha\lambda\cos\vartheta}{p}$$

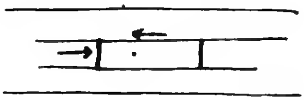
$$\left\langle G_0' - G_0^x = \left(1-\frac{1}{p}\right)G_0 + \langle p\rangle\frac{\partial G}{\partial z}\alpha\lambda\cos\vartheta\right\rangle$$

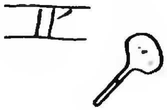
$$G_0' = G_0^x - \left(1-\frac{1}{p}\right)G_0^x + \left(1-\frac{1}{p}\right)G_0 + \frac{\partial G}{\partial z}\alpha\lambda\cos\vartheta$$

$$\left\langle G_0' - G_0^x = \left(1-\frac{p}{p}\right)(G_0^x - G_0^{\langle x\rangle}) + \frac{\partial G}{\partial z}\alpha\lambda\cos\vartheta\right\rangle$$

G

Ordinary viscosity flow [101] [p. 54]





$$-\pi\langle\rho\rangle r^2\frac{\partial p}{\partial 2} + 2\pi\langle\rho\rangle r\eta\frac{\partial u}{\partial r} = 0$$

$$\langle\gamma = 2\rangle\frac{\partial u}{\partial r} = \frac{\gamma}{2\eta}r$$

$$u = \frac{\gamma}{4\eta} - \langle r^2 + \text{konst}\rangle (R^2 - r^2)^{[102]}$$

$$\cancel{2\pi r\,dr\cdot u = \frac{\pi}{2}\frac{\gamma}{\eta} r^3\,dr + \alpha r\,dr}$$

$$\cancel{F_{\text{tot. Vol}} = \frac{\pi}{2}\frac{\gamma}{\eta}\left(\frac{r^4}{4} + \alpha\frac{r^2}{2}\right)}$$

$$\cancel{= \frac{\pi}{2}\frac{\gamma}{\eta}\left(\frac{r^4}{4}\langle + \rangle - \frac{r^2R^2}{4}\right)}$$

$$2\pi r\,dr\,u = \frac{\pi}{2}\frac{\gamma}{\eta}(R^2 - r^2) r\,dr$$

$$F_{\text{Vol tot}} = \frac{\pi}{2}\frac{\gamma}{\eta}R^4\left(\frac{1}{2} - \frac{1}{4}\right) = \frac{\pi}{4}\frac{\gamma}{\eta}R^4 \quad [103]$$

$$\eta = \frac{1}{3}\rho c\lambda$$

$$\left[F_{m\,\text{tot}} = \frac{\pi}{4}\frac{\gamma}{\frac{\eta}{\rho}}R^4 = \frac{3\pi}{4}\frac{\gamma}{c\lambda}R^4 = \left\langle\frac{9}{16}\quad R^4\gamma\right\rangle \pi^2\frac{\sigma^2 n}{c\lambda}\right] [104]$$

$$\lambda = \frac{1}{\frac{4}{3}\pi\sigma^2 n}$$

$$qu = F_{\text{Vol.}} = \frac{4q^2}{\rho c P}\gamma \approx \frac{\cancel{4}\pi^{\cancel{2}}R^4}{\cancel{2}\cancel{\pi}\rho c R}\gamma = \frac{2\pi R^3}{\rho c}\gamma$$

$$F_{[\text{masse}]} = \frac{4q^2}{\langle\rho\rangle c P}\gamma$$

Quotient[105]

[again] $F_{\langle\text{masse}\rangle\text{Vol}} = \dfrac{\pi}{4}\dfrac{R^4}{\eta}\gamma$

$$\frac{8}{R}\frac{\eta}{\rho c} = \frac{8}{P}\frac{1}{\rho c}\cdot\frac{1}{3}\rho c\lambda = \frac{8\lambda}{3P}$$

[p. 55] $(dt)_{\text{const}}$ z vanishes in infinity.

Application of Probabilistic Analysis to Processes of Motion.[106]

1) Point moving with constant velocity along a closed curve Two elements chosen, ds_1 & ds_2, where $ds_1 = ds_2$. We say: it is equally probable that we will find the point in ds_1 & ds_2

2) The same case considered, but $v = \varphi(\lambda)$. Now it is no longer equally probable that we will find the point in ds_1 as in ds_2. What is meant by this?

In order to traverse ds_1, it requires the time $\frac{ds_1}{v_1}$, and to traverse ds_2, it requires the time $\frac{ds_2}{v_2}$ It turns out that $\frac{ds_1}{v_1} \neq \frac{ds_2}{v_2}$ We will consider $\frac{ds_1}{v_1}$ & $\frac{ds_2}{v_2}$ as a relative measure for the probability of finding the point on ds_1. We divide in the following way

$$\frac{ds_1}{v_1} = \tau_1 \qquad \frac{ds_2}{v_2} = \tau_2$$

T = duration of one orbit

Probability for region $ds_1 = \frac{\tau_1}{T}$

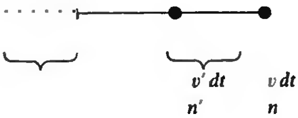
" " " $ds_2 = \frac{\tau_2}{T}$

Thus, by the probability associated with the region *ds* we understand the fraction of time during which a point is found in ds_1, divided by the time of a whole orbit.

$$W = \frac{\text{length of time the condition applies}}{\text{total time.}}$$

This definition can also be understood in another way. We imagine that very many (∞ many)[107] points are traversing our curve in the same direction and according to the same law. Can they be distributed in such a way that every region *ds* always contains the same number of points, i.e., that for a specific region this number does not change with time? Let nds be the number of points in the element ds; n is to be conceived as a function of s. We calculate the change of nds in the very short time dt

$$\text{Increase in } n = vn\,dt - v'n'dt$$



$$= -\frac{\partial(vn)}{\partial s}ds\,dt$$

but now according to our assumption this number must remain constant, hence

$$v \cdot n \text{ independent of } s = \text{const}$$

$$n = \frac{\text{const.}}{v}$$

$$nds = \text{const}\frac{ds}{v} = \text{const} \cdot \tau$$

But up to a constant common to all line elements, the right-hand side, and thus also the left-hand side, is nothing else but the statistical probability of the element ds. Thus, the point density n is proportional to the probability of the element in question. Since $\sum W = 1$ for the whole orbit,

$W = \frac{nds}{n_{\text{tot}}}$, if n_{total} denotes the total number of the points distributed over the curve.

Doc. 5

On the Theory of Light Quanta and the Question of the Localization of Electromagnetic Energy

by A. Einstein

[*Archives des sciences physiques et naturelles* 29 (1910): 525–528]

What we understand by the "theory of light quanta" may be formulated in the following fashion: a radiation of frequency ν can be emitted or absorbed only in a well-defined quantum of magnitude $h\nu$[1] (and not in a smaller quantum). With the aid of this theory, several groups of hitherto unexplained phenomena can be considered from one and the same point of view. This is true of Stokes's law of phosphorescence and the principal laws for the emission of cathode radiation produced by visible and ultraviolet light (as well as by X rays). As a matter of fact, the kinetic energy L of photoelectrically [2] produced cathode radiation increases proportionally—or *at least approximately* proportionally—to the frequency of the exciting light according to the formula $L = c + h\nu$, where c is a negative constant that depends on the nature of the body under consideration. In general, it can be said that the theory of light quanta is the quantitative expression of the experimental fact that the energy of the molecular phenomena produced by light increases with the refrangibility of the light employed. [3]

It is now generally accepted that molecular mechanics, with the aid of the Maxwell-Lorentz equations, leads to the radiation formula $\rho = K\nu^2 T$, as it has been demonstrated in particular by Mssrs. Jeans and H. A. Lorentz. This formula is [4] contradicted by experiment and does not contain the constant h: this leads to the conclusion that the foundations of the theory must be modified in such a way that the constant h will play a role in it. Only in this way will it be possible to establish a theory of radiation and to understand the fundamental laws of radiation cited above. This modification of the foundations has not yet been possible to accomplish. The theoreticians have not yet even come to an agreement in regard to the following question: Can the light quanta be accounted for entirely by a characteristic of the emitting or absorbing substance, or should the electromagnetic radiation itself be assigned, besides a wave structure, a second kind of structure, such that the energy of the radiation itself is already divided in definite quanta? I believe that I have proven that [5]

---

[1] h is a universal constant that enters the radiation equations of Wien $\left(\rho = h\nu^3 e^{-\frac{h\nu}{RT}}\right)$ and Planck $\left(\rho = h\nu^3 \frac{1}{e^{\frac{h\nu}{RT}} - 1}\right)$. [1]

this latter view should be adopted.[2] The considerations on which I based myself rest on a principle of Boltzmann's, according to which the entropy S and the statistical probability W of a state of an isolated system are connected by the relation

$$S = \frac{R}{N}\log W,$$

where R is the gas constant for a perfect gas and N the number of molecules in one gram-molecule. If a complete molecular picture of the system considered is given, one can calculate the statistical probability W for each state of the system, and from this one can calculate S with the aid of the formula. If, conversely, the system is known thermodynamically, then one will know S, and from this one will be able to derive the statistical probability of each state of the system. To be sure, one cannot establish an elementary theory (e.g., a molecular theory) of the system from W in a unique and well-defined fashion; but, still, any theory giving the wrong values of W for any of the states can be considered unacceptable. One can then find the entropy of radiation in empty space by means of thermodynamics, using the law of black-body radiation, and solve the following problem: consider two spaces enclosed within impermeable walls and connected by a tube that can be closed; let V be the volume of one of the spaces and V_0 the total volume; assume that these spaces are filled with a radiation whose frequency lies between ν and $\nu + d\nu$, and whose total energy is E_0. We seek to calculate the entropy S of the system for every possible distribution of the energy E_0 between the two spaces. From the entropy S of each of these possible distributions, one can deduce the statistical probability corresponding to each of them. In this way one finds for a sufficiently dilute radiation the following expression for the probability that at a given moment all of the energy E_0 is contained in the volume V:

$$W = \left(\frac{V}{V_0}\right)^{\frac{E_0}{h\nu}}.$$

[7] It can easily be shown that this expression is not compatible with the principle of superposition. As regards the distribution between the two spaces, the radiation behaves as if its energy were localized in $E_0/h\nu$ points moving independently of each other. From this it follows—unless one wants to admit that the use of impermeable walls in these considerations is inadmissible—that, regarding the localization of its energy, the radiation must itself have a structure not given by the ordinary theory.

In conclusion, let me say that the commonly accepted localization of energy (just like the momentum in the electromagnetic field) is by no means a necessary consequence of the Maxwell-Lorentz equations. Furthermore, one can give, for example, a distribution compatible with the mentioned equations that, for static and stationary states, coincides completely with the one given by the old theory of action at a distance. [8]

---

[6] [2] A. Einstein, *Ann. d. Phys.*, 4, 17 (1905): 139ff.

Doc. 6

On the Ponderomotive Forces Acting on Ferromagnetic Conductors Carrying a Current in a Magnetic Field

by A. Einstein

[*Archives des sciences physiques et naturelles* 30 (1910): 323–324]

A current-carrying conductor in a magnetic field H is subjected to a ponderomotive force, the formula for which is

$$F = [iH], \tag{1}$$

where i is the vector of the current density, and the expression in the brackets is the vector product.

This formula is applicable, in particular, in the case where the body conducting the current is not magnetizable, i.e., where the magnetic induction B is equal to the intensity of the magnetic field H. If the conductor of the current is magnetizable and its magnetic state is consequently characterized by two vectors H and B that are different from each other, then one has to ask oneself which of these two vectors gives rise to the ponderomotive force sought.

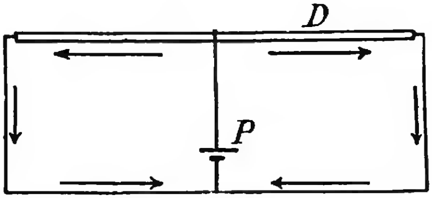
Till now this role has been attributed to B, and it has been accepted that

$$F = [iB]. \tag{2}$$

But we will show in a simple special case that, even in the case of a magnetic conductor, formula (1) is the correct one.

Let D be a metal disk through which a current flows from its center to its periphery. This current is provided by a battery P; the other lines in the diagram complete the circuit.

By virtue of the principle of equality of action and reaction, regardless of the material of which the disk is made, the resultant of all the electrodynamic forces acting on the different parts of the system is zero. This must be especially so if the disk D is made of a nonmagnetic substance ($B = H$).



Let us examine, in the second place, the case when the disk is made of a hard magnetic metal, e.g., steel, and constitutes a permanent magnet, with circular lines of force distributed around its center. In this case, the magnetic field produced by the passage of the current through the disk superposes on the magnetic field resulting from this magnetization of the disk. If we let H_{m} denote the strength of the latter field, and B_{m} its induction, then reasons of symmetry permit us to conclude from Maxwell's equations that

$$H_{\mathrm{m}} = 0,$$

but obviously B_{m} is not equal to zero.

On the other hand, the additional magnetization we have considered cannot give rise to a corresponding additional ponderomotive force, for the latter would be the only ponderomotive force that would appear, and the system would violate the law of equality of action and reaction.

Thus, the additional ponderomotive force vanishes together with H_{m}, even if B_{m} is different from zero. It follows that it is formula (1), and not formula (2), that satisfies the principle of equality of action and reaction.

Doc. 7

On a Theorem of the Probability Calculus and Its Application in the Theory of Radiation

by A. Einstein and L. Hopf [1]

[*Annalen der Physik* 33 (1910): 1096–1104]

§1. The Physical Problem as the Point of Departure

If one wants to calculate any effect of radiation in the theory of radiation, say the force acting on an oscillator, then one always uses Fourier series of the general form

$$\sum^{n} \left(A_n \sin 2\pi n \frac{t}{T} + B_n \cos 2\pi n \frac{t}{T}\right)$$ [2]

as the analytical expression for the electric or the magnetic force. The problem is here immediately specialized to a given spatial point, which is of no importance to what follows; t denotes the variable time, and T the very long time period for which the series applies. When calculating any average values—and, in general, only such values occur in the theory of radiation—one takes the individual coefficients A_n, B_n to be independent of each other, one assumes that each coefficient follows the Gaussian error law independently of the numerical values of the other coefficients, so that the probability[1] dW of a combination of values A_n, B_n must simply be the product of the probabilities of the individual coefficients.

(1) $$dW = W_{A_1} \cdot W_{A_2} \ldots W_{B_1} \cdot W_{B_2} \ldots dA_1 \ldots dB_1 \ldots$$

Since the theory of radiation, in the form in which it follows exactly from the generally accepted foundations of the theory of electricity and statistical mechanics leads, as we know, to irresolvable conflicts with experience, it is natural to mistrust this simple assumption of independence and to blame it for the failures of the theory of radiation. [4]

We shall show in what follows that this way out is impossible, and that, on the contrary, the physical problem can be reduced to a purely mathematical problem that leads to the statistical law (1).

That is to say that if we consider a radiation arriving from a certain direction,[2] then this radiation is certain to have a higher degree of order than the total radiation acting at a point. But the radiation arriving from a specific direction can always be conceived as arising from a great number of emission centers, i.e., the surface that emits the radiation can be subdivided into very many surface elements that emit radiation

---

[1] By "probability of a coefficient" we obviously have to understand the following: We imagine that the electrical force is expanded in Fourier series for very many moments of time. That fraction of these expansions in which a coefficient lies within a specified range of values is the probability of this range of values for the coefficient considered. [3]

[2] More accurately: "corresponding to a certain elementary angle $d\kappa$."

independently of each other; for since there is no limit to the distance of this surface from the test point, there is no limit to its total extension either. Into these radiation elements arising from the individual surface elements we again introduce a higher principle of order in that we conceive of all of these radiation elements as having the same form and differing only in their temporal phases; or, in mathematical terms: The coefficients of the Fourier series that represent the radiation of the individual surface elements shall be the same for all the surface elements, and only the initial times shall differ from element to element. If equation (1) can be proven on the basis of these [5] principles of order, then it will hold a fortiori in the case where these principles have been dropped. If the index s denotes the individual surface element, then the radiation emitted there will be of the form

$$\sum^{(n)} a_n \sin 2\pi n \frac{t - t_s}{T}.$$

Hence, the total radiation we are considering will be represented by the double sums

$$\sum^{s} \sum^{n} a_n \left(\sin 2\pi n \frac{t}{T} \cos 2\pi n \frac{t_s}{T} - \cos 2\pi n \frac{t}{T} \sin 2\pi n \frac{t_s}{T} \right). \tag{2}$$

Comparison of (2) and (1) leads thus to the expressions

$$\left\{ \begin{aligned} A_n &= a_n \sum^{s} \cos 2\pi n \frac{t_s}{T}, \\ -B_n &= a_n \sum^{n} \sin 2\pi n \frac{t_s}{T}, \end{aligned} \right. \tag{3}$$

where n is a very large number, and t_s can assume any value between 0 and T, so that the individual summands

$$\cos 2\pi n \frac{t_s}{T} \quad \text{and} \quad \sin 2\pi n \frac{t_s}{T}$$

are randomly distributed between -1 and $+1$, and are as likely to be positive as negative. If we can prove the general validity of our equation (1) for a combination of sums of such quantities, we will also have proved thereby the impossibility of introducing any order principle into the radiation propagated in empty space.

[6]

§2. Formulation of the General Mathematical Problem

We thus set ourselves the following mathematical problem: We are given a very large number of elements whose numerical values α (corresponding to t_s) follow a known statistical law. From each of these numerical values we build certain functions

$f_1(\alpha), f_2(\alpha)$ . . . (corresponding to $\sin 2\pi n \frac{t_s}{T_n}$, $\cos 2\pi n \frac{t_s}{T}$). We must subject these functions to an additional restriction: namely, from the probability that one of the quantities α lies between $\alpha + d\alpha$ there follows a statistical law for the f; let the probability $\varphi(f)df$ that f has a numerical value between f and $f + df$ be always such a function that the average value

$$\overline{f} = \int_{-\infty}^{+\infty} f\varphi(f)df = 0.$$

(It can easily be seen that our functions sin and cos indeed fulfill this assumption, because if every value of t_s between 0 and T is equally probable, the average values

$$\overline{\sin 2\pi n \frac{t_s}{T}}$$

and

$$\overline{\cos 2\pi n \frac{t_s}{T}}$$

vanish.)

We now assemble a (very large) number Z of such elements α into one system. To such a system belong certain sums

$$\sum_{(z)} f_1(\alpha), \quad \sum_{(z)} f_2(\alpha)\ldots$$

(corresponding to the coefficients A_n/a_n, B_n/a_n). We set ourselves the task of finding the statistical law that a combination of these sums obeys.

First we must be clear about a fundamental point.

The statistical law obeyed by the sums Σ themselves will not at all be independent of the number Z of the elements. This we can easily see in the simple special case when $f(\alpha)$ can assume only the values +1 and −1. Then we evidently have

$$\sum_{(Z+1)} = \sum_{(Z)} \pm 1$$

and

$$\overline{\sum^2_{(Z+1)}} = \overline{\sum^2_{(Z)}} + 1.$$

Thus, the mean square value of the sum increases proportionally to the number of elements. Hence, if we wish to arrive at a statistical law that is independent of Z, we must not consider the Σ but rather, since $\overline{\Sigma^2}/Z$ stays constant, the quantities

$$S = \frac{\Sigma}{\sqrt{Z}}.$$

§ 3. The Statistical Law of the Individual S

Before investigating a combination of all quantities

$$S^{(n)} = \frac{\Sigma_{(Z)} f_n(\alpha)}{\sqrt{Z}},$$

we will formulate the probability law for one single such quantity.

We consider a manifold of N-systems of the kind defined above. To each system belongs a numerical value S. Because of the statistical distribution of the α, these quantities obey a specific probability law, so that the number of systems whose numerical values lies between S and $S + dS$:

$$dN = F(S)dS. \tag{4}$$

[7] If we now add one more element to each system consisting of Z elements, i.e., if we pass from S_Z to S_{Z+1}, the individual members of our manifold will change their numerical values and will enter into another region dS. If it is to be possible, nevertheless, to arrive at a statistical law that is independent of Z, then the number dN must not change in this transition. Thus, the number of systems entering a given (in our simplest case, one-dimensional) region dS must be the same as the number of systems leaving it. If Φ denotes the number of systems passing through a given numerical value S_0 in their transition from Z to $Z + 1$ elements, both as regards magnitude and direction, then we must have

$$\operatorname{div} \Phi = 0, \tag{5}$$

hence

$$\frac{d\Phi}{dS} = 0$$

and, since, indeed, Φ must always be zero for $S = \infty$, we will also have

$$\Phi = 0. \tag{6}$$

Now we have

$$S_{(Z+1)} = \frac{\Sigma_{(Z+1)} f(\alpha)}{\sqrt{Z+1}} = S_{(Z)} \sqrt{\frac{Z}{Z+1}} + \frac{f(\alpha)}{\sqrt{Z+1}},$$

or, since Z is to be a very large number,

$$S_{(Z+1)} = S_{(Z)} - \frac{S_{(Z)}}{2Z} + \frac{f(\alpha)}{\sqrt{Z}}. \tag{7}$$

Thus, the number Φ is composed of two parts: a Φ_1, which comes from the summand $-S/2Z$, and a Φ_2, which comes from $f(\alpha)/\sqrt{Z}$.

Φ_1 contains all those S that have been at a positive distance $\leqq S_0/2Z$ from the value S_0; and, to be sure, these members cross S_0 in the negative direction. Since $S_0/2Z$ is a very small number, the number of these members is, up to infinitely small quantities of higher order,

$$\Phi_1 = -\frac{S_0}{2Z}F(S_0). \tag{8}$$

Contributions to the number Φ_2 come from every arbitrary positive and negative distance Δ from S_0; indeed, the contributions are positive or negative, depending on whether Δ is negative or positive. The number dN at the distance Δ is given by

$$F(S_0 + \Delta)dS = F(S_0 + \Delta)d\Delta,$$

or, since only small values of Δ are, after all, of importance, by

$$\left\{F(S_o) + \Delta\left(\frac{dF}{d\Delta}\right)_{S_0}\right\}d\Delta.$$

Of this number, all those cross the value S_0 in the positive direction which, coming from a negative Δ, have a $f(\alpha)$ so large that

$$\frac{f(\alpha)}{\sqrt{Z}} \geqq |\Delta|,$$

and thus the number will be

$$\int\limits_{-\Delta\sqrt{Z}}^{+\infty} \varphi(f)df.$$

Analogously, the number going in the negative direction will be

$$\int\limits_{-\infty}^{-\Delta\sqrt{Z}} \varphi(f)df.$$

We will then have

$$\Phi_2 = \int\limits_{-\infty}^{0} d\Delta \left\{ F(S_0) + \Delta \left(\frac{dF}{d\Delta} \right)_{S_0} \right\} \int\limits_{-\Delta\sqrt{Z}}^{+\infty} \varphi(f)df$$

$$- \int\limits_{0}^{\infty} d\Delta \left\{ F(S_0) + \Delta \left(\frac{dF}{d\Delta} \right)_{S_0} \right\} \int\limits_{-\infty}^{-\Delta\sqrt{Z}} \varphi(f)df$$

Integrated by parts, this becomes

$$\Phi_2 = - \int\limits_{-\infty}^{0} d\Delta \left\{ \Delta . F(S_0) + \frac{\Delta^2}{2} \left(\frac{dF}{d\Delta} \right)_{S_0} \right\} \varphi\left(-\Delta\sqrt{Z}\right) . \sqrt{Z}$$

$$- \int\limits_{0}^{\infty} d\Delta \left\{ \Delta F(S_0) + \frac{\Delta^2}{2} \left(\frac{dF}{d\Delta} \right)_{S_0} \right\} \varphi\left(-\Delta\sqrt{Z}\right) . \sqrt{Z} .$$

Now, since by the assumption,

$$\int\limits_{-\infty}^{+\infty} f\varphi(f)df = 0,$$

we get, if we introduce $\Delta\sqrt{Z} = f$ as a variable,

$$(9) \qquad \left\{ \begin{aligned} \Phi_2 &= - \frac{1}{2Z} \left(\frac{dF}{d\Delta} \right)_{S_0} \int\limits_{-\infty}^{+\infty} f^2 \varphi(f)df \\ &= - \frac{1}{2Z} \left(\frac{dF}{d\Delta} \right)_{S_0} \cdot \overline{f^2} . \end{aligned} \right.$$

(8) and (9) inserted in (6) yield the differential equation

$$SF + \overline{f^2} \frac{dF}{ds} = 0,$$

the solution of which,

$$(10) \qquad F = \text{const.} e^{-\frac{s^2}{2\overline{f^2}}},$$

expresses the Gaussian law of errors.

§ 4. The Statistical Law for a Combination of All $S^{(n)}$

Let us now extend the considerations of the preceding section from the one-dimensional case to the case of arbitrarily many dimensions. This time we have to consider a combination of many quantities $S^{(n)}$. Let the number of systems in an infinitesimally small region

(11) $dS^{(1)}dS^{(2)}\ldots$ be $dN = F(S^{(1)}, S^{(2)}\ldots)dS^{(1)}dS^{(2)}\ldots$

We require again that the dN shall not change when we pass from $S^{(n)}_{(Z)}$ to $S^{(n)}_{(Z+1)}$, and this leads again to the differential equation (5),

$$\operatorname{div} \Phi = 0.$$

In the present case, however, the number Φ has components in every direction $S^{(1)}, S^{(2)}\ldots$, and we will denote them by $\Phi^{(1)}, \Phi^{(2)}\ldots$ Hence, (5) assumes the form

$$\sum_n \frac{\partial \Phi^{(n)}}{\partial S^{(n)}} = 0.$$

$S^{(n)}_{(Z)}$ and $S^{(n)}_{(Z+1)}$ are related by equation (7) as before, and the arguments of the preceding section are therefore fully applicable to the calculation of the individual $\Phi^{(n)}$. Consequently, we have

$$\Phi^{(n)} = S^{(n)}F + \overline{f_n^2}\frac{\partial F}{\partial S^{(n)}}.$$

We can simplify this expression further by assuming that all $\overline{f_n^2}$ are equal. All that this obviously amounts to is thinking of the individual functions f_n as multiplied by appropriate constants. (In the special case of our sin and cos this simplification is automatically satisfied.)

In this way we finally obtain the following differential equation for the function F:

$$\sum_n \frac{\partial}{\partial S^{(n)}}\left(S^{(n)}F + \overline{f^2}\,\frac{\partial F}{\partial S^{(n)}}\right) = 0. \tag{12}$$

We arrive at the solution of this differential equation by considering the following integral extended over the whole space:

$$(13)\quad \begin{cases} \displaystyle\int \frac{1}{F}\sum_{0}^{n_1}{}_{n} \left\{\left(S^{(n)}F + \overline{f^2}\frac{\partial F}{\partial S^{(n)}}\right)^2\right\} dS^{(1)}\ldots dS^{(n_1)} \\[2ex] \displaystyle = \int \sum_{0}^{n_1}{}_{n} \left\{\left(S^{(n)}F + \overline{f^2}\frac{\partial F}{\partial S^{(n)}}\right)\left(S^{(n)} + \overline{f^2}\frac{\partial \log F}{\partial S^{(n)}}\right)\right\} dS^{(1)}\ldots (n_1). \end{cases}$$

But

$$\int \sum_{0}^{n_1}{}_{n} \left\{\left(S^{(n)}F + \overline{f^2}\frac{\partial F}{\partial S^{(n)}}\right) S^{(n)}\right\} dS^{(1)}\ldots dS^{(n_1)}$$

$$= \int \left(F \sum_{0}^{n_1}{}_{n}\, S^{(n)2} + \overline{f^2} \sum_{0}^{n_1}{}_{n}\, S^{(n)} \frac{\partial F}{\partial S^{(n)}}\right) dS^{(1)}\ldots dS^{(n_1)},$$

or, if we integrate by parts the second summand and take into account that at infinity we must have $F = 0$,

$$= \int F \left(\sum_{0}^{n_1}{}_{n}\, S^{(n)2} - \overline{f^2}\cdot n_1\right) dS^{(1)}\ldots dS^{(n_1)}.$$

However, this expression vanishes because

$$\int F S^{(n)2} dS^{(1)}\ldots dS^{(n_1)}$$

is nothing else but the average value $\overline{S^{(n)2}}$, which we derived in the last section, for the case when only a single S is being considered; for the latter it follows from equation (10) that

$$\overline{S^2} = \overline{f^2}.$$

On the other hand, integration by parts yields

$$\int \sum \left\{ \left(S^{(n)}F + \overline{f^2}\frac{\partial F}{\partial S^{(n)}} \right) \overline{f^2}\frac{\partial \log F}{\partial S^{(n)}} \right\} dS^{(1)} \ldots dS^{(n_1)}$$

$$= \int \overline{f^2} \log F \sum \left\{ \frac{\partial}{\partial S^{(n)}} \left(S^{(n)}F + \overline{f^2}\frac{\partial F}{\partial S^{(n)}} \right) \right\} dS^{(1)} \ldots dS^{(n_1)}$$ [8]

which, according to (12), also vanishes.

This proves that integral (13) vanishes; however, because of the quadratic character of the integrand, this is possible only if we have everywhere for every n

(14) $$S^{(n)}F + \overline{f^2}\frac{\partial F}{\partial S^{(n)}} = 0.$$

Thus, we arrive at a statistical law for F that is identical with Gauss's law of errors with respect to every $S^{(n)}$:

(15) $$F = \mathrm{const.}e^{-\frac{S^{(1)^2}}{2\overline{f^2}}}.e^{-\frac{S^{(2)^2}}{2\overline{f^2}}}.$$ [9]

The probability of a combination of values $S^{(n)}$ is thus simply the product of the probabilities of the individual $S^{(n)}$.

It is clear that, if equation (15) holds for $S^{(1)}, S^{(2)} \ldots$, then the same equation is satisfied for a combination of quantities

$$S^{(n)'} = \alpha_n S^{(n)}.$$ [10]

In that case, instead of $\overline{f^2}$, we have the quantities $\alpha_n^2\overline{f^2}$ in the exponents. But the coefficients A_n, B_n in our physical problem are of the type $S^{(n)'}$; and, indeed, we have to set

$$S^{(n)} = \frac{A_n}{a_n\sqrt{Z}},$$

and hence

$$\alpha_n = a_n\sqrt{Z}.$$

Therewith is also proved the validity of equation (1) and the impossibility of constructing a probability-theoretical relation between the coefficients of the Fourier series that describes the thermal radiation.

(Received on 29 August 1910)

Doc. 8

Statistical Investigation of a Resonator's Motion in a Radiation Field

by A. Einstein and L. Hopf

[*Annalen der Physik* 33 (1910): 1105–1115]

§1. Train of Thought

It has already been shown in a variety of ways and it is now generally accepted that, when correctly applied in the theory of radiation, our current views on the distribution and propagation of electromagnetic energy on the one hand, and on the statistical distribution [1] of energy on the other hand, can lead to no other but the so-called Rayleigh (Jeans) radiation law. Since this law is in complete contradiction with experiment, it is necessary to undertake a modification of the foundations of the theories used for its derivation; and it has often been suspected that the application of the statistical energy distribution laws [2] to radiation or to rapidly oscillating motions (resonators) is not flawless. The following investigation shall now show that such a dubious application is not required at all, and that it suffices to apply the equipartition theorem for energy solely to the *translational* motion of the molecules and oscillators in order to arrive at the Rayleigh radiation law. The applicability of the law to translatory motion has been adequately proved by the successes of the kinetic theory of gases; we may therefore conclude that only a more radical and more profound change in our fundamental conceptions can lead to a law of [3] radiation that is in better agreement with experiment.

We consider a mobile electromagnetic oscillator[1] that is, on the one hand, subjected to the effects of a radiation field and, on the other hand, possesses a mass m and enters into interaction with the molecules present in the radiation-filled space. If the above interaction were the only one present, then the mean square value of the momentum associated with the oscillator's translatory motion would be completely determined by statistical mechanics. In our case there also exists the interaction of the oscillator with the radiation field. For a statistical equilibrium to be possible, this latter interaction must not produce any change in that mean value. In other words: The mean square value of the momentum associated with the translatory motion that the oscillator assumes under the influence of *the radiation alone* must be the same as that which it would assume, in accordance with statistical mechanics, under the mechanical influence of the molecules alone. This reduces the problem to the task of determining the mean square value $\overline{(mv)^2}$ of the momentum assumed by the oscillator under the sole influence of the radiation [4] field.

This mean value must be the same at time $t = 0$ as at time $t = \tau$, so that we have

---

[1] For the sake of simplicity we will assume that the oscillator oscillates only in the z-direction and moves only in the x-direction.

$$\overline{(mv)^2_{t=0}} = \overline{(mv)^2_{t=\tau}}.$$

For what follows, it is expedient to distinguish two kinds of dynamical effects through which the radiation field influences the oscillator, namely

1. The resistive force K, with which the radiation pressure opposes the rectilinear motion of the oscillator. Neglecting the terms of the order of magnitude of $(v/c)^2$ (c = velocity of light), this is proportional to the velocity v, and we can therefore write: $K = -Pv$. If we further assume that the velocity v does not change markedly during time τ, then the momentum deriving from this force $= -Pv\tau$.

2. The fluctuations Δ of the electromagnetic momentum that arise in the disordered radiation field owing to the motion of the electric masses. These can be positive just as well as negative, and are independent—in first approximation—of the circumstance that the oscillator is in motion.

These momenta superpose themselves on the momentum $(mv)_{t=0}$ during time τ, and our equation becomes

$$\overline{(mv)^2_{t=0}} = \overline{(mv_{t=0} + \Delta - Pv\tau)^2}. \tag{1}$$

By increasing the mass m, we can always bring it about that the term multiplied by τ^2, which appears on the right-hand side of equation (1), can be neglected. Further, the term multiplied by $\overline{v\Delta}$ vanishes because v and Δ can become positive or negative quite independently of each other. If, in addition, we replace $\overline{mv^2}$ by the temperature Θ using the equation known from the theory of gases

$$\overline{mv^2} = \frac{R}{N}\Theta$$

(R = the absolute gas constant, N = Loschmidt's number), then equation (1) assumes the form

$$\overline{\Delta^2} = 2\frac{R}{N}P\Theta\tau. \tag{2}$$

Thus, we only have to find Δ^2 and P (or $\overline{K}$) by means of electromagnetic arguments, and equation (2) will yield the radiation law.

§2. *Calculation of the Force* $\overline{K}$.[2]

To calculate the force with which the radiation opposes a moving oscillator, we calculate first the force on an oscillator at rest, and then transform this force by means of the formulas that follow from the theory of relativity.

Let the oscillator with proper frequency ν_0 freely oscillate in the z-direction of an orthogonal coordinate system x, y, z. If $\mathfrak{E}$ and $\mathfrak{h}$ denote the electric and the magnetic

---

[2] Cf. M. Abraham, *Ann. d. Phys.* 14 (1904): 273 ff. [5]

force of the external field, then, according to Planck,[3] the momentum f of the oscillator obeys the differential equation

$$(3) \qquad 16\pi^4\nu_0^3 f + 4\pi^2\nu_0 \ddot{f} - 2\sigma \dddot{f} = 3\sigma c^3 \mathfrak{E}_z.$$

Here σ denotes a constant characteristic of the damping of the oscillator through radiation emission.

Now let a plane wave impinge on the oscillator; let the ray form the angle φ with the z-axis, and its projection on the xy-plane the angle ω with the x-axis. If we decompose this wave into two waves polarized perpendicularly to each other, with the electric force of the first lying in the ray-oscillator plane, and that of the second perpendicular to the first, then it is obvious that only the first one imparts a certain momentum to the oscillator. If we write down the electric force of this first wave as a Fourier series

[7] $$(4) \qquad \mathfrak{E} = \sum_n A_n \cos\left\{\frac{2\pi n}{T}\left(t - \frac{\alpha x + \beta y + \gamma z}{c}\right) - \vartheta_n\right\},$$

where T denotes a very long time period, then the direction cosines α, β, γ of the ray are expressed through φ and ω in the following way:

$$\alpha = \sin\varphi \cos\omega, \quad \beta = \sin\varphi \sin\omega, \quad \gamma = \cos\varphi,$$

and the components of the electric and magnetic forces that are of relevance to our further calculations are

$$(5) \qquad \begin{cases} \mathfrak{E}_x = \mathfrak{E}\cos\varphi \cos\omega, \\ \mathfrak{E}_z = -\mathfrak{E}\sin\varphi, \\ \mathfrak{H}_y = \mathfrak{E}\cos\varphi \sin\omega \end{cases}$$ [8]

The ponderomotive force exerted on the oscillator is

[9] $$k = f\frac{\partial \mathfrak{E}}{\partial z} + \frac{1}{c}\left[\frac{df}{dt}\mathfrak{H}\right].$$

For this equation as well as equation (3) to be valid, it must be assumed that the dimensions of the oscillator are always small compared with the relevant wavelengths of the radiation. The x-component k_x of the ponderomotive force is

$$(6) \qquad k_x = \frac{\partial \mathfrak{E}_x}{\partial z} f - \frac{1}{c}\mathfrak{H}_y \frac{df}{dt}.$$

---

[6] [3] M. Planck, *Vorl. über die Theorie der Wärmestrahlung*, p. 113.

If we solve (3)[4] and take into account (4) and (5), we obtain

$$f = -\frac{3c^3}{16\pi^3}T^3\sin\varphi\sum^n A_n\frac{\sin\gamma_n}{n^3}\cos(\tau_n - \gamma_n),$$

$$\dot{f} = \frac{3c^3}{8\pi^2}T^3\sin\varphi\sum^n A_n\frac{\sin\gamma_n}{n^2}\sin(\tau_n - \gamma_n),$$

where

$$\tau_n = 2\pi n\frac{t}{T} - \vartheta_n,$$

has been assumed for the sake of brevity, and γ_n is given by the equation

$$\operatorname{cotg}\gamma_n = \frac{\pi\nu_0\left(\nu_0^2 - \frac{n^2}{T^2}\right)}{\sigma\frac{n^3}{T^3}}.$$

Since, further,

$$\frac{\partial\mathfrak{E}_x}{\partial z} = \frac{2\pi}{ct}\cos^2\varphi\cos\omega\sum^n nA_n\sin^2\tau_n,$$ [5] [11]

k_x appears as the double sum

$$k_x = -\frac{3c^2}{8\pi}T^2\cos^2\varphi\sin\varphi\cos\omega\sum^n\sum^m A_n\frac{\sin\gamma_n}{n^3}A_m\, m\cos(\tau_n - \gamma_n)\sin\tau_m$$

$$-\frac{3c^2}{8\pi}T^2\sin\varphi\cos\omega\sum^n\sum^m A_n\frac{\sin\gamma_n}{n^2}A_m\sin(\tau_n - \gamma_n)\cos\tau_m.$$ [12]

Because the phase angles ϑ are independent of each other, only the terms $n = m$ need be considered in forming the average value,[6] and we get[7]

---

[4] M. Planck, loc. cit., p. 114. [10]

[5] In fact, this expression for $\partial\mathfrak{E}_x/\partial_z$ as well as the one for $\mathfrak{H}_y$, would have to be supplemented by the components of the wave that is polarized perpendicularly to the wave that excites the oscillator; however, it is obvious that these expressions do not contribute anything to the mean value of the force because their phases are independent of those of the oscillator.

[6] This independence follows from the result of the preceding paper. [13]

[7] M. Planck, l.c., p. 122. [15]

[14] (7)
$$\left\{\begin{aligned}\overline{K_x} &= \frac{3c^2}{16\pi^2}T^2\sin^3\varphi\cos\omega\sum_n A_n^2\frac{\sin\gamma_n}{n^2}\\ &= \frac{3c^2}{16\pi^2}\overline{A^2_{\nu_0,T}}\,T\frac{\sigma}{2\nu_0}\sin^3\varphi\cos\omega.\end{aligned}\right.$$

This is the mean value of the x-component of the force that a wave incident in the φ,ω direction exerts on the oscillator at rest.

If the oscillator is moving in the x-direction with the velocity v, then it is more practical to replace the angles φ,ω with the angle φ_1 between the ray and the x-axis, and the angle ω_1 between the projection of the ray on the yz-plane and the y-axis. The following relations then hold:

$$\cos\varphi_1 = \sin\varphi\cos\omega,$$

$$\sin\varphi_1\cos\omega_1 = \sin\varphi\sin\omega,$$

$$\sin\varphi_1\sin\omega_1 = \cos\varphi.$$

We are led to the value of the force $\overline{K'_x}$ acting on the moving oscillator by the transformation formulas of the theory of relativity[8]

$$A' = A\left(1 - \frac{v}{c}\cos\varphi_1\right),$$

$$T' = T\left(1 + \frac{v}{c}\cos\varphi_1\right),$$

$$\nu' = \nu\left(1 - \frac{v}{c}\cos\varphi_1\right),$$

$$\cos\varphi_1' = \frac{\cos\varphi_1 - \frac{v}{c}}{1 - \frac{v}{c}\cos\varphi_1}, \qquad \omega_1' = \omega_1.$$

We then get

$$\overline{K'_x} = \frac{3c^2}{16\pi^2}\overline{A'^2_{\nu_0'T'}}\cdot T'\frac{\sigma}{2\nu_0'}(1 - \sin^2\varphi_1'\sin^2\omega_1')\cos\varphi_1'.$$

Neglecting the terms with $(v/c)^2$, we get

---

[16] [8] A. Einstein, *Ann. d. Phys.* 17 (1905): 914.

$$\overline{A'^{2}_{\nu_0'T'}} = \overline{A^{2}_{\nu_0 T}}\left(1 - 2\frac{v}{c}\cos\varphi_1\right),$$

or, since we have to relate everything to the proper frequency ν_0' of the moving oscillator

$$\overline{A'^{2}_{\nu_0'T'}} = \overline{A^{2}_{\nu_0'\left(1 + \frac{v}{c}\cos\varphi_1\right)T}}\left(1 - 2\frac{v}{c}\cos\varphi_1\right)$$

$$= \left\{\overline{A^{2}_{\nu_0'T}} + \nu_0'\frac{v}{c}\cos\varphi_1\left(\frac{d\overline{A^2}}{d\nu}\right)_{\nu_0 T}\right\}\cdot\left(1 - 2\frac{v}{c}\cos\varphi_1\right).$$

Furthermore, we express the quantity $\overline{A^2 T}$ in terms of the mean radiation density ρ. We set the mean energy of a plane wave coming from a given direction equal to the energy density in a cone with a solid angle $d\kappa$. If, in addition, we also keep in mind that the magnetic and the electric forces are equal, and take into account the two planes of polarization, we arrive at the relation

$$\rho\frac{d\kappa}{4\pi} = \frac{1}{8\pi}\frac{\overline{A^2 T}}{2}\cdot 2\cdot 2.$$

Our expression for the force becomes

$$(8)\qquad \left\{\begin{aligned} \overline{k_x^T} = \frac{3c^2}{16\pi^2}\cdot\frac{\sigma}{2\nu_0'}\left\{\rho_{\nu_0'} + \nu_0'\,\frac{v}{c}\cos\varphi_1\left(\frac{d\rho}{d\nu}\right)_{\nu_0'}\right\}\left(\cos\varphi_1 - \frac{v}{c}\right)\\ \left(1 - \frac{\sin^2\varphi_1}{1 - 2\frac{v}{c}\cos\varphi_1}\sin^2\omega_1\right)d\kappa. \end{aligned}\right.$$

Finally, integrating also over all solid angles, we obtain the total force that we were seeking:

$$(9)\qquad \bar{K} = -\frac{3c\sigma}{10\pi\nu_0'}v\left\{\rho_{\nu_0'} - \frac{\nu_0'}{3}\left(\frac{d\rho}{d\nu}\right)_{\nu_0'}\right\}.$$

§3. Calculation of the Fluctuations of Momentum $\overline{\Delta^2}$

The calculation of the momentum fluctuations can be made considerably more simple than the calculation of the force because a transformation according to the theory of relativity is not necessary.[9] It suffices to expand the electric and magnetic forces at the origin in a Fourier series, dependent on time alone, as long as one can prove that the individual components of force appearing in this expression are independent of each other.

The momentum that the oscillator experiences in the x-direction during time τ is

$$J = \int_0^\tau k_x dt = \int_0^\tau \left(\frac{\partial \mathfrak{E}_x}{\partial z} f - \frac{1}{c} \mathfrak{H}_y \frac{df}{dt} \right) dt .$$

Integration by parts yields

$$\int_0^\tau \mathfrak{H}_y \frac{df}{dt} dt = \left[\mathfrak{H}_y f \right]_0^\tau - \int_0^\tau \frac{\partial \mathfrak{H}_y}{\partial t} f dt.$$

The first summand vanishes if τ is appropriately chosen, i.e., if τ is large enough. If, in accordance with Maxwell's equation, one also puts

$$\frac{1}{c} \frac{\partial \mathfrak{H}_y}{\partial t} = \frac{\partial \mathfrak{E}_z}{\partial x} - \frac{\partial \mathfrak{E}_x}{\partial z},$$

one arrives at the simple expression

$$J = \int_0^\tau \frac{\partial \mathfrak{E}_z}{\partial x} f dt. \tag{10}$$

Now only the component E_z and its derivative $\partial \mathfrak{E}_z / \partial x$ appear in our expression. However, their independence can easily be proved. For if we just consider two wave trains (with identical solid angles) approaching each other, we can write

$$\begin{aligned} E_z = \sum \Big\{ & a_n \sin \frac{2\pi n}{T} \left(t - \frac{\alpha x + \beta y + \gamma z}{c} \right) \\ & + b_n \cos \frac{2\pi n}{T} \left(t - \frac{\alpha x + \beta y + \gamma z}{c} \right) \\ & + a_n' \sin \frac{2\pi n}{T} \left(t + \frac{\alpha x + \beta y + \gamma z}{c} \right) \\ & + b_n' \cos \frac{2\pi n}{T} \left(t + \frac{\alpha x + \beta y + \gamma z}{c} \right) \Big\} \end{aligned}$$

and

---

[9] Because the momenta with variable signs, which result from the irregularities of the radiation process, can be determined for a resonator at rest.

$$\frac{\partial \mathfrak{E}_z}{\partial x} = \sum \left\{ \frac{2\pi n\alpha}{Tc} \left[-a_n \cos\frac{2\pi n}{T}(\ldots) + b_n \sin\frac{2\pi n}{T}(\ldots) \right.\right.$$

$$\left.\left. + a_n' \cos\frac{2\pi n}{T}(\ldots) - b_n' \sin\frac{2\pi n}{T}(\ldots) \right]\right\}.$$

But the quantities $a_n + a_n'$, $a_n - a_n' \ldots$ are mutually independent and of the same type as the quantities denoted by S in the preceding paper; for such quantities it has been [17] shown that the probability for a combination is represented as the product of the Gaussian error functions of the individual quantities. It is easy to see from what has been said that no probability relation of any kind can exist between the coefficients of the expansions of $\mathfrak{E}_x$ and $\partial\mathfrak{E}_z/\partial x$. Now we write down $\mathfrak{E}_z$ and $\partial\mathfrak{E}_z/\partial x$ in the form of Fourier series:

$$\mathfrak{E}_x = \sum^m B_n \cos\left(2\pi n \frac{t}{T} - \vartheta_n\right),$$

$$\frac{\partial \mathfrak{E}}{\partial x_z} = \sum^n C_m \cos\left(2\pi m \frac{t}{T} - \xi_m\right). \qquad [18]$$

We then get

$$f = \frac{3c^3}{16\pi^3} T^3 \sum^n B_n \frac{\sin\gamma_n}{n^3} \cos\left(2\pi n \frac{t}{T} - \vartheta_n - \gamma_n\right)$$

and

$$J = \frac{3c^3}{16\pi^3} T^3 \int_0^\tau dt \sum^m \sum^n C_m B_n \frac{\sin\gamma_n}{n^3} \left[\cos\left\{2\pi(n+m)\frac{t}{T} - \xi_m - \vartheta_n - \gamma_n\right\}\right. \qquad [19]$$

$$\left. - \cos\{2\pi(n-m)t + \xi_m - \vartheta_n - \gamma_n\}\right].$$

Integration over t yields two summands with the factors $1/n+m$ and $1/n-m$; since n and m are very large numbers, the first summand is very small and can be neglected. Thus one arrives at the expression

$$(11) \qquad J = -\frac{3c^3}{32\pi^4} T^4 \sum^m \sum^n C_m B_n \frac{\sin\gamma_n}{n^3} \frac{1}{n-m} \cos\delta_{mn} \sin\pi(n-m)\frac{\tau}{T},$$

where, for brevity,

$$\delta_{mn} = \pi(n-m)\frac{\tau}{T} + \xi_m - \vartheta_n - \gamma_n.$$

J^2 then appears as the fourfold sum over n, m and two further variables n' and m'. If we calculate the mean value $\overline{J^2}$, we must take into account the fact that the angles δ_{mn} and $\delta_{m'n'}$ are totally independent of each other, and, thus, that only those terms in which this independence does not occur are to be considered in the averaging. Obviously, this is only the case if

$$m = m' \quad \text{and} \quad n = n',$$

and we arrive at the desired mean value:

[20]

$$\overline{J^2} = \left(\frac{3c^3T^4}{32\pi^4}\right)^2 \sum^m \sum^n \frac{1}{2} C_m^2 B_n^2 \left(\frac{\sin\gamma_n}{n^3}\right)^2 \frac{1}{(n-m)^2} \sin^2\pi(n-m)\frac{t}{T}.$$

Since

$$\sum^m \frac{1}{(n-m)^2} \sin^2\pi(n-m)\frac{t}{T} = \frac{1}{T}\int_0^\infty \frac{1}{(\nu-\mu)^2} \sin^2(\nu-\mu)\pi\tau \cdot d\mu = \frac{\pi^2\tau}{T}$$

and

[21]

$$\sum^n \frac{\sin^2\gamma_n}{n^6} = \frac{1}{T^5}\int_0^\infty \frac{\sin^2\gamma_n}{\nu^6} d\nu = \frac{1}{T^5}\frac{\sigma}{2\nu_0^5},$$

$\overline{J^2}$ becomes

(12)

$$\overline{J^2} = \left(\frac{3c^3}{32\pi^3}\right)^2 \frac{\sigma\tau}{4\nu_0^5} \overline{B_{\nu_0 T}^2}\ \overline{C_{\nu_0 T}^2} T^2.$$

Now

$$\overline{J^2} = \overline{(\bar{J} + \Delta)^2} = \overline{J^2} + 2\bar{J}\bar{\Delta} + \overline{\Delta^2},$$

and since the mean values $\bar{J}$ and $\bar{\Delta}$ vanish, expression (12) itself gives the value of the momentum fluctuations $\overline{\Delta^2}$. It only remains to express the mean values of the amplitudes $B_{\nu_0 T}^2$ and $C_{\nu_0 T}^2$ through the radiation density $\rho_{\nu_0'}$.

To that end we must again consider the radiation coming from different directions and, as above, relate the amplitude of the radiation coming from a specific direction to the energy density by means of the equation

$$\overline{A_{\nu_0 T}^2} T = \rho_{\nu_0} d\kappa.$$

The amplitude

$$B_{\nu_0 T} = \sum A_{\nu_0 T} \sin\varphi$$

over all angles of incidence, hence

$$\overline{B^2_{\nu_0 T}} \cdot T = \overline{A^2_{\nu_0 T}} \cdot T \sum \sin^2 \varphi = \frac{8}{3} \pi \rho_{\nu_0}. \tag{13}$$

In the same way we get

$$\overline{C^2_{\nu_0 T}} T = \left(\frac{2\pi\nu}{c}\right)^2 \overline{A^2_{\nu_0 T}} T \sum \sin^4 \varphi \cos^2 \omega = \frac{64}{15} \frac{\pi^3 \nu_0^2}{c^2} \rho_{\nu_0}. \tag{14}$$

By inserting (13) and (14) into (12) we finally obtain

$$\overline{\Delta^2} = \frac{c^4 \sigma \tau}{40 \pi^2 \nu_0^3} \rho^2_{\nu_0}. \tag{15}$$

§5. *The Radiation Law*

Now we only have to insert the values (9) and (15) found above into our equation (2), and we arrive at the differential equation containing the radiation law:

$$\frac{c^3 N}{24 \pi R \theta \nu^2} \rho^2 = \rho - \frac{\nu}{3} \frac{d\rho}{d\nu},$$

which yields by integration

$$\rho = \frac{8 \pi R \theta \nu^2}{c^3 N}. \tag{16}$$

This is the well-known Rayleigh radiation law, which is in the most glaring contradiction with experiment. The foundations of our derivation must therefore contain an assertion that does not agree with what really takes place in thermal radiation.

Let us therefore place these foundations under closer critical scrutiny.

One has wanted to find the reason why all exact statistical analyses in the field of radiation theory lead to Rayleigh's law in the application of this approach to the radiation itself. With some justification, Planck[10] brings up this argument against Jeans's derivation. However, in the above derivation there is no question whatsoever of a somehow arbitrary transference of statistical considerations to radiation; the energy equipartition theorem was applied only to the translatory motion of oscillators. But the successes of the kinetic theory of gases demonstrate that this law can be considered as thoroughly proved for translatory motion.

The theoretical foundation we used in our derivation, which is certain to contain an unfounded assumption, is thus nothing else but that underlying the theory of light dispersion in completely transparent bodies. The actual phenomena differ from the results deducible from this foundation owing to the fact that additional kinds of

---

[10] M. Planck, l.c., p. 178.

[22]

momentum fluctuations are discernible in the former, which, in the case of short-wave radiation of low density, enormously overwhelm those obtained from the theory.[11]

Zurich, August 1910. (Received on 29 August 1910)

---

[23]

[11] Cf. A. Einstein, *Phys. Zeitschr.* 10 (1909): 135 ff. What is essentially new in the present paper is the circumstance that the momentum fluctuations have been exactly calculated for the first time.

Doc. 9

The Theory of the Opalescence of Homogeneous Fluids and Liquid Mixtures near the Critical State

by A. Einstein.

[*Annalen der Physik* 33 (1910): 1275–1298]

In an important theoretical paper,[1] Smoluchowski has shown that the opalescence of fluids near the critical state as well as the opalescence of liquid mixtures near the critical mixing ratio and the critical temperature can be explained in a simple way from the point of view of the molecular theory of heat. This explanation is based on the following [2] general implication of Boltzmann's entropy-probability principle: In the course of an infinitely long period of time, an externally closed system passes through all the states that are compatible with the (constant) value of its energy. However, the statistical probability of a state is noticeably different from zero only when the work that would have to be expended according to thermodynamics to produce the state in question from the state of ideal thermodynamic equilibrium is of the same order of magnitude as the kinetic energy of a monatomic gas molecule at the temperature under consideration. [3]

If such a small amount of work suffices to bring about, in volumes of fluid of the order of magnitude of the cube of a wavelength, a density that deviates markedly from the average density of the fluid or a mixing ratio that deviates markedly from the average, then, obviously, the phenomenon of opalescence (the Tyndall phenomenon) must take place. Smoluchowski has shown that this condition is actually fulfilled near the critical [4] state; however, he did not provide an exact calculation of the quantity of light given off laterally through opalescence. This gap shall be filled in the following.

§1. General Remarks about the Boltzmann Principle [5]

Boltzmann's principle can be expressed by the equation

$$S = \frac{R}{N}\lg W + \text{const.}, \tag{1}$$

where

R is the gas constant,
N is the number of molecules in one gram-molecule,
S is the entropy,
W is the quantity customarily designated as the "probability" of the state with which the entropy value S is associated.

---

[1] M. v. Smoluchowski, *Ann. d. Phys.* 25 (1908): 205–226. [1]

W is commonly equated with the number of different possible ways (complexions) in which the state considered—which is incompletely defined in the sense of a molecular [6] theory by observable parameters of a system—can conceivably be realized. In order to be able to calculate W, one needs a *complete* theory (perhaps a complete molecular-mechanical theory) of the system under consideration. Given this kind of approach, it therefore seems questionable whether Boltzmann's principle *by itself* has any meaning whatsoever, i.e., without a *complete* molecular-mechanical or other theory that completely represents the elementary processes (elementary theory). If not supplemented by an elementary theory or—to put it differently—considered from a phenomenological point of view, equation (1) appears devoid of content.

However, Boltzmann's principle does acquire some content independent of any elementary theory if one assumes and generalizes from molecular kinetics the proposition that the irreversibility of physical processes is only apparent.

For let the state of a system be determined in the phenomenological sense by the variables $\lambda_1 \ldots \lambda_n$ that are observable in principle. To each state Z there corresponds a combination of values of these variables. If the system is externally closed, then the energy—and, indeed, in general, no other function of the variables—is constant. Let us think of all the states of the system that are compatible with the energy value of the system, and let us denote them by $Z_1 \ldots Z_l$. If the irreversibility of the process is not one of principle, then, in the course of time, the system will pass through these states $Z_1 \ldots Z_l$ again and again. On this assumption, one can speak of the probability of the individual states in the following sense: Suppose we observe the system for an immensely long period of time θ and determine the fraction τ_1 of the time θ during which the system is in the state Z_1; then τ_1/θ represents the probability of the state Z_1. The same holds for the probability of the other states Z. According to Boltzmann, the apparent irreversibility must be attributed to the fact that the states differ in their probabilities, and that the system is probably going to assume states of higher probability, if it happens to find itself in a state of relatively low probability. That which appears to be completely law governed in irreversible processes is to be attributed to the fact that the probabilities of the individual states Z are of *different orders of magnitude*, so that a given state Z will practically always be followed by *one* state, from among all the states bordering on Z, because of this one state's enormous probability as compared with the probabilities of the other states.

It is this probability we have just described, for the definition of which no elementary theory is needed, which is related to the entropy in the way expressed by equation (1). It can easily be recognized that equation (1) must really be valid for the probability so defined. For entropy is a function that does not decrease in any process in which the system is isolated (within the range of validity of thermodynamics). There are other functions, too, that have this property; however, if the energy E is the only function of the system that does not vary with time, then all of these functions are of the form $\varphi(S, E)$, where $\partial\varphi/\partial S$ is always positive. Since the probability W is, as well, a function that does not decrease in any process, then W is also a function of S and E alone, or—if only states of the same energy are being compared—a function of S alone. That the relation between S and W given in equation (1) is the only possible one can be deduced, as we know, from the theorem that the entropy of a total system that is composed of subsystems equals the sum of the entropies of the subsystems. Equation (1) can thus be proved for all of the states Z that belong to the same value of energy.

The following objection can be raised against this interpretation of Boltzmann's principle: one cannot speak of the statistical probability of a *state*, but only of that of a *state region*. The latter is defined by a portion g of the "energy surface" $E(\lambda_1 ... \lambda_n) =$ [7] 0. Obviously, W tends toward zero along with the size of the chosen portion of the energy surface. For this reason, equation (1) would be totally meaningless if the relation between S and W were not of a quite special kind. That is to say, lg W appears in the equation (1) multiplied by the very small factor R/N. If one imagines that W has been obtained for a region G_w just large enough that its dimensions lie on the border of the perceptible, then lg W will have a certain value. If the region is reduced perhaps e^{10} times, then the right-hand side will only be diminished by the vanishingly small quantity $10(R/N)$ on account of the reduction in the size of the region. Thus, if the dimensions of the region are indeed chosen small compared with perceptible dimensions, but nevertheless large enough for R/N lg G_w/G to be a numerically negligible quantity, then equation (1) will have a sufficiently exact meaning.

We have assumed so far that $\lambda_1 ... l_n$ determines *completely*, in the phenomenological sense, the state of the system in question. However, equation (1) also retains its meaning undiminished if we seek the probability of a state that is incompletely determined in the phenomenological sense. For let us seek the probability of a state that is defined by specific values of $\lambda_1 \ldots \lambda_\nu$ (where $\nu < n$), while the values of $\lambda_\nu \ldots \lambda_n$ are left [8] indeterminate. Among all the states with the values $\lambda_1 \ldots \lambda_\nu$, those values of $\lambda_\nu \ldots \lambda_n$ will be far and away the most frequent which make the entropy of the system at constant [9] $\lambda_1 \ldots \lambda_\nu$ a maximum. In that case, equation (1) will hold between this maximum value of the energy and the probability of *this* state.

§ 2. On the Deviations from a State of Thermodynamic Equilibrium [10]

Let us now draw conclusions from equation (1) regarding the relation between the thermodynamic and statistical properties of a system. Equation (1) yields immediately the probability of a state if its entropy is given. We have seen, though, that this relation is not exact; instead, only the order of magnitude of the probability W of the state in question can be determined from a known S. Nevertheless, it is possible to derive exact relationships concerning the statistical behavior of a system from equation (1) in cases where the range of the state variables for which W has values for the kind under consideration can be regarded as infinitely small.

It follows from equation (1) that

$$W = \text{const.}\, e^{\frac{N}{R}S} .$$

This equation is valid to an order of magnitude if each state Z is assigned a small region of the order of magnitude of perceptible regions. The order of magnitude of the constant is determined by taking into account that for the state of maximum entropy

(entropy S_0) W is of the order of magnitude one, so that we then have, with order-of-magnitude accuracy,

$$W = e^{\frac{N}{R}(S - S_0)}.$$

From this we can conclude that the probability dW that the quantities $\lambda_1 \ldots \lambda_n$ lie between λ_1 and $\lambda_1 + d\lambda_1, \ldots, \lambda_n$ and $\lambda_n + d\lambda_n$ is given, in order of magnitude, by the equation[2]

$$dW = e^{\frac{N}{R}(S - S_0)} . d\lambda_1 \ldots d\lambda_n$$

in the case when the system is determined only incompletely (in the phenomenological sense) by $\lambda_1 \ldots \lambda_n$[3]. To be exact, dW still differs from the given expression by a factor f, so that we must set

$$dW = e^{\frac{N}{R}(S - S_0)} . f . d\lambda_1 \ldots d\lambda_n,$$

where f will be a function of $\lambda_1 \ldots \lambda_n$ and its order of magnitude will be such that it does not affect the order of magnitude of the factor on the right-hand side.[4]

We now form dW for the immediate vicinity of an entropy maximum. If the Taylor expansion converges in the region considered, we may put

$$S = S_0 - \frac{1}{2} \sum \sum s_{\mu\nu} \lambda_\mu \lambda_\nu + \ldots$$

$$f = f_0 + \sum \lambda_\nu \left(\frac{\partial f}{\partial \lambda_\nu} \right) + \ldots,$$

if, for the state of maximum entropy, $\lambda_1 = \lambda_2 = \ldots \lambda_n = 0$. Since we are dealing with an entropy maximum, the double sum in the expression for S is essentially positive. One can therefore introduce new variables in place of the λ's, so that the above double sum is transformed into a simple sum in which only squares of the new variables, which are again denoted by λ, will appear. We get

$$dW = \text{const.}\, e^{-\frac{N}{2R} \sum s_\nu \lambda_\nu^2 + \ldots} \cdot \left[f_0 + \sum \left(\frac{\partial f}{\partial \lambda_\nu} \lambda_\nu \right) \right] d\lambda_1 \ldots d\lambda_n .$$

---

[2] We will assume that regions with extensions of observable size have a finite extension in λ.

[3] The manifold of possible states would otherwise be only $(n - 1)$-dimensional on account of the energy principle.

[4] We do not know anything about the order of magnitude of the derivatives of the function f with respect to λ. But we will assume in what follows that the derivatives of f are of the same order of magnitude as the function f itself.

The terms entering in the exponent appear multiplied by the very large number N/R. For that reason the exponential factor will, in general, practically vanish already for such values of λ which, because of their smallness, do not correspond to states of the system that deviate significantly from the state of thermodynamic equilibrium. For such small values of λ, the factor f can always be replaced by the value f_0 that it has in the state of thermodynamic equilibrium. Hence, in all those cases in which the variables deviate only slightly from their values at the ideal thermal equilibrium, the last formula can be replaced by

$$(2) \qquad dW = \text{const.} e^{-\frac{N}{R}(S-S_0)} d\lambda_1 \ldots d\lambda_n . \qquad [11]$$

For deviations from the thermodynamic equilibrium as small as those considered in our case, the quantity $S - S_0$ has an intuitive meaning. If we imagine that the states of interest to us in the vicinity of thermodynamic equilibrium are produced in a reversible manner by external influences, then, according to thermodynamics, every elementary process obeys the energy equation

$$dU = dA + TdS,$$

if one denotes by U the energy of the system, and by dA the elementary work applied to it. We are interested only in those states that an externally closed system can assume, namely states belonging to the same energy value. For the transition of such a state to a neighboring state, we will have $dU = 0$. Further, we will cause only a negligible error if we substitute the temperature T_0 of the thermodynamic equilibrium for T in the above equation. The latter will then have the form

$$dA + T_0 dS = 0$$

or

$$(3) \qquad \int dS = S - S_0 = \frac{1}{T_0} A,$$

where A denotes the work that has to be expended, according to thermodynamics, in order to transfer the system from the state of thermodynamic equilibrium to the state under consideration. We can therefore write equation (2) in the form

$$(2a) \qquad dW = \text{const.}\, e^{\frac{N}{RT_0}A} d\lambda_1 \ldots d\lambda_n .$$

Let us now imagine that the parameters λ are chosen so that they vanish just at thermodynamic equilibrium. In a certain region it will then be possible to expand A in terms of λ according to the Taylor theorem, and, given an appropriate choice of λ's, this expansion will have the form

$A + \frac{1}{2}\Sigma a_\nu \lambda_\nu^2$ + terms of higher than the second power in the λ, [12]

where all of the a_ν are positive. Further, since the quantity A enters the exponent of equation (2a) multiplied by the very large factor N/RT_0, the exponential factor will, in general, deviate markedly from zero only at very small values of A, and hence also at very small values of λ. For such small values of λ, the terms of higher than the first power in the expression for A will generally make a negligible contribution as compared with that of the second-power terms. If that is the case, we can substitute for equation (2a)

[13] (2b)
$$dW = \text{const.} e^{-\frac{N}{2RT_0}\sum a_\nu \lambda_\nu^2} d\lambda_1 \ldots d\lambda_n,$$

which has the form of the Gaussian error law.

In this paper we shall confine ourselves to this most important special case. It follows directly from (2b) that the mean value of the fluctuation work A_ν allotted to the parameter λ_ν is

(4)
$$\overline{A_\nu} = \overline{\frac{1}{2} a_\nu \lambda_\nu^2} = \frac{RT_0}{2N}.$$

Thus, this average work is equal to one-third of the mean kinetic energy of a monatomic gas molecule.

§ 3. On Deviations in the Spatial Distribution of Fluids and Liquid Mixtures from a Uniform Distribution

We denote by ρ_0 the mean density of a homogeneous substance or the mean density of one component of a binary liquid mixture. Because of the irregularity of the thermal motion, the density ρ at a point in the fluid will generally differ from ρ_0. If the liquid is enclosed in a cube characterized, with respect to a coordinate system, by

$$0 < x < L$$
$$0 < y < L$$
and
$$0 < z < L,$$

we can put, for the interior of this cube,

(5)
$$\left\{\begin{aligned} \rho &= \rho_0 + \Delta \\ \Delta &= \sum_\rho \sum_\sigma \sum_\tau B_{\rho\sigma\tau} \cos 2\pi\rho\frac{x}{2L} \cos 2\pi\sigma\frac{y}{2L} \cos 2\pi\tau\frac{z}{2L}. \end{aligned}\right.$$

The quantities ρ, σ, τ denote positive integers. However, the following needs to be noted.

Strictly speaking, we cannot speak of the density of a fluid at a spatial point, but only of the mean density in a volume whose dimensions are large compared with the mean distance between neighboring molecules. For this reason, the terms of the series in which one of the quantities ρ, σ, τ exceeds certain limits will have no physical meaning. However, we will see from the following that this circumstance is of no importance to us.

The quantities $B_{\rho\sigma\tau}$ will change over time such that on the average they will be zero. Let us now seek the statistical laws that underlie the quantities B. The latter play the part of the parameters λ in the preceding section, which determine the state of our system in the phenomenological sense.

In accordance with the preceding section, we obtain these statistical laws by determining the work A as a function of the quantities B. This can be done in the following way. If $\varphi(\rho)$ denotes the work one must do to bring a unit mass isothermally from the mean density ρ_0 to the density ρ, then this work has the value

$$\rho\varphi d\tau$$

for the mass $\rho d\tau$ contained in the volume element $d\tau$, and hence the value

$$A = \int \rho . \varphi . d\tau$$

for the whole fluid cube. We will have to assume that the deviations Δ of the density from the mean value are very small, and will set

$$\rho = \rho_0 + \Delta,$$

$$\varphi = \varphi(\rho_0) + \left(\frac{\partial\varphi}{\partial\rho}\right)_0 \Delta + \frac{1}{2}\left(\frac{\partial^2\varphi}{\partial\rho^2}\right)_0 \Delta^2 + \dots.$$

From this it follows that, because $\varphi(\rho_0) = 0$ and $\int \Delta d\tau = 0$,

$$A = \left(\frac{\partial\varphi}{\partial\rho} + \frac{1}{2}\rho\frac{\partial^2\varphi}{\partial\rho^2}\right)\int \Delta^2 d\tau,$$

where the index "0" has been omitted for the sake of simplicity. The fourth-order and higher terms have been omitted from the integrand, which is obviously only permitted if

$$\frac{\partial\varphi}{\partial\rho} + \frac{1}{2}\rho\frac{\partial^2\varphi}{\partial\rho^2}$$

is not too small, and the terms multiplied by Δ^4 etc. are not too large. But according to (5) we have

$$\int \Delta^2 d\tau = \frac{L^3}{8}\sum_\rho \sum_\sigma \sum_\tau B^2_{\rho\sigma\tau},$$

because the volume integrals of the double products of the Fourier summation terms vanish. Hence we have

$$A = \left(\frac{\partial\varphi}{\partial\rho} + \frac{1}{2}\rho\frac{\partial^2\varphi}{\partial\rho^2}\right)\frac{L^3}{8}\sum_\rho \sum_\sigma \sum_\tau B^2_{\rho\sigma\tau}.$$

If the work per unit mass that needs to be done to attain a state with a certain ρ from the state of thermodynamic equilibrium is expressed as a function of the specific volume $1/\rho = \nu$, i.e., if we thus put

$$\varphi(\rho) = \psi(\nu),$$

we obtain the even simpler expression

$$A = \frac{L^3}{16}\nu^3\frac{\partial^2\psi}{\partial\nu^2}\sum_\rho\sum_\sigma\sum_\tau B^2_{\rho\sigma\tau}, \tag{6}$$

where one has to substitute the values of the quantities ν and $\partial^2\psi/\partial\nu^2$ for the state of ideal thermodynamic equilibrium. We note that the coefficients B appear only as squares in the expression for A, and not as double products. Thus, the quantities B are parameters of the system of the same kind as those seen in equations (2b) and (4) in the preceding section. The quantities B obey therefore (independently of each other) the Gaussian error law, and equation (4) yields directly

$$\frac{L^3}{8}\nu^3\frac{\partial^2\psi}{\partial\nu^2}\overline{B^2_{\rho\sigma\tau}} = \frac{RT_0}{N}. \tag{7}$$

The statistical properties of our system are thus completely determined, or reduced to the thermodynamically ascertainable function ψ.

It should be noted that the omission of the terms with Δ^3 etc. is permissible only if $\partial^2\psi/\partial\nu^2$ for the ideal thermodynamic equilibrium is not too small or even vanishes. The latter occurs in the case of fluids or liquid mixtures that are exactly in the critical state. Within a certain (very small) region about the critical state, the formulas (6) and (7) become invalid. However, there is no difficulty, *in principle*, in completing the theory by [15] taking into consideration the terms of higher order in the coefficients.[5]

[16] § 4. *A Calculation of the Deflection of Light from an Infinitely Slightly Inhomogeneous, Absorption-free Medium*

Now that we have obtained from Boltzmann's principle the statistical law according to which the density of a uniform substance or the mixing ratio of a mixture varies with position, we will investigate the influence of the medium on a ray of light traveling through it.

Again let $\rho = \rho_0 + \Delta$ denote the density at some point of the medium, or, if we are dealing with a mixture, the spatial density of one of the components. Let the light ray under consideration be monochromatic. As regards this ray, the medium can be characterized by the refractive index g or by the apparent dielectric constant ε that corresponds to the frequency involved, and which is connected with the refractive index by the relation $g = \sqrt{\varepsilon}$. We put

$$\varepsilon = \varepsilon_0 + \left(\frac{\partial\varepsilon}{\partial\rho}\right)_0\Delta = \varepsilon_0 + \iota, \tag{8}$$

where both ι and Δ should be treated as infinitesimally small quantities.

---

[14] [5] Cf. M. v. Smoluchowski, l.c., p. 215.

Maxwell's equations hold at every point of the medium, and since we can neglect the influence that the rate of the change of ε over time has on the light, they take the form

$$\frac{\varepsilon}{c}\frac{\partial \mathfrak{E}}{\partial t} = \operatorname{curl} \mathfrak{H} \qquad \operatorname{div} \mathfrak{H} = 0,$$

$$\frac{1}{c}\frac{\partial \mathfrak{H}}{\partial t} = -\operatorname{curl} \mathfrak{E}, \operatorname{div} \varepsilon\mathfrak{E} = 0.$$

Here $\mathfrak{E}$ denotes the electric, and $\mathfrak{H}$ the magnetic field strength, while c is the vacuum velocity of light. By eliminating $\mathfrak{H}$, we get

$$\frac{\varepsilon}{c^2}\frac{\partial^2 \mathfrak{E}}{\partial t^2} = \Delta\mathfrak{E} = \operatorname{grad} \operatorname{div} \mathfrak{E}, \tag{9}$$

$$\operatorname{div} (\epsilon\mathfrak{E}) = 0. \tag{10}$$

Now let $\mathfrak{E}_0$ be the electric field of a light wave, as it would behave if ε did not vary with position. We shall call it "the field of the excitory light wave." The actual field (total field) $\mathfrak{E}$ will differ infinitesimally from $\mathfrak{E}_0$ by the opalescence field $\mathfrak{e}$, so that we can write

$$\mathfrak{E} = \mathfrak{E}_0 + \mathfrak{e}. \tag{11}$$

Substituting the expressions for ι and $\mathfrak{E}$ from (8) and (11) in (9) and (10), we obtain, if we neglect the inifinitesimally small second-order terms and take into account the fact that $\mathfrak{E}_0$ satisfies Maxwell's equations with constant dielectric constant ε_0,

$$\frac{\varepsilon_0}{c^2}\frac{\partial^2 \mathfrak{e}}{\partial t^2} - \triangle\,\mathfrak{e} = -\frac{1}{c^2}\,\iota\,\frac{\partial^2 \mathfrak{E}_0}{\partial t^2} - \operatorname{grad} \operatorname{div} \mathfrak{e}, \tag{9a}$$

$$\operatorname{div}(\iota\,\mathfrak{E}_0) + \operatorname{div}(\varepsilon_0\,\mathfrak{e}) = 0. \tag{10a}$$

If we expand (10a) and take into consideration that div $\mathfrak{E}_0 = 0$ and grad $\varepsilon o = 0$, we get

$$\operatorname{div} \mathfrak{e} = -\frac{1}{\varepsilon_0}\mathfrak{E}_0 \operatorname{grad} \iota.$$

Substituting this into (9a), we obtain

$$\frac{\varepsilon_0}{c^2}\frac{\partial^2 \mathfrak{e}}{\partial t^2} - \triangle\,\mathfrak{e} = -\frac{1}{c^2}\,\iota\,\frac{\partial^2 \mathfrak{E}_0}{\partial t^2} + \frac{1}{\varepsilon_0}\operatorname{grad}\{\mathfrak{E}_0 \operatorname{grad} \iota\} = \mathfrak{a}, \tag{9b}$$

where the right-hand side is a vector that is to be regarded as known, which we denote for brevity by "$\mathfrak{a}$." Thus, between the opalescence field $\mathfrak{e}$ and the vector $\mathfrak{a}$ there is a relationship of the same form as that between the vector potential and the electric current. As we know, the solution reads

(12)
$$\mathfrak{e} = \frac{1}{4\pi}\int \frac{\{\mathfrak{a}\}_{t_0 - \frac{r}{V}}}{r\,d\tau},$$

where r denotes the distance of $d\tau$ from the test point, and $V = c/\sqrt{\varepsilon_0}$ is the propagation velocity of the light waves. The volume integral has to be extended over the entire volume in which the exciting light field $\mathfrak{E}_0$ is different from zero. If the integral is extended only over a part of this volume, one obtains that part of the opalescence field that the exciting light wave produces by traversing the partial volume in question.

We now set ourselves the task of finding the part of the opalescence field that is produced by a plane, monochromatic excitory light wave inside the cube

$$\begin{gathered} 0 < x < l, \\ 0 < y < l, \\ 0 < z < l, \end{gathered}$$

where the edge length l of this cube shall be small compared with the edge length L of the cube considered earlier.

The exciting plane light wave shall be given by

(13)
$$\mathfrak{E}_0 = \mathfrak{A}\cos 2\pi n\left(t - \frac{\mathfrak{n}\mathfrak{r}}{V}\right),$$

where $\mathfrak{n}$ denotes the unit vector of the wave normal (components α, β, γ) and $\mathfrak{r}$ the radius vector (components x, y, z) drawn from the coordinate origin. For the sake of simplicity, we choose a point of incidence on the X-axis of our coordinate system at a distance D that is infinitely great in comparison with l. For such a point of incidence, the equation (12) takes the form

(12a)
$$\mathfrak{e} = \frac{1}{4\pi D}\int \{\mathfrak{a}\}_{t_1 + \frac{x}{V}} d\tau.$$

For we have to set

$$t_0 - \frac{r}{V} = t_0 - \frac{D - x}{V},$$

where, for brevity, we have set

$$t_0 - \frac{D}{V} = t_1,$$

and one can replace the factor $1/r$ in the integrand by the constant factor $1/D$, which equals $1/r$ to within a relatively infinitely small amount.

Now we have to calculate the volume integral occurring in (12a), which extends over our cube of edge length l, by inserting the expression for $\mathfrak{a}$ from (9b). We make this calculation easier for ourselves by introducing the following symbol: If φ is a scalar or a vector that is a function of x, y, z with t, we set

$$\varphi\left(x,\ y,\ z,\ t_1\ +\ \frac{x}{V}\right) = \varphi*,$$

so that φ^x depends, therefore, solely on x, y, z. For a scalar φ we immediately obtain [17] from this the equation

$$\text{grad}\ \varphi^{\bullet} = (\text{grad}\ \varphi)^{\bullet} + \mathfrak{i}\frac{1}{V}\left(\frac{\partial\varphi}{\partial t}\right)^{\bullet} d\tau,$$

from which it follows that

$$\int(\text{grad}\ \varphi)^{\bullet}d\tau = \int\text{grad}\ \varphi^{\bullet}d\tau - \mathfrak{i}\frac{1}{V}\int\left(\frac{\partial\varphi}{\partial t}\right)^{\bullet}d\tau,$$

where $\mathfrak{i}$ denotes the unit vector in the direction of the X-axis. The first integral on the right side can be transformed through integration by parts. If $\mathfrak{n}$ denotes the outer unit normal to the surface of the integration volume, and ds the surface element, then

$$\int\text{grad}\ \varphi^{\bullet}d\tau = \int\varphi^{\bullet}\mathfrak{n}ds.$$

Thus we have

$$(14)\qquad \int(\text{grad}\,\varphi)^{\bullet}d\tau = \int\varphi^{\bullet}\mathfrak{n}ds - \mathfrak{i}\frac{1}{V}\int\left(\frac{\partial\varphi}{\partial t}\right)^{\bullet}dt.$$

If φ is a function of undulatory character, then the surface integral on the right side of our equation will not make any contribution proportional to the volume of the integration space, or any contribution whatsoever of interest to us. Thus, in this case, an integral of the form

$$\int(\text{grad}\ \varphi)^{\bullet}d\tau$$

can make a contribution only to the X-component.

If we form the two integrals produced by the substitution of $\mathfrak{a}$ (equation (9b)) in the integral

$$\int\mathfrak{a}^{\bullet}dt$$

occurring in (12a), then we see that the second of these integrals has the form of the left side of (14), wherein $\varphi = \mathfrak{E}_0\,\text{grad}\ \iota$. Since this is really a function of undulatory character, which, moreover, vanishes if grad ι vanishes on the surface, therefore, according to (14), this second integral can make a significant contribution only to the X-component of $\mathfrak{e}$. A more exact calculation shows that this second integral exactly compensates the X-component of the first integral. We do not need to prove this expressly, because $\mathfrak{e}_x$ must vanish owing to the transverse nature of light. By virtue of what we have just said, it follows from (12a) and (9b) that

$$(12\text{b})\qquad \begin{cases} \mathfrak{e}_x = 0\,, \\ \mathfrak{e}_y = -\dfrac{1}{4\pi Dc^2}\displaystyle\int \iota \left(\frac{\partial^2 \mathfrak{E}_{0y}}{\partial t^2}\right)^{*} d\tau\,, \\ \mathfrak{e}_z = -\dfrac{1}{4\pi Dc^2}\displaystyle\int \iota \left(\frac{\partial^2 \mathfrak{E}_{0z}}{\partial t^2}\right)^{*} d\tau\,. \end{cases}$$

Now we calculate $\mathfrak{e}_y$ by substituting, in the second of these equations,

$$\left(\frac{\partial \mathfrak{E}_{0y}}{\partial t^2}\right)^{*} = -\mathfrak{A}(2\pi n)^2 \cos 2\pi n \left(t_1 + \frac{x}{V} - \frac{\alpha x + \beta y + \gamma z}{V}\right)$$

from equation (13). Further, we replace ι by means of equations (8) and (5). Interchanging the summation and integration signs, we thus obtain

$$\mathfrak{e}_y = \frac{\mathfrak{A}_y (2\pi n)^2}{4\pi Dc^2} \frac{\partial \varepsilon}{\partial \rho} \sum_\rho \sum_\sigma \sum_\tau B_{\rho\sigma\tau} \iiint \cos 2\pi n \left(t_1 + \frac{(1-\alpha)x - \beta y - \gamma z}{V}\right)$$

$$.\cos\left(2\pi\rho \frac{x}{2L}\right) .\cos\left(2\pi\sigma \frac{y}{2L}\right) .\cos\left(2\pi\tau \frac{z}{2L}\right) dxdydz,$$

where the volume integral has to be extended over the cube of the length of the edge l. The volume integral is of the form

$$J_{\rho\sigma\tau} = \iiint \cos(\lambda x + \mu y + \nu z) \cos \lambda' x \cos \mu' y \cos \nu' z \; dxdydz,$$

where one should bear in mind that $\lambda, \mu, \nu, \lambda', \mu', \nu'$ are to be regarded as very large numbers.[6] In that case one has to put

$$(15)\qquad \begin{cases} J_{\rho\sigma\tau} = \left(\dfrac{1}{2}\right)^3 l^3 \dfrac{\sin(\lambda - \lambda')\frac{l}{2}}{\frac{(\lambda - \lambda')l}{2}} \dfrac{\sin(\mu - \mu')\frac{l}{2}}{\frac{(\mu - \mu')l}{2}} \dfrac{\sin(\nu - \nu')\frac{l}{2}}{\frac{(\nu - \nu)l}{2}} \\[2ex] \cos\left(2\pi n t_1 + \dfrac{(\lambda - \lambda')l}{2} + \dfrac{(\mu - \mu')l}{2} + \dfrac{(\nu - \nu')l}{2}\right). \end{cases}$$

---

[6] In what follows, the calculations are based on the assumption that λ, μ, ν are *positive*. If this is not the case, then one or more signs in equation (15) will be different. But the final result is always the same.

When we integrate, we neglect, besides this expression, those expressions containing one or more of the very large quantities $(\lambda + \lambda')$ etc. in the denominator. We see that J deviates appreciably from zero only for such $\rho\ \sigma\ \tau$, for which the differences $(\lambda - \lambda')$ etc. are not very large. We note that we have here set

$$(15a)\qquad \begin{cases} \lambda = 2\pi n \dfrac{1-\alpha}{V}, & \lambda' = \dfrac{\pi\rho}{L}, \\[2ex] \mu = -2\pi n \dfrac{\beta}{V}, & \mu' = \dfrac{\pi\sigma}{L}, \\[2ex] \nu = -2\pi n \dfrac{\nu}{V}, & \nu' = \dfrac{\pi\tau}{L}. \end{cases}$$

If, for the sake of brevity, we set

$$\frac{\mathfrak{A}_y (2\pi n)^2}{4\pi D c^2} \frac{\partial \mathfrak{e}}{\partial \rho} = A,$$

then

$$(12c)\qquad \mathfrak{e}_y = A \sum_\rho \sum_\sigma \sum_\tau B_{\rho\sigma\tau} J_{\rho\sigma\tau}.$$

In conjunction with (15) and (15a), this equation yields the instantaneous value of the opalescence field at the point $x = D, y = z = 0$ for every instant $t_0 = t_1 + D/V$. We are especially interested in the mean intensity of the opalescence light, where the mean is to be taken both with respect to time and with respect to the opalescence-producing density fluctuations that may appear. As the measure of this mean intensity we can use the mean value of $\mathfrak{e}^2 = \mathfrak{e}_y^2 + \mathfrak{e}_z^2$. We have

$$\mathfrak{e}_y^2 = A^2 \sum_\rho \sum_\sigma \sum_\tau \sum_{\rho'} \sum_{\sigma'} \sum_{\tau'} B_{\rho\sigma\tau} B_{\rho'\sigma'\tau'} J_{\rho\sigma\tau} J_{\rho'\sigma'\tau'},$$

where the sum is to be extended over all combinations of the indices ρ, σ, τ, ρ', σ', τ'—always for the same value of t_1. We now form the mean value of this quantity with respect to the different distributions of density. We see from (15) that the quantities $J_{\rho\sigma\tau}$ do not depend on the density distribution, and neither does the quantity A. Hence, if we indicate the mean value of a quantity by a bar set over it, we get

$$\overline{\mathfrak{e}_y^2} = A^2 \sum \sum \sum \sum \sum \sum \overline{B_{\rho\sigma\tau} B_{\rho'\sigma'\tau'}} J_{\rho\sigma\tau} J_{\rho'\sigma'\tau'}.$$

However, since according to §3 the quantities B satisfy the Gaussian error law (at least in the approximation we are using) independently of each other, we will have, if it is not the case that $\rho = \rho'$, $\sigma = \sigma'$ *and* $\tau = \tau$,

$$\overline{B_{\rho\sigma\tau} B_{\rho'\sigma'\tau'}} = 0.$$

Our expression for $\overline{\mathfrak{e}_y^2}$ reduces therefore to

$$\overline{\mathfrak{e}_y^2} = A^2 \sum \sum \sum \overline{B^2_{\rho\sigma\tau}} J^2_{\rho\sigma\tau}.$$

However, this is not yet the average value we seek. We must also take the average value with respect to *time*. This appears only in the last factor of the expression for $J_{\rho\sigma\tau}$. If we take into account that the time average of this factor has the value ½, and put for brevity

$$(16) \qquad \begin{cases} \dfrac{(\lambda - \lambda')l}{2} = \xi, \\[2ex] \dfrac{(\mu - \mu')l}{2} = \eta, \\[2ex] \dfrac{(\nu - \nu')l}{2} = \zeta, \end{cases}$$

we obtain the following expression for the final average value of $\overline{\overline{\mathfrak{e}_y^2}}$:

$$\overline{\overline{\mathfrak{e}_y^2}} = \frac{1}{2} A^2 \left(\frac{l}{2}\right)^6 \sum \sum \sum \overline{B^2_{\rho\sigma\tau}} \frac{\sin^2 \xi}{\xi^2} \frac{\sin^2 \eta}{\eta^2} \frac{\sin^2 \zeta}{\zeta^2}.$$

Further, according to (7), $\overline{B^2_{\rho\sigma\tau}}$ is independent of $\rho\ \sigma\ \tau$, and can therefore be placed before the summation signs. Also, according to (16) and (15a), the ξ belonging to consecutive values of ρ differ from each other by $\dfrac{\pi}{2} \dfrac{l}{L}$ and, thus, by an infinitesimally small quantity. The triple sum that appears can therefore be turned into a triple integral. Since, according to the aforesaid, the interval $\Delta\xi$ between two consecutive ξ-values in the triple sum is described by the relation

$$\Delta\xi \frac{2}{\pi} \frac{L}{l} = 1,$$

we have

$$\sum \sum \sum \frac{\sin^2 \xi}{\xi^2} \frac{\sin^2 \eta}{\eta^2} \frac{\sin^2 \zeta}{\zeta^2}$$

$$= \left(\frac{2}{\pi} \frac{L}{l}\right)^2 \sum \sum \sum \frac{\sin^2 \xi}{\xi^2} \frac{\sin^2 \eta}{\eta^2} \frac{\sin^2 \zeta}{\zeta^2} \Delta\xi \, \Delta\eta \, \Delta\zeta,$$

where the last sum can immediately be written as a triple integral. From (16) and (15a) we can conclude that for all practical purposes this integral can be taken between the limits $-\infty$ and $+\infty$, so that it decomposes into a product of three integrals, each of which has the value π. With this taken into account, we finally obtain for $\overline{\overline{\mathfrak{e}_y^2}}$, with the help of (7) and substituting the expression for A,

$$\overline{\mathfrak{z}_y^2} = \frac{RT_0}{N} \frac{\left(\frac{\partial\varepsilon}{\partial\rho}\right)^2}{\nu^2 \frac{\partial^2\psi}{\partial\nu^2}} \left(\frac{2\pi n}{c}\right)^4 \frac{l^3}{(4\pi D)^2} \frac{\mathfrak{A}_y^2}{2},$$ [18]

or, if we consistently introduce the specific volume ν and replace c/n by the wavelength λ of the excitory light,

(17)
$$\overline{\mathfrak{z}_y^2} = \frac{RT_0}{N} \frac{\nu\left(\frac{\partial\varepsilon}{\partial\nu}\right)^2}{\frac{\partial^2\psi}{\partial\nu^2}} \left(\frac{2\pi}{\lambda}\right)^4 \frac{\Phi}{(4\pi D)^2} \frac{\mathfrak{A}_y^2}{2}.$$

Here Φ denotes the opalescence-producing volume traversed by the light, the shape of which volume is of no consequence. An analogous formula holds for the z-component, while the x-component of $\mathfrak{z}$ vanishes. From this we see that, for determining the intensity and the polarization state of the opalescence light emitted in a given direction, the decisive factor is the projection of the electric vector on the plane normal to the opalescence ray, no matter in what direction the exciting light may propagate.[7] If J_e denotes the intensity of the exciting light, J_0 the intensity of the opalescent light at a distance D in a specified direction from the place of excitation, and φ the angle between the electric vector of the exciting light and the plane normal to the opalescence ray under consideration, then we will have, according to (17),

(17a)
$$\frac{J_0}{J_e} = \frac{RT_0}{N} \frac{\nu\left(\frac{\partial\varepsilon}{\partial\nu}\right)^2}{\frac{\partial^2\psi}{\partial\nu^2}} \left(\frac{2\pi}{\lambda}\right)^4 \frac{\Phi}{(4\pi D)^2} \cos^2\varphi.$$

In addition, we will calculate the apparent absorption due to opalescence by integrating the opalescent light over all directions. If the thickness of the layer traversed by the light is denoted by δ and the absorption constant by α ($e^{-\alpha\delta}$ = intensity attenuation factor), we get

(18)
$$\alpha = \frac{1}{6\pi} \frac{RT_0}{N} \frac{\nu\left(\frac{\partial\varepsilon}{\partial\nu}\right)^2}{\frac{\partial^2\psi}{\partial\nu^2}} \left(\frac{2\pi}{\lambda}\right)^4.$$

---

[7] It is not surprising that our opalescence light shares this property with the opalescent light produced by suspended particles that are small compared with the wavelength of the light. After all, both cases involve irregular disturbances of the homogeneity of the irradiated substance, the locations of which are rapidly changing.

It is significant that the main result of our investigation, given by the formula (17a), [19] permits the exact determination of the constant N, i.e., the absolute size of molecules. In what follows, this result will be applied to the special case of a homogeneous substance, as well as to binary liquid mixtures in the vicinity of the critical state.

§ 5. *Homogeneous Substances*

In the case of a homogeneous substance we have

$$\psi = -\int p dv,$$

hence

$$\frac{\partial^2 \psi}{\partial v^2} = -\frac{\partial p}{\partial v}.$$

Further, according to the Clausius-Mosotti-Lorentz relation,

$$\frac{\varepsilon - 1}{\varepsilon + 2} v = \text{const.},$$

hence

$$\left(\frac{\partial \varepsilon}{\partial v}\right)^2 = \frac{(\varepsilon - 1)^2 (\varepsilon + 2)^2}{9 v^2}.$$

Substituting these values in (17a), we get

$$\frac{J_0}{J_e} = \frac{R T_0}{N} \frac{(\varepsilon - 1)^2 (\varepsilon + 2)^2}{9 v \left(-\frac{\partial p}{\partial v}\right)} \left(\frac{2\pi}{\lambda}\right)^4 \frac{\Phi}{(4\pi D)^2} \cos^2 \varphi. \tag{17b}$$

In this formula, which gives the ratio of the intensity of the opalescent light to that of the excitory light, in case the latter is measured at distance D from the volume Φ originally traversed by the light, we use the following notation:

R is the gas constant,
T is the absolute temperature,
N is the number of molecules in one gram-molecule,
ε is the square of the refraction exponent for wavelength λ,
v is the specific volume,
$\partial p/\partial v$ is the isothermal derivative of the pressure with respect to the volume,
φ is the angle between the electric field vector of the exciting wave and the plane normal to the opalescence ray under consideration.

That $\partial p/\partial v$ is the isothermal, and not, say, the adiabatic derivative, has to do with the fact that of all the states belonging to a given density distribution, the state of constant temperature is the state of greatest entropy, and thus the greatest statistical probability, at a given total energy.

If the substance in question is an ideal gas, then we can set $\varepsilon + 2 = 3$. For this case we obtain

$$\frac{J_0}{J_e} = \frac{RT_0}{N}\frac{(\varepsilon - 1)^2}{p}\left(\frac{2\pi}{\lambda}\right)^4 \frac{\Phi}{(4\pi D)^2}\cos^2\varphi \tag{17c}$$

As a rough calculation shows, this formula might very well explain why the light given off by the irradiated atmosphere is predominantly blue.[8] In this connection it is worth noting that our theory does not make any *direct* use of the assumption of the discrete distribution of matter.

§ 6. Liquid Mixtures [21]

The derivation according to equation (17a) is also valid in the case of a liquid mixture if one sets

v = specific volume of the unit mass of the first component,
ψ = work needed to bring the unit mass of the first component along a reversible path from the specific volume it has in thermal equilibrium to some other given specific volume along a reversible path at constant temperature.

If the vapor coexisting with the liquid mixture under consideration can be regarded as a mixture of ideal gases, and the mixture can be regarded as incompressible, then the quantity ψ can be replaced by quantities accessible to experience. We then find ψ by the following elementary argument.

Let the mass k of the second component be mixed with the unit mass of the first component. In that case, k is a measure of the composition of the mixture, the total mass of which is $1 + k$. Let this mixture have a vapor phase, and let p'' be the partial pressure, and v'' the specific volume of the second component in the vapor phase. Let this system be enclosed in a container with a semipermeable section of wall through which the second, but not the first component can be taken in and out in gaseous form. Let a second, relatively infinitely large container enclose a relatively infinitely large amount of the mixture with that composition (characterized by k_0) for which we wish to calculate the opalescence. This second mixture shall also occupy a vapor space with a semipermeable wall, and the partial pressure and specific volume of the second component in the vapor space shall be denoted by p_0'' and v_0'', respectively. Let the temperature inside both containers be T_0. We shall now calculate the work $d\psi$ that is necessary to increase the concentration measure k in the first container by dk by transporting, in gaseous form and in a reversible way, the mass dk of the second

---

[8] Equation (17c) can also be obtained by summing the radiations of the individual gas molecules, which are considered to be completely randomly distributed. (Cf. Rayleigh, *Phil. Mag.* 47 [1899]: 375, and *Papers* 4, p. 400.) [20]

component from the second to the first container. This work is composed of the following three parts:

$-\frac{dk}{M''}p_0''v_0''$ (work of removal from the second container)

$\frac{dk}{M''}RT_0\lg\frac{p''}{p_0''}$ (isothermal compression to the partial pressure in the first container)

$+\frac{dk}{M''}p''v''$ (work of input into the first container).

The volume of the liquid is neglected here compared with the volume of the gas. M'' is the molecular weight of the second component in the vapor phase. Since the first and the third terms cancel out according to Mariotte's law, we get

$$d\psi = \frac{RT_0}{M''}dk\ \lg\frac{p''}{p_0''}.$$

The function ψ can thus be calculated directly from the concentrations and partial pressures. Now we have to find $\partial^2\psi/\partial v^2$ for the state we denoted by the index "0." We have

$$\lg\left(\frac{p''}{p_0''}\right) = \lg\left(1 + \frac{p'' - p_0''}{p_0''}\right) = \lg\ (1 + \pi) = \pi - \frac{\pi^2}{2}\ \ldots,$$

where π is the relative pressure change of the second component with respect to the original state. From the last two equations there follows

$$\frac{\partial\psi}{\partial v} = \frac{RT_0}{M''}\frac{\pi - \frac{\pi^2}{2} + \ldots}{\frac{\partial v}{\partial k}}.$$

Differentiating one more time with respect to v, and considering that

$$\frac{\partial}{\partial v} = \frac{\frac{\partial}{\partial k}}{\frac{\partial v}{\partial k}}\ ,$$

we obtain, if we set $\pi = 0$ in the result,

$$\left(\frac{\partial^2\psi}{\partial v^2}\right)_0 = \frac{RT_0}{M''}\frac{\frac{\partial\pi}{\partial k}}{\left(\frac{\partial v}{\partial k}\right)^2} = \frac{RT_0}{M''}\frac{\frac{1}{p''}\frac{\partial p''}{\partial k}}{\left(\frac{\partial v}{\partial k}\right)^2}.$$

Taking this into account, along with the fact that

$$\frac{\partial \varepsilon}{\partial v} = \frac{\dfrac{\partial \varepsilon}{\partial k}}{\dfrac{\partial v}{\partial k}} \quad ,$$

we can rewrite the formula (17a) in the form

$$\text{(17d)} \qquad \frac{J_0}{J_e} = \frac{M''}{N} \frac{v\left(\dfrac{\partial \varepsilon}{\partial k}\right)^2}{\dfrac{\partial (\log p'')}{\partial k}} \left(\frac{2\pi}{\lambda}\right)^4 \frac{\Phi}{(4\pi D)^2} \cos^2 \varphi .$$

This formula, which now contains only quantities accessible to experiment, completely determines the opalescent properties of binary liquid mixtures—to the extent that their saturated vapors can be treated as ideal gases—up to a small region in the immediate vicinity of the critical point. But because of the strong absorption of light and its great dependence on the temperature, a quantitative investigation might well be ruled out here anyhow. Let us repeat here the meanings of the symbols that appear in this formula insofar as they have not been explained along with formula (17b):

M''	is the molecular weight of the second component in the vapor phase,
v	is the volume of the liquid mixture in which the unit mass of the first component is contained,
k	is the mass of the second component which falls to the share of the unit mass of the first component,
p''	is the vapor pressure of the second component.

Lest it not look peculiar that the two components play different roles in (17d), let me mention the well-known thermodynamic relation

$$\frac{1}{M''} \frac{dp''}{p} = -\frac{1}{M'} \cdot \frac{1}{k} \frac{dp'}{p'} .$$

From this relation one can conclude that it does not matter which component is treated as the first, and which as the second.

A quantitative experimental investigation of the phenomena here considered would be of great interest: on the one hand, it would be valuable to know whether Boltzmann's principle gives indeed a correct account of the phenomena here considered, and on the other hand, such investigations could lead to accurate values for the number N.

Zurich, October 1910. (Received on 8 October 1910)

Doc. 10

Comments on P. Hertz's Papers: "On the Mechanical Foundations of Thermodynamics"[1]

by A. Einstein.

[*Annalen der Physik* 34 (1911): 175–176]

In his superb papers titled as above, Mr. Hertz has criticized two passages in my papers
[3] on the same topic. In the following, I will briefly comment on these criticisms, noting that what is said here is the result of an oral discussion with Mr. Hertz, in which we came to a perfect agreement regarding both points in question.

1. In the penultimate section of §13 of his second paper, Hertz criticizes a derivation
[4] that I gave of the entropy law for irreversible processes. I consider this criticism totally valid. I was not satisfied with my derivation even then, which is why I soon thereafter
[5] produced a second derivation, also cited by Mr. Hertz.

2. The comments contained in §4 of his first paper that are directed against an argument about thermal equilibrium contained in my first paper in question[2] are based on a misunderstanding caused by an all-too terse and insufficiently careful formulation
[6] of that argument.

However, since the topic has been adequately elucidated in works by other authors, and since, moreover, a detailed discussion of this specific point is not likely to claim much interest, I do not wish to elaborate on it here. I only wish to add that the road taken by Gibbs in his book, which consists in one's starting directly from the canonical ensemble,
[7] is in my opinion preferable to the road I took. Had I been familiar with Gibbs's book
[8] at that time, I would not have published those papers at all, but would have limited myself to the discussion of just a few points.

Zurich, October 1910. (Received on 30 November 1910)

---

[1] [1] A. Einstein, *Ann. d. Phys.* 9 (1902): 425 and 11 (1903): 176.

[2] [2] P. Hertz, *Ann. d. Phys.* 33 (1910): 225 and 537.

Doc. 11

Lecture Notes for Course on Electricity and Magnetism at the University of Zurich, Winter Semester 1910/11[1]

Electrostatics [p. 1]

If one rubs glass, sealing wax, or other bodies with other bodies, then after this procedure they will (temporarily) exert forces on each other that were not observable before, without their having been otherwise influenced in a perceptible manner. One says that they are "electrified," where by this word one does not denote anything but what has been said. Metals & many other bodies can be electrified only if affixed to a prop of glass or sealing wax etc., or suspended by a silk thread. A body can be electrified not only by rubbing but also by bringing it into contact with an electrified body.

Let us examine the laws according to which electrified bodies act upon one another, assuming for the sake of simplicity that the bodies are small compared with the distances between them. The forces exerted by these bodies on each other act in the direction of the connecting lines (equality of action & reaction, we can measure them absolutely by the methods of mechanics, for example in the following way:

$$\mathfrak{F}_e = mg \sin x$$

Consider now many bodies, say small metal balls suspended by silk threads, and let [p. 2] us suppose that we have determined that the forces that *any two of them* exert on each other, and assume, for the time being, that they are at a distance R that always stays the same. We designate attractive forces as negative, repulsive as positive.

If we combine the bodies 1 2 3 .. with the body a of our group, we obtain the forces F_{1a}, F_{2a}, F_{3a} .... .. If we combine the same bodies 1 2 3 .. with the body b, we obtain the forces

$$F_{1b}\, F_{2b}\, F_{3b}$$

Experience shows that $F_{1a} : F_{2a} : F_{3a}.. = F_{1b} : F_{2b} : F_{3b}..$ Thus, the effects of the bodies 1 2 3 .. another body always stand in the same ratio no matter how that other body has been chosen. Hence we can characterize the electrical influence of *one* el. body by means of a number, if we have assigned an arbitrarily chosen number, for example the number 1, to the influence of one of the bodies.[2] This number is called the quantity of electricity. It follows from this definition that the force f exerted by two bodies on each other is directly proportional to their quantities of electricity.

$$F = k \cdot e_1 e_2.$$

However, k also depends on the distance.

Further, it follows from experiments that this force is inversely proportional to the square of the distance, so that we have, with another interpretation of the constant k,

$$F = k\frac{e_1 e_2}{r^2},$$

where k no longer depends on the distance but only on our choice of the body in our [p. 3] group to which we have assigned the quantity of electricity 1.

The sign of k is determined by our earlier stipulation in conjunction with experience. That is to say, it has been found that quantities of electricity that are alike according to the above definition repel each other. Thus, k is a positive constant. Its value depends on what we stipulate as the unit of the quantity of electricity. However, we may also freely choose k and thereby define the unit of the quantity of electricity. We do that by setting $k = 1$. We have then

$$F = \frac{e_1 e_2}{r^2}$$

In order to measure a quantity of electricity absolutely after according to this <kind of> definition,[3] one has to measure, in principle, a force and a length, which quantities occur in the form

$$e = \sqrt{\text{force}} \cdot \text{length} = M^{1/2} L^{3/2} T^{-1}$$

This is the "dimension" of the electrostatically measured quantity of electricity.

We must mention a few more facts that are of fundamental importance for the foundations of the theory.

If a quantity of electricity e_a is subjected to the action of two quantities of electricity [p. 4] e_1 & e_2, one finds the force acting on e_a from the law of the parallelogram of forces. In the special case where e_1 & e_2 are very close to each other, their effects on e_a will add up algebraically; in other words: the quantity of electricity of a system of bodies is equal to the sum of the quantities of electricity of the system's individual bodies.

This principle can be further extended, given the character of our experience with electrified bodies. If bodies with quantities of electricity e_1 & e_2 are brought into contact with one another, then, in general, their electric state will change. But their action at a distance on a third e. q. e_a will not change upon the contact, and so the sum of the electrical quantities will not change either. (Important law of the constancy of the sum of quantities of electricity, an exception to which has never been found.)

We endow these two laws with a tangible, physical meaning by imagining that the substrate for the quantity of electricity is some sort of indestructible matter, which, however, must be thought of as being present in a positive and a negative modification, because the experiments alluded to above show the existence of positive as well as negative electrical quantities (in the case of attractive forces).

One more thing has to be added to complete what has been said so far, for there is no way to decide which sign to ascribe to a specific given electrical quantity, because the [p. 5] interaction between two e. q. only makes it possible to decide whether the two have to be assigned *like* or *opposite* signs. But all that is needed, therefore, is to fix the sign in a specific case (glass rubbed with wool is positive), in order to fix signs for all other quantities of electricity.

In completing what has been said about the auxiliary representation of positive and negative electricity, it should be added that one imagines that the interactive forces act between the electricities and are transferred from them to the carriers of electricity (bodies) to which they are bound. We further complete the picture by the assumption that not only the algebraic sum of the electrical quantities, but also the sum of the electricities of each of the signs is constant—a proposition that is part of the picture and that cannot be either directly confirmed or directly discomfirmed by experiment.

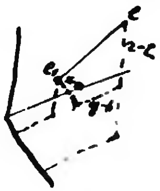
The action of a system of electric masses $(e_1\, e_2\, \ldots.)$ on a pointlike quantity of electricity (e).

An electrical quantity $e_1(x\, y\, z)$ exerts the force K on a quantity of electricity $e(a,b,c)$.[4] We have

$$K = \frac{e_1 e}{r^2}, \quad \text{where} \quad r^2 = (x-a)^2 + (y-b)^2 + (z-c)^2.$$

The direction cosines of this force are $\frac{x-a}{r}$, $\frac{y-b}{r}$, $\frac{z-c}{r}$,

so that its components are



[p. 6]

$$\left.\begin{aligned} K_{x1} &= \frac{e_1 e}{r_1^2}\,\frac{x-a_1}{r_1} \\ K_{y1} &= \frac{e_1 e}{r_1^2}\,\frac{y-b_1}{r_1} \\ K_{z1} &= \frac{e_1 e}{r_1^2}\,\frac{z-c_1}{r_1} \end{aligned}\right\} \quad (1)$$

If several masses e_1 e_2 .... act simultaneously on mass e, we get[5]

$$K_x = \Sigma K_{x1} = \frac{e_1 e}{r_1^2}\,\frac{x-a_1}{r_1} + \frac{e_2 e}{r_2^2}\,\frac{x-a_2}{r_2} \ldots\ldots = e \cdot \sum_1^n \frac{e_1}{r_1^2}\,\frac{x-a_1}{r_1}$$

– – – – – – – – – – – – – – – – – –

– – – – – – – – – – – – – – – – – –

For a given distribution of the masses e_1 etc., and a given position for e, these force components are proportional to the e.q. e. But the sums appearing on the right-hand side depend only on e_1 e_2 ... & the test point. These sums

$$\sum_1^n \frac{e_1}{r_1^2}\,\frac{x-a_1}{r_1} = X \quad \text{(other components } Y\ Z\text{)}$$

are called the X-component of the electric force or field strength. It is equal to the force exerted on the unit of electricity. $X\ Y\ Z$ is a vector which is related to the vector of the force acting upon the e quantity e in the following way:

$$K_x = eX \quad K_y = eY \quad K_z = eZ \ldots\ldots (2)$$

If one draws from every spatial point a directed straight line in the direction of the field intensity, one gets a picture of the course of the field intensity, of the vector field $X\ Y\ Z$ that brings about the (possible) actions of forces deriving from the quantities e_1 e_2 etc. [p. 7] This field is determined chiefly by 3 spatial functions (X Y and Z). However, these can be reduced to a single spatial function. For we have

$$X = \sum \frac{e_1}{r_1^2}\,\frac{x-a_1}{r_1} = \sum \frac{e_1}{r_1^2}\,\frac{\partial r_1}{\partial x} = -\frac{\partial}{\partial x}\left\{\sum\left(\frac{e_1}{r}\right)\right\},$$

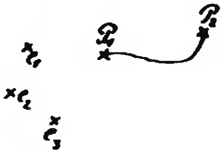
since because $r_1^2 = (x-a_1)^2 + \cdot + \cdot$; $r_1 dr_1 = (x-a_1)dx + \cdot + \cdot$ Hence, if we set $\sum \frac{e_1}{r_1} = \varphi$, we get

$$\left.\begin{aligned} X &= -\frac{\partial\varphi}{\partial x} \\ Y &= -\frac{\partial\varphi}{\partial y} \qquad \varphi = \sum \frac{e_1}{r_1} \\ Z &= -\frac{\partial\varphi}{\partial z} \end{aligned}\right| \quad (3)$$

Thus, $X\ Y Z$ can be described as derivatives of *one* spatial function φ. We call φ the potential of the masses in question.

The Physical Meaning of the Potential

We consider the electrical unit mass in the field of the e. q. $e_1, e_2, e_3 \ldots$. We move the unit m. from the point P_1 to the point P_2. For an infinitesimally small portion of the path with projections *dx dy dz*, the work performed by the forces of electric origin equals $Xdx + Ydy + ZdZ$. The total work is therefore $A = \int_{P_1}^{P_2} Xdx + Ydy + Zdz$



With the help of (3), this work can be given the form

$$A = -\int \frac{\partial\varphi}{\partial x}dx + \frac{\partial\varphi}{\partial y}dy + \frac{\partial\varphi}{\partial z}dz = -\int d\varphi,$$

where $d\varphi$ denotes the total change of φ when the element *dx dy dz* is traversed. Hence we obtain

$$A = \varphi_1 - \varphi_2 \ldots . \quad (4)$$

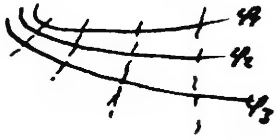
Thus, the work done on the unit electr. mass between two points is equal to the potential [p. 8]
drop between these two points φ is independent of the choice of the coordinate system. This quantity is totally independent of the shape of the path. Hence, if the unit pole

describes a closed curve, i.e., if P_1 & P_2 coincide, $\varphi_1 = \varphi_2$, and hence the work $A = 0$. This fact contains the more profound interpretation of the reason why the vector XYZ of the el. field strength is derivable from a potential. If the integral were not to vanish for a closed curve, it would be possible to produce work from nothing, without limit, by means of electrical quantities.

The Theorems of Laplace and Gauss. Lines of Force

<Here give a little kiss to his poor!>[6]

The funct φ provides a graphic overview of the course of the el. field. If one thinks of a surface φ = const., then the field vector XYZ will be perpendicular to the surface φ = const. Because every derivative $-\dfrac{\partial\varphi}{\partial s}$ taken in the direction of a line element in the surface vanishes. If we think of two adjacent surfaces $\varphi = \varphi_0$ & $\varphi = \varphi - \varepsilon$, we will have $-\dfrac{\partial\varphi}{\partial n} = \dfrac{\varepsilon}{\delta}$, and since ε is everywhere constant along the two surfaces, $\dfrac{1}{\delta}$ is a relative measure for $-\dfrac{\partial\varphi}{\partial n}$, i.e., for the el. field strength, or—as we will call it in brief for the el. force. An additional aid for intuitive visual representation is provided by the concept of lines of force, i.e., of lines that at each point have the same direction as the electric force. According to what we have said, these lines of force everywhere intersect the surfaces of equal potential perpendicularly. Beyond this, we will see that the density of these lines of force is proportional to the field intensity. But in order to do [p. 9] this, we must first derive a few laws.



The Theorems of Laplace & Gauss

If only one charge is pres., then $\varphi = \dfrac{e}{r}$, where

$$r = +\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}\,.$$

Diff, we obt.

$$\frac{\partial\varphi}{\partial x} = -\frac{e}{r^2}\frac{x-a}{r} = -\frac{e}{r^3}(x-a)$$

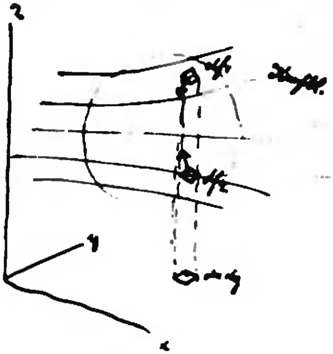
$$\frac{\partial^2\varphi}{\partial x^2} = -\frac{e}{r^3} + \frac{3e}{r^4}\frac{(x-a)^2}{r}$$

From this, $\Delta\varphi = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2} = 0 \ldots.$ (5) also holds for an arbitrary number of masses (Laplace's theorem)

We can express this theorem in still another form if we use the field intensity instead of the derivatives of φ.

$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} = 0 \ldots. (5a)$$

We can give this theorem a new form by integrating over a volume bounded by a closed surface that contains no electric masses.



$$\int \frac{\partial Z}{\partial z} dxdydz$$

Portion of an element

$$dxdy = dxdy\int \frac{\partial Z}{\partial z}\,dz = dxdy(Z_1 - Z_2)$$

if n_1 and n_2 are the inwardly oriented normals, then[7]

$$dxdy = -\,df_1 \cos(n_1z) = df_2 \cos(n_2z)$$

We can set $-\Sigma Z \cos nz\, df$ over the two elements Every other element $dx\,dy$ has the same form, so that when one finally replaces the sum with the integral, one obtains[8] [p. 10]

$$\int \frac{\partial Z}{\partial z} d\tau = -\int Z \cos nz\, ds$$

Applying this theorem three times, one obtains

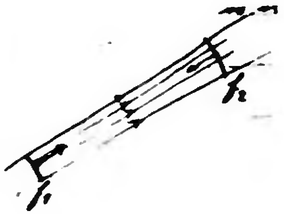
$$0 = \int\left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z}\right)d\tau = -\int (X \cos nx + Y \cos ny + Z \cos nz)ds.$$

Considering that the expression in the brackets is in fact the field component N in the direction of the inward normals, we obtain

$$\left.\begin{aligned} &\int N ds = 0. \\ \text{which we can also write in the form}\quad &\int \frac{\partial \phi}{\partial n} ds = 0 \end{aligned}\right\} (6)$$

This theorem yields us a further property of the field of electric lines of force. Define tube of force & write down above theorem for it. Integral vanishes on the surface. On the initial and terminal cross section we have

$$N_1 f_1 = N_2 f_2.$$

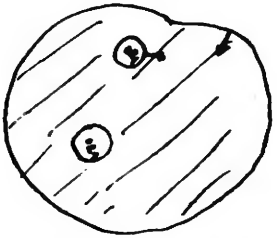


This vanishes. $\frac{N_2}{N_1} = \frac{f_1}{f_2}$. Thus, the field strengths vary inversely as the surfaces of the tubes of force. If one draws a number of lines of force through f_1 and continues them up to f_2, then the density of these lines of force will likewise be inversely proportional to the surface areas, and thus directly proportional to the field intensities N. Thus, one can draw unending lines of force in the field, so that line density = field strength. This is why the lines of force afford a quite complete & direct intuitive visual representation of a field.

[p. 11] Equation (6) expresses the one special case of the so-called Gauss's theorem.[9] This equation can easily be extended to the case where the closed surface encloses electric masses e_1 e_2 ....

We extend the surface integral to the volume bounded by the giv. surface F and the auxiliary spherical surf. K_1 K_2 etc.

$$\int_F N ds + \int_{K_1} N ds + \int_{K_2} N ds \ldots = 0.$$



We seek integral extended over sphere K_1 We divide the total field into 1. X_1 Y_1 Z_1 N_1, which derives from e_1, & second, the rest $X'Y'Z'N'$

The surface integral $\int_{K_1} N' ds$ vanishes, $\int_{K_1} N_1 ds = \frac{e_1}{r^2} 4\pi r^2 - 4\pi e_1$ We thus obtain

$\int N ds = -4\pi \sum e_1$ (General form of Gauss's theorem.)

Continuously Distributed Electricity

So far we have assumed that electricity is unalterably bound to small bodies (treated as points). But the character of experience favors the assumption that electricity is spatially distributed. We must generalize our investigations in this sense. To begin with, we think of electricity as continuously distributed, $\rho d\tau$ being the quantity of electricity in the space element $d\tau$. ρ is the difference between the densities of positive and negative electricity [p. 12] at one locus, as we imagine it. We assume that the electricities are movable relative to ponderable matter, and that they cannot undergo any other changes except those of position. This model is suggested by the earlier-mentioned empirical law of the constancy of the quantity of electricity in the electrical balance between two small bodies.

The following should be noted here. We have seen how experience led to the introd. of the concept of the quantity of electricity. it was defined by means of the forces that small electrified bodies exert on each other. But now we extend the application of the concept to cases in which this definition cannot be applied directly as soon as we conceive the el. forces as forces exerted *on electricity* rather than on material particles. We set up a conceptual system the individual parts of which do not correspond directly to empirical facts. Only a certain totality of theoretical material corresponds again to a certain totality of experimental facts.[10]

We find that such an el. continuum is always applicable only for the representation of el. states of affairs in the interior of ponderable bodies. Here too we define the vector [p. 13] of el. field strength as the vector of the mech. force exerted on the unit of pos. electr. quantity inside a body. But the force so defined is no longer directly accessible to exp. It is one part of a theoretical construction that can be correct or false, i.e., consistent or not consistent with experience, only *as a whole*. The laws that we found empirically for small electrified bodies we now apply to the fictional electricity itself.

We invest. the pot. of cont. distribution

$\varphi = \int \rho \frac{d\tau}{r}$ R small radius sphere about the test point region decomposed polar coordinates introduced

$$c - z = r\cos\vartheta$$
$$a - x = r\sin\vartheta\cos\omega \quad \text{volume el}\cdot r^2 \sin\vartheta\, dr\, d\omega\, d\vartheta$$
$$b - y = r\sin\vartheta\sin\omega$$

In small sphere $\int_K \frac{\rho d\tau}{r}$ replaceable by $\int \rho_0 r \sin\vartheta\, dr\, d\omega\, d\vartheta$ always finite. Thus, the integral is not infinite.

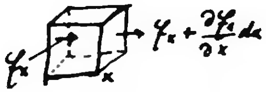
$$\frac{\partial\varphi}{\partial z} = \int \frac{\rho\, d\tau}{r^2}\frac{c-z}{r} = \int_R + \int_K \rho \cos\vartheta \,\sin\vartheta \, dr\, d\omega\, d\vartheta.$$

The second int. is finite.[11] Hence field strength always finite. One proves that when ρ with all derivatives is continuous, the same must be true of φ.

The equation $\Delta\varphi = 0$ is not valid here. We find the corresponding theorem by [p. 14] applying Gauss's theorem to an arb. closed surface inside the continuum.

$$\boxed{\int \mathfrak{E}_n d\sigma = -\int 4\pi\rho\, d\tau,}$$

where $\mathfrak{E}_n$ denotes the component of the el. field strength along the inward normal. First we apply the theorem to the special case where the surface is the boundary of an elementary parallelepiped. The right side becomes $-4\pi\rho\, d\tau$. The left side



$$\mathfrak{E}_x\, dy\, dz + -\left(\mathfrak{E}_x + \frac{\partial \mathfrak{E}_x}{\partial x} dx\right) dy\, dz$$

— — — — — — — — — —

— — — — — — — — — —

or $$-\left(\frac{\partial \mathfrak{E}_x}{\partial x} + \frac{\partial \mathfrak{E}_y}{\partial y} + \frac{\partial \mathfrak{E}_z}{\partial z}\right) d\tau$$

If the two sides are set equal, one obtains

$$\boxed{\frac{\partial \mathfrak{E}_x}{\partial x} + \frac{\partial \mathfrak{E}_y}{\partial y} + \frac{\partial \mathfrak{E}_z}{\partial z} = 4\pi\rho}$$

If one replaces $\mathfrak{E}_x$ etc. by the derivatives of the potential, one obtains

$$\boxed{+\Delta\varphi = -4\pi\rho}$$

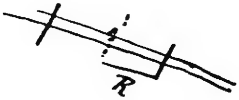
This is Poisson's theorem.

Distribution of the Electricity on Conductors

A conductor is a substance in which the electricity is freely movable. Equilibrium possible only if no forces act on el. in the interior. $\mathfrak{E}_x$ etc. vanish. Poisson's theorem applied to a point in the interior of the conductor yields $\rho = 0$. Thus, the electric masses sit only on the surface, & in the interior of the conductor φ = const.

[p. 15] Since the electricity is distributed two-dimensionally on the surface, we must consider a two-dimensionally distributed potential.

1) Potential is uniform over surface. A little piece of the surface is cut out by a cylinder around the spot under investigation. That which derives from the external part of the covering is uniform. That which derives from the internal part vanishes for small radius; for[12]



$$\varphi_i = \int \frac{\eta\, d\sigma}{r} = \eta_0 \int_0^R \frac{2\pi r\, dr}{r} = 2\pi R,$$

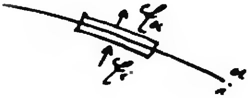
which decreases with decreasing R.
From the constancy of φ it follows that the tangential components of $\mathfrak{E}$ on the two sides of the layer are equal.

$$\left.\begin{aligned} \varphi_1 &= \varphi_2 \\ \varphi_1' &= \varphi_2' \end{aligned}\right\} \varphi_1 - \varphi_1' = \varphi_2 - \varphi_2'$$

$$\frac{\varphi_1 - \varphi_1'}{\delta} = \frac{\varphi_2 - \varphi_2'}{\delta} \qquad \text{or} \qquad \boxed{\mathfrak{E}_{t2} = \mathfrak{E}_{t1}}$$

From this it follows $\mathfrak{E}_t$ vanishes on the external surface of a conductor, i.e., that the lines of force must intersect the surface of the conductor perpendicularly.
2) How does the normal component behave on the two sides?
This follows at once from Gauss's theorem.[13]

$$4\pi\sigma\, df = \mathfrak{E}_i\, df - \mathfrak{E}_a\, df$$

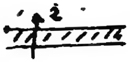


or

$$\mathfrak{E}_{ni} - E_{na} = -4\pi\sigma. \quad \text{special case} \quad \mathfrak{E}_{ni} = 0 \quad \mathfrak{E}_{na} = 4\pi\sigma$$

or $\left(\frac{\partial\varphi}{\partial n}\right)_i - \left(\frac{\partial\varphi}{\partial n}\right)_a = 4\pi\sigma$, if both normals are taken toward the *external* side.

Force on piece of the conductor surf. [p. 16]



$$\int \frac{\partial \mathfrak{E}_2}{\partial z}\, dz = \mathfrak{E}_{2a} - \mathfrak{E}_{2i} = 4\pi \int \rho\, dz = 4\pi\sigma$$

$$+\,4\pi\rho = \frac{\partial \mathfrak{E}_2}{\partial z}.$$

$$4\pi \int \rho \mathfrak{E}_2\, dz = \int \mathfrak{E}_2 \frac{\partial \mathfrak{E}_2}{\partial z}\, dz = \frac{1}{2}(\mathfrak{E}_{2a}^2 - \mathfrak{E}_{2i}^2) \Big| \mathfrak{Kraft} = \frac{1}{8\pi}\mathfrak{E}_2^2$$

The problem of finding the distribution of electricity on a conductor is now easy to formulate mathematically if we further stipulate that the potential should be constant at ∞. If all effective el. masses are at a finite distance, its value there is zero. For φ can be determined from the following conditions:

1) $\varphi = \text{const} = P_0$ inside the body
2) $\Delta\, \varphi = 0$ outside the body.
3) φ constant on the surface of the body. φ together with the derivatives in the external region.
4) φ vanishes at ∞.

We prove later that these conditions are *sufficient.*

For the difference φ_1 of two solutions φ_1 must vanish ~~*outside*~~ on the surface. <Thus, if there existed a closed surface anywhere in the external region>

We now choose a closed surface in the external region

$$\int\left(\frac{\partial\varphi}{\partial x}\right)^2 + \left(\frac{\partial\varphi}{\partial y}\right)^2 + \left(\frac{\partial\varphi}{\partial z}\right)^2 = -\int \varphi \frac{\partial\varphi}{\partial n} ds$$

If φ is determined in accordance with these conditions, one obtains the surface density η by means of the relation $4\pi\eta = \mathfrak{E}_n = -\frac{\partial\varphi}{\partial n}$, where the normal is directed toward the outer side of the conductor. One obtains the total charge by integrating η over the surface.

[p. 17] Example. Let the given body be a sphere. We show that the solution $\varphi = \frac{\alpha}{r}$ in the external region and $\varphi = P$ in the internal region satisfies all the conditions.

1) satisfied

2) satisfied, because $\Delta\left(\frac{\alpha}{r}\right) = 0$

3) satisfied, if $\frac{\alpha}{R} = P$

4) satisfied.

We determine the charge e.

$$e = \int \eta \, d\sigma = \int \frac{\alpha}{4\pi R^2} R^2 d\kappa = \alpha \qquad \eta = -\frac{1}{4\pi}\left(\frac{\partial\varphi}{\partial r}\right)_R = +\frac{1}{4\pi}\frac{\alpha}{r^2}$$

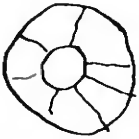
Thus, we obtain

$$\boxed{\begin{array}{l}\varphi = \frac{e}{r} \\ e = RP\end{array}}$$

This shows that e is proportional to the potential difference P. This holds not only for a sphere but quite generally. For let the problem be solved for a specific P. One then finds the solution for a $P^x = \lambda_{const} P$ by using the function $\varphi^x = \lambda\varphi$ instead of φ. Thus, $\frac{e}{P}$ depends only on the shape of the conductor and is called the capacity of the latter. The capacity of the sphere is equal to its radius.

Instead of a single conductor, let us think of one surrounded by a conducting casing.

1) $\varphi = P_1$ in the interior $\varphi = P_2$ in casing
2) $\Delta\varphi = 0$ bet. body & casing
3) constancy req.



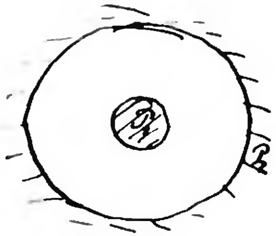
Then $\lambda P_1 \lambda P_2 \lambda\varphi$ solution λe el. quantity on body as well as on casing
Charge dep. only on pot diff.

$$\frac{P_1 - P_2}{e} = \frac{\lambda P_1 - \lambda P_2}{\lambda e} = c \text{ capacity, (mutual)}$$

Example parallel plate condenser[14] [p. 18]

$$E = -\frac{\partial\varphi}{\partial x} = 4\pi\sigma = \frac{P}{\delta} \quad e = \sigma f = \frac{Pf}{4\pi\delta} \quad C = \frac{f}{4\pi\delta}$$

Example concentric hollow spheres.



$$\varphi = \frac{\alpha}{r} + \beta$$

$$\frac{\alpha}{R_1} + \beta = P_1 \qquad \alpha\left(\frac{1}{R_1} - \frac{1}{R_2}\right) = P_1 - P_2$$

$$\frac{\alpha}{R_2} + \beta = P_2 \qquad \alpha = (P_1 - P_2)\frac{R_1R_2}{R_2 - R_1}$$

β does not interest us.

$$\varepsilon = \int \eta\, d\sigma = \frac{1}{4\pi}\int -\frac{\partial\varphi}{\partial r} d\sigma = \frac{1}{4\pi}\frac{\alpha}{R^2}4\pi R^2 = \alpha$$

$$\frac{\varepsilon}{P_1 - P_2} = \frac{R_1R_2}{R_2 - R_1} = \text{mutual capacity}$$

Even simpler derivation (center)

Example concentric cylinders, φ depends only on $\delta = \sqrt{x^2 + y^2}$. One could set $\Delta\varphi = 0$ for this special case & integrate. Even simpler, apply Gauss's theorem directly. e = el. charge per unit length

$$4\pi e = 2\pi r \mathfrak{E}_r$$

$$\mathfrak{E}_r = \frac{2e}{r} = -\frac{\partial\varphi}{\partial r}$$

$$\varphi = -2e \lg r + \text{const} = -2e \lg \frac{r}{c}$$

Boundary conditions yield

$$P_1 = -2e \lg \frac{R_1}{c}$$

$$P_2 = -2e \lg \frac{R_2}{c}$$

$$\overline{P_1 - P_2 = -2e \lg \frac{R_1}{R_2} = 2e \lg \frac{R_2}{R_1}}$$

Capacitance $= c = \dfrac{e}{P_1 - P_2} = \dfrac{1}{2 \lg \dfrac{R_2}{R_1}}$

Becomes zero when $R_2 = \infty$. Only slightly dependent on the ratio $\cdot \dfrac{R_2}{K_1}$

Electrical reflection of two spheres. [p. 19]
Uniqueness of the solution. Green's theorem.

$$\int\left(\frac{\partial U}{\partial x}\frac{\partial V}{\partial x} + \cdot + \cdot\right)d\tau = -\int U\frac{\partial V}{\partial n}d\sigma - \int U \Delta V d\tau$$

$$\int dy\,dz \int \left(\frac{\partial U}{\partial x}\frac{\partial V}{\partial x}dx = \int \underbrace{dy\,dz}_{-\alpha\sigma\cos nx}\left(U\frac{\partial V}{\partial x}\right) - \int U\frac{\partial^2 U}{\partial x^2}d\tau$$

$$-\int U\left(\frac{\partial V}{\partial x}\cos nx + \cdot + \cdot\right) = \int U\frac{\partial V}{\partial n}d\sigma$$

The above equation is a form of Green's theorem. If we set $U = V$ & $\Delta U = 0$ & on the surface $U = 0$, then $\int\left(\dfrac{\partial U^2}{\partial x} + \cdot + \cdot\right)d\tau = 0$ Provides the proof of uniqueness. It is easy to calc. U in a point if one knows U & $\dfrac{\partial U}{\partial n}$ on the boundary surface of a space.

Electrical Energy

We start again from system of small electrified bodies. First two bodies a & b. Mutual force $\dfrac{e_a \cdot e_b}{r^2} = F$

Components

$$\begin{array}{c|c|c|c}
F\dfrac{x_b - x_a}{r} & dx_b & -F\dfrac{x_b - x_a}{r} & dx_a \\[2ex]
F\dfrac{y_b - y_a}{r} & dy_b & -F\dfrac{y_b - y_a}{r} & dy_a \\[2ex]
F\dfrac{z_b - z_a}{r} & dz_b & -F\dfrac{z_b - z_a}{r} & dz_a \\
\hline
\end{array}$$

$$dA + F\left\{\frac{(x_b - x_a)(dx_b - dx_a) + \cdot + \cdot}{r}\right\} = F\frac{r\,dr}{r} = F\,dr$$

Can also be understood geometrically

But $F = -\dfrac{\partial\Phi_{ab}}{\partial r}$, where $\Phi_{ab} = \dfrac{e_a e_b}{r_{ab}}$.

$$dA = -\frac{\partial\Phi_{ab}}{\partial r_{ab}}dr_{ab} = -d\Phi_{ab}$$

If many masses are present, one obtains the analogous expression, but one has to sum over all combinations.

[p. 20]
$$dA = -\sum\sum d\Phi_{ab} = -d\left\{\sum\sum\Phi_{ab}\right\} = -d\Phi \qquad \Phi = \sum\sum\frac{\varepsilon_a\varepsilon_b}{r_{ab}}.$$

The elementary work is equal to the decrease in the function Φ, which we may call the potential energy of the electric forces or simply potential energy. When doing the double sum, each combination should be counted *once*.

But if one proceeds by first comb. the mass 1 with all the other masses, then mass 2 with all the others etc., then one counts each combination twice; hence one has to set

$$\Phi = \frac{1}{2}\sum\sum\frac{e_a e_b}{r} \quad \text{or} \quad \varphi = \frac{1}{2}\sum e_a\varphi_a$$

The potential energy of a system with continuously distributed masses is to be built in the same way, except that the sums have to be replaced by integrals. One obtains

$$\Phi = \frac{1}{2}\iint\frac{\rho d\tau\,\rho'\,d\tau'}{r}$$

$$\text{or} \quad \Phi = \frac{1}{2}\int\varphi\rho d\tau$$

This expr. is very important, for it permits the calculation of the forces that electrified bodies exert on each other.

We attach to this expression a theoretical analysis. Φ can be decomposed in such a manner that one assigns the energy $\frac{1}{2}\varphi\rho d\tau$ to the individual volume element. Then energy is to be assumed only where el. masses are present, e.g., on the surface. However, [p. 21] the energy can also be localized in another way. That is to say, we have

$$\Phi = \frac{1}{2}\int\varphi\rho d\tau = -\frac{1}{2}\int\varphi\cdot\frac{1}{4\pi}\left(\frac{\partial^2\varphi}{\partial x^2} + \cdot + \cdot\right) = -\frac{1}{8\pi}\int\varphi\Delta\varphi\, d\tau$$

Now, $\varphi\frac{\partial^2\varphi}{\partial x^2} = \frac{\partial}{\partial x}\left(\varphi\frac{\partial\varphi}{\partial x}\right) - \left(\frac{\partial\varphi}{\partial x}\right)^2$

If we integrate & take into account that φ & its derivatives vanish at the limits of int., we get

$$\Phi = \frac{1}{8\pi}\int\left(\frac{\partial\varphi^2}{\partial x} + \cdot + \cdot\right)d\tau = \frac{1}{8\pi}\int(\mathfrak{E}_x^2 + \mathfrak{E}_y^2 + \mathfrak{E}_z^2)d\tau = \frac{1}{8\pi}\int\mathfrak{E}^2 d\tau$$

Here a volume element contributes the term $\mathfrak{E}^2 d\tau$. The energy appears localized in space. Of course, all these expressions for the total energy are equally valid. We find easily the electr. energy of an electrified conductor It is

$$\Phi = \frac{1}{2}\int\varphi\rho d\tau = \frac{1}{2}P\int\rho d\tau = \frac{1}{2}Pe = \frac{1}{2}P^2c = \frac{1}{2}\frac{E^2}{c}$$

Application of the energy law. E in the conductor experiences an infinitely small change

1) through addition of the quantity of electricity dE, along with el. work PdE.
2) through change of shape. mech. work taken up $= dA$

The energy principle yields the equation

$$PdE + dA = d\Phi = \frac{1}{2}(PdE + EdP)$$

The mechanical work $-dA$ done by the system is

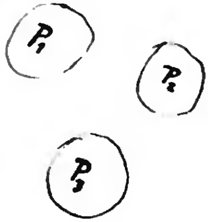
$$-dA = \frac{1}{2}(PdE - EdP)$$

If $dP = 0$, then PdE is el. work supplied. Half of it is converted into mech work. But if $dE = 0$, then

$$dA = \frac{1}{2}EdP = \frac{1}{2}Ed\left(\frac{E}{c}\right) = d\left(\frac{1}{2}\frac{E^2}{c}\right) = d\left(\frac{1}{2}EP.\right)$$

[p. 22] *Some Properties of a System of Conductors*

We imagine that the conductors 1 2 3 .. are charged How do individual potentials depend on the charges?



If φ is a solution, then $\alpha\varphi$ is also such a solution, with the surface densities, and thus also the total charges, being multiplied by α.

We start from the case

$P_1 = 1 \quad P_2 = 0 \ldots\ldots$ let φ_1 be this solution.

$\varphi = P_1\varphi_1$ is then also a solution.

If one defines φ_2 analogously, then

$\varphi = P_2\varphi_2$ is a solution. The φ_1 φ_2 etc. are determined by the conductors alone

$\varphi = P_1\varphi_1 + P_2\varphi_2 \ldots.$ is also a solution.

Thus, φ is homogeneous & linear in the P. The same holds for $\frac{\partial\varphi}{\partial n}$, hence also for the individual $E_1 \ldots E_n$

Thus, we get

$$\left.\begin{aligned} E_1 &= a_{11}P_1 + a_{12}P_2 + \ldots. \\ E_2 &= a_{21}P_1 + a_{22}P_2 + \ldots. \end{aligned}\right\} a$$

Solving for P, we get

$$\left.\begin{aligned} P_1 &= b_{11}E_1 + b_{12}E_2 + \ldots. \\ P_2 &= b_{21}E_1 + b_{22}E_2 + \ldots\ldots \end{aligned}\right\} b$$

$$a_{11}P_1^2 + 2a_{12}P_1P_2 + a_{22}P_2^2$$

must not be negative[15]

$a_{11} \cdot a_{22} - 2a_{12} > 0$

sometimes another form more convenient

$$E_1 = \frac{a_{11} + a_{12}}{2}(P_1 + P_2) + \frac{a_{11} - a_{12}}{2}(P_1 - P_2)$$

$$E_2 = \frac{a_{21} + a_{22}}{2}(P_1 + P_2) + \frac{a_{21} - a_{22}}{2}(P_1 - P_2)$$

[p. 23] The coefficients satisfy a condition that we must derive. If the coefficients are constant, i.e., the position of the body remains unchanged, then ΣPdE must be a total differential. This is the case only when $b_{ik} = b_{ki}$ and $a_{ik} = a_{ki}$. This means ....

One obtains the equivalent expressions

$$\Phi = \frac{1}{2}\sum\sum a_{ik}P_iP_k$$

$$\Phi = \frac{1}{2}\sum\sum b_{ik}E_iE_k$$

From this we get

$E_1 = \dfrac{\partial\Phi}{\partial P_1}$ ...... where Φ is a funct of the P

$P_1 = \dfrac{\partial\Phi}{\partial E_1}$ ...... " " " " " " " E.

We again investigate the work performed. The latter is accompanied by a change in the coefficients. Work supplied $dA = -d\Phi$ at constant E

$$dA_m = -\frac{1}{2}(E_1dP_1 + E_2dP_2 \ldots\ldots)$$

The quantity of work is equally large if P is constant. But one has to increase the potential by dP_1 ..... In doing this el. work is supplied $dA_e = P_1dE_1 + P_2dE_2$ ......

$dA_e - dA_m = d\Phi \qquad \sum PdE - \frac{1}{2}\sum EdP = \frac{1}{2}\sum PdE + \frac{1}{2}\sum EdP$ [16]

Here $P_1dE_1 + \ldots\ldots = \frac{1}{2}(P_1dE_1 \ldots\ldots + EdP_1 \ldots\ldots)$

or $\sum E_1dP_1 = \sum P_1dE_1$
Hence the supplied
el. work is[18]
$dA_e = \Sigma E_1dP_1$
$dA_e = 2dA$.

better this way:[17]
$dA_e = d\Phi + dA_m$
$dA_m = \Sigma PdE - \frac{1}{2}\Sigma PdE - \frac{1}{2}\Sigma EdP$
$dA_m = \frac{1}{2}\Sigma PdE - \frac{1}{2}\Sigma EdP$
For constant potentials
$dA_m = \frac{1}{2}\, dA_e$

4. Example Motion of a Conductor. Plate Condenser

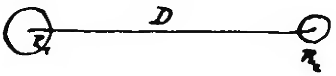
[p. 24]

Examples.

Two spheres whose distance from each other is large compared with their radii. We calculate potentials as funct. of the quantities of electricity (approximately).

$$P_1 = \frac{1}{R_1}E_1 + \frac{1}{D}E_2 = b_{11}E_1 + b_{12}E_2$$

$$P_2 = \frac{1}{D}E_1 + \frac{1}{R_2}E_2 = b_{21}E_1 + b_{22}E_2$$



When the second sphere was moved closer, the potential of the first sphere was increased if the two were similarly charged, while the reverse took place in the opposite case.

$$\Phi = \frac{1}{2}(P_1E_1 + P_2E_2) = \frac{1}{2}(b_{11}E_1^2 + 2b_{12}E_1E_2 + b_{22}E_2^2)$$

We also calculate the constants a, which permit the evaluation of the capacities. To do this, we need only solve the above equations for E_1 & E_2. Setting $\Delta = b_{11}b_{22} - b_{12}^2$, we obtain

$$E_1 = \frac{1}{\Delta}\{b_{22}P_1 - b_{12}P_2\} = a_{11}P_1 + a_{12}P_2$$

$$E_2 = \frac{1}{\Delta}\{-b_{12}P_1 + b_{11}P_2\} = a_{12}P_1 + a_{22}P_2$$

From this we obtain the quantities of electricity for given P. If, for example, $P_2 = 0$ (the second sphere permanently grounded or connected to a casing, we have[19]

$$\frac{E_1}{P_1} = \frac{b_{22}}{\Delta} = \frac{\frac{1}{b_{11}}}{1 - \frac{b_{12}^2}{b_1b_2}} = \frac{R_1}{1 - \frac{R_1R_2}{D}}$$

The presence of the second sphere increases the capacity of the first.[20]

$$\frac{E_2}{P_1} = -\frac{b_{12}}{\Delta} \propto -\frac{b_{12}}{b_{11}b_{22}} = -\frac{R_1R_2}{D}$$

[p. 25] On the second sphere the opposite charge is produced of the approximate magnitude $P_1\frac{R_1R_2}{D}$

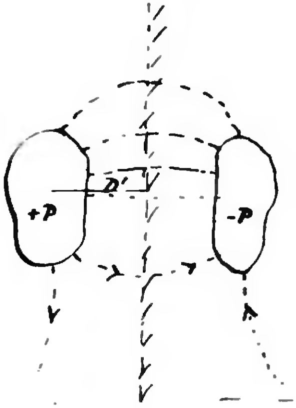
One can reduce the problem of the interaction between a sphere & a conducting plane to the problem of the interaction between two spheres using the principle of the electric mirror-image, which consists in the following: One sees that the case body-plane can always be reduced to the case body-symmetric body.[21] We have

$E_1 = a_{11}P - a_{12}P = (a_{11} - a_{22})P \; D = 2D'$

In our case we have, for example[22]

$$\frac{E_1}{P} = \frac{R_1}{1 - \frac{R_1R_2}{2D'}}$$

The case of two wires at a distance from each other can be treated in a manner very similar to the case of two spheres at a distance from each other, even if a conducting plane is present. All one needs for this is the potential of the electrified line. We shall not go deeper into this.



In those cases in which the field is known, another way of treating the problem, by means of $\Phi = \frac{1}{8\pi}\int \mathfrak{E}^2 d\tau$, is often more advantageous.

Plate Condenser

Only the field between the plates away from the edge is considered.[23]

$$\frac{\partial \mathfrak{E}_x}{\partial x} = 0 \qquad \mathfrak{E}_x = \text{const.}$$

$$\Phi = \frac{1}{8\pi}\mathfrak{E}_x^2 \cdot f\delta$$ [p. 26]

$$\Delta\varphi = -\int \frac{\partial \varphi}{\partial x} dx = \int \mathfrak{E}_x dx = \mathfrak{E}_x \delta$$

Quantity of charge $E = \frac{\mathfrak{E}_x}{4\pi} \cdot f.$ $\qquad 4\pi\eta = \mathfrak{E}_x$

$$\text{Capacity} = \frac{E}{\Delta\varphi} = \frac{f}{4\pi\delta}$$

We find $\Phi = \frac{1}{2}E\Delta\varphi = \frac{1}{2}E(P_1 - P_2) = \frac{1}{2}(P_1 - P_2)^2 C = \frac{1}{2}\frac{E}{c}$

We find the attractive force by constructing

$$\left(\frac{\partial \Phi}{\partial \delta}\right)_E = \frac{1}{2}E^2 \frac{\partial}{\partial \delta}\left(\frac{4\pi\delta}{f}\right) = \frac{1}{2}E^2 \frac{4\pi}{f} = \frac{1}{2}E^2 \cdot \frac{1}{c\delta} = \frac{1}{2}EP \cdot \frac{1}{\delta} = \frac{1}{2}P^2 \frac{f}{4\pi\delta^2} = \frac{1}{8\pi}\mathfrak{E}_x^2 f$$

Thus, the force per unit surface area is $\frac{1}{8\pi}\mathfrak{E}_x$.

This last law can be derived quite generally. We seek the force that acts on the charge of unit surface. We think of the latter as being of finite thickness.

$$\text{Force} = \int \rho \mathfrak{E}_x dx \qquad \text{where} \quad \frac{\partial \mathfrak{E}_x}{\partial x} = 4\pi\rho$$

$$= \frac{1}{4\pi}\int \mathfrak{E}_x \frac{\partial \mathfrak{E}_x}{\partial x}\,dx = \frac{1}{8\pi}\mathfrak{E}_x^2 \approx \frac{1}{2}\eta\mathfrak{E}_x$$

Thus, it is as if the outer normal force acted on the whole layer.

How does P change if one varies δ at constant E.

$P = \frac{E}{c} = E \cdot \frac{4\pi\delta}{f}$. The potential difference varies as δ. Means of increasing the potential difference by expenditure of work. How does the potential change in the case of circular plates if one increases the distance to ∞ for constant quantities of electricity?

$$E = P_1 C_1$$

$$E = P_2 C_2$$

$$\frac{P_2}{P_1} = \frac{C_1}{C_2} = \frac{\dfrac{f}{4\pi\delta}}{\dfrac{2R}{\pi}} = \frac{f}{8\delta R} = \frac{\pi}{8}\frac{R}{\delta}$$ [p. 27]

$$\Phi = \frac{1}{8\pi}\mathfrak{E}^2 f\delta = \frac{f}{8\pi}\frac{P^2}{\delta}^{[24]}$$

$$\left(\frac{P}{\delta}\right)^2$$

$$\frac{\partial\Phi}{\partial\delta} = -\frac{f}{8\pi}\frac{P^2}{\delta^2}$$

Application for the detection of small potential differences by electrostatic means. (Volta's experiment.)

Absolute measurement of potentials by the "guard ring" electrometer.[25]

$$\text{force} = \frac{\partial}{\partial\delta}\left(\frac{1}{2}\frac{E^2}{C}\right) = \frac{1}{2}\mathfrak{E}^2\frac{4\pi}{f}$$

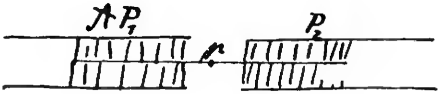
$$= 2\pi \cdot P^2 \cdot \frac{f^2}{(4\pi\delta)^2} \cdot \frac{4\pi}{f} = \frac{1}{8\pi}\frac{f}{\delta^2}P^2$$

Since the force can be measured absolutely, & so too f and δ, the same is true of P.

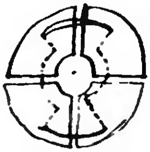
Kelvin's Quadrant Electrometer for the Measurement of Voltages and Small Quantities of Electricity[26]

$$\Phi = \frac{1}{2}\kappa(P_1 - p)^2(a - x)$$

$$+ \frac{1}{2}\kappa(P_2 - p)^2(a + x)^{[27]}$$



Interests us only insofar as it is dependent on x



$$\Phi = \kappa\{(P_2 - p)^2 - (P_1 - p)^2\}x$$

$$D = \left(\frac{\partial \Phi}{\partial x}\right)_{P_1} = \kappa\{(P_2 - p)^2 - (P_1 - p)^2\}$$

Most important circuits 1) $p = P_1$ $D = \kappa(P_2 - P_1)^2$ Quadratic instrument

2) Needle at auxiliary potential p. $P_2 = P - \frac{\alpha}{2}$ $P_1 = P + \frac{\alpha}{2}$

$$D = 2\kappa\alpha(p - P)$$

If p large compared with P, then instrument is linear.

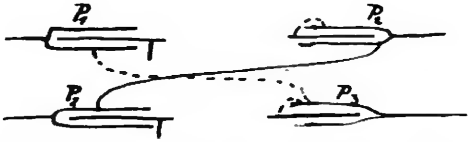
$$\frac{f}{8\pi}\frac{P^2}{\delta} \sim \frac{1}{12}\cdot\frac{\left(\frac{1}{3}\right)^2}{0.1}$$
$$\sim \frac{1}{10}$$

Maschinchen[28] *& Thomson's Multiplier.* [p. 28]

$$e = P_1\, C$$
$$e = P\, C' \quad P = P_1 \frac{C}{C_1}$$

Even stronger amplification if stirrup $P = P_1 \frac{C}{c}$

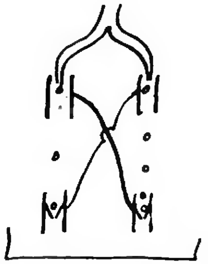
Repeat *Maschinchen*



$$P_3 = P_1 \left(\frac{C}{c}\right)^2$$

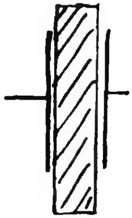
If one more connection, then Thomson's multiplier[29]

Drop multiplier



All induction machines are based on this principle.

Dielectrics



Experience. Voltage on condenser plates drops if a nonconductor is inserted between them. Conversely, if the voltage is constant, the quantity of electricity increases in a specific ratio. This ratio is characteristic of the (homogeneous) nonconductor in question. It is called the dielectric constant. <The theory that has been put forward thus far can be maintained for this case if one imagines that electricity [p. 29] possesses limited mobility in the dielectric. Neutral molecules become dipoles>

We can now distinguish two kinds of field strength

1) Field strength between plates & dielectric or $\perp$ lines of force in an arbitrary gap. ($\mathfrak{D}$)

2) Field strength in a channel connecting the plates $\perp$. The latter is equal to $\frac{P_1 - P_2}{\delta} = -\frac{\partial\varphi}{\partial x}$, if x is the direction of the axis. As before, we denote this kind of field strength by $\mathfrak{E}$. The relation $\mathfrak{D} = \varepsilon\mathfrak{E}$ is generally valid everywhere in the dielectric.

It is easy to calculate the energy of such a system. We have $d\Phi = PdE$, if the plates are immovable.

But according to the special form of Gauss's law

$$\mathfrak{D} = 4\pi\sigma = 4\pi\frac{E}{f} \qquad dE = \frac{f}{4\pi}d\mathfrak{D}$$

$$P = -\int_0^\delta \mathfrak{E}dx = \mathfrak{E}\cdot\delta$$

$$d\Phi = \frac{f\delta}{4\pi}\mathfrak{E}d\mathfrak{D}$$

Integrating, we get $\Phi = \frac{V}{8\pi}\varepsilon\mathfrak{E}^2 = \frac{V}{8\pi}\mathfrak{E}\mathfrak{D}$

Accordingly, we generalize the earlier expression for the energy to

$$\Phi = \frac{1}{8\pi}\int\mathfrak{E}\mathfrak{D}d\tau = \frac{\varepsilon}{8\pi}\int\mathfrak{E}^2d\tau$$

Intuitive representation by means of dipoles which strive to bond with one another by elastic forces. [p. 30]

$$\left.\begin{aligned}\mathfrak{D}_x &= \varepsilon\mathfrak{E}_x\\ \mathfrak{D}_y &= \varepsilon\mathfrak{E}_y\\ \mathfrak{D}_z &= \varepsilon\mathfrak{E}_z\end{aligned}\right\}1 \qquad \text{In vector notation abbreviated } \mathfrak{D} = \varepsilon\mathfrak{E}$$

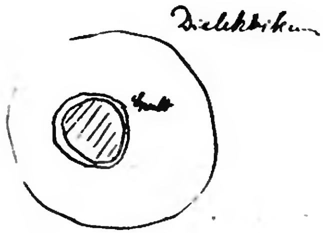
$\mathfrak{D}_x$ number of the electric lines of force (field strength) through gap perpendicular to the X axis etc. $\mathfrak{E}_x$ field strength in channel parallel to X-axis. What kinds of laws hold for the vectors $\mathfrak{D}$ and $\mathfrak{E}$ inside a dielectric?

1) $\mathfrak{E}$ derivable from the potential.

$$\left.\begin{aligned}\mathfrak{E}_x &= -\frac{\partial\varphi}{\partial x}\\ \mathfrak{E}_y &= -\frac{\partial\varphi}{\partial y}\\ \mathfrak{E}_z &= -\frac{\partial\varphi}{\partial z}\end{aligned}\right\}(2) \qquad \text{Intuitive model.}$$

or, as verified by differentiation

$$\left.\begin{aligned} \frac{\partial \mathfrak{E}_z}{\partial y} - \frac{\partial \mathfrak{E}_y}{\partial z} &= 0 \\ \frac{\partial \mathfrak{E}_z}{\partial x} - \frac{\partial \mathfrak{E}_x}{\partial z} &= 0 \\ \frac{\partial \mathfrak{E}_x}{\partial y} - \frac{\partial \mathfrak{E}_y}{\partial x} &= 0 \end{aligned}\right\} \quad (2a)$$



2) A surface containing a great number of unbroken dipoles set up in the gap. Here Gauss's law holds.

$$\int (\text{normal component of the field strength in the gap}) \, . \, d\sigma = 0$$

$$\int \mathfrak{D}_n \, d\sigma = 0 \, ... \, (3)$$

[p. 31] If one applies this law to a parallelepiped that is enclosed in the gap, one obtains

$$\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} = 0 \quad .... \, (3a)$$

What conditions hold on the boundary between two dielectrics? The constancy of φ, and hence also the constancy of the tangential components of $\mathfrak{E}$, holds here as well

$$\mathfrak{E}_{t1} = \mathfrak{E}_{t2} \, ... \quad (2c)$$

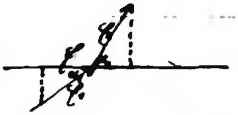
If one chooses a relatively infinitely low cylinder whose bases are separated by the boundary surface, and applies to its boundary the generalized form of Gauss's law, one obtains

$$\mathfrak{D}_{n1} = \mathfrak{D}_{n2} \, ... \quad (3b)$$

Refraction of the lines of force at the boundary between two media

$$\mathfrak{E}_{t1} = \mathfrak{E}_{t2}$$

$$\varepsilon_1 \mathfrak{E}_{n1} = \varepsilon_2 \mathfrak{E}_{n2} \qquad \varepsilon_1 \operatorname{tg} \alpha_1 = \varepsilon_2 \operatorname{tg} \alpha_2$$



$$\frac{\operatorname{tg} \alpha_2}{\operatorname{tg} \alpha_1} = \frac{\varepsilon_1}{\varepsilon_2}$$

Case where movable electric quantities of spatial density ρ are also present.
(1) and (2) are valid here too

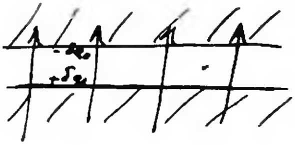
3) becomes $\int \mathfrak{D}_n d\sigma = -4\pi \int \rho d\tau$ (3^x)

$$\frac{\partial \mathfrak{D}_x}{\partial x} + \frac{\partial \mathfrak{D}_y}{\partial y} + \frac{\partial \mathfrak{D}_z}{\partial z} = 4\pi\rho \quad (3a^x)$$

The Meaning of Dielectric Displacement According to the Electron Theory[30] [p. 32]

The circumstance that the dielectric displacement and the electrical field strength are different in the interior of insulators has been attributed to the limited motion of elastically bound electricity. We investigate the meaning of $\mathfrak{D}$ according to this conception. Positive as well as negative el. in nonelectrified state density ρ_0. $\mathfrak{D} - \mathfrak{E}$ is produced by the field in the gap through the action of bound electricity. If δ is now the displacement of the positive el. in the insulator, then

$\delta \rho_0$ is posit. coating below
$-\delta \rho_0$ neg. coating on top.



Each sends out $4\pi\delta\rho_0$ lines of force, hence, $2\pi\delta\rho_0$ to one side Both together $4\pi\delta\rho_0$ in the gap. Thus, we have

$$\mathfrak{D} - \mathfrak{E} = 4\pi\delta\rho_0$$

Nothing changes here if we assume that the electricity in the dielectric is distributed in discrete quantities ± e. Then $\rho_0 = ne$

$$\delta\rho_0 = n\underbrace{e\delta}_{\mu} = n\mu = \mathfrak{P}$$

Thus one obtains

$$\mathfrak{D} - \mathfrak{E} = 4\pi\mathfrak{P}.$$

[p. 33]

Electrostatic Energy

$$\frac{1}{8\pi}\int(\mathfrak{E}_x\mathfrak{D}_x + \mathfrak{E}_y\mathfrak{D}_y + \cdot)d\tau = \Phi = \frac{1}{2}\int\varphi\rho d\tau$$ | Extended over ∞ space.

Another form.

$$-\int\left(\frac{\partial\varphi}{\partial x}\mathfrak{D}_x + \cdot + \cdot\right) = +\int\left(\varphi\frac{\partial\mathfrak{D}_x}{\partial x} + \cdot + \cdot\right) = 4\pi\int\varphi\rho\, d\tau.$$

Thus, the second form of the energy expression also holds unchanged. The uniqueness proof for the conductor problem in the case where arbitrary uncharged dielectrics are present is also easy to carry out. We think of the dielectrics as being *uniformly* distributed. In that case φ & $\frac{\partial\varphi}{\partial x}$ etc. are constant in the entire space except on the conductor surfaces.

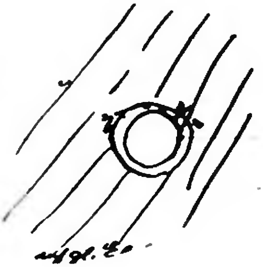
$$\int(\mathfrak{E}_x\mathfrak{D}_x + \cdot + \cdot) = +\int(\varphi\mathfrak{D}_n)d\sigma + \int\varphi\rho d\tau$$

In the domain of integration $\rho = 0$. on the boundaries $\varphi = 0$ for difference solution. Thus, the left side = 0, which is a sum of positive magn.

Charged sphere. Generalization of Coul. law. Forces calculable from Φ using the energy principle.

$$K = \frac{ee'}{r^2}\cdot\frac{1}{\varepsilon}$$

<Energy of a> charged sphere <*P*> in the dielectric.



$$4\pi\sigma = \mathfrak{E}_{PL} = \mathfrak{D}_{\text{dielectr.}} = \mathfrak{E}_D\varepsilon$$

$$\mathfrak{E}_D = \mathfrak{E}_P\cdot\frac{1}{\varepsilon} \qquad \varphi_D = \varphi_L\cdot\frac{1}{\varepsilon}$$

$$\Phi = \frac{1}{\varepsilon}\Phi_L$$

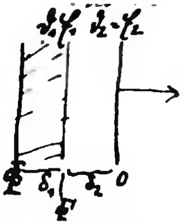
It requires less energy to charge a sphere in the dielectric to the same quantity of electricity, more energy to charge it to the same voltage.

$$\mathfrak{D}_1 = \mathfrak{D}_2 \qquad \cancel{4}\pi R^2 \mathfrak{D} = \cancel{4}\pi e$$

$$\mathfrak{D} = \frac{e}{r^2} \qquad \mathfrak{E}_{(1)} = \frac{1}{\varepsilon}\frac{e}{r^2} \qquad \mathfrak{E}_{(2)} = \frac{e}{r^2} = -\frac{\partial\varphi}{\partial r}$$

$$\Phi = \int_{\infty}^{R'} \frac{e}{r^2}dr + \int_{R'}^{R} \frac{1}{\varepsilon}\frac{e}{r^2}dr = e\left\{\frac{1}{R'} + \varepsilon\left(\frac{1}{R} - \frac{1}{R'}\right)\right\}$$

Plate condenser partly with air, partly with dielectric [p. 34]



$$\mathfrak{E}_1 = -\frac{\partial\varphi}{\partial x} \qquad \mathfrak{E}_2 = \frac{\partial\varphi}{\partial x}$$

$$\mathfrak{D}_1 = \varepsilon\mathfrak{E}_1$$

<$4\pi\sigma = \mathfrak{D}_1$> boundary condition[31] $\varphi = \Phi$ for $x = 0$

$\varphi = 0$ for $x = \delta_1 + \delta_2$

$\mathfrak{D}_1 = \mathfrak{E}_2$ for $x = \delta$

Vectors spatially constant

$$\Phi - \Phi' = -\int_0^{\delta_1} \frac{\partial\varphi}{\partial x}dx = \mathfrak{E}_1\delta_1$$

$$\Phi' - 0 = -\int \frac{\partial\varphi}{\partial x}dx = \mathfrak{E}_2\delta_2$$

$$\Phi = \mathfrak{E}_1\delta_1 + \mathfrak{E}_2\delta_2$$

$$\mathfrak{D}_1 = \mathfrak{E}_2 = \varepsilon\mathfrak{E}_1$$

$$\Phi = \mathfrak{E}_2\left(\delta_2 + \frac{\delta_1}{\varepsilon}\right) = \mathfrak{D}_1\left(\delta_2 + \frac{\delta_1}{\varepsilon}\right)$$

Charges $-\frac{1}{4\pi}\mathfrak{E}_2 \quad \frac{1}{4\pi}\mathfrak{D}_1 = \eta$, thus, equally large.

$$\Phi = 4\pi\eta\left(\delta_2 + \frac{\delta_1}{\varepsilon}\right) = 4\pi\frac{E}{f}\left(\delta_2 + \frac{\delta_1}{\varepsilon}\right)$$

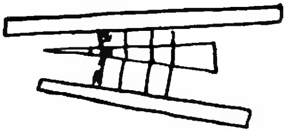
$$C = \frac{f}{4\pi\left(\delta_2 + \frac{\delta_1}{\varepsilon}\right)}$$

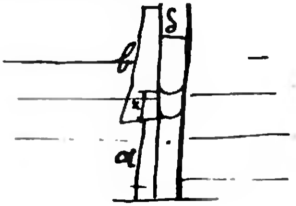
static methods for the determination of the dielectric const. Comparison of condenser potentials when the charge is the same.

Force proportional to ε when voltage given. From this, the dielectric constant of liquids.

[p. 35] Rising of liquids between plates. Perot, refraction of the lines of force.[32]

$$\frac{tg\ \alpha}{tg\ \beta} = \frac{\varepsilon}{\varepsilon_0}$$





α = height of rise without field.
$\alpha + x$ = height of rise with field

Pot. energy of gravity $\delta \cdot (<\alpha + >x) \cdot \rho \cdot \frac{\alpha + x}{2g}$

$$\Phi_g = \delta \frac{(<\alpha> + x)^2}{2} \rho g$$

$$\Phi_e = \frac{1}{8\pi} \mathfrak{E}^2 (\delta(\alpha + x)\varepsilon + \delta(b - x)\varepsilon_0)$$

$d\Phi_e$ = work of the el. forces
$-d\Phi_e$ = work of the grav. forces
Sum must be zero.

$$\delta(<\alpha + >x)\rho g - \frac{1}{8\pi} \mathfrak{E}^2 <\delta> (\varepsilon - \varepsilon) = 0$$

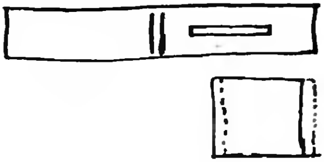
absolute measurement of $(\varepsilon - 1)$.
Better directly with force.

Volta Effect. Electric Double Layer.

Magnetism.

Coul.'s Law Unit of Pole Strength. Potential Laplace's Theorem

Intuitive meaning of the magnetization constants. Let us have a homogeneous isotropic material in the shape of a bar. Let it be uniformly magnetized. Displacement δ



Density of the polarization electricity[33] of either sign ρ_0 How large are H and $\mathfrak{B}$? Material surface perpendicular to the x axis. $\rho_0\delta$ positive <electricity> magnetism has traversed unit surface area Coatings of density $\rho_0\delta$. When surface slanted, then $\rho_0\delta\cos\varphi$ exit per unit surface. If molecular model, then $\rho_0 = \mu \cdot n$ Density of coating $\mu n\delta\cos\varphi = \mathfrak{Q}\cos\varphi$ $\mathfrak{Q}$ polarization. Channel walls do not have magnetic covering. End surfaces can [p. 36] be neglected. Thus, magnetic field str. in the interior of the channel the same as in the channel.

But in the gap the coverings do send out lines of force. From Gauss's law directly $B - \mathfrak{H} = 4\pi<\mathfrak{Q}>$. Exactly as with dielectrics. There are no true magnetic masses. From this it follows that

$$\int \mathfrak{B}_n d\sigma = 0. \quad or \quad \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0$$

$\mathfrak{H}$ derivable from a potential. Potential of the densities of bound magnetism.

$$\mathfrak{H} = -\frac{\partial\varphi}{\partial x} \text{ etc.}$$

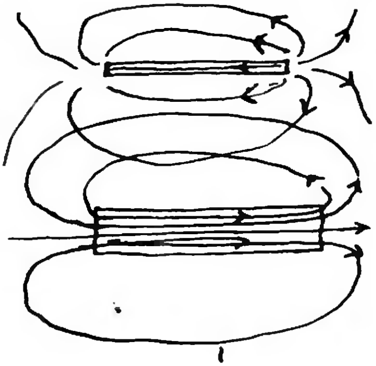
These are our fundamental laws.
Where is the density of the bound magnetism located? Surface in subst.

$$4\pi \int \rho_g d\tau = \int \mathfrak{H}_n d\sigma = -\int \mathfrak{P}_n d\sigma$$

$$\rho_g = -\left(\frac{\partial \mathfrak{P}_x}{\partial x} + \frac{\partial \mathfrak{P}_y}{\partial y} + \frac{\partial \mathfrak{P}_z}{\partial z}\right)$$

Parallel magnetized iron bar 

Magnetic coverings of bound magnetism $= f\mathfrak{P}_x$
The fields $\mathfrak{H}$ and $\mathfrak{B}$ are here independent of each other in the magnet $\mathfrak{H}$ and $\mathfrak{B}$ are differently oriented in the magnet.



[p. 37] Magnetic force exerted on each other by very closely neighboring surfaces.

$$\eta_g = \frac{1}{4\pi}(\mathfrak{H}_a - \mathfrak{H}_i)$$

$$4\pi\rho = \frac{\partial\mathfrak{H}_x}{\partial x} + \underset{0}{\frac{\partial\mathfrak{H}_y}{\partial y}} + \underset{0}{\frac{\partial\mathfrak{H}_z}{\partial z}}$$

$$\int\rho\mathfrak{H}_x\,dx = \frac{1}{4\pi}\int\mathfrak{H}_x\frac{d\mathfrak{H}_x}{dx}dx = \frac{1}{8\pi}(\mathfrak{H}_a^2 - \mathfrak{H}_i^2)$$

$$K = \frac{1}{8\pi}(\mathfrak{B}_i^2 - \mathfrak{H}_i^2)$$

If iron, then $\mathfrak{B}_i$ larger than $\mathfrak{H}_i$, so that approx. $K = \frac{1}{8\pi}\mathfrak{B}_i^2$.

This must be eq. to the magnetic energy of the unit volume.
Can be very large. $\mathfrak{B}_i = 20{,}000$ $K \sim 2.10^2$ = ca 20 kg per cm^2.

Energy of the Magnetic Field

a) in vacuum $\sum\sum\frac{\mu\mu'}{r} = \frac{1}{2}\sum\mu\varphi$

This is $g\frac{1}{8\pi}\mathfrak{H}^2$

b) If $\mu \neq 1$, then this energy has another value. One can calculate this value by taking into account that work must also be expended for the displacement.

The posit. magn. masses of the unit volume are subjected to the total force $\mathfrak{H}\rho_0$ The work expended on the displ. $d\mathfrak{d}$ is $\mathfrak{H}\rho_0 d\mathfrak{d}$. This is equal to $\mathfrak{H}\,d\mathfrak{P}$ because one can set $\rho_0\mathfrak{d} = \mathfrak{P}$.

In the process, magnetic energy in the vacuum increases by $\frac{1}{4\pi}\mathfrak{H}\,d\mathfrak{H}$.

The two combined

$$\frac{1}{4\pi}\mathfrak{H}d(\mathfrak{H} + 4\pi\mathfrak{P}) = \frac{1}{4\pi}\mathfrak{H}\,d\mathfrak{B}.$$

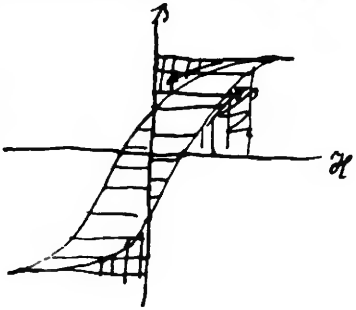
If $\mathfrak{B} = \mu\mathfrak{H}$, then this is integrable $\frac{1}{4\pi}\int\mu\mathfrak{H}\,d\mathfrak{H}$

If μ = const, then

$$\frac{1}{8\pi}\mu H^2$$ [p. 38]

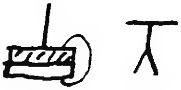
per unit volume. As a matter of fact, one may designate this energy as "magnetic."

But things are different in the case where no relat. exists between $\mathfrak{H}$ and $\mathfrak{B}$. In that case, too, $Hd\mathfrak{B}$ is the work supplied to unit volume. But this work need not represent an available store. Surface of the "hysteresis curve" represents the energy lost in a cyclic process. This energy is converted to heat.

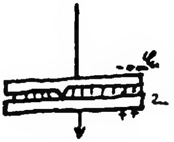


The Volta Effect - Electromotive Forces

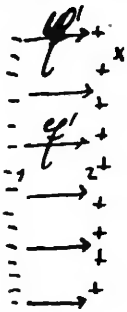
In the arrangement shown in the accompanying sketch, one observes an el. field between the plates. Such a field would not to be expected according to the theory employed so far. Potential diff. cannot arise in the interior of metals. Hence they must arise on the boundary surfaces. Let us first assume that the potential jump occurs more or less at the contact surface of the metals—later on this will turn out not to be valid. Volta discovered.



Mechanics of the effect



Contradiction with the theory employed so far. How can the latter be extended so [p. 39] that in agreement with experience At the surface of separation, electricities are acted upon by a force that separates them. We will conceive of this as a field that has an external source, not originating from el. masses.



This "impressed" force seeks to move positive electr. to the right.[34] Equilib. can exist only if the effect of $\mathfrak{E}'$ is compensated by an opposite electrost. field.

$$\mathfrak{E} + \mathfrak{E} = 0 \quad \mathfrak{E}' - \frac{\partial\varphi}{\partial x} = 0 \quad \varphi_2 - \varphi_1 = \int \mathfrak{E}'\,dx.$$

Thus, there is a potential jump at the surface. How is it produced?

$$\frac{\partial\mathfrak{E}_x}{\partial x} = 4\pi\rho = -\frac{\partial\mathfrak{E}'}{\partial x}$$

Thus, we have two opposite coverings. If $\mathfrak{E}'$ is constant inside the layer, then these coverings are planar $(\eta)\,|\mathfrak{E}| = 4\pi\eta$ Gauss's law.

$$\int \mathfrak{E}'dx = \Delta\varphi = \mathfrak{E}'\delta = 4\pi\underbrace{\eta\delta}_{\zeta}\ \zeta \text{ moment of the unit surface area of the double layer}$$

Double layer <corresponds> not an arbitr. theory but demanded directly by experience. Is very dependent on the constitution of the surface—especially water layer. Can be removed almost completely with removal of the latter. Thus, is located in the surface facing the air.

[p. 40] If instead of air, water between the plates, then also field.

But because water conductor, el. moves in water. Current arises. In accordance with the law of conservation of el., such a current must also flow in the metals, so that no excess would arise anywhere. Chem. processes on electr. Since we have already meas. the el. unit, the unit of the el. current is also def. (Number of electrost. units flowing through the conductor per second. The direction of the current is the direction in which the positive electricity flows.

Magnetic Field of Currents

Current acts on magnetic needle. What is the constitution of the magnetic field outside the conductor?

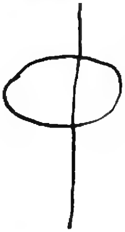
Let the conductors be surrounded by vacuum (or air). For such a case we have found that

$$\mathfrak{H}_x = -\frac{\partial\varphi}{\partial x}\cdot\cdot$$

$$\frac{\partial\mathfrak{H}_x}{\partial x} + \frac{\partial\mathfrak{H}_y}{\partial y} + \frac{\partial\mathfrak{H}_z}{\partial z} = 0$$

If the concept of magnetic field has a general meaning, then these equations must hold here as well.

$\mathfrak{H}$ derivable from a potential. In such a case we have seen until now that the line integral of the (magn.) field strength over a closed curve always vanished.



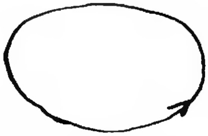
But it does not take much to see that the magn. lines of force surround [p. 41] an el. current. Thus, if we form the line integral $\int(\mathfrak{H}_x dx + \mathfrak{H}_y dy + \mathfrak{H}_z dz) = \int \mathfrak{H} ds \cos(\mathfrak{H} ds)$, we certainly do not obtain zero.

This notwithstanding, our above formulas may be right

$$\int(\mathfrak{H}_x dx + \cdot + \cdot) = -\int d\varphi = \varphi_1 - \varphi_2$$

This quantity must vanish, then, for a closed path only if φ is a *single-valued* spatial function. How must fields be constituted for φ to become multivalued? In order to resolve this, we investigate the closed line integral of an arbitrary vector.

Stokes's Theorem

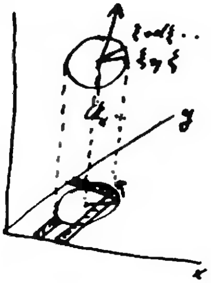


Vector $\mathfrak{A}_x \mathfrak{A}_y \mathfrak{A}_z$

Line integral $\int \mathfrak{A}_x dx + \mathfrak{A}_y dy + \mathfrak{A}_z dz$

Decomposed into so many such integrals over ∞ small surfaces which can be regarded as planar.

Over this $\int \mathfrak{A}_x dx$



$$\int \mathfrak{A}_x\, dx = \int \left(\mathfrak{A}_{x_0} + \frac{\partial \mathfrak{A}_x}{\partial \xi_0}\xi + \frac{\partial \mathfrak{A}_x}{\partial \eta_0}\eta + \frac{\partial \mathfrak{A}_x}{\partial \zeta_0}\zeta \right) d\xi$$

$$= \mathfrak{A}_{x_0} \underset{\substack{| \\ 0}}{\int d\xi} + \frac{\partial \mathfrak{A}_x}{\partial \xi} \int \underbrace{\xi\, d\xi}_{0} + \frac{\partial \mathfrak{A}_x}{\partial \eta} \underset{-d\sigma \cos nz}{\int \eta\, d\xi} + \frac{\partial \mathfrak{A}_x}{\partial \zeta} \underset{+d\sigma \cos ny}{\zeta\, d\xi}$$

$$\begin{array}{l|ll}
\int_{d\sigma} \mathfrak{A}\, dx = +d\sigma & \frac{\partial \mathfrak{A}_x}{\partial \zeta}\cos ny - \frac{\partial \mathfrak{A}_x}{\partial \eta}\cos nz & \left(\frac{\partial \mathfrak{A}_z}{\partial y} - \frac{\partial \mathfrak{A}_y}{\partial z} \right) \cos nx \\
\text{— — — — — — —} & \frac{\partial \mathfrak{A}_y}{\partial \xi}\cos nz - \frac{\partial \mathfrak{A}_y}{\partial \zeta}\cos nx & \text{— — — — — — — — — —} \\
\text{— — — — — — —} & \frac{\partial \mathfrak{A}_z}{\partial \eta}\cos nx - \frac{\partial \mathfrak{A}_z}{\partial \xi}\cos ny & \text{— — — — — — — — — —}
\end{array}$$

The integral is thereby converted into a surface integral.

[p. 42]

Elementary Derivation of the Properties of the Magnetic Field

For a field of permanent magnets we have

$$\frac{\partial \mathfrak{H}_x}{\partial x} + \frac{\partial \mathfrak{H}_y}{\partial y} + \frac{\partial \mathfrak{H}_z}{\partial z} = 0$$

and

$$\mathfrak{H}_x = -\frac{\partial \varphi}{\partial x} \qquad \mathfrak{H}_y = -\frac{\partial \varphi}{\partial y} \qquad \mathfrak{H}_z = -\frac{\partial \varphi}{\partial z}$$

If φ is a single-valued function, this means that the line integral of $\mathfrak{H}$ along a closed curve is zero.

We now assume that the lines of force around a rectilinear current path are circles. How must then the field decrease with the distance? In the space that is simply connected & outside the circuit, the field of magn. & the field of current shall not be distinguishable from each other. How must then the field strength depend on the dist?

$$\text{Line integral} = \mathfrak{H}(\overbrace{r+dr}^{r'})\cdot(\overbrace{r+dr}^{r'})\,d\varphi - \mathfrak{H}(r)r\,d\varphi = 0$$

$$\mathfrak{H}(r')\cdot r' = \mathfrak{H}(r)\cdot r$$

If we let r vary for constant r', then

we obtain $\mathfrak{H}(r)\cdot r = \text{const.}$ $\mathfrak{H}(r) = \dfrac{\text{const.}}{\text{r}}$

This law is confirmed by experience. The const. depends on the strength of the current. It can serve as a measure of the current strength. We stipulate const $= 2i$, and thereby obtain a definition for the current strength

$$\mathfrak{H} = \frac{2i}{r}$$

i is then equal to 1 if the current produces field strength 2 at dist. of 1 cm. This dependence on r is confirmed by experience.

If we integrate $\int \mathfrak{H}ds$ along a circle around the current path, we obtain $\int \frac{2i}{r}\cdot r d\varphi = 4\pi i$, [p. 43]

thus independent of r. But this is valid not only for a circular path but for any arbitrary path.

$$\mathfrak{H}_s ds = \mathfrak{H}_t \cdot r d\varphi = \frac{2i}{r} r d\varphi = 2i d\varphi$$

$$\int \mathfrak{H}_s ds = 4\pi i.$$

Potential[35] $\int \mathfrak{H}_s ds = -\int \frac{\partial\varphi}{\partial s} ds = 2i \int d\vartheta$

$$d\varphi = -2i d\vartheta$$

$$\varphi = -2i\vartheta + \text{const.}$$

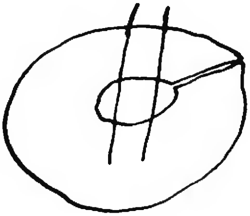
Now, φ is not a single-valued function, because ∞ many angles ϑ belong to one location. If many currents, then pot ∞ multivalued Differential equations of the magnetic force derived from that.

Arbitrarily Distributed Currents (Rigorous Analysis)

We start off from Stokes's theorem

$$\int \mathfrak{H}_x dx + \mathfrak{H}_y dy + \mathfrak{H}_z dz = \int \left\{ \left(\frac{\partial \mathfrak{H}_z}{\partial y} - \frac{\partial \mathfrak{H}_y}{\partial z} \right) \cos nx + \cdot + \cdot \right\} d\sigma$$

If in all points of the plane $\frac{\partial \mathfrak{H}_z}{\partial y} - \frac{\partial \mathfrak{H}_y}{\partial z}$ etc $= 0$, then the integral vanishes over every closed curve. But this is not at all the case if the current is twisted around. In that case, however, the integral over the curve shown in the sketch vanishes. $\int \mathfrak{H} ds$ cos $\mathfrak{H} ds$ is independent of the integration path. In engineering this quantity is called the "magnetomotive force". We set this quantity equal to $4\pi i$. We set the current density to be $\mathfrak{i}_x \mathfrak{i}_y \mathfrak{i}_z$, then the electricity flowing through $d\sigma$ per unit time is $(\mathfrak{i}_x \cos nx + \mathfrak{i}_y \cos ny + \mathfrak{i}_z \cos nz)\, d\sigma$, so that we have



$$4\pi \int (\mathrm{i}_x \cos nx + \cdot + \cdot) d\sigma = \int \left\{ \left(\frac{\partial \mathfrak{H}_z}{\partial y} - \frac{\partial \mathfrak{H}_y}{\partial z} \right\} \cos nx + \cdot + \cdot \right\} d\sigma$$

Holds also for an infinitely small surface.

$$4\pi \mathrm{i}_x = \frac{\partial \mathfrak{H}_z}{\partial y} - \frac{\partial \mathfrak{H}_y}{\partial z}$$

\- - - - - - - - - - - - - - -

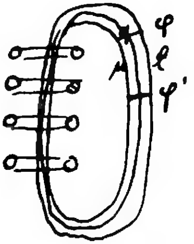
\- - - - - - - - - - - - - - -

[p. 44]

Applications of the Integral Law

$4\pi\mathfrak{i} = \int \mathfrak{H}_s\, ds$, if current wound once

$4\pi n\mathfrak{i} = \int \mathfrak{H}_s\, ds$ " " " n times



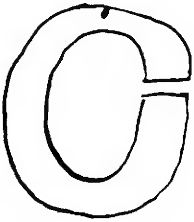
$$\mathfrak{H}_s = \frac{\varphi - \varphi'}{l}$$

$$\Sigma\,(\varphi - \varphi') = 4\pi ni$$

$$= \Sigma\, \mathfrak{H}_s l = \Sigma \frac{1}{\mu} \mathfrak{B}_s l$$

$$= \Sigma \frac{F}{f} \cdot \frac{l}{\mu} = F \cdot \Sigma \frac{l}{\mu f} = 4\pi ni$$

Σ is called the magnetic resistance of the line-of-force tube. $F \int \mathfrak{B}_n df$ = flux



Solenoid inside & outside. Pole str. of the solenoid

$4\pi ni = \mathfrak{H} \cdot l$

$\frac{4\pi ni}{l} q$ = number of lines of force.

$\frac{niq}{l}$ = pole strength.

Determination of the field when the position of the currents is given. $\mathfrak{H}_1$ & $\mathfrak{H}_2$ two solutions. Difference $\mathfrak{H}$.

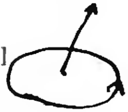
Everywhere $\frac{\partial \mathfrak{H}_z}{\partial y} = \frac{\partial \mathfrak{H}_y}{\partial z}$ ...

Then $\int (\mathfrak{H}_x dx + \cdot + \cdot)$ indep. of integr path $= -\varphi$. Then $\mathfrak{H}_x = -\frac{\partial \varphi}{\partial x}$ ... Thus, in the entire space φ dependent on single-valued potential.[36] $\int \left(\frac{\partial \varphi^2}{\partial x} + \cdot + \cdot \right) d\tau = -\int \varphi \Delta \varphi d\tau = 0$ (at least if no iron pres) Thus, φ = const. Thus, uniquely determined holds also if bodies with $\mu \neq 1$ are present.

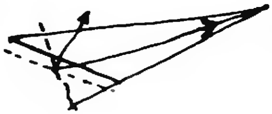
Magnetic Potential of a Current (Ampère)

What is sought is the potential function, which changes by $4\pi i$ for one circle around the current.

[p. 45]



Double layer considered[37]



$$\int d\sigma \cdot \eta \cdot \frac{\partial \frac{1}{r}}{\partial n}\delta = \zeta \int d\sigma \frac{1}{r^2} \cos nr$$
$$= \zeta d\kappa$$

For finite angle $\zeta\kappa$

Potential changes by $4\pi\xi$ for one revolution, no matter at which point one starts and to which point one gets. Current replaceable by double layer of moment ζ. Holds only outside the double layer.

Action-at-Distance of Circuits from Maxwell's Equations without Iron

$$4\pi i_x = \frac{\partial \mathfrak{H}_z}{\partial y} - \frac{\partial \mathfrak{H}_y}{\partial z} = -\Delta\Gamma_x$$

$$4\pi i_y = \frac{\partial \mathfrak{H}_x}{\partial z} - \frac{\partial \mathfrak{H}_z}{\partial x}$$

$$4\pi i_z = \text{---------}$$

$$\Delta\Gamma_x = -4\pi i_x$$

$$\mathfrak{H}_x = \frac{\partial \Gamma_z}{\partial y} - \frac{\partial \Gamma_y}{\partial z}$$

$$\frac{\partial}{\partial z}\mathfrak{H}_y = \frac{\partial \Gamma_x}{\partial z} - \frac{\partial \Gamma_z}{\partial x}$$

$$\frac{\partial}{\partial y}\mathfrak{H}_z = \frac{\partial \Gamma_y}{\partial x} - \frac{\partial \Gamma_x}{\partial y}$$

Rel. same as betw. pot. & el. density. Thus,

$$\Gamma_x = \int \frac{i_x d\tau}{r}$$

$$\Gamma_y = \int \frac{i_y d\tau}{r}$$

In fact $\frac{\partial \Gamma_x}{\partial x} + \cdot + \cdot = 0$

as consequence of $\frac{\partial i_x}{\partial x} + \cdot + \cdot = 0$

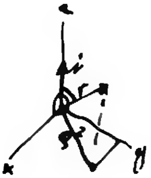
Represented as the sum of the distant-actions of elements.

 $q\,ds = d\tau$

$\mathfrak{H}_x d\tau = d\tau\frac{\mathfrak{i}_y z - \mathfrak{i}_z y}{r^3}$ $\mathfrak{H}_x ds = ds\frac{\mathfrak{i}_y z - \mathfrak{i}_z y}{r^3} = \mathfrak{i}\,ds\frac{\beta z - \gamma y}{}$

$\mathfrak{H}_y d\tau = d\tau$ -

- -



Is ⊥ to $\mathfrak{i}$ & r Choose $\mathfrak{i}$ and r as in Fig.

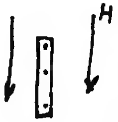
Then $\mathfrak{H} = \frac{d\tau}{r^3}\mathfrak{i}\cdot\rho = \frac{d\tau\,\mathfrak{i}}{r^2}\cdot\sin(\mathfrak{i}r)$ Interpretation of the vector product.

Media with permeability present. $\mathfrak{B}_x = \frac{\partial\Gamma_z}{\partial y} - \frac{\partial\Gamma_y}{\partial z} - \frac{\partial\varphi}{\partial x}$ [p. 46]

- - - - - - - - - - - - - - - -

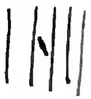
- - - - - - - - - - - - - - - -

Galvanometer with Earth Field Intensity of the Latter Magnetometer for the Determination of μ[38]



$-MH\sin x = I\frac{d^2x}{dt^2}$ $x = A\sin 2\pi\frac{t}{T}$

For small oscillations $\frac{d^2x}{dt^2} + \frac{MH}{I}x = 0$



$\left(\frac{2\pi}{T}\right)^2 = \frac{MH}{I}$ For meas [---] I & $I + I'$

$\frac{M}{H}$ can be determined

From this M & H separately (Gauss).
If H is known, then current strength with tangent galvanometer.[39]

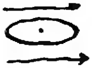
$$\frac{2\pi R}{R^2}i = H_i$$

From this H_i & thus also i.

Magnetometer

If infinitely thin magnetizable bar inside, then deflection owing to transverse field. $\frac{4\pi i n}{l} = \mathfrak{H}_s$ known.

$$\mathfrak{B} = \mu \mathfrak{H}_s \qquad \mathfrak{Q} = (\mu - l)\mathfrak{H}_s$$



$$\mathfrak{m} = \frac{(\mu - 1)}{4\pi}\mathfrak{H}_s \cdot f$$

$$\mathfrak{m}\left(\frac{1}{r^2} - \frac{1}{r'^2}\right) = \mathfrak{H}_a$$

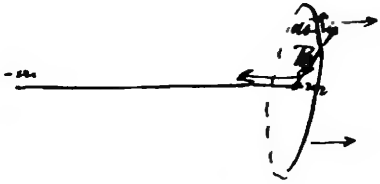
If the little bar is of finite thickness, then demagnetization factor

$$H_r = H - F\mathfrak{Q} = H - F\,\frac{\mu - 1}{4\pi}H_r \qquad H = H_r\left(1 + \frac{\mu - 1}{4\pi}F\right)$$

$\mathfrak{Q} = \frac{\mu - 1}{4\pi}H_r \qquad M = \mathfrak{Q} \cdot \mathrm{V}$ connection more indirect.

[p. 47] $\mathfrak{Q} = \frac{\kappa}{1 + \kappa F} . H \qquad I = \mathfrak{Q} \cdot \mathrm{Vol.} \qquad F =$

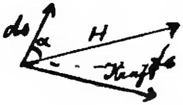
Ponderomotive Force on Element of Current



System cannot start moving by itself. Action & reaction are equal to one another.

$\frac{i.2\pi R}{R^2}\cdot m$ = force on magnetic pole, and thus, conversely, force on current. Thus, force on element of current $= i\frac{m}{R^2}ds = \underline{iHds}$.

No force in the direction of the element. General formulation. Force $\perp$ to $i \perp$ to H. If no right angle between H & ds, then only the component of H perpendicular to ds effective.



$H\cdot ids\cdot \sin\alpha$

We have to form the so-called vect. product of ds and H

$dK_x = i(dy\,\mathfrak{H}_z - dz\,\mathfrak{H}_y)$ $\qquad$ $dK_x = dt(\mathfrak{l}_y\mathfrak{H}_z - \mathfrak{l}_z\,\mathfrak{H}_y)$

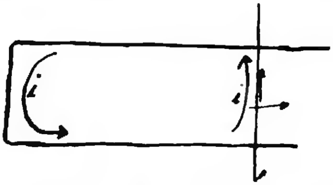
– – – – – – – – – – – – – – – – – – $\qquad$ –

– – – – – – – – – – – – – – – – – – $\qquad$ –

If instead of air or vacuum subst with permeability μ, then ponderomotive force dep. on $\mathfrak{B}$. Again *force* $= \frac{i2\pi R}{R^2}m$, but $\frac{\mu\pi m}{R^2} = \mathfrak{B}$.

Deprez-D'Arsonval instruments.[40]

Total force on finite conductor by integration.[41] [p. 48]



$l\cdot i\cdot \mathfrak{H}\cdot \delta = i\cdot \Delta N.$

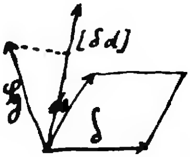
Work of the ponderomotive forces = increase in the number of lines of force · current strength. Flexible circuit seeks maximum extension. In general, work on element of current = number of the lines of force intersected. Force vector[42]

$$\begin{array}{l|l} i(dy\,\mathfrak{H}_z z - dz\,\mathfrak{H}_y)\cdot & \delta_x \\ i(dz\,\mathfrak{H}_x - dx\,\mathfrak{H}_z) & \delta_y \\ i(dx\,\mathfrak{H}_y - dy\,\mathfrak{H}_x) & \delta_z \end{array}$$

Multiplied by components of displacement $\delta_x\delta_y\delta_z$ yields work.

This can also be arranged in the following way

$$\begin{array}{l|l} i(\delta_y dz - \delta_z dy) & \mathfrak{h}_x \\ i(\text{- - - - - - - - - -}) & \\ \text{- - - - - - - - - - - -} & \end{array}$$

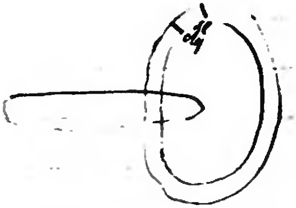


This proves the theorem.

A circuit with given i seeks to orient and deform itself in such a way that the number of [p. 49] lines of force it cuts becomes a maximum. Thus, the forces acting on circuits have <potential> funct. that plays role of pot. en equal to iN, where N thus directed is positive, like the field generated by the current.

Magnetic energy of a circuit.

$$\frac{1}{8\pi}\sum \mathfrak{B}\mathfrak{h}\, q dl = \frac{N}{8\pi}\int \mathfrak{h}\, dl = \frac{Ni}{2}$$



Electrostatic & Electromagn. Measure of the Current Strength & Quantity of El.

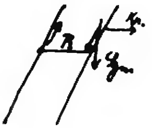
In electrostatics we derived an absolute measure for the quantity of el. & potential difference/electrostatic measure. $\mathfrak{E} = 4\pi\sigma$

$$\text{Force} = \mathfrak{E}\cdot\frac{\sigma}{2}\cdot f = \frac{1}{8\pi}\mathfrak{E}^2 f = 2\pi\sigma^2 f = 2\pi\frac{E^2}{f}$$

$$E_s = \frac{1}{\sqrt{2\pi}}\sqrt{\text{force}\cdot f} \approx M^{\frac{1}{2}} l^{\frac{3}{2}} t^{-1} \qquad i_s = M^{\frac{1}{2}} l^{\frac{3}{2}} t^{-2}$$

The quantity of electricity can also be measured electrodynamically as $\int i_m dt = E_m$

Dimension of the magnetically measured current: [Force]



$$\frac{2i_m}{R} = H \qquad H_m i_m l = \text{Force} = \frac{2i_m^2 l}{R} \qquad M\frac{L}{T^2} = |i_m^2|$$

$$\{i_m\} = M^{1/2}L^{1/2}T^{-1}$$

$$\left\{\frac{i_s}{i_m}\right\} = \frac{L}{T}$$

Deprez-D'Arsonval.[43]  $\mathfrak{H} li \cdot 2nR$ = D = Θx equilibrium. [p. 50]
abs. measurement of quantities of electricity

$$\mathfrak{H} = \frac{2\pi i}{R} \quad \mathfrak{H}M = I\frac{d^2x}{dt^2} \quad \int_t^0 \mathfrak{H}dt = \frac{I}{M}\left|\frac{dx}{dt}\right|_T = \frac{2\pi}{R}\int idt$$

$$\int idt = \frac{R}{2\pi}\frac{I}{M}\left|\frac{dx}{dt}\right|_T$$

From here on undamped sinusoidal oscillation according to the equation

$$MH_e x = -I\frac{d^2x}{dt^2}$$

$$x = A \sin 2\pi\frac{t}{T}$$

Inserted

$$MH_e = \left(\frac{2\pi}{T}\right)^2 I$$

$$\left(\frac{dx}{dt}\right)_{t-T} = \frac{2\pi}{T}A$$

$\int idt = \frac{R}{2\pi}H_e\left(\frac{T}{2\pi}\right)^2\cdot\frac{2\pi}{T}A = \frac{RT}{(2\pi)^2}H_eA$, where A is the maximum deflection in absol. angular measure.

According to Deprez $ki = I\frac{d^2x}{dt^2}$ initial period.

$$\int idt = \frac{I}{\kappa}\left\{\frac{dx}{dt}\right\}_T \qquad (1)$$

For the process thereafter $\frac{d^2x}{dt^2} = -\frac{\Theta}{I}x \quad A\sin 2\pi\frac{t}{T}$

$$\left(\frac{2\pi}{T}\right)^2 = \frac{\Theta}{I} \qquad (2)$$

$$\kappa i_0 = \Theta x_0 \qquad \kappa\Theta^{-1} = \text{susceptibility } \eta \qquad (3)$$

$$\int i\,dt = \frac{I}{\kappa}\cdot\frac{2\pi}{T}A = \frac{I}{\Theta\eta}\frac{2\pi}{T}A = \frac{T}{2\pi\eta}A.$$

$$/$$

$$\frac{1}{\kappa} = \frac{1}{\Theta\eta}$$

Thus, when the susceptibility η for the direct current is known, then quantities of electr. can be measured abs. e.g., with Deprez. Damping can also be calc.

[p. 51] Then one has

$$I\frac{d^2x}{dt^2} = -\Theta x - R\frac{dx}{dt}$$

Such lin equations with const. coefficients are most conveniently treated using imaginary quantities

$$e^{j\omega t} = \cos\omega t + j\sin\omega t \qquad e^{(-\alpha+j\omega)t} = e^{-\alpha t}e^{j\omega t} = e^{-\alpha t}(\cos\omega t + j\sin\omega t)$$

Instead of $A\cos\omega t$ and $Ae^{-\alpha t}(\cos\omega t)$, one inserts $Ae^{\gamma t}$, where γ can be compl. The real part of this solution is then the solution sought.
Now, again

$$\int i\,dt = \frac{I}{\kappa}\left\{\frac{dx}{dt}\right\}_T$$

Now, the above equation

$$\frac{d^2x}{dt^2} + \frac{R}{I}\frac{dx}{dt} + \frac{\Theta}{I}x = 0 \qquad e^{\lambda t} \text{ solution}$$

$$\lambda^2 + \frac{R}{I}\lambda + \frac{\Theta}{I} = 0 \qquad \left(\lambda + \frac{1}{2}\frac{R}{I}\right)^2 = \frac{1}{4}\left(\frac{R}{I}\right)^2 - \frac{\Theta}{I}$$

$$\lambda = -\frac{1}{2}\frac{R}{I} \pm \sqrt{\frac{1}{4}\left(\frac{R}{I}\right)^2 - \frac{\Theta}{I}}$$

Let discriminant be negative $\lambda = -\frac{R}{\underbrace{2I}_{\alpha}} \pm i\underset{\omega}{\sqrt{\frac{\Theta}{I} - \frac{1}{4}\left(\frac{R}{I}\right)^2}}$

$x = Ae^{-\alpha t}\sin\omega t$ solution.

Discussion of the solution. Damped oscillation $\left\{\frac{dx}{dt}\right\}_{t=0} = A\omega$

Oscillation period: $\frac{2\pi}{T} = \omega$

Damping $e^{\alpha T}$ = ratio $\frac{x_1}{x^2}$

Calculation of the first point of reversal

$$\frac{dx}{dt} = A\{-\alpha e^{--}\sin + \omega e^{-}\cos\} = Ae^{-\alpha t}\sqrt{\alpha^2+\omega^2}\sin(\varphi - \omega t) \quad \sin\varphi = \frac{\omega}{\sqrt{}}$$

$$\frac{dx}{dt} = 0 \text{ für } t = \frac{\varphi}{\omega} = \frac{1}{\omega}\operatorname{arctg}\frac{\omega}{\alpha} \qquad \cos\varphi = \frac{\alpha}{\sqrt{}} \quad \operatorname{tg}\varphi = \frac{\omega}{\alpha}$$

$$x_{max} = Ae^{-\alpha(\varphi/\omega)}\sin\varphi, \text{ where } \operatorname{tg}\varphi = \frac{\omega}{\alpha}$$

$$x_{max.} = \frac{1}{\omega}\left|\frac{dx}{dt}\right|_{t=0} e^{--}\sin\varphi = \frac{1}{\omega}\frac{\kappa}{I}\int i\,dt\cdot e^{--}\sin\varphi$$

$$= \frac{1}{\omega}\eta\sqrt{\left(\left(\frac{2\pi}{T}\right)^2 + \alpha^2\right)}Ee^{-\alpha(\varphi/\omega)}\sin\varphi^{[44]} \quad \operatorname{arctg}\varphi = \frac{\omega^{[45]}}{\alpha}$$

Thus, we can measure the quantity of electricity of a current impulse absolutely by [p. 52] electromagnetic means. We have seen earlier that the electr. quantities can be measured absolutely by static means. Since voltages can be measured absolutely with Thomson's balance,[46] & the capacities calculated.

Dimensions $\frac{E_s^2}{L} = M\frac{L^2}{T^2}$ $\quad E_s = M^{1/2}L^{3/2}T^{-1}$

$$\mathfrak{H} = M^{1/2}L^{-1/2}T^{-1}$$

$$\frac{i\,ds}{R^2} = \frac{i}{l} = \mathfrak{H} = M^{1/2}L^{-1/2}T^{-1}$$

$$i = M^{1/2}L^{+1/2}T^{-1}$$

$$E_m = \int i\,dt = M^{1/2}L^{1/2}$$

$\frac{E_s}{e_m} \approx \frac{L}{T}$ experiment yielded $3 \cdot 10^{10}$ = velocity of light = c.

This result led to Maxwell's theory of light. Remark. The fact that $\frac{E_s}{E_m}$ is indep. of the experimental design justifies the assumption that i_m is equal to the quantity of static electricity transported per unit time. $\frac{i_s}{i_m} = c$

Unit for Voltage. Ohm's Law

We consider a piece of conductor that is not acted upon by any electromot. forces (Expl.) The electric energy supplied to this piece per second is $p_1 i - p_2 i$ (electrostatically) = $\Delta p_s \cdot i_{st}$.

$$\frac{\text{Effekt}}{\text{erg}} = \frac{\boxed{c\Delta p_s}}{\Delta p_m} \cdot i_m = \Delta p_m \cdot i_m$$

We have thus obtained a new absolute unit for the voltage. Calorimetrically, if no effect other than heat is produced.

[p. 53] Practical unit constructed, which is 10^8 greater

$$10^{-8}\Delta p_m = \Delta p_{pr} \quad 10 i_m = i_{pr.}$$

$$\text{Effect} = \Delta p_m \cdot i_m = \underline{\Delta p_{pr} \cdot 10^7}$$

$\int i_{pr.} dt = E_{pr.}$ Coul. unit $E_{pr} = 10 E_m$

It turns out that for metallic & electrolytic conductors at constant temperature $\frac{\Delta p}{i}$ is a constant; one calls it the resist w of the conductor.

$$\Delta p = iw. \text{ (Ohm's law)}$$

w depends on geometrical conditions and on constants characteristic of the material. For [large] hom rod[47]

$$w = \omega\frac{l}{1} \; \omega \text{ spec. resist.} \quad \frac{1}{\omega} = \sigma \text{ conductivity of the material.}$$

It is possible to calculate the resist. of solid conductors if the current distr. is known. For a linear current we have

$$\Delta p = i_V \cdot \int_V \frac{\omega dl}{q} \quad (\omega \;\&\; q \text{ are funct. of } l$$

$$i = \sum i_V = \Delta p \sum \frac{1}{\int_V \frac{\omega dl}{q}} = \Delta p \sum \frac{1}{w_V}$$

The practical unit of resistance, the Ohm, is so defined that we get the equation

$$\Delta p_{pr.} = i_{pr.} \cdot w_{pr} 10^{-8} \Delta p_m = 10 i_m \cdot w_{pr.} 10^{-9} w_m = w_{pr.}$$

Determ. of the current flow in solid conductors.

$$i_x = -\frac{\partial\varphi}{\partial x} \cdot \sigma \qquad \frac{\partial}{\partial x}\left(\sigma \frac{\partial\varphi}{\partial x}\right) + \cdot + \cdot = 0$$

$$\text{------}$$

if homogen $\underline{\Delta\varphi = 0}$

$$\text{------}$$

On surface [--] $i_x \cos nx + \cdot + \cdot = 0$

$$\underline{\frac{\partial\varphi}{\partial n} = 0}$$

Mathematical problem the same as in electrostatics.

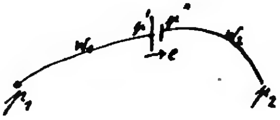
Math. Relationship between Resist. & Capacity[48]

[p. 54]

$$i = \sigma \int -\frac{\partial\varphi}{\partial n} d\sigma = \sigma \cdot 4\pi E_s$$

$$w = \frac{\Delta\varphi}{i} = \frac{1}{4\pi\sigma}\frac{\Delta\varphi}{E_s} = \frac{1}{4\pi\sigma} \cdot \frac{1}{C_s} = \frac{\omega}{4\pi}\frac{1}{C_s}$$

The capacity problem and resistance problem are, thus, identical. We give the mater. resistance wherever electrost. cap. has been calculated.
Ohm's Law, if electrom. forces[49]

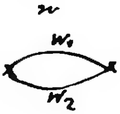


$$\begin{array}{lcll} p_1 - p' & = & iw_1 & 1 \\ p'' - p_2 & = & iw_2 & 1 \\ \underline{p'' - p'} & = & \underline{e} & -1 \end{array}$$

$$i(w_1 + w_2) - e = p_1 - p_2$$

$$iw = e + (p_1 - p_2)$$

Also applicable if electromot. forces are uniformly distributed. Special case starting point & end point coincide. Then $e = iw$, if w total resist. of the circuit. Resistances connected in parallel.

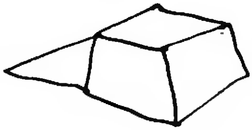


$$\Delta p = i_1 w_1 = i_2 w_2$$

$$i_1 + i_2 = i = \Delta p\left(\frac{1}{w_1} + \frac{1}{w_2}\right) = \frac{\Delta p}{w}$$

$$\frac{1}{w} = \frac{1}{w_1} + \frac{1}{w_2}$$

Kirchhoff's Laws for Current Networks



1) In junction $\sum i = 0$, because otherwise incess. accumulation of charge.
2) Any polygon considered

[p. 55]

$$p_2 - p_1 + e_1 = i_1 w_1$$

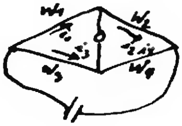
$$p_3 - p_2 + e_2 = i_2 w_2$$

- - - - - - - - - - - - - - - - -

$$p_1 - p_4 + e_4 = i_4 w_4$$

$$\sum e = \sum iw$$

Application to Wheatst. bridge[50]



$$i_2 = i_1 \qquad i_4 = i_3$$

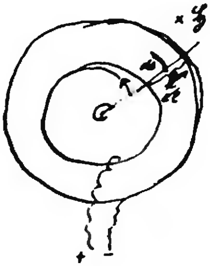
$$i_1 w_1 - i_3 w_3 = 0$$

$$i_1 w_2 - i_3 w_4 = 0$$

$$\frac{w_1}{w_3} = \frac{w_2}{w_4} \qquad \frac{w_1}{w_2} = \frac{w_3}{w_4}$$

Solid conductors only here.[51]

Electric Induction [force work on the path ds

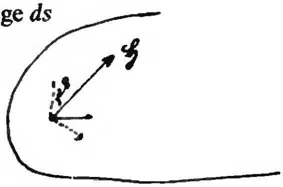


$il\mathfrak{H}$ = Kraft

$il\mathfrak{H}\,ds$ = Arbeit auf dem Wege ds

$= ei\,dt$

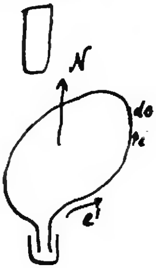
$$e = l\mathfrak{H}\frac{ds}{dt}.$$



Motional Field

If a circuit is displaced in a magn field, the system expends work that is equal to idN, where N is the number of the lines of force traversing the field and the lines of force originating from the current itself are neglected. The result holds also if magnetization constant μ.[52]

[p. 56] If the field originates from magnets and the total energy of the field does not change in the course of the displacement, then an electromagnetic force must counteract the current, a force against which we must apply electrical work that is equal to the ponderomotive work.



$$idN = e'idt = -\,eidt$$

$$e = -\frac{dN}{dt}$$

If e is measured in practical units, then $e_{\text{volt}} = -10^{-8}\dfrac{dN}{dt}$.

This is the induction law of Faraday. Since the origin of the magnetic field is obviously not important, the law is generally valid, no matter how the field might be produced.

Extension to the case where the magnet is in motion and the conductor at rest.

Provides a method for the determination of magnetic fields and their changes.

The current field also acts on this current itself, if the current is changed. From this it follows that every linear current can be conceived as a bundle of linear currents. Insofar as one can view N as defined, one has again

$$e = -\frac{dN}{dt}$$

But now one has to set $N = L \cdot i$, where i is the instantaneous current strength, thus also [p. 57]

$$e = -\frac{dL_i}{dt},$$

or, if L is independent of the time:

$$e = -L\frac{di}{dt} \quad e_{pr.} \cdot 10^8 = -L\ 10^{-1}\frac{di_{pr.}}{dt} \quad (L \cdot 10^{-9}) = L_{pr.}$$

L is the coefficient of self-induction. The practical unit of self-induction is the Henry = 10^9 abs. The equations are then valid for pr. un. as well.
Solenoid:[53]

$$N = \frac{4\pi n i}{l} f$$

$$e = -n\frac{dN}{dt} = -\underset{L}{\boxed{\frac{4\pi n^2 f}{l}}}\frac{di}{dt}$$

Ring analogous. If permeability μ, then $L\mu$ times greater

Linear conductor through which variable current flows.

$$\Delta p + e = iw$$

$$-L\frac{di}{dt}.$$

If self-ind. the only electromot. force, then

$$\Delta p = iw + L\frac{di}{dt}$$

Conductor in zero-current state suddenly connected to potential difference.

How does current increase? $P = iw + L\frac{di}{dt}$

$$i = \frac{P}{w} + i_1$$

$$i_1 w = -L\frac{di}{dt} \quad \frac{d\lg i}{dt} = -\frac{w}{L} \quad i = \text{const } e^{-(w/L)t}.$$

$$i = \frac{P}{w}\left(1 - e^{-(w/L)t}\right)$$

Time: $T\frac{w}{L} = 5 \quad T = \frac{5L}{w}$ Practically very short time.

Fading away of the current analogous.

[p. 58] Sine current

$$\Delta p = iw + L\frac{di}{dt}$$

$\Delta p = A \cos \omega t$ given

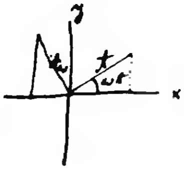
$i = B \cos(\omega t - \varphi)$ φ is then the phase difference between voltage & current. If φ pos, then current lags.

$$\begin{cases} B(\underset{w\cos\alpha}{\underset{|}{w}}\cos(\omega t - \varphi) - \underbrace{\omega L}_{\sin\alpha}\sin(\omega t - \varphi)) \\ A\cos\omega t = B\sqrt{w^2 + \omega^2 L^2}\cos(\omega t - \varphi + \alpha) \end{cases} \qquad \operatorname{tg}\alpha = \frac{\omega L}{w}$$

$$B = \frac{A}{\sqrt{w^2 + \omega^2 L^2}} \qquad \varphi = \alpha$$

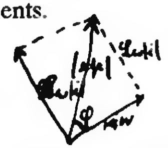
Graphic Illustration with Rotating Vectors

$A \cos \omega t$



Derivative
$-A \cos \omega t$

Thus, rule for differentiation.
Rule for summing parallelogram, because projection of the resultant always sum of projections of the components.



$$\mathrm{tg}\,\varphi = \frac{\omega L}{w} \qquad P^2 = i^2(w^2 + (\omega L)^2)$$

Comes to the same as the replacement of the trigonometric funct by exp. with complex [p. 59] arg.[55] $A\cos(\omega t - \varphi)$ is the real portion of $Ae^{j(\omega t - \varphi)} = Ae^{-j\varphi}e^{j\omega t}$, $= \mathfrak{A}e^{j\omega t}$, where $\mathfrak{A}$ complex $= Ae^{-j\varphi}$. Thus, phase angle & amplitude known if $\mathfrak{A}$ known.

$$\Delta p = \mathfrak{P}e^{j\omega t}$$

$$i = \mathfrak{J}e^{j\omega t}$$

$$\mathfrak{P} = \mathfrak{J}w + j\omega L\mathfrak{J} = I(w + j\omega L)$$

$$\frac{\mathfrak{P}}{\mathfrak{J}} = \frac{Pe^{-\varphi_p}}{Ie^{-\varphi_i}} = \underbrace{(w + j\omega L)}_{\sqrt{w^2 + (\omega L)^2}\,e^{j\,\mathrm{arctg}\frac{\omega L}{w}}}$$

$$\frac{P}{I} = \sqrt{w^2 + (\omega L)^2} \qquad \varphi_i - \varphi_p = \mathrm{arctg}\frac{\omega L}{w}$$

$$\Delta p = Pe^{j\omega t}$$

$$i = Ie^{j(\omega t - \varphi)}$$

$$P = Iwe^{-j\varphi} + j\omega LIe^{-j\varphi}$$

$$= I(w + j\omega L)e^{j\varphi}$$

etc.

Calculation considerably simpler than with sin & cos. Therefore almost always applied nowadays.
The calculation is simplest if by variables one immediately understands corresp. compl. Then we get at once

$$\Delta p = i(w + j\omega L)$$

i-vector to be multiplied by $(w + j\omega L)$ vector in order to have Δp. Coincides with theory of rotating vectors. Naturally, these methods are applicable only to harmonic functions.

Earth Inductor—Measurement of Self-Induction

Magnetic energy of a circuit

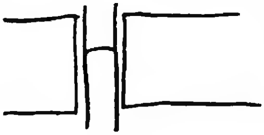
$$\int e'\,idt = -\int eidt = \int\frac{dN}{dt}idt = L\frac{i^2}{2}.$$

[p. 60]

Remark about Ponderomotive Effects on Magnetizable Bodies in the Field of a Current[56]

$$dA = e'\,idt = -eidt = \frac{dN}{dt}idt$$

The bodies seek to move in such a way that N becomes a maximum. This fact forms the basis for the measurement of small <dielectric> magnetization constants.[57]



Homogeneous field $\mathfrak{H}$ in the body $H\mu$

If column rises by δh, then the change in the energy is

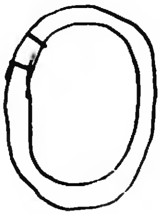
$$\frac{1}{8\pi}\mu\mathfrak{H}^2 q\,\delta h - \frac{1}{8\pi}\mathfrak{H}^2 q\,\delta h.$$

Thus force on column $= \frac{\mu-\mu_0}{8\pi}\mathfrak{H}^2 q \Big| = qh\rho g,$

where h is the height of the rise produced by the magnetic force.

Energy & Energy Principle

Previously: For a circuit we have



$$E = \frac{1}{8\pi}\int \mathfrak{H}\mathfrak{B}\,\underset{q\,dl}{dt} = \frac{1}{8\pi}\int \mathfrak{H}\,dl\,dN = \frac{1}{2}iN$$

Since by definition $N = Li$, one obtains $E = \frac{Li^2}{2}$ in agreement with the above analysis.

Application of the energy principle to current of constant intensity.

$$e'\,idt = dE + dA$$

$$+\,i\frac{dN}{dt} = \frac{1}{2}\frac{d(iN)}{dt} + dA$$

Thus, $$dA = \frac{1}{2}i\frac{dN}{dt} - \frac{1}{2}N\frac{di}{dt} = \frac{1}{2}i\frac{d(Li)}{dt} - \frac{1}{2}iL\frac{di}{dt}$$

If i is const., we obtain $dA = \frac{1}{2}idN = dE$ [p. 61]

Work is equal to the increase in energy. The expression differs from that for the work of the current in an external magnetic field by the factor $\frac{1}{2}$. Example. Parallel currents.

Measurement of an EMF of Short Duration. Earth Inductor[58]

$$e = iw - L\frac{di}{dt}$$

At the start $i = 0$. At the end $i = 0$

$$\int e\,dt = w\int i\,dt - \underset{\overset{\|}{0}}{L}|i|_0^t$$

For the earth inductor $e = n\frac{dN}{dt}$ $\int e\,dt = 2Nn$.

Quantity of electricity measured with ballistic instrument. Analogous method for the investigation of hysteresis.

Interaction between Permanent Magnets & Current

N_m = circuit-traversing flux that originates from the magnet
N_i = " " " " " " " current. L self-ind.

$$dA = idN_m + \frac{1}{2}i^2dL$$

$$e = -\frac{dN_m}{dt} - \frac{d(iL)}{dt}$$

Ohm's equation

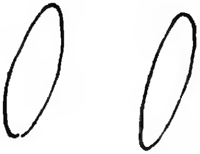
$e' + e = iw$

Energy principle
$$\begin{aligned} i^2 w\,dt &= -i\,dN_m - i\,diL + e'i\,dt \\ &= -dA + \underbrace{\frac{1}{2} i^2\,dL - id(iL)}_{-\frac{1}{2} i^2\,dL + \frac{1}{2} L\,di^2} + e'i\,dt \\ &= -dA + e'i\,dt + d\left(\frac{1}{2} L i^2\right)^{[59]} \end{aligned}$$

[p. 62]

Interaction between Two Circuits

The circuits are immobile.



L_1 = flux that current 1 of strength 1 yields through its surface
M_{12} = " " " 1 " " " " boundary of cur. 2
M_{21} = " " 2 " " " " " cur. 1
L_2 = " " " 2 " " " " its boundary.

Total flux through 1): $L_1 i_1 + M_{21} i_2 = N_1$
" " " 2): $M_{12} i_1 + L_2 i_2 = N_2$

The equation for the two circuits is

$$e_1 - \frac{dN_1}{dt} = i_1 w_1$$

$$e_2 - \frac{dN_2}{dt} = i_2 w_2$$

or

$$e_1 = i_1 w_1 + L_1 \frac{di_1}{dt} + M_{21} \frac{di_2}{dt}$$

$$e_2 = i_2 w_2 + M_{12} \frac{di_1}{dt} + L_2 \frac{di_2}{dt}$$

What form does the energy principle take?[60]

[p. 63]

$$e_1 i_1\,dt = i_1^2 w\,dt + \frac{d}{dt}\left(\frac{1}{2}L_1 i_1^2\right) + M_{21} i_1 \frac{di_2}{dt}\,dt$$

$$e_2 i_2\,dt = i_2^2 w\,dt + M_{12} i_2 \frac{di_1}{dt}\,dt + \frac{d}{dt}\left(\frac{1}{2}L_2 i_2^2\right)$$

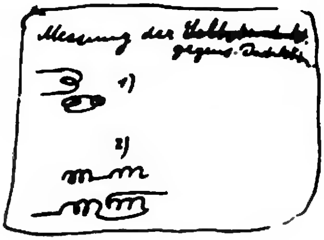
$$dA_e = G + dE_{m1} + dE_{m2} + M_{21} i_1\,di_2 + M_{12} i_2\,di_1$$

$dA_e - G$ must be tot. differential. Thus $M_{21} = M_{12} = M$.

$E = \frac{1}{2}(L_1 i_1^2 + 2M i_1 i_2 + L_2 i_2^2)$ must never be negative

$$L_1 + 2Mx + L_2 x^2 | M + L_2 x = 0 \qquad L_1 - 2\frac{M^2}{L_2} + \frac{M^2}{L_2} > 0$$

Measurement of mutual induction.



One can also resolve the magnetic field in another way.

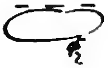
Φ number of lines of force traversing both circuits.
Φ_1 " " " " " " only 1)
Φ_2 " " " " " " " 2).

Model

This resolution is especially advantageous when almost all lines of force traverse both circuits. How do we determ Φ_1, Φ_2 & Φ?

$$\Phi_1 = \frac{N_1}{n_1} = \frac{L_1}{n_1} i_1 + \frac{M}{n_1} i_2 = \underbrace{\left(\frac{L}{n_1} - \frac{M}{n_2}\right)}_{\Phi_1} i_1 + \overbrace{M\left(\frac{i_2}{n_1} + \frac{i_1}{n_2}\right)}^{\Phi}$$

$$\Phi_2 \qquad \text{--------------} = \text{----------------------}$$



Transformer, with Resist. & Leakage Neglected [61]

$$\Delta p_1 = n_1 \frac{d\Phi}{dt}$$

$$\Phi = \frac{1}{w}(i_1 n_1 + i_2 n_2)$$

$$\Delta p_2 = n_2 \frac{d\Phi}{dt}$$

The phase of the current depends on what is switched on. If only resist. then [p. 64] i_2 phase of Δp_2.

Two mobile circuits
Work el. force. Energy

$$p_1 = i_1 w_1 + \frac{dL_1 i_1}{dt} + \frac{dM i_2}{dt}$$

$$p_2 = i_2 w_2 + \frac{dM i_1}{dt} + \frac{dL_2 i_2}{dt}$$

$$d'A_1 = \frac{1}{2} i_1 d(L_1 i_1) + i_1 d(M i_2)$$

$$d'A_2 = \frac{1}{2} i_2 d(L_2 i_2) + i_2 d(M i_1)$$

$$d'A_{e1} = p_1 i_1 \, dt$$

$$d'A_{e2} = p_2 i_2 \, dt$$

$$E_m = \frac{1}{2}(L_1 i_1^2 + 2M i_1 i_2 + L_2 i_2^2)$$

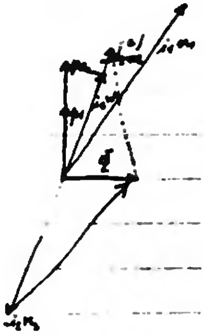
$$d'A_e = d'G + dE_m + d'A_e \text{ [62]}$$

$$p_1 = n_1 \frac{d\Phi}{dt}$$

$$p_2 = n_2 \frac{d\Phi}{dt}$$

$$p_2 = i_2 w_2 + L_2 \frac{di_2}{dt}$$

$$\Phi = \frac{1}{w}(i_1 n_1 + i_2 n_2)$$



Transformer with imaginary.[63] By $e_1 i_1 \ldots$ one immediately understands imaginary vectors.

$$\begin{array}{ll|l|l} e_1 = i_1(w_1 + j\omega L_1) & + i_2 \cdot j\omega M & w_2 + j\omega L_2 & j\omega M \\ 0 = i_1\, j\omega M & + i_2(w_2 + j\omega L_2) & -j\omega M & -(w_1 + j\omega L_1) \end{array}$$

$$e_1 = i_1 \frac{[(w_1 + j\omega L_1)(w_2 + j\omega L_2) + \omega^2 M^2]}{(w_2 + j\omega L_2)}$$

$$= i_1 \left[w_1 + j\omega L_1 - \frac{\omega^2 M^2}{w_2 + j\omega L_2} \right]$$[64]

$$e_1 = i_2 \frac{-(w_1 + j\omega L_1)(w_2 + j\omega L_2) - \omega^2 M^2}{j\omega M}$$

In the absence of leakage $(L_1 L_2 - M_2 = 0)$,[65] the second equation becomes

$$e_1 = -i_2 \frac{w_1 w_2 + j\omega(L_1 w_2 + L_2 w_1)}{j\omega M}$$

& if, in addition, $w_1 = 0$, then $e_1 = -i_2 \dfrac{L_1 w_2}{M} = -i_2 w_2 \dfrac{n_1}{n_2} \quad \dfrac{n_1^2}{n_1 n_2}$

$$\text{The } [--] \text{ in } e_1 = i_1 \frac{j\omega L_1 w_2}{w_2 + j\omega L_2} = i_1 \boxed{w_2 \frac{n_1^2}{n_2^2}}$$

or if w_2 negligible compared with $j\omega L_2$

[p. 65]

Capacity

$E_m = C_m p_m$ If this equation is to be valid in electromagnetic units, then the unit of capacity is fixed thereby. What is the relation between this unit and the static unit?

$$E_m = \frac{1}{c} \cdot E_s$$

$$p_m = c \cdot p_s$$

$$\text{hence } \frac{1}{c} E_s = C_m c p_s$$

$$E_s = \underset{C_s}{\boxed{C_m c^2}}\, p_s \qquad C_{st} = C_m \cdot c^2$$

The static unit is c^2 times smaller than the electromagnetic one
There is also a practical unit

$$E_{pr} = C_{pr} p_{pr.}$$

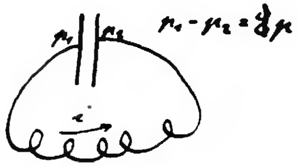
$$10E_m = C_{pr.}\, 10^{-8} p_m \qquad E_m = \underset{C_m}{\boxed{10^{-9} C_{pr.}}}\; p_m \qquad \underset{\text{Farad}}{C_{pr.} = 10^9 C_m}$$

Practical unit 10^{-9} of the absolute magnet. unit
This is $9 \cdot 10^{20}$ electrostatic units
Practical unit (farad) $9 \cdot 10^{11}$ electrostatic units In addition, microfarad 10^{-6} of the farad. $9 \cdot 10^5$ electrostatic units.

Circuit with Capacitance and Self-Induction. Electric Oscillations

$$pC = E \qquad -\frac{dE}{dt} = -C\frac{dp}{dt} = i$$

$$p - L\frac{di}{dt} = iw$$



differentiated once again
[if...then..or solution; oscillation period easily realizable]

$$\frac{dp}{dt} = \frac{di}{dt}w + L\frac{d^2i}{dt^2} = -\frac{i}{C}$$

$$\text{also } \frac{1}{C}i + w\frac{di}{dt} + L\frac{d^2i}{dt^2} = 0$$

$$i + wC\frac{di}{dt} + LC\frac{d^2i}{dt^2} = 0$$

Wenn $w = 0$, dann $\mathfrak{J}\cos\omega t$ oder $Ie^{j\omega t}$ Lösung.

$$I + (j\omega)^2 LCI = 0 \quad \omega = 2\pi n = \sqrt{\frac{1}{CL}}$$

$$n = \frac{1}{2\pi}\sqrt{\frac{1}{CL}}$$

$$\omega 10^{[2]}\, 10^{[+6]}\, 10^{[+8]}$$
$$10^8\, \omega 10^{-10}\, 10^{-8}$$

Schwingungsdau 10^{-4} Sek. wohl realisierbar,

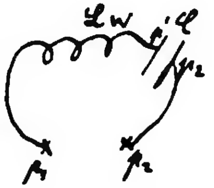
$$1 + \alpha wC + \alpha^2 LC = 0$$

$$\alpha^2 + \alpha\frac{w}{L} + \frac{1}{CL} = 0$$

$$\alpha = -\frac{w}{2L} \pm \sqrt{-\frac{1}{LC} + \left(\frac{w}{2L}\right)^2}$$

[p. 66] Frequency somewhat influenced (reduced) by resistance. Amplitude decreases with $e^{-\frac{w}{2L}t}$ $\left(W = 1 \text{ ohm } \& \ L = \frac{1}{100} \ T = \frac{1}{50}.\right)$

We shall also discuss in particular the case of sinusoidal currents.



$$p_1 - p' = iw + L\frac{di}{dt}$$

$$p' - p_2 = \frac{1}{C}E = \frac{1}{C}\int i\,dt$$

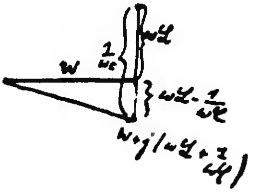
$$p = p_1 - p_2 = iw + L\frac{di}{dt} + \frac{1}{C}\int i\,dt$$

Solution by means of imag. $i = \mathfrak{J}_0 e^{j\omega t} \int i dt = \frac{I_0}{j\omega}e^{j\omega t} = \frac{i}{j\omega}$

Inserting this, one obtains

$$p = i\left(w + j\omega L + \frac{1}{j\omega C}\right) = i\left(w + j\left(\omega L - \frac{1}{\omega C}\right)\right)$$

$$-j\frac{1}{wC}$$



$$\mathfrak{P} = \mathfrak{J}\sqrt{w^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}\, e^{j\varphi} \qquad \operatorname{tg}\varphi = \frac{\omega L - \frac{1}{wC}}{\omega}$$

If $i = I\cos\omega t$

$$p = I\sqrt{\qquad}\cos(\omega t + \varphi)$$

$$\text{Ampl. } I = \frac{P}{\sqrt{w^2 + \left(\omega L - \frac{1}{\omega c}\right)^2}}$$

Resonance when I maximum $\omega = \frac{1}{\sqrt{cL}}$ proper oscillations. For the latter, I becomes infinite when $w = 0$ at given voltage. Capacity compensates self-ind. But only for specif. perm. With very weak terminal voltage considerable current.

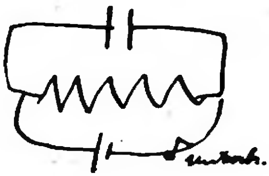
When there is resonance, voltage on the condenser $\frac{1}{C}\int idt = p' - p_2 = \frac{p}{Cw}\cos.$[66] [p. 67]
May become enormously large if C small & W small.

Energy Principle in Oscillations

$$p = iw + L\frac{di}{dt}\Big|\, i\,dt$$

$$pi\,dt = i^2 w\,dt + d\left(\frac{L}{2}i^2\right)$$

$$\Big| \quad -c\frac{dp}{dt}$$



If $w = 0$ $\quad \frac{d}{dt}\left(c\frac{p^2}{2} + \frac{L}{2}i^2\right) = 0$

The to-and-fro oscillation of energy[67] $\quad p_m^2 = \frac{L}{c}i_m^2$

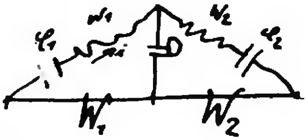
$$p_m = i_m\sqrt{\frac{L}{c}} \qquad \text{If} \qquad L = 10^{-2}\ \text{Henry}$$

$$C = 10^{-8}\ \text{Farad}$$

$$p_m = 10^3\, i_m$$

Comparison of capacities[68]

$$\frac{p_1}{p_2} = \frac{W_1}{W_2} = \frac{\not{i}\left(w_1 - j\frac{1}{wC_1}\right)}{\not{i}\left(w_2 - j\frac{1}{wC_2}\right)}$$



From this the relationship (independent of period).
Rapid oscillations if L small. Not coils but simple wires.
Back-and-forth loop[69]

$$\int_0^{R_1} 2i\frac{r^2}{R^2}\cdot\frac{1}{r}dr = i\Big|$$

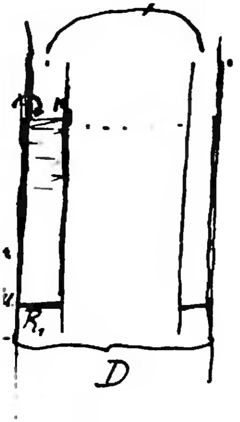
$$\int_{R_1}^{D}\frac{2i}{r}dr = 2i\lg\frac{D}{R_1}$$

Thus, all in all

$$2 + 2\lg\frac{D^2}{R_1R_2} \text{ too large}$$

$$2\lg\frac{D^2}{R_1R_2} \text{ too small}$$

$$\underline{1 + 2\lg\frac{D^2}{R_1R_2}}$$



[p. 68]

$$L = l\left(1 + 2\lg\frac{D^2}{R_1R_2}\right)$$

If we introduce the total length $l' = 2l$, & set $R_1 = R_2$, then

$$L = l'\left(\frac{1}{2} + 2\lg\frac{D}{R}\right)$$

We obtain the approximate value of L for a square.

$$L = 2l'\left(\frac{1}{4} + \lg\frac{s}{R}\right) = \underline{2l'\left(\lg\frac{l'}{R} - 1.13\right)}^{[70]}$$

$$\lg\frac{l'}{4R}$$

Is too large, because field calculated too large. In reality, according to rigorous calculation[71] $\underline{L = 2l'\left(\lg\frac{l'}{R} - 1.9\right)}$.

For circle the same formula but −1.5.

Waves in a Wire (Distributed Capacity)

═══ c = capacity per unit length.

p pot. e el. quant. " " " i current str.

$$-\frac{\partial i}{\partial x} = c\frac{\partial p}{\partial t} \quad \text{(continuity equation for electricity}$$

$$iw + l\frac{\partial i}{\partial t} = -\frac{\partial p}{\partial x}\text{[72]} \;\Big|\; c\frac{\partial}{\partial t}$$

These are differential equations for i & p.
p eliminated

$$cw\frac{\partial i}{\partial t} + c\ell\frac{\partial^2 i}{\partial t^2} = \frac{\partial^2 i}{\partial x^2}$$

p can then be determined from the first equation. If w neglected, then

$$cl\frac{\partial^2 i}{\partial t^2} = \frac{\partial^2 i}{\partial x^2} \qquad i = f(x - Vt) \text{ is solution}$$

$$clV^2 = 1$$

$$V = \frac{1}{\sqrt{cl}}$$

Two parallel wires whose radius is negligible compared with the distance between them[73] [p. 69]

$$l = 2\lg\frac{D}{R}$$

$$\text{cap} = \frac{1}{2\lg\frac{D}{R}}\cdot\frac{1}{c^2}$$

$V = c$ Such electric waves propagate with the speed of light. For other c & l different results.

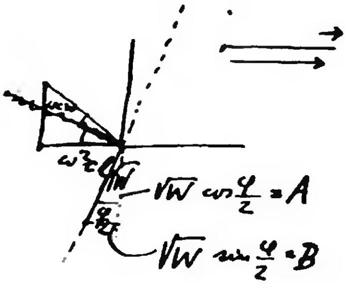
W not neglected. ∞ long wire. Sinusoidal solution. Influence of ε & μ.

$$i = Xe^{j\omega t}$$

$$\underbrace{(j\omega cw - \omega^2 cl)}X = X'' \qquad X = Ae^{\gamma x}$$

$$\gamma^2 = -\omega^2 cl + j\omega cw \qquad \sqrt{(\omega^2 cl)^2 + (\omega cw)^2} = W$$

$$\gamma = \sqrt{}$$



The solution is

$$i = Ae^{j\omega t}e^{(-Aj-B)x}$$

$$= Ae^{-Bx}e^{j\omega(t-(A/\omega)x)} \qquad \frac{\omega}{A} \text{ velocity}$$

B = damping constant.

A = current amplitude at start.

$$W = \omega c\sqrt{w^2+\omega^2 l^2} \qquad \frac{\omega}{A} \sim \frac{\omega}{\sqrt{w}} \sim \frac{1}{\sqrt{cl}}$$

$$\operatorname{tg}\varphi = \frac{w}{\omega l}$$

$$W = \sqrt{(\omega^2 cl)^2 + (\omega cw)^2}$$

$$B = \sqrt{W}\,\sin\frac{\varphi}{2}$$

$$A = \sqrt{W}\,\cos\frac{\varphi}{2}$$

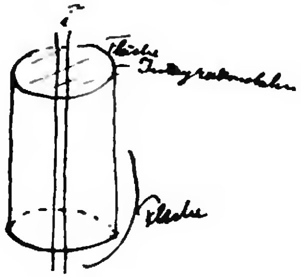
Damping coefficient $B \propto \omega\sqrt{cl}\,\dfrac{w}{2\omega l} = \dfrac{w}{2}\sqrt{\dfrac{c}{l}}$ From this, telephone transmission range.

Pupin's system.[74] | Extreme $w > \omega l$.

[p. 70]

Maxwell's Equations

1) We have reason to assume finite propagation. Conduction with distributed capacity.



$4\pi i = \int \mathfrak{h}\, ds$ can then no longer be valid for arbitrary surfaces. The law can therefore be strictly maintained only for surface *elements*

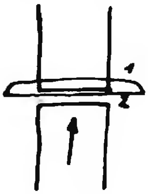
$$4\pi j_x = \frac{\partial \mathfrak{H}_z}{\partial y} - \frac{\partial \mathfrak{H}_y}{\partial z}$$

\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-

\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-

is therefore surely more exact than the above law in integral form if the currents in question are not constant.

2) Open currents



Conductor interrupted by a dielectric of arbitrary dielectric constant. Condenser. There also seem to be exceptions to $4\pi i = \int \mathfrak{h}\, ds$ for slow currents if one places the surface across the intervening space. This would apply to any intervening space, no matter how narrow. But we can maintain the law in general if we assume that the temporal change of the dielectric associated with the current acts magnetically like a conduction current.

$$\eta = \frac{1}{4\pi}\mathfrak{D}$$

$$E = \int \eta\, d\sigma = \frac{1}{4\pi}\int \mathfrak{D}\, d\sigma$$

$$i = \frac{dE}{dt} = \frac{1}{4\pi}\int \frac{d\mathfrak{D}_n}{dt}\, d\sigma$$

We assume that the right side is equivalent to a current. The X-component of this vector: [p. 71]

$$\frac{1}{4\pi}\frac{\partial \mathfrak{D}_x}{\partial t}$$

Acts like the x-component of a current density (displacement current) Conduction current & displacement current can be present together.

$j_x + \frac{1}{4\pi}\frac{\partial \mathfrak{D}_k}{\partial t} = X$ component of the total current).

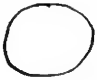
If one corrects the above differential equations in this manner, one obtains

$$4\pi j_x + \frac{\partial \mathfrak{D}_x}{\partial t} = \frac{\partial \mathfrak{H}_z}{\partial y} - \frac{\partial \mathfrak{H}_y}{\partial z}$$

$$4\pi j_y + \frac{\partial \mathfrak{D}_y}{\partial t} = \frac{\partial \mathfrak{H}_x}{\partial z} - \frac{\partial \mathfrak{H}_z}{\partial x} \qquad \text{in vector not.} \quad 4\pi j + \frac{d\mathfrak{D}}{dt} = \text{curl } \mathfrak{H}$$

$$4\pi j_z + \frac{\partial \mathfrak{D}_z}{\partial t} = \frac{\partial \mathfrak{H}_y}{\partial x} - \frac{\partial \mathfrak{H}_x}{\partial y}$$

These equations are joined by a fourth one, that of Gauss's theorem



$$4\pi E = \int \mathfrak{D}_n d\sigma \quad \leftarrow\!\square\!\rightarrow$$

$$4\pi\rho = \frac{\partial \mathfrak{D}_x}{\partial x} + \frac{\partial \mathfrak{D}_y}{\partial y} + \frac{\partial \mathfrak{D}_z}{\partial z} \qquad 4\pi\rho = \text{div } \mathfrak{D}$$

If one bears in mind that $\frac{\partial \rho}{\partial t} = -\left(\frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z}\right) = -\text{div } j$, then one has $4\pi \text{ div } j + \frac{\partial}{\partial t}$ $(\text{div } \mathfrak{D}) = 0$.

But this equation is contained in the ones above, as one can see by differentiating with respect to x,y,z and adding.
As usual, it is assumed that j and $\mathfrak{D}$ are determined by $\mathfrak{E}$. The simplest hypothesis is

$$j_x = \sigma \mathfrak{E}_x \qquad \mathfrak{D}_x = \varepsilon \mathfrak{E}_x$$

- - - - - - - - - - - - - - -

- - - - - - - - - - - - - - -

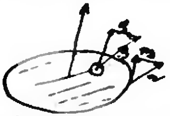
However, the relation can be a more complicated one.

[p. 72] 3) This was the law that defined the magnetic fields determined by electric currents. We have also become acquainted with a law for the production of electromotive effects by the alteration of magnetic fields.

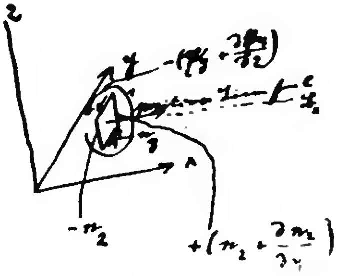
$$e = -\frac{\partial N}{\partial t}$$

This holds first of all for closed circuits. If we think of the EMF as a line integral of an EMF field $\mathfrak{e}$, then the law takes the form

$$\int \mathfrak{e}_s ds = -\frac{d}{dt}\int \mathfrak{B}_n d\sigma$$



Because of the finite propagation velocity of electric effects, this law, too, will only hold for ∞ *small* surface elements. We apply it to the following surface.



$$-\frac{\partial \mathfrak{B}_x}{\partial t} = \frac{\partial \mathfrak{e}_z}{\partial y} - \frac{\partial \mathfrak{e}_y}{\partial z}$$

$$-\frac{\partial \mathfrak{B}_y}{\partial t} = \text{- - - - - - -}$$

$$\text{- - - - - - - - - - - - - - - .}$$

Besides this electromotive field we also have an *electric* field $\mathfrak{E}_x\mathfrak{E}_y\mathfrak{E}_z$. This has been taken over from electrostatics. We shall therefore call it $\mathfrak{E}_{sx}$ etc. The following equations hold for it

$$0 = \frac{\partial \mathfrak{E}_{sz}}{\partial y} - \frac{\partial \mathfrak{E}_{sy}}{\partial z}$$

$$\text{- - - - - - - - - - - - - -}$$

$$\text{- - - - - - - - - - - - - -}$$

Electromotive & electrostatic field are both def. by the force exerted on the el. unit. We [p. 73] have therefore no reason a priori to consider them as being of different nature. The formal laws also require that the sum $\mathfrak{e}_x + \mathfrak{E}_{sx} \ldots$ be considered simply as the elec. field str. $\mathfrak{E}_x \ldots$ For if one adds these equations, one obt. [or]

$$-\frac{\partial \mathfrak{B}_x}{\partial t} = \frac{\partial \mathfrak{E}_z}{\partial y} - \frac{\partial \mathfrak{E}_y}{\partial z}$$

$$\text{- - - - - - - - - - - - - - - - -} \qquad \text{or} \quad -\frac{\partial \mathfrak{B}}{\partial t} = \operatorname{curl} \mathfrak{E}.$$

$$\text{- - - - - - - - - - - - - - - - -}$$

These equations give $\frac{\partial}{\partial t}\left(\frac{\partial \mathfrak{B}_x}{\partial x} + \cdot + \cdot\right) = 0$. Thus, they are compatible with the condition div $\mathfrak{B} = 0$ (There is no real magnetism).

Plane waves.

Let $\mathfrak{B} = \mu\mathfrak{H}$ & $\mathfrak{D} = \varepsilon\mathfrak{E}$, & let μ & ε be indep. of the location. Then the equations read:

We are looking for waves propagating in the X-direction. Everything dep. only on x & t. Let $\langle F(x - vt)\rangle$ be the dependence of all components

$$\begin{array}{ll}
-\frac{\mu}{c}\frac{\partial\mathfrak{H}_x}{\partial t}=0 & \frac{\varepsilon}{c}\frac{\partial\mathfrak{E}_x}{\partial t}=0 \\[2ex]
-\frac{\mu}{c}\frac{\partial\mathfrak{H}_y}{\partial t}=-\frac{\partial\mathfrak{E}_z}{\partial x} & \frac{\varepsilon}{c}\frac{\partial\mathfrak{E}_y}{\partial t}=-\frac{\partial\mathfrak{H}_z}{\partial x} \\[2ex]
-\frac{\mu}{c}\frac{\partial\mathfrak{H}_z}{\partial t}=\frac{\partial\mathfrak{E}_y}{\partial x} & \frac{\varepsilon}{c}\frac{\partial\mathfrak{E}_z}{\partial t}=\frac{\partial\mathfrak{H}_y}{\partial x}
\end{array}$$

If at one location $\mathfrak{H}_x = \mathfrak{E}_x = 0$ initially, then it will also be so in the future Diagonal pairs mutually independent.

$$\begin{array}{l|l}
-\frac{\mu}{c}\frac{\partial\mathfrak{H}_z}{\partial t}=\frac{\partial\mathfrak{E}_y}{\partial x} & \frac{\partial}{\partial x} \\[2ex]
\frac{\varepsilon}{c}\frac{\partial\mathfrak{E}_y}{\partial t}=-\frac{\partial\mathfrak{H}_z}{\partial x} & -\frac{\mu}{c}\frac{\partial}{\partial t}
\end{array}
\qquad \frac{\partial^2\mathfrak{E}_y}{\partial x^2}-\frac{\varepsilon\mu}{c^2}\frac{\partial^2\mathfrak{E}_y}{\partial t^2}=0$$

$$\mathfrak{E}_y=\mathfrak{F}(x-vt) \qquad \mathfrak{H}_z=\alpha\mathfrak{F}(x-vt)$$

$$v=\frac{c}{\sqrt{\varepsilon\mu}} \qquad \frac{\mu}{c}\alpha\frac{c}{\sqrt{\varepsilon\mu}}\mathfrak{F}=\mathfrak{F}$$

$$\alpha=\sqrt{\frac{\varepsilon}{\mu}}$$

[p. 74] In the case of vacuum $v = c$ $\mathfrak{H}_z = \mathfrak{E}_y$
For a dielectric $\mathfrak{H}_z\sqrt{\mu} = \mathfrak{E}_y\sqrt{\varepsilon}$, further $v = \frac{c}{\sqrt{\varepsilon\mu}}$
For light waves $\mu = 1$ $v = \frac{c}{\sqrt{\varepsilon}} = \frac{c}{n}$ $n = \sqrt{\varepsilon}$ holds for the majority of simple gases and for some liquids. In general, more complicated relations because the connection between $\mathfrak{D}$ and $\mathfrak{E}$ not so simple. On the generation of electric waves later on.
General differential equation of wave propagation in transparent media

$$\begin{array}{l|l}
-\frac{\mu}{c}\frac{\partial\mathfrak{H}_y}{\partial t}=\frac{\partial\mathfrak{E}_x}{\partial z}-\frac{\partial\mathfrak{E}_z}{\partial x} & \frac{\partial}{\partial z} \\[2ex]
-\frac{\mu}{c}\frac{\partial\mathfrak{H}_z}{\partial t}=\frac{\partial\mathfrak{E}_y}{\partial x}-\frac{\partial\mathfrak{E}_x}{\partial y} & -\frac{\partial}{\partial y}
\end{array}$$

$$\frac{\mu}{c}\frac{\partial}{\partial t}\underbrace{\left(\frac{\partial \mathfrak{H}_z}{\partial y}-\frac{\partial \mathfrak{H}_y}{\partial z}\right)}_{\frac{\varepsilon}{c}\frac{\partial \mathfrak{E}_x}{\partial t}}=-\frac{\partial}{\partial x}\left(\frac{\partial \mathfrak{E}_x}{\partial x}+\frac{\partial \mathfrak{E}_y}{\partial y}+\frac{\partial \mathfrak{E}_z}{\partial z}\right)+\Delta\mathfrak{E}_x$$

$$\frac{\mu\varepsilon}{c^2}\frac{\partial^2 \mathfrak{E}_x}{\partial t^2}-\Delta\mathfrak{E}_x=0.\ \text{etc}$$

These are the fundamental equations of the wave theory.

The Energy Principle and the Law of Conservation of Momentum

$$\begin{array}{l|c|c}
-\frac{\mu}{c}\frac{\partial \mathfrak{H}_x}{\partial t}=\frac{\partial \mathfrak{E}_z}{\partial y}-\frac{\partial \mathfrak{E}_y}{\partial z} & -c\frac{\mathfrak{H}_x}{4\pi} & \\
-\frac{\mu}{c}\frac{\partial \mathfrak{H}_y}{\partial t}=\frac{\partial \mathfrak{E}_x}{\partial z}-\frac{\partial \mathfrak{E}_z}{\partial x} & -c\frac{\mathfrak{H}_y}{4\pi} & \frac{\mathfrak{E}_z}{4\pi} \\
-\frac{\mu}{c}\frac{\partial \mathfrak{H}_z}{\partial t}=\frac{\partial \mathfrak{E}_y}{\partial x}-\frac{\partial \mathfrak{E}_x}{\partial y} & -c\frac{\mathfrak{H}_z}{4\pi} & -\frac{\mathfrak{E}_y}{4\pi}
\end{array}$$

$$\begin{array}{l|c|c}
4\pi j_x+\frac{\varepsilon}{c}\frac{\partial \mathfrak{E}_x}{\partial t}=\frac{\partial \mathfrak{H}_z}{\partial y}-\frac{\partial \mathfrak{H}_y}{\partial z} & c\frac{\mathfrak{E}_x}{4\pi} & \\
4\pi j_y+\frac{\varepsilon}{c}\frac{\partial \mathfrak{E}_y}{\partial t}=\frac{\partial \mathfrak{H}_x}{\partial z}-\frac{\partial \mathfrak{H}_z}{\partial x} & c\frac{\mathfrak{E}_y}{4\pi} & \frac{\mathfrak{H}_z}{4\pi} \\
4\pi j_z+\frac{\varepsilon}{c}\frac{\partial \mathfrak{E}_z}{\partial t}=\frac{\partial \mathfrak{H}_y}{\partial x}-\frac{\partial \mathfrak{H}_x}{\partial y} & c\frac{\mathfrak{E}_z}{4\pi} & -\frac{\mathfrak{H}_y}{4\pi}
\end{array}$$

$$\frac{\partial P_E}{\partial t}+c(\mathfrak{E}_x j_x+\mathfrak{E}_y j_y+\cdot)=-\frac{c}{4\pi}\left\{\frac{\partial}{\partial x}(\mathfrak{E}_y\mathfrak{H}_z-\mathfrak{E}_z\mathfrak{H}_y)+\cdot+\cdot\right\}$$ [75]

Is canceled by what comes from the right-hand side of the second system.

$$\begin{array}{lll}
\pm\frac{c}{4\pi}\frac{\partial}{\partial x}(\mathfrak{E}_y\mathfrak{H}_z-\mathfrak{E}_z\mathfrak{H}_y) & +\frac{c}{4\pi}\left(\mathfrak{E}_y\frac{\partial \mathfrak{H}_z}{\partial x}-\mathfrak{E}_z\frac{\partial \mathfrak{H}_y}{\partial x}\right) & \\
\text{-------------------} & \text{-------------------} & \\
\text{-------------------} & \text{-------------------} & \\
 & S_x \qquad\qquad S_y & S_z
\end{array}$$

[p. 75] Vector of the energy flow $\frac{c}{4\pi}(\mathfrak{E}_y\mathfrak{H}_z - \mathfrak{E}_z\mathfrak{H}_y)$

$$\frac{dE}{dt} + \text{heat loss} = \int S_n \cdot d\sigma.$$

Thus, the energy principle has been satisfied, with the expression for the energy being the same as in electrostatics.

The law of conservation of momentum; radiation pressure.

$$m_\nu \frac{d^2x_\nu}{dt} = X_\nu$$

Law of the equality of action & reaction $\sum X_n = 0$

From this $\sum m_1 \frac{d^2x_1}{dt^2} = 0$ for complete system. $\frac{d}{dt}\left[\sum m_1 \frac{dx_1}{dt}\right] = 0 \quad \sum m_1 \frac{dx_1}{dt} = \text{const.}$

If external forces $\frac{d}{dt}\sum m_1 \frac{dx_1}{dt} = \sum X_a$

Can the momentum of a system be increased by internal electrom. processes (Can a system start moving by itself?) We must calculate the sum of the ponderomotive forces acting on the system. Per unit volume

$$j_y\mathfrak{H}_z - j_z\mathfrak{H}_y = -\frac{1}{4\pi c}\left\{\frac{\partial\mathfrak{E}_y}{\partial t}\mathfrak{H}_z - \frac{\partial\mathfrak{E}_z}{\partial t}\mathfrak{H}_y\right\} - \frac{1}{8\pi}\frac{\partial}{\partial x}(\mathfrak{H}_y^2 + \mathfrak{H}_z^2)$$

$$+\frac{1}{4\pi}\left\{\frac{\partial}{\partial z}(\mathfrak{H}_x\mathfrak{H}_z) + \frac{\partial}{\partial y}(\mathfrak{H}_x\mathfrak{H}_y)\right\}\left(-\frac{1}{4\pi}\mathfrak{H}_x\underbrace{\left(\frac{\partial\mathfrak{H}_y}{\partial y} + \frac{\partial\mathfrak{H}_z}{\partial z}\right)}_{-\frac{\partial\mathfrak{H}_x}{\partial x}}\right) + \frac{1}{8\pi}\frac{\partial\mathfrak{H}_x^2}{\partial x}$$

$$0 = \frac{1}{4\pi c}\left\{\frac{\partial\mathfrak{H}_y}{\partial t}\mathfrak{E}_z - \frac{\partial\mathfrak{H}_z}{\partial t}\mathfrak{E}_y\right\} - \frac{1}{8\pi}\frac{\partial}{\partial x}(\mathfrak{E}_y^2 + \mathfrak{E}_z^2) + \frac{1}{4\pi}\{\text{-------}\}$$

$$-\mathfrak{E}_x\rho + \frac{1}{8\pi}\frac{\partial(\mathfrak{E}_x^2)}{\partial x}$$

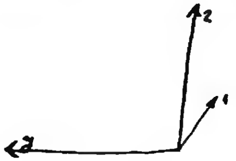
$$\begin{array}{l|l} \mathfrak{E}_x\rho + j_y\mathfrak{H}_z - j_z\mathfrak{H}_y & \frac{\partial}{\partial x}\left(-\frac{1}{2}\mathfrak{H}^2 + \mathfrak{H}_x^2\right) + \frac{\partial}{\partial y}(\mathfrak{H}_x\mathfrak{H}_y) + \frac{\partial}{\partial z}(\mathfrak{H}_x\mathfrak{H}_z) \\ = -\frac{\partial}{\partial t}\left\{\frac{1}{4\pi c}(\mathfrak{E}_y\mathfrak{H}_z - \mathfrak{E}_z\mathfrak{H}_y)\right\} & \quad + \frac{\partial}{\partial x}\left(-\frac{1}{2}\mathfrak{E}^2 + \mathfrak{E}_x^2\right) + \frac{\partial}{\partial y}(\mathfrak{E}_x\mathfrak{E}_y) + \frac{\partial}{\partial z}(\mathfrak{E}_x\mathfrak{E}_z) \\ \quad + \frac{1}{4\pi} & \end{array}$$

integrated over the whole system [p. 76]

$$\frac{d\mathfrak{J}_x}{dt} = -\frac{\partial}{\partial t}\underbrace{\int \frac{1}{4\pi c}(\mathfrak{E}_y\mathfrak{H}_z - \mathfrak{E}_z\mathfrak{H}_y)}_{J_x}$$

$$\mathfrak{J}_x + \mathfrak{J}_x = \text{konst.}$$

If we call $\mathfrak{J}_x$ the momentum of the electromagnetic field, then this tells us that the sum of the mechanical and the electromagnetic momentum in a complete system is constant.
Application to a plane wave ‖ x-axis
Electric force in Y-direction.



$$\mathfrak{E}_y = \mathfrak{H}_z \qquad \mathfrak{J}_{[-]=1} = \frac{1}{8\pi c}(\mathfrak{E}_y^2 + \mathfrak{H}_z^2) = \frac{1}{c}E_1$$

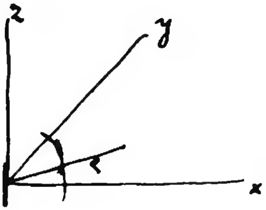
Momentum that impinges on the surface l per unit time = $\frac{I}{c}$ This is equal to radiation pressure.

Terms that were canceled out in integration. Maxwell stresses. The momentum transmitted by them to a unit of volume per unit time = momentum transm. to mech. system + increase of momentum in the element.

$$\frac{1}{c}\frac{\partial\mathfrak{E}}{\partial t} = \operatorname{curl}\mathfrak{H} - \frac{1}{c}\frac{\partial\mathfrak{H}}{\partial t} = \operatorname{curl}\mathfrak{E} \quad \frac{\partial\mathfrak{H}_x}{\partial z} - \frac{\partial\mathfrak{H}_z}{\partial x}$$

[p. 77]

Hertz's Oscillator[76]



$$\left.\begin{array}{l|l|l} \mathfrak{H}_x = \frac{\partial\psi}{\partial y} & \frac{\partial\psi}{\partial\rho}\cdot\frac{y}{\rho} & x \\ \mathfrak{H}_y = -\frac{\partial\psi}{\partial x} & \frac{\partial\psi}{\partial\rho}\cdot-\frac{x}{\rho} & y \end{array}\right\} 0$$

$$\left\langle\frac{1}{c}\right\rangle\frac{\partial\mathfrak{E}_x}{\partial t} = \frac{\partial^2\psi}{\partial x\partial z} = \frac{\partial}{\partial t}\frac{\partial^2 F}{\partial x\partial z} \qquad \frac{\partial F}{\partial t} = c\overset{\partial}{\psi}$$

$$\left\langle\frac{1}{c}\right\rangle\frac{\partial\mathfrak{E}_y}{\partial t} = \frac{\partial^2\psi}{\partial y\partial z} = \text{-----}$$

$$\left\langle\frac{1}{c}\right\rangle\frac{\partial\mathfrak{E}_z}{\partial t} = -\frac{\partial^2\psi}{\partial y^2} - \frac{\partial^2\psi}{\partial z^2} = \text{----}$$[77]

$$\mathfrak{E}_x = \frac{\partial^2 F}{\partial x\partial z} \qquad \mathfrak{H}_x = \frac{1}{c}\frac{\partial^2 F}{\partial y\partial t}$$

$$\mathfrak{E}_y = \frac{\partial^2 F}{\partial y\partial z} \qquad \mathfrak{H}_y = -\frac{1}{c}\frac{\partial^2 F}{\partial x\partial t}$$

$$\mathfrak{E}_z = -\frac{\partial^2 F}{\partial y^2} - \frac{\partial^2 F}{\partial z^2}$$

$$= \frac{\partial^2 F}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 F}{\partial t^2}$$

F depends only on r

$$\frac{\partial F}{\partial x} = \frac{dF}{dr}\cdot\frac{x}{r}$$

$$\frac{\partial^2 F}{\partial x^2} = \frac{d^2 F}{dr^2}\frac{x^2}{r^2} + \frac{dF}{dr}\left(\frac{1}{r} - \frac{x^2}{r^2}\right)$$[78]

$$\Delta F = \frac{d^2 F}{dr^2} + \frac{2}{r}\frac{dF}{dr} = \frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dF}{dr}\right)$$

Here $$\frac{1}{c^2}\frac{\partial^2 F}{\partial t^2} = \Delta F$$

$$\frac{1}{c^2}\frac{\partial^2 F}{\partial t^2} = \frac{1}{r^2}\frac{d}{dt}\left(r^2\frac{dF}{dr}\right)$$

Solution $F = \frac{1}{r}f\left(t - \frac{r}{2}\right)$

$$\frac{1}{c^2}\cdot\frac{1}{r}\frac{d^2 f}{dt^2} = \frac{1}{r^2}\cdot\frac{1}{c^2}r\ddot{f} \quad \text{q.e.d.}$$

$$\frac{\partial F}{\partial r} = -\frac{1}{r^2}f - \frac{1}{cr}\dot{f}\,\Big|\, r^2\frac{\partial F}{\partial r} = -f - \frac{1}{c}$$

$$\frac{d}{dr}(\quad) = \frac{1}{c}\dot{f} - \frac{1}{c}\dot{f} + \frac{1}{c^2}r\ddot{f}$$

Behavior in immediate vicinity of the oscillator.

$$\mathfrak{E} = -\frac{\partial\varphi}{\partial z} \qquad \varphi = -\frac{\partial\frac{f}{r}}{\partial z} \text{ Pot. of a dipole}$$

$$e \quad \begin{matrix} +\varepsilon \\ -\varepsilon \end{matrix} \qquad \underbrace{\frac{\varepsilon}{r} - \frac{\varepsilon}{r'}} \qquad \frac{1}{r'} = \frac{1}{r} - \frac{\partial\frac{1}{r}}{\partial z}l$$

$$\frac{\partial\frac{\varepsilon l}{r}}{\partial z} \qquad \varepsilon l = f$$

f is the moment of the dipole

Thus, process oscillation of dipole, which is ∞ small compared with the wave length. Calculation of the energy radiated outwards.[79]

[p. 78]

$$\mathfrak{E}_x = \frac{xz}{c^2r^3}\ddot{f} \qquad \mathfrak{H}_x = -\frac{y}{c^2r^3}\ddot{f} \qquad z^2x^2$$

$$\mathfrak{E}_y = \frac{yz}{c^2r^3}\ddot{f} \qquad \mathfrak{H}_y = \frac{x}{c^2r^3}\ddot{f} \qquad z^2y^2$$

$$z^4 - 2z^2r^2 + r^4$$

$$\mathfrak{E}_z = -\frac{x^2+y^2}{c^2r^3}\ddot{f} \qquad \mathfrak{H}_z = 0$$

$$r^4 - z^2r^2$$

$$\sqrt{\quad} = r^2\left(1 - \frac{z^2}{r^2}\right) =$$

$$\mathfrak{E} \perp \mathfrak{H} \qquad \mathfrak{E} \,\&\, \mathfrak{H} \perp xyz \qquad |\mathfrak{E}| = \frac{1}{c^2r}\ddot{f}\sin^2\vartheta = |\mathfrak{H}|$$[80]

Rad[iation of?] energy[81]

$$\int_T \int |\mathfrak{E}||\mathfrak{H}|r^2\,d\Omega = \frac{1}{c^4}\int \ddot{f}^2\,dt \underbrace{\int \sin^2\vartheta\,dw \quad 2\pi\sin\vartheta\,d\vartheta}_{2\pi\int_0^\pi (1-\cos^2\vartheta)\sin\vartheta\,d\vartheta}$$

$$= \left| -\cos\vartheta + \frac{\cos^3\vartheta}{3} \right|_0^\pi = 2 - \frac{2}{3} = \frac{4}{3}$$

In unit time $\frac{c}{4\pi}\cdot\frac{4\cdot 2\pi}{3c^4}\overline{\ddot{f}^2} = \frac{2}{3c^3}\overline{\ddot{f}^2}$

If excitory sines[82] $f = f_0\cos(2\pi nt)$, then $\ddot{f} = f_0(2\pi n)^2\cos(\ )$

$$\overline{\ddot{f}^2} = \frac{1}{2}f_0^2(2\pi n)^4$$

$$A = \frac{1}{3c^3}(2\pi n)^4 f_0^2$$

$$l = 100 \qquad p_{max} = 3 \qquad C = 30 \qquad n = 10^8$$

$$f_0 = 10^4 \qquad 2\pi n = 6\ 10^8$$

$$A = \frac{10^{35}\ 10^8}{3\cdot 27\cdot 10^{30}} \approx 10^{11} = 2000 \text{ cal. per sec.}$$

Doc. 12

Comment on Eötvös's Law

[1] by A. Einstein

[*Annalen der Physik* 34 (1911): 165–169]

Eötvös established empirically the following lawlike regularity for liquids, which, as we know, proves to be correct to a remarkable approximation:

$$\gamma v^{2/3} = k(\tau - T). \tag{1}$$

Here γ is the surface tension, v the molecular volume, k a universal constant, T
[2] the temperature, and τ a temperature that deviates only slightly from the critical temperature.

γ is the free energy per unit surface area, hence

$$\gamma - T\frac{d\gamma}{dT}$$

is the energy per unit surface area. Taking into account that, in comparison with γ, v depends only slightly on the temperature, one can write in similarly good approximation:

$$\left(\gamma - T\frac{d\gamma}{dT}\right)v^{2/3} = k\tau. \tag{1a}$$

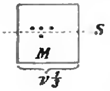
But according to the rule of corresponding states, the boiling temperature at atmospheric
[3] pressure is approximately equal to a fixed fraction of the critical temperature, and on the
[4] other hand, the boiling temperature is proportional to the heat of evaporation (Trouton's rule).

From this it follows that equation (1a) also entails the approximate validity of the equation

[5]

$$\left(\gamma - T\frac{d\gamma}{dT}\right)v_s^{2/3} = k'(D_s - RT_s). \tag{1b}$$

[6] Since γ is to a high degree of approximation a linear function of the temperature, the bracket on the left-hand side does not need to be calculated for the boiling temperature
[7] at atmospheric pressure. The left-hand side of the equation equals the energy Uf necessary to increase the surface of the substance by an area equal to a lateral surface of a gram-molecular cube. $D_s - RT_s$ is the internal energy U_i to be expended in the evaporation of one gram-molecule. Equation (1b) can therefore be written in the form



$$\text{(1c)} \qquad \frac{U_f}{U_i} = k'$$

Let us now interpret the last equation. Let S (see the figure) be a cross-section of a gram-molecular cube parallel to a lateral face. $2U_f$ is then equal to the potential energy (taken as negative) that corresponds to the totality of the interactions between the molecules on the one side of S and those on the other side of S. U_i is the potential energy (taken as negative) that corresponds to the interactions of all the molecules of the cube.[1]

The most obvious fundamental hypothesis concerning the molecular forces that leads to a simple relation between U_f and U_i is the following one:

The radius of the molecule's sphere of action is large compared with the molecule but is of the same size for different kinds of molecules. At a distance r two molecules exert on each other a force whose negative potential energy is given by $c^2 f(r)$, where c is a constant characteristic of the molecule, $f(r)$ is a universal function of r, and $f(\infty)$ is equal to zero. For this case to lead to simple relations, $f(r)$ must be so constituted that the sums representing U_f and U_i can be written as integrals; we will assume this as well (with van der Waals). We then obtain by simple calculation [8]

$$U_f = c^2 N^2 K_2 \nu^{-\frac{4}{3}},$$
$$U_i = c^2 N^2 K_1 \nu^{-1}.$$ [9]

Here we have

$$K_1 = \int f(r) d\tau,$$ [10]

extended over the entire volume, and

$$K_2 = \frac{1}{2} \int_0^\infty \psi(\Delta)\, d\Delta,$$

where

$$\psi(\Delta) = \int_\Delta^\infty dx \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(r)\, dy\, dz.$$ [11]

---

[1] There is a noteworthy inaccuracy here, inasmuch as certainly not all of the energy U_i can be designated as potential energy in the sense of mechanics; this would be permissible only if the specific heat at constant volume would be the same in the liquid and in the gaseous state. It would surely be more correct to introduce the heat of evaporation extrapolated to absolute zero.

Thus, K_1 and K_2 are universal constants that depend only on the elementary law of molecular forces. From this one obtains

$$\frac{U_f}{U_i} = \frac{K_2}{K_1} v^{-1/3}, \tag{2}$$

[12] in contradiction to (1c), the equation to be regarded as expressing the results of experience. Even without any calculation one can see that, neglecting the universal factors, the ratio of U_f to U_i must be the same as the ratio of the radius of the molecular sphere of action to the edge of the gram-molecular cube ($v^{1/3}$). Thus, if the radius of the sphere of action is universal, one cannot arrive at equation (1c) but only at equation (2).

It can easily be seen that, in case equation (2) were valid, it would be impossible to [13] draw a conclusion about the molecular weight of a liquid from the capillarity constant.

To arrive at equation (1c), one must start from the assumption that the radius of the molecular sphere of action is proportional to the quantity $v^{1/3}$ or, what amounts to the same, to the distance between neighboring molecules of the liquid. This assumption seems rather absurd at first sight, because what should the radius of the sphere of action of a molecule have to do with how far away the neighboring molecules are situated? The supposition becomes reasonable only in the case when *only the neighboring molecules*, but [14] not those farther removed, are within the region of action of a molecule. In that case, in accordance to what has been said above, equation (1a) must be obtained, and we are even in a position to estimate the value of the constant k'. The argument to that effect, which I am now going to present, could probably be replaced by a more exact one; but I have chosen it because it makes do with a minimum of formal elements.

Let me conceive of the molecules as being regularly distributed in a quadratic lattice. In this lattice I consider an elementary cube, the edges of which contain three molecules each, so that the entire cube contains $3^3 = 27$ molecules. One of them is in the center. The other 26, and only these, I consider as neighboring the molecule in the center, and [15] I make my calculation by assuming that they are equidistant from the central molecule. If the potential energy (taken as negative) of a molecule with respect to one of its neighbors is denoted by φ, then its potential energy with respect to all the neighboring molecules is equal to 26φ, and hence

$$U_i = \frac{1}{2} N \cdot 26\varphi.$$

If we, further, imagine that our central molecule M lies immediately below the plane S in the figure, and that the boundary surfaces of the gram-molecular cube depicted there are parallel to the lateral faces of the elementary cube of the molecular lattice, then our molecule M is in interaction with 9 molecules of the next-upper layer. Since $N^{2/3}$

such molecules M lie directly below the surface S, the potential energy, which we denoted above by $2U_f$, will be given by

$$2U_f = 9 \cdot N^{2/3} \cdot \varphi.$$

Thus, we get

$$\frac{U_f}{U_i} = \frac{9}{26} N^{-1/3},$$

or, if we substitute the value $7 \cdot 10^{23}$ for N,

$$\frac{U_f}{U_i} = 3 \cdot 10^{-9}.$$

On the other hand, I calculated the constant k'—which, according to (1c), should be identical to the value just obtained—for mercury and benzene from experimental data by means of the empirical equation (1b), and obtained the values

$$5.18 \cdot 10^{-9}$$
$$5.31 \cdot 10^{-9}.$$

This order-of-magnitude agreement with the value obtained by the rough theoretical argument presented above is very remarkable.

Stimulated by an oral remark by my colleague G. Bredig, I also pondered what the [16] order of magnitude of the theoretically obtained value of U_f/U_i would be if one assumed that the molecule interacts not only with its immediate neighbors, but also with molecules that are farther removed. The cube that contains the molecules in interaction with one molecule will then have n^3 rather than 3^3 molecules. In that case U_f/U_i comes out nearly proportional to n. Thus, one still obtains a value for U_f/U_i of the right order [17] of magnitude for $n = 5$ or $n = 7$. Nevertheless, it is highly probable that a molecule interacts only with its immediate neighbors, because it must be considered as very improbable that the radius of the molecular sphere of action would be proportional to the cubic root of the molecular volume without depending otherwise on any physical constant of the molecule.

There is one more remark that comes to mind in connection with this argument. We know that substances with very small molecules deviate considerably from the law of corresponding states. Should this not be related to the fact that the radius of the molecular sphere of action of such substances is more than three times as large as the molecular radius?

(Received on 30 November 1910)

Doc. 13

A Relationship between Elastic Behavior and Specific Heat in Solids with a Monatomic Molecule

by A. Einstein

[*Annalen der Physik* 34 (1911): 170–174]

My colleague, Professor Zangger, drew my attention to an important note published [1] recently by Sutherland.[1] Sutherland asked himself whether the elastic forces of solids are forces of the same kind as the forces that drive the carriers of infrared proper oscillations back to their position of rest, and hence give rise to their proper frequencies. He found that the question should very probably be answered in the affirmative on the basis of the following fact: the infrared proper frequencies are of the same order of magnitude as the frequencies one would have to apply in order to send through the body elastic transverse vibrations whose half-wavelength is equal to the distance between the [3] neighboring molecules of the body.

Notwithstanding the importance of Sutherland's analysis, it is clear that this way leads at best to a rough order-of-magnitude relation, particularly because it must be assumed that the known infrared proper frequencies are chiefly to be viewed as oscillations of differently charged ions of a molecule with respect to one another, while elastic [4] oscillations are to be viewed as oscillations of whole molecules with respect to one another. It seems to me therefore that a more exact test of Sutherland's idea is possible only for substances with monatomic molecules, which, according to both experience and the theoretical model, do not manifest optically detectable proper oscillations of the [5] known kind. However, according to my theory of the specific heat of solids,[2] which I based on Planck's theory of radiation, it is possible to determine the proper frequencies of heat-carrying monatomic bodies from the dependency of the specific heat on the temperature. These proper frequencies can be used to test Sutherland's conception by comparing them with those obtained from elasticity. One way in which this can be done is given below, and let me say here right away that, in this way, Sutherland's conception of the essential identity of the elastic forces and those determining the proper frequency was satisfactorily confirmed in the case of silver.

An *exact* calculation of the proper oscillation frequencies from the elasticity constants is out of the question for the time being. Instead, we are using here a rough method of

---

[2] [1] W. Sutherland, *Phil. Mag.* 20 (1910): 657.

[6] [2] A. Einstein, *Ann. d. Phys.* 22 (1907): 180.

calculation that is similar to the one used in the preceding paper, but that should in essence hit the mark. [7]

First we envision the molecules of the substance arranged in a quadratic spatial lattice. Each molecule has then 26 neighboring molecules, which, to be sure, are not equally distant from it. However, we will do our calculation as if these 26 neighboring molecules were equidistant from the molecule in question when they are at rest.

Now we have to choose some plausible representation of the molecular forces that is as simple as possible. First we introduce the assumption that is fundamental for what follows and that has been proved correct for liquids in the preceding communication, namely that each molecule interacts only with its neighbor molecules, but not with those farther away. Let two neighboring molecules exert a central force on each other that [8] vanishes when the distance between the molecules equals d. When the distance between the molecules is $d - \Delta$, a repulsive force of the magnitude $a\Delta$ will be exerted.

We now calculate the force with which the 26 neighboring molecules oppose the displacement of one molecule. In order to do this, we imagine that, rather than being distributed over the surface of a cube, the 26 neighboring molecules are distributed over the surface of a sphere of the same volume, the radius of which is to be chosen to be equal to d, so that we have

$$\frac{4}{3}d^3\pi = 8\frac{v}{N}, \tag{1}$$

if v denotes the molecular volume of the substance, and N the number of molecules in one gram-molecule. We imagine that the molecule lying in the center of the sphere is displaced by a distance x in an arbitrary direction, x being small compared with d, and calculate the force that opposes the displacement, proceeding as if the mass of the 26 molecules were distributed uniformly over the surface of the sphere. On the solid angle element $d\kappa$ that is drawn from the molecule with an axis that forms the angle ϑ with the direction of the displacement x, there will then lie $26 . (d\kappa/4\pi)$ molecules, which exert the force

$$-\frac{26}{4\pi}d\kappa \cdot a \cdot x\cos\vartheta \cdot \cos\vartheta$$

in the direction of the displacement x. Integrating, we obtain for the force acting on the displaced molecule the value

$$-\frac{26}{3}ax.$$

From this we obtain, if we take into account that M/N is equal to the mass of one molecule (M = molecular weight of the substance), the proper frequency ν and the corresponding vacuum wave length λ of the molecule. We have

(2)
$$\nu = \frac{1}{2\pi}\sqrt{\frac{26}{3}a \cdot \frac{N}{M}}$$

and

[9] (2a)
$$\lambda = 2\pi e \sqrt{\frac{3}{26}\frac{M}{aN}}.$$

Based on the same approximative assumptions, we now calculate the coefficient of compressibility of the substance. To this end, we express in two different ways the work A that must be applied in uniform compression, and set the two expressions equal to one another.

The work that must be applied to reduce the distance between two neighboring molecules by Δ is $(a/2)\Delta^2$. Since each molecule has 26 neighboring molecules, the work to be applied to reduce its distance from the neighboring molecules is $26 . (a/2)\Delta^2$. Since [10] there are N/v molecules in a unit volume, and each term $(a/2)\Delta^2$ belongs to two molecules, we obtain

$$A = \frac{26}{4} \cdot \frac{N}{v} a \Delta^2.$$

On the other hand, if κ denotes the compressibility, and Θ the contraction of the unit volume, then $A = 1/2\kappa . \Theta^2$, or, since $\Theta = 3\Delta/d$,

$$A = \frac{9}{2}\frac{\Delta^2}{\kappa \cdot d^2}.$$

Equating these two values of A, we obtain

(3)
$$\kappa = \frac{18}{26}\frac{v}{a \cdot d^2 \cdot N}.$$

Eliminating a and d from equations (1), (2a), and (3) we obtain

[11]
$$\lambda = \frac{2\pi}{\sqrt{6}}\left(\frac{6}{\pi}\right)^{1/3}\frac{c}{N^{1/3}}M^{1/3}\rho^{1/6}\sqrt{\kappa} = 1.08 \cdot 10^3 \cdot M^{1/3}\rho^{1/6}\sqrt{\kappa}.$$

Of course, the formula assumes that no polymerization takes place. In what follows, I used this formula to calculate the proper wavelengths (as a measure of proper frequencies) of those metals for which the cubic compressibilities were determined by Grüneisen.[3] These are the results:[4]

---

[12] [3] E. Grüneisen, *Ann. d. Phys.* 25 (1908): 848.

[4] The temperature dependence of cubic compressibility has been neglected.

Substance	$\lambda \cdot 10^4$	Substance	$\lambda \cdot 10^4$
Aluminum	45	Palladium	58
Copper	53	Platinum	66
Silver	73	Cadmium	115
Gold	79	Tin	102
Nickel	45	Lead	135
Iron	46	Bismuth	168

According to the theory of specific heats derived from Planck's radiation theory, specific heat is supposed to decline as one approaches absolute zero according to the following law:

$$C = 3R \frac{e^{-\frac{a}{T}}\left(\frac{a}{T}\right)^2}{\left(e^{-\frac{a}{T}} - 1\right)^2},$$ [13]

where C denotes the specific molar heat and it has been stipulated that

$$\frac{h\nu}{k} = a = \frac{h.c}{k.\lambda}.$$

Here h and κ are the constants of Planck's radiation formula. Thus, one can use the [14] behavior of specific heats to determine λ for a second time. Among the substances listed above, silver is the only one to have had its specific heat at low temperatures determined with adequate accuracy. Nernst[5] found for silver that $a = 162$, which yields $\lambda . 10^4 = 90$, while we calculated from the constants of elasticity that $\lambda . 10^4 = 73$. This close agreement is really surprising. A still more exact test of Sutherland's conception is likely to be realized only by perfecting the molecular theory of solids.

(Received on 30 November 1910)

---

[5] Cf. W. Nernst, *Bulletin des Séances de la Société franç. de Phys.* (1910): 19–48. [15]

Doc. 14

Correction to My Paper: "A New Determination of Molecular Dimensions"[1]

by A. Einstein

[*Annalen der Physik* 34 (1911): 591–592]

Mr. Bacelin, who has done an experimental investigation on viscosities of suspensions at
Mr. Perrin's instigation, wrote me a few weeks ago that according to his results the
coefficient of viscosity of suspensions is considerably greater than that corresponding to
[2] the formula I developed in §2 of my paper. I therefore asked Mr. Hopf to check my
[3] calculations, and he indeed found a mathematical error that considerably affects the
result. In what follows, I will correct this error.

[4] The expressions for the pressure components X_y and X_z on p. 296 of the paper mentioned above are wrong because of an error in the differentiation of the velocity components u, v, w. These expressions should read:

$$X_x = -2kA + 10kP^3\frac{A\xi^2}{\rho^5} - 25kP^3\frac{M\xi^2}{\rho^7},$$

$$X_y = 5kP^3\frac{(A+B)\xi\eta}{\rho^5} - 25kP^3\frac{M\xi\eta}{\rho^7},$$

$$X_z = 5kP^3\frac{(A+C)\xi\zeta}{\rho^5} - 25kP^3\frac{M\xi\zeta}{\rho^7},$$

where

$$M = A\xi^2 + B\eta^2 + C\zeta^2.$$

If one then calculates the energy transmitted per unit time by the pressure forces to the
[5] liquid contained in the sphere of radius R, one obtains, instead of equation (7) on p.
296,

$$W = 2\delta^2 k(V + \tfrac{1}{2}\Phi). \tag{7}$$

Using this corrected formula, one obtains the equation

[6] $$k^* = k(1 + 2.5\varphi).$$

instead of the equation $k^* = k(1 + \varphi)$, developed in §2. Thus, the viscosity coefficient k^* of the suspension is 2.5 times more strongly influenced by the total volume φ of the spheres suspended in a unit volume than according to the formula found there.

---

[1] [1] A. Einstein. *Ann. d. Phys.* 19 (1906): 289 ff.

If one takes the corrected formula as a basis, then, instead of 2.45 cm^3, the value given in §3, one obtains for the volume of 1 g sugar dissolved in water, and thus a value [7] that deviates considerably less from 0.61 cm^3, the volume of 1 g solid sugar. Finally, from the viscosity and the diffusion of dilute sugar solutions, one obtains the value $N = 6.56 \cdot 10^{23}$ for the number of molecules in one gram-molecule, instead of the value $N = 4.15 \cdot 10^{23}$, given in the appendix of that paper. [8]

Zurich, January 1911. (Received on 21 January 1911)

Doc. 15

Comment on My Paper:[1] "A Relationship between Elastic Behavior..."

by A. Einstein.

[*Annalen der Physik* 34 (1911): 590]

In the paper cited above I mentioned Sutherland as the discoverer of the connection between the elastic and optical behavior of solids. It escaped my notice that [3] E. Madelung was the first to draw attention to this fundamentally important connection.[2] Madelung found a quantitative relationship between the elasticity and (optical) proper frequency of diatomic compounds that corresponds exactly to the one that I derived for monatomic substances and that shows quite satisfactory agreement with experience. It should be especially emphasized that Madelung is able to arrive at his relation only by assuming that the forces acting between the atoms of a molecule are of the same order [4] of magnitude as the forces acting between like atoms of adjacent molecules; in other words, the molecular bond does not seem to persist in the solid state of the substances investigated by Madelung; these substances seem to be completely dissociated. This [5] agrees totally with the pictures to which the investigation of molten salts has led.

Zurich, January 1911. (Received 30 January 1911)

---

[1] [1] A. Einstein, *Ann. d. Phys.* 34 (1911): 170 ff.

[2] E. Madelung, *Nachr. d. kgl. Ges. d. Wissensch. zu Göttingen. Math-phys. Kl.* (20 Feb. 1909 and 29 Jan. 1910); *Physik. Zeitschr.* 11 (1910): 898–905. [2]

Doc. 16

Comment on a Fundamental Difficulty in Theoretical Physics

Zurich, 2 January 1911

Our present physical world picture rests on the fundamental equations of point mechanics and on Maxwell's equations for the electromagnetic field in a vacuum. It becomes more and more apparent that all those consequences of this foundation that refer to slow, i.e., not rapid periodical processes, are in excellent agreement with experience. We have [1] succeeded in arriving at a general formulation of the limits of validity of thermodynamics with the help of point mechanics, and in deriving from the latter the fundamental laws of thermodynamics. We have succeeded in entirely different ways in determining the absolute sizes of atoms and molecules with undreamed-of accuracy. We have also been able to derive the law of thermal radiation for long wavelengths and high temperatures from statistical mechanics and electrodynamics. But the foundations of the theory leave [2] us in the lurch when it comes to all those phenomena that involve the transformation of energy of rapid periodical processes. We know of no flawless derivation of the law of radiant heat for short wavelengths and low temperatures. We do not know the reason [3] why high molecular temperatures are needed for the generation of short-wave radiation, and why the absorption of the latter can produce elementary processes of relatively great energy. We do not know why the specific heat at low temperatures is smaller than predicted by the Dulong-Petit law. We know just as little about why those mechanical degrees of freedom of matter that must be postulated in order to comprehend the optical properties of transparent bodies make no contribution to the specific heats of these bodies. [4]

But one thing has been done. M. Planck has shown that one arrives at a radiation formula that is in agreement with experience if one modifies the formulas resulting from our theoretical foundations as though the energy of oscillations of frequency ν could only occur in integral multiples of the quantity $h\nu$. This modification also leads to a [5] modification of the consequences of mechanics that has thus far proved useful if rapid oscillations are involved. A proper theory has not yet come into being, but it can be said [6] with certainty: point mechanics is not valid for rapid periodic processes, and the customary conception of the distribution of radiant energy in space can also not be maintained. [7]

A. Einstein

Doc. 17

The Theory of Relativity[1]

[1] by A. Einstein

[*Naturforschende Gesellschaft in Zürich. Vierteljahrsschrift* 56 (1911): 1–14]

The one basic pillar upon which the theory designated as the "theory of relativity" rests is the so-called principle of relativity. First I will try to make clear what is understood by the principle of relativity. Picture to yourself two physicists. Let both physicists be equipped with every physical instrument imaginable; let each of them have a laboratory. Suppose that the laboratory of one of the physicists is arranged somewhere in an open field, and that of the second in a railroad car traveling at constant velocity in a given direction. The principle of relativity states the following: if, using all their equipment, these two physicists were to study all the laws of nature, one in his stationary laboratory and the other in his laboratory on the train, they would discover exactly the same laws of nature, provided that the train is not shaking and is traveling in uniform motion. Somewhat more abstractly, we can say: according to the principle of relativity, the laws of nature are independent of the translational motion of the reference system.

Let us consider the role that this principle of relativity plays in classical mechanics. Classical mechanics is based first and foremost on Galileo's principle, according to which a body not subjected to the influences of other bodies finds itself in uniform, rectilinear motion. If this principle holds for one of the laboratories mentioned above, then it holds for the other one as well. This we can deduce directly from intuition; however, we can also deduce this from the equations of Newtonian mechanics if we transform these equations to a reference system that moves uniformly relative to the original reference system.

All I have been talking about is laboratories. However, in mathematical physics, it is customary to relate things to coordinate systems and not to a specific laboratory. What is essential in this relating-to-something is the following: when we state anything whatsoever about the location of a point, we always indicate the coincidence of this point with some point of a specific other physical system. If, for example, I choose myself as this material point, and say, "I am at this location in this hall," then I have brought myself into spatial coincidence with a certain point of this hall, or rather, I have asserted this coincidence. This is done in mathematical physics by using three numbers, the so-called coordinates, to indicate with which points of the rigid system, called the coordinate system, the point whose location is to be described coincides.

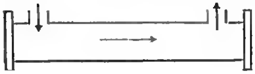
---

[1] Lecture given at the meeting of the Zurich *Naturforschende Gesellschaft* on 16 January 1911.

This would be the most general description of the principle of relativity. If a physicist of the 18th century or the first half of the 19th century would have been asked whether he had any doubts about this principle, he would have answered the question with a resolute "no." He had no grounds for doubting it, since at that time people were convinced that whatever happened in nature could be reduced to the laws of classical mechanics. I shall now analyze how experience has led physicists to the formulation of physical laws that contradict this principle. In order to do so, we must briefly consider the gradual development of optics and thermodynamics during the last few decades from the standpoint of the principle of relativity.

Light displays interference and diffraction exactly like sound waves do, so that one felt compelled to view light as a wave motion, or, generally, as a periodically changing state of some medium. This medium was named "ether." Until recently, the existence [2] of such a medium seemed absolutely certain to physicists. The theory to be outlined below is incompatible with the ether hypothesis, but for the time being we shall hold on to it. Let us now see how ideas about this medium developed, and what kind of problems the introduction of this physical theory based on the ether hypothesis engendered. We already said that one thought that light consists in vibrations of a medium, i.e., that the medium takes over the propagation of light and heat vibrations. As long as one dealt solely with optical phenomena of bodies *at rest*, one had no reason to wonder about motions of this medium other than the motion presumed to constitute light. It was simply assumed that—except for the oscillatory motions that were supposed to constitute light—both this medium and the material bodies under consideration were in a state of rest.

Once one started to consider the optical phenomena of *moving* bodies and simultaneously, in connection with it, the electromagnetic properties of moving bodies, one had to ask how the luminiferous ether behaves if we impart different velocities to the bodies in a physical system under our observation. Does the luminiferous ether move together with the bodies, so as to move at each location in the same manner as the matter situated there, or is this not the case? The simplest assumption is that the luminiferous ether moves everywhere exactly like matter. The second possible assumption, which is also very simple, is that the luminiferous ether does not participate at all in the motions of matter. And then intermediate cases would also be possible, these intermediate cases being characterized by a motion of ether in space that is to a certain degree independent of the motion of the matter. Let us now see how one tried to get an answer to this question. The first important clarification one obtained came from a very important experiment conducted by the French physicist Fizeau. This experiment was set up to answer the following question: [3]



The front and rear ends of the tube in the accompanying sketch are closed by a glass plate. Pipe connections attached at the two ends make it possible for a liquid to flow through the tube in the direction of its axis. How does the velocity with which the liquid flows through the tube affect the propagation velocity of a light ray passing axially through the tube? If it is true that the luminiferous ether moves with the matter that flows through the tube, then the following picture obtains. If we assume that in water at rest light propagates with velocity V, V thus being the velocity of light relative to water, while v is the velocity of the water relative to the tube, then we must say: If the luminiferous ether adheres to the water, then the velocity of light relative to the water is always the same, regardless of whether the water is in motion or not. Accordingly, one should expect that the velocity of the propagation of light relative to the tube is greater by v if the liquid is in motion than if it is at rest. In Fizeau's experiment, one of two beams of light capable of interference traversed the tube in the manner described. From the influence of the known velocity of motion of the liquid on the position of the interference fringes, it was possible to calculate the influence that the water moving with velocity v exerted on the velocity of the propagation of light relative to the tube at rest. Fizeau found that the motion of the liquid did not increase the velocity of light relative to the tube by v, but only by a fraction of this value ($v\left(1 - \frac{1}{n^2}\right)$, if n denotes the refractive capacity of the liquid). If this refractive capacity is very close to 1, i.e., if light propagates almost as fast in the liquid as in empty space, then the motion of the liquid has practically no influence. From this it had to be concluded that the conception according to which light always propagates with the same velocity V relative to water is not compatible with experience.

The next simplest hypothesis was that the luminiferous ether does not participate in the motion of the matter. From this hypothesis, as a basis, we cannot deduce in such a simple manner how the optical phenomena are influenced by the motion of matter. But in the mid-'90s H. A. Lorentz succeeded in formulating a theory based on the assumption [4] of a completely stationary luminiferous ether. His theory provides a completely correct account of almost all known phenomena in the optics and electrodynamics of moving bodies, including the experiment of Fizeau we have just discussed. Let me add at once that a theory fundamentally different from that of Lorentz, which would be based on simple and intuitive assumptions and would accomplish the same ends, could not be formulated. For that reason, the theory of the stationary luminiferous ether had to be accepted for the time being as the only theory compatible with the totality of experience.

Let us now consider this theory of the stationary ether from the standpoint of the principle of relativity. If we designate as acceleration-free all those systems with respect to which material points not subjected to external forces move uniformly, the principle of relativity states the following: The laws of nature are identical in all acceleration-free systems. On the other hand, Lorentz's fundamental hypothesis of the stationary luminiferous ether distinguishes from among all possible acceleration-free moving systems those in a state of motion, namely systems that are at rest relative to this light medium. Thus, even though one cannot say according to this conception that there exists an absolute motion in the philosophical sense—because this is completely out of the question, we can only conceive of relative changes of position of bodies—an absolute motion in the physical sense is affirmed insofar as we have privileged one state of motion, namely, the state of rest relative to the ether. Every body that is at rest relative to ether can be designated, in a certain sense, as absolutely at rest. Reference systems that are at rest relative to the ether are distinguished from all other acceleration-free systems. In that sense Lorentz's basic concept of the stationary luminiferous ether does not satisfy the principle of relativity. The basic concept of the stationary luminiferous ether leads to the following general argument: Let a reference system k be at rest relative to the luminiferous ether. Let another reference system k' be in uniform motion relative to the luminiferous ether. It is to be expected that the relative motion of k' with respect to the ether will have an influence on the laws of nature that are valid with respect to k'. Hence, it was to be expected that the laws of nature with respect to k' would differ from those with respect to k on account of the motion of k' in the luminiferous ether. One had to tell oneself, furthermore, that the earth and its laboratories could not possibly be at rest relative to this luminiferous medium throughout the entire year, i.e., that the earth must, therefore, play the role of a reference system k'. One, therefore, had to assume that some phenomenon could be found in which the influence of this motion on the experiments in our laboratories would come to the fore. One would think that, because of this relative motion, our physical space in the form we find it on the earth would show different behavior in different directions. But there was not a single case in which it was possible to prove something like that.

One was, thus, in an awkward position with regard to the ether. Fizeau's experiment says: The ether does not move with matter, i.e., there does exist a motion of the light medium relative to matter. But all attempts to detect this relative motion yielded negative results. These are two results that seem to contradict one another, and physicists found it enormously distressing that they could not get rid of this unpleasant conflict. They were bound to ask themselves whether it might not be possible, after all, to reconcile the principle of relativity—to which no exception could be found in spite of all the searching—with Lorentz's theory. Before delving into this question, we will extract from Lorentz's theory of the stationary luminiferous ether the following aspects most essential to us. What is the physical meaning of the statement that there exists a stationary luminiferous ether? The most important content of this hypothesis can be expressed as follows: There exists a reference system (called in Lorentz's theory "a system at rest relative to the ether") with respect to which every light ray propagates in a vacuum with the universal velocity c. This ought to hold independently of whether the light-emitting body is in motion or at rest. We shall designate this proposition as the [5] principle of constancy of the velocity of light. Thus, the question we have just asked can also be formulated as follows: Is it impossible to reconcile the principle of relativity, which seems to be satisfied without exception, with this principle of constancy of the velocity of light?

To begin with, the following obvious consideration argues against this possibility: If every ray of light propagates with the velocity c relative to the reference system k, then the same cannot be true with respect to the reference system k', if k' is in motion relative to k. For if k' is moving with velocity v in the direction of the propagation of a light ray, then the propagation velocity of the light ray relative to k' would have to be set equal to $c - v$ according to our customary views. The laws of propagation of light with respect to k' would then differ from those with respect to k, which would mean a violation of the principle of relativity. That is a frightful dilemma. But it turned out that nature is not responsible for this dilemma; rather, this dilemma stems from the fact that we have been making tacit and arbitrary assumptions in our arguments, and thus also in the argument just given, and that these have to be dropped in order to arrive at a consistent and simple interpretation of things.

Let me try to analyze these arbitrary assumptions, which permeated the foundations of our thinking in physics. The first and most important of these arbitrary assumptions concerned the concept of time, and I will try to explain in what this arbitrariness consists. To be able to do this well, I will start by discussing space, in order to draw a parallel between space and time. If we wish to describe the position of a point in space, i.e., the position of a point relative to a coordinate system k, we specify the point's orthogonal coordinates x, y, z. The meaning of these coordinates is as follows: According to familiar rules, we construct perpendiculars to the coordinate planes, and check how many times a given unit measuring rod can be laid along these perpendiculars. The coordinates are the results of this counting. Thus, the specifying of spatial position by means of coordinates is the result of specific manipulations. Accordingly, the coordinates I specify have a completely determinate physical meaning; one can verify whether a specific, given point really has the indicated coordinates or not.

Where do we stand with time in this respect? As we shall see, we are not so well off when it comes to time. Up to now, people always contented themselves with saying: Time is the independent variable of events. The measurement of the time value of an actually occurring event can never be based on such a definition. Hence we must try to define time in a way that will make it possible to measure time on the basis of this definition. Let us imagine a clock (a balance wheel clock, for example) at the origin of a coordinate system k. Using this clock we can evaluate the time of events occurring directly at this point or in its immediate vicinity. However, events occurring at another point of k cannot be evaluated directly with this clock. If an observer standing next to the clock at the origin of k notes the time at which he received notice of the event in question by means of a ray of light, this time will not be the time of the event itself, but a time greater than the latter by the velocity of propagation of the light ray from the event to the clock. If we knew the velocity of propagation of light relative to the system k in the direction under consideration, it would be possible to determine the time of the event using the above clock; but the velocity of light can be measured only if the problem of the determination of time, which we are now discussing, has been solved. To measure the velocity of light in a given direction, we would have to measure the distance between points A and B, between which the light ray propagates, and further, the time of the emission of the light at A and the time of the arrival of the light at B. Thus, time would have to be measured at different locations; however, this can be done only if the definition of time we are seeking has already been given. But if it is impossible in principle to measure a velocity, in particular the velocity of light, without recourse to arbitrary stipulations, then we are justified in making further arbitrary stipulations regarding the velocity of light. We shall now stipulate that the velocity of the propagation of light in vacuum from some point A to some point B is the same as that from B to A. By virtue of this stipulation we are indeed in a position to regulate identically constructed clocks that we have arranged at various points at rest relative to the system k. For example, we will set the clocks at the points A and B in such a manner that the following will obtain: If a ray of light sent from A toward B at time t (measured by the clock at A) arrives at B at time $t + a$ (measured by the clock at B), then, conversely, a ray sent from B toward A at time t (measured by the clock at B) must arrive at A at time $t + a$ (measured by the clock at A). This is the rule according to which all clocks arranged in the system k must be regulated. If we follow this rule, we achieve a determination of time from the standpoint of the measuring physicist. That is to say, the time of an event is equal to the readings of the clocks located at the place of the event that are regulated according to the rule we just described.

Since all this sounds self-evident, one may wonder what is particularly remarkable about the result we have obtained. What is remarkable is the fact that, in order to obtain time readings with a perfectly definite meaning, this rule refers to a system of clocks that is at rest relative to an exactly specified coordinate system k. We have not merely obtained a time, but a time that refers to the coordinate system k, or to the system k together with the clocks set up at rest relative to k. Of course, we can carry out exactly the same operations if we have another system k' that is moving uniformly relative to k. We can distribute throughout space a system of clocks relative to this coordinate system k′, but in such a way that all of them move together with k'. We can then regulate these clocks, which are at rest relative to k', exactly according to the rule described before. If we do this, we obtain a time with respect to the system k' as well.

But this does not say at all a priori that, when two events are simultaneous with respect to the reference system k—by that I mean the coordinate system together with the clocks—they are also simultaneous as understood with respect to the system k'. This does not say that time has an absolute meaning, i.e., a meaning independent of the state of motion of the reference system. This is an arbitrariness that was contained in our kinematics.

And now we come to a second factor that was also arbitrary in kinematics up until now. We speak about the shape of a body, the length of a rod, for example, and believe that we know exactly what the length of the rod is, even when it is in motion with respect to the reference system from which we are describing the events. A brief reflection shows, however, that these concepts are not at all as simple as we instinctively believe them to be. Consider a rod moving in the direction of its axis relative to the reference system k. We ask: What is the length of this rod? This question can have only the following meaning: What experiments do we have to perform in order to learn what the length of the rod is? We can take a man with a measuring rod and give him such a push that he assumes the same velocity as the rod; in that case he will be at rest relative to the rod, and will be able to determine its length by repeated application of his measuring rod, in the same way the lengths of bodies at rest are actually determined. He will obtain a perfectly definite number and will be able to declare with some degree of justification that he has measured the length of this rod.

However, if only such observers are available who do not move along with the rod, but instead all of them are at rest relative to a reference system k, we can proceed in the following manner: We imagine that very many clocks, with an observer assigned to each of them, are distributed along the route traveled by the axially moving rod. The clocks are regulated by means of light signals according to the procedure described before, such that in their totality they indicate the time associated with the reference system k. These observers determine the two positions with respect to the system k at which the beginning and the end of the rod are found at a given time t, or, what amounts to the same, those two clocks that the beginning and the end of the rod just pass by when the clock in question indicates the time t. The distance between the two positions (or clocks) so obtained is then measured by repeatedly applying a measuring rod, which is at rest relative to the reference system k, along the connecting line. The results of the two procedures can justly be designated as the length of the moving rod. However, it should be noted that these two manipulations do not necessarily lead to the same result, or, in other words, the geometrical dimensions of a body do not need to be independent of the state of motion of the reference system with respect to which the dimensions are determined.

If we do not make these two arbitrary assumptions, then we are at first no longer capable of solving the following problem: Given are the coordinates x, y, z and the time t of an event with respect to the system k; find the space-time coordinates x',

y', z', t' of the same event referred to another system k', which is in a known, uniform translational motion relative to k. It turns out that the customary simple solution of this problem is based on the two assumptions we have just identified as arbitrary.

How to put kinematics back on its feet? The answer is self-evident: the very same circumstances that led us into so many embarrassing difficulties in the past lead us to a negotiable path now that we have gained more room to maneuver by putting aside the arbitrary assumptions mentioned above. For it turns out that precisely those two seemingly incompatible axioms, which were imposed on us by experience, namely the principle of relativity and the principle of constancy of the velocity of light, lead us to a perfectly definite solution of the space-time transformation problem. One arrives at results that, in part, run very much counter to our customary conceptions. The mathematical considerations leading to these results are very simple; this is not the place to dwell on them.[2] It will be better if I deal with the most important consequences that were reached in this way by a quite logical procedure, without additional assumptions.

First, things purely kinematic. Since we defined the coordinates and the time in a definite way in physical terms, all relationships betwen spatial and temporal quantities will have a perfectly definite physical content. We obtain the following: If we have a solid body that is moving uniformly with respect to the coordinate system k, which we take as the basis for our analysis, then this body appears contracted by a perfectly definite ratio in the direction of its motion, as compared with the shape it has when it is in a state of rest with respect to this system. If we denote the velocity of motion of the body by v and the velocity of light by c, then each length measured in the direction of motion, and equal to l when the body is in a motionless state, will be diminished because of the body's motion relative to the noncomoving observer to the length

$$l \cdot \sqrt{1-\frac{v^2}{c^2}}\ .$$

If the body has a spherical shape in the state of rest, it will have the shape of a flattened ellipsoid if we move it in a certain direction. When its velocity reaches the

---

[2] If x, y, z, t, and x', y', z', t' are space and time coordinates with respect to the two reference systems k and k', then the two underlying principles demand that the transformation equations be such that each of the two equations

$$x^2+y^2+z^2 = c^2t^2$$
$$x'^2+y'^2+z'^2 = c'^2t'^2$$

have the other equation as its consequence. Since, for reasons into which I shall not enter here, the substitution equations must be linear, this determines the transformation law, as a brief analysis shows (cf., e.g., *Jahrbuch der Radioaktivität und Elektronik* 4 [1907]: 418ff).

velocity of light, it will collapse to a plane. However, as judged by a comoving observer, the body retains, before and after, its spherical shape; on the other hand, to the observer moving with the body, all noncomoving objects appear, in exactly the same way, contracted in the direction of the relative motion. This result loses very much of its oddness if one considers that this assertion about the shape of a moving body has quite a complicated meaning since, according to what has been said above, this shape can be ascertained only with the aid of determinations of time.

The feeling that this concept, "the shape of the moving body," has an immediately obvious meaning is due to the fact that in our day-to-day experience we are accustomed to encountering only such velocities of motion that are practically infinitely small compared with the velocity of light.

And now a second purely kinematic consequence of the theory that strikes us as even more peculiar. We imagine that there is given a clock capable of indicating the time of a reference system k, provided that it is arranged at rest relative to this system. It can be proved that this same clock, when set into uniform motion relative to the reference system k, runs slower, as judged from the system k, in such a way that when the time reading of the clock has increased by 1, the clocks of the system k indicate that, with respect to the system k, there has elapsed the time

$$\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Thus, the moving clock runs more slowly than the same clock when in state of rest with respect to k. One must imagine that one determines the rate of the clock in a state of motion by comparing, from time to time, the position of the hands of this clock with the positions of the hands of those clocks at rest relative to k that measure the time relative to k and that the moving clock under consideration is just then passing by. Were we to succeed in making the clock move with the velocity of light—we would be able to make it move with a velocity approximating the velocity of light if we had sufficient force—the hands of the clock, as judged from k, would move forward infinitely slowly.

The thing is at its funniest when one imagines that the following is being done: One imparts to this clock a very great velocity (almost equal to c), then lets it fly on in uniform motion, and after the clock has covered a long stretch, one imparts to it a momentum in the opposite direction, so that it returns to the point from which it has been launched. It then turns out that the positions of the clock's hands have hardly changed during the clock's entire trip, while an identically constituted clock that remained at rest at the launching point during the entire time changed the setting of its hands quite substantially. It should be added that whatever holds for this clock, which we introduced as a simple representation of all physical phenomena, holds also for closed physical systems of any other constitution. Were we, for example, to place a living organism in a box and make it perform the same to-and-fro motion as the clock discussed above, it would be possible to have this organism return to its original starting point after an arbitrarily long flight having undergone an arbitrarily small change, while identically constituted organisms that remained at rest at the point of origin have long since given way to new generations. The long time spent on the trip represented only an instant for the moving organism if the motion occurred with approximately the velocity of light! [7] This is an inevitable consequence of our fundamental principles, imposed on us by experience.

And now a few more words on the significance of the theory of relativity for physics. This theory demands that the mathematical expression of a law of nature valid for arbitrary velocities does not change its form if one introduces with the help of transformation equations new space-time coordinates into the formulas that express that law. This substantially narrows the manifoldness of possibilities. By means of a simple transformation it is possible to derive laws for bodies moving arbitrarily fast from the laws that are already known for bodies at rest or in slow motion. In this way one can derive, for example, the laws of motion for fast cathode rays. At the same time it turned out that Newton's equations do not hold for material points moving with arbitrarily great velocity, but have to be replaced by equations of motion of a somewhat more complicated structure. These laws of cathode ray deflection turned out to be in quite satisfactory agreement with experience. [8]

Of the physically important consequences of the theory of relativity we ought to mention the following. We saw earlier that, according to the theory of relativity, a moving clock runs more slowly than does the same clock in the state of rest. It will probably never be possible to verify this by experiments with a pocket watch, because the velocities we can impart to the latter are vanishingly small compared with the velocity of light. But nature provides us with objects that have quite the same character as clocks and can be moved exceedingly fast. These are atoms which emit spectral lines, and to which we can impart velocities of several thousand kilometers per second by means of electric fields (canal rays). In accordance with theory, it is to be expected that the oscillation frequencies of these atoms should appear to be affected by their motions in exactly the same way in which this is to be deduced for moving clocks. Even though the experiments in question face great difficulties, we do hope that in the next few decades we will obtain an important confirmation or refutation of the theory of relativity in this way. [9]

The theory further leads to the important result that the inertial mass of a body depends on its energy content, though to such a small extent that a direct proof seems absolutely hopeless. If the energy of a body increases by E, the inertial mass increases by $\frac{E}{c^2}$. This theorem overturns the principle of the conservation of mass, or, rather, fuses it with the principle of the conservation of energy into a single principle. However odd this result might seem, still, in a few special cases, one can unequivocally conclude from empirically known facts, and even without the theory of relativity, that the inertial mass increases with energy content. [10]

And now let me say just a few words about the highly interesting mathematical elaboration that the theory has undergone, thanks, mainly, to the sadly so prematurely deceased mathematician Minkowski. [11] The transformation equations of the theory of relativity are so constituted that they possess the expression

$$x^2 + y^2 + z^2 - c^2t^2$$

as an invariant. If we introduce the imaginary variable $ct \cdot \sqrt{-1} = \tau$ as the time variable instead of the time t, then this invariant will assume the form

$$x^2 + y^2 + z^2 + \tau^2.$$

Here the spatial coordinates and the temporal coordinate play the same role. The further pursuance of this formal equivalence of the space and time coordinates in the theory of relativity led to a very perspicuous representation of the theory, which makes its application substantially easier. Physical events are represented in a 4-dimensional space, and the spatio-temporal relations of what results appear as geometrical theorems in this 4-dimensional space. [12]

Doc. 18

"Discussion" Following Lecture Version of "The Theory of Relativity"

[IN: *Naturforschende Gesellschaft in Zürich. Sitzungsberichte* (1911): II-IX. Published in vol. 4 of *Vierteljahrschrift der Naturforschenden Gesellschaft in Zürich* 56 (1911). Minutes of the meeting of 16 January 1911.]

Discussion

After a few warm words of regret on account of the impending departure of the speaker, *Prof. Kleiner* presents his opinion on the relativity principle in the following way: [1]

[*Prof. Kleiner*:] As far as the principle of relativity is concerned, it is being called revolutionary. This is being done especially with regard to those postulates that are uniquely Einsteinian innovations in our physical picture. This concerns most of all the formulation of the concept of time. Until now we were accustomed to view time as something that always flows, under all circumstances, in the same direction, as something that exists independently of our thoughts. We have become accustomed to imagine that somewhere in the world there exists a clock that categorizes time. At least one thought it permissible to imagine the thing in such a way. But according to the relativity principle, time turns out to be dependent on velocities, on coordinates, on spatial magnitudes. This is what is supposed to constitute the revolutionary character of the new conception of time. If we examine the issue more closely, it turns out that we are dealing with improvements in precision that were greatly needed, because if we recall how we arrive at the determinations of time, we realize that everything is very simple as long as we are dealing with the determination of events in our immediate vicinity. We have our good old clocks and can fix the instant of time at which something happens. Things are quite different in regard to this certainty about time when it comes to the temporal determination of events that are distant from us. We know that the light from certain fixed stars reaches us only after many years, so that we can say that by virtue of this fact we can look into the past. We can also quite easily imagine that we are looking into the future, so that this stability in the conception of time has now already been undermined to some extent by the facts. Let us imagine a man who is accustomed to rely on his hearing devices for his orientation. Such would be the case with a blind man. Let us assume that he suddenly regains his sight and sees a man driving in nails with a hammer. He will then have the peculiar experience of seeing the fall of the hammer first, and only thereafter hearing the blow. But he has been conditioned to regard hearing as that which corresponds to the phenomenon, and according to his way of thinking, he now has in his eyes an organ with which he looks into the future. He sees an event before it has actually happened. I mention this because it shows precisely how the interpretation of the time concept also depends on the manner in which we explain to ourselves the perception of time. The difficulties start only when it comes to the fixing of temporal events at places removed from us. In view of this circumstance Einstein introduced and carried out the radical expedient of measuring and fixing times by making them measurable by means of light paths, because, in the end, he always comes to perceive the world around us by way of light signals. He makes times measurable by light paths and lays down the postulate, which has recently emerged from our experience, that equal distances must be traveled in equal times. This postulate makes it possible to compare clocks with each other, and this in turn makes it possible to resolve the question: How do clocks run if one of them is located in a system at rest, and the other in a moving system? Quite stringent arguments show that these clocks do not run synchronously. It turns out that the notion of time as something absolute in the old sense cannot be maintained, but that, instead, that which we designate as time depends on the states of motion.

Something similar obtains for the spatial coordinates by means of which we usually represent spatial relations. They prove to be dependent on the state of motion. This also seems to be of a revolutionary character insofar as we used to think of length as something absolute, i.e., something independent of velocity. Upon closer examination, the matter of this fixity and peculiar definiteness of the spatial coordinates is not all that simple.

I would say that the relativity principle brings us only a clarification and not something that is fundamentally new. Now Mr. Einstein has shown that, based on the assumption of the constancy of the velocity of light and the relativity principle, some simple relations exist between the coordinates of space and time for systems moving relative to each other. If we introduce into the mathematical expressions of laws that are valid with respect to a coordinate system k the space and time coordinates of another reference system k', which are connected with those of k by the simple equations peculiar to the theory of relativity, we arrive at laws of the same form. This is the property that, above all else, made the relativity principle creditable to the mathematicians. They recognized that this invariability for these systems involves something with which they are familiar, a special case of the invariance they occasionally observe in the structures of projective geometry. The observation that something well known in its mathematical formulation is already finding application in reality has helped gain credit for the relativity principle.

As for the physicist, when the admissibility of such a principle is being discussed, he is wont to keep just to arguments of a more physical character. The consequence of the relativity principle that motion results in a change of shape is to us of much greater importance. In other words, this consequence yields the result that rigid bodies in the usual sense of the word do not exist. A body moving in a certain direction gets flattened, it becomes an ellipsoid in the direction of motion. Thus, rigid bodies do not exist, because all bodies are in motion. This is something that runs counter to the naive conception, and it is what disturbs many physicists, namely, to have to assume that rigid bodies do not exist. But I think that this should not be interpreted as meaning that a body must deviate from rigidity in all possible directions because motions take place in all possible directions, but, rather, that the invariance holds only when considering a certain direction of motion. It will be the business of the mathematicians to formulate more exactly the conditions of rigidity in these systems.

As for the rest, it is difficult to decide whether all the consequences of the relativity principle agree with experience, because the deviation from what mechanics yields for systems at rest is always of the relative magnitude $\frac{v^2}{V^2}$. This is a quantity that always remains small. It shows up in the discussion of the electromagnetic masses of electrons of longitudinal and transverse mass. These have been calculated according to the ideas of the relativity principle, but it must be said that this matter has not yet really been decided. But it is to be expected, as my colleague Mr. Einstein indicated, that, in the course of time, there will be experimental findings that will decide the matter.

Perhaps only the following remains to be said in this area, something which proves to be a difficult matter for the physicist in discussions of this relativity principle. You have heard that we must give up the notion that an ether exists. Perhaps we should say that it is not too bad about the ether. We came to know it as an ad hoc hypothesis by means of which to explain all sorts of phenomena. One has saddled it with more and more incomprehensible properties. But the fact still remains that we are supposed to discuss propagations without having any idea in what these propagations consist. The velocity of propagation of light, wave motions that are propagated, the whole theory of interference, all these were hitherto based on certain conceptions that are now gone. We are supposed to speak about propagation in a medium which is not a medium, and about which we do not know a thing. I think that this is a gap that must be filled, because scientific, and especially physical, discussions that operate with formulas that cannot be linked to any mental image cannot be maintained forever.

So, as far as the principle designated as the principle of relativity is concerned, I think that this is something that has been needed, that, for once, simply stipulates some things, some unclarities to which we have not given any thought at all, and puts them in certain order. Time will probably show what sort of difficulties are involved, but whatever they are, they will probably find their solutions.

Prof. Einstein: First of all, I wish to thank Prof. Kleiner for his kind words. For the rest, I wish to say a few things in response to the things he brought up. According to the theory of relativity, a rigid body cannot exist at all. Let us imagine a rod of a certain [2] length. If we pull on one side, the other end will start to move at once. This would be a signal that moves with infinite speed and that could be used to define time, which leads to highly improbable consequences for reasons that cannot be explained here in greater detail. The ether was of real value for the intuitive representation of optical processes only as long as one actually reduced these processes, with all their peculiarities, to mechanical processes. After the concept of fields of lines of force had been put into the foreground, the ether hypothesis has come in fact to play only a fictitious role.

[3] *Fritz Müller*: If there are two synchronous clocks at point A, and one of them is moved with a given velocity from this point to point B, then, according to the reasoning of the speaker, this second clock will run slower, if only by a tiny fraction. What happens now if this clock returns by a polygonal or circular path to point A? According to the reasoning given in the lecture, at the moment of its meeting with the other clock at point A, the second clock will not be running in synchrony again. How can this be possible, since, on the other hand, Prof. Einstein says that a rod of a specific length L in a system at rest, which he holds in his hand, will become shorter by a definite amount when it is set in motion? But as soon as the rod is brought to halt by a sudden jerk, its length is once again $= L$, i.e., the rod is no longer deformed. If this latter argument holds good for length, i.e., for a specific dimension, and if what the mathematician Minkowski asserted, which was termed acceptable by Prof. Einstein, is correct, namely, that we can speak of a 4-dimensional geometry, so that we can compare length with time, then how do things stand with the clock? Must it not then, exactly like the rod, run synchronously again from the moment it is brought to rest at point A? This reasoning would suit me better, whereas I cannot grasp the other one.

Prof. Einstein: It is not the clock's indication of time that is to be likened to the rod, but its rate. After having completed its motion and returned, the rod has the same length. In the same way, the clock has again the same rate. We can designate the rod as the carrier of the space differential, and the clock as the carrier of the time differential. It is impossible to assume that, after having traveled along a polygonal path and returned to point A, the clock will again be running synchronously with the clock that has been at rest at point A. The clock runs slower if it is in uniform motion, but if it undergoes a change in direction as a result of a jolt, then the theory of relativity does not tell us what happens. The sudden change of direction might produce a sudden change in the position of the hands of the clock. However, the longer the clock is moving rectilinearly and uniformly with a given speed of forward motion, i.e., the larger the dimensions of the polygon, the smaller must be the effect of such a hypothetical sudden change.

[4] *Prof. Prašil*: In his famous essay, "Space and Time," Minkowski wrote about the nature of dilation, that the latter is a concomitant circumstance of the state of motion.
[5] He makes it absolutely independent of any physical influence. Lorentz, on the other

hand, when he was explaining Michelson's experiment, stated outright his conjecture that it may well be good to assume that such a change in length is brought about by the influence of the ether or of molecular forces. These are two things that I cannot [6] reconcile.

Prof. Einstein: Allow me to answer with a comparison. It has to do with the second law of thermodynamics, the law of the limited convertibility of thermal energy. If one takes the assumption of the impossibility of a perpetuum mobile of the second kind as the starting point of the argument, then our law appears as almost an immediate consequence of the basic premise of the theory. But if one bases the theory of heat on the equations of motion of molecules, then our law appears as the result of a long series of most subtle arguments. Just as here both of these routes have their undeniable justification, so the above-mentioned points of view of Minkowski on the one hand, and of H. A. Lorentz, on the other, also seem to me completely justified.

Prof. Meissner: As far as I know, Minkowski used the theory of relativity in order to [7] derive the general equations for moving bodies from the fundamental equations of the electrodynamics of bodies at rest. He set up a system of formulas that does not coincide either with the formulas of Cohn nor with those of H. A. Lorentz. Objections have been raised against this new system. Since I know the whole theory of relativity more from the [8] mathematical point of view, I would very much like to know the reasons that prompt the physicists to decide against Minkowski's equations and in favor of those of Cohn and Lorentz. It seems to me, from the mathematical point of view, that there must exist only one system of equations, namely that of Minkowksi.

Prof. Einstein: If one starts out from the theory of bodies at rest, then one can derive only the laws of electrodynamics for uniformly moving bodies by means of the relativity transformation. Do the equations of electrodynamics for uniformly moving bodies also hold for bodies in spatially and temporally non-uniform motion? This is possible but not certain. To that extent, Minkowski's equations are a hypothetical extension of those existing before.

Regarding the theories of Cohn and Lorentz, the following should be noted. The theory of Lorentz deviates from that of Minkowski insofar as a small inaccuracy crept into it, due to the much more difficult method of derivation. In fact, there are no fundamental differences between Minkowski's and Lorentz's theory. On the other hand, Cohn's electrodynamics must be viewed as fundamentally different. [9]

Fritz Müller: According to the explanations given in the lecture, when a clock is set up at the North Pole, and a synchronously running clock is located at the equator, then, if we consider the rotation of the earth, the clock at the North Pole is at rest, while the other one is moving with the rotational velocity of the earth. If it were possible to make the position of the hands of the clock at the North Pole visible at the equator, the latter would have to run slower. Perhaps this could be used as the basis for a practical experiment, because a measurable amount of time might perhaps emerge from it.

According to Prof. Einstein's arguments, a velocity greater than the velocity of light is inconceivable because it contradicts our experience, since the necessary consequence of this would be that we could then in the future perceive the consequences of events before the event itself has occurred. My question is, are not these equations based on the fact that one simply inserts the velocity of light V and builds everything else on that. Suppose there are people who have an additional sense organ that enables them to perceive velocities greater than the velocity of light; in that case it certainly would be conceivable that if these people were to set up these same equations, they would arrive again at the theory that there does not exist a greater velocity than that which they perceive with their senses. Perhaps Prof. Einstein can declare himself in agreement if we reduce the theorem he propounded to the following: A velocity greater than the velocity of light is out of the question for the organs available to humans.

Prof. Einstein: In answer to the first question, I would only like to note that time is a very poor multiplying factor. It is absolutely impossible to obtain a useful result within a practical time period, e.g., within a human lifespan, for the simple reason that a human's life consists of relatively few seconds.

[10] I did not say that a superluminal velocity is impossible; it is not impossible from the logical point of view, rather one can only say: if there existed a velocity that could really be conceived as the velocity of propagation of a physical stimulus, then it would be possible to construct an arrangement that would allow us to see at some place consequences of actions before we had innervated the thing by acts of volition. This seems to me to be something that has to be ruled out until proven otherwise, because it does not seem to be in accord with our experience. Physical propagation velocities have nothing to do with the character of our sensory organs.

[11] *Dr. Lämmel*: There is something that is even faster than light: gravitation. We would be faced with a great difficulty if we had to settle on the view that there can be no talk of velocity in the case of an attraction between two masses, that there can be only an instantaneous effect. Thus, gravitation too must possess a certain velocity. But it has not yet been possible to detect this velocity. It seems very probable that this velocity is much greater than that of light. If we were to subsitute gravitational signals for light signals, we would have a new world picture, on the basis of which we would be able to prophesy: There is no greater velocity than the velocity of gravitation.

A second question that interests me is the following: Is the world picture resulting from the conceptions of the relativity principle an inevitable one, or are the assumptions arbitrary and expedient but not necessary? If we are forced to give up the ether, then we must consider light as a substance that possesses the velocity of light. As for several remarks that we heard here, I would like to point out that the analogy between the coordinates of space and of time is only a mathematical one, obtained by way of definition. For a mathematician, things can emerge that defy physical representation. Thus, for example, $\sqrt{-1}$ occurs in this formula.

Prof. Einstein: If we had electrostatic forces instead of gravitation, what would be the result? Would you find a velocity of propagation? You would only find that things go infinitely fast because the question has been wrongly put. The thing has been calculated as if the particles had been hurled out from the from the center of gravitation. It is very possible, and it is even to be expected, that gravitation propagates with the velocity of light. If there existed a universal velocity which, like the velocity of light, were so constituted with respect to a single system that a stimulus would propagate with a universal velocity independent of the velocity of the emitting body, the theory of relativity would be impossible. If gravitation were to propagate with a (universal) superluminal velocity, this would suffice to bring down the principle of relativity once and for all. If it propagated infinitely fast, this would provide us with a means to determine the absolute time.

The comparison of light with other "stuff" is not permissible. At small velocities, material stuff in the usual sense of the word moves according to Newton's equations of motion. This is not the case with light; the parallel is therefore not permissible.

The principle of relativity is a principle that narrows the possibilities; it is not a model, just as the second law of thermodynamics is not a model.

Dr. Lämmel: The question is whether the principle is inevitable and necessary or merely expedient.

Dr. Einstein: The principle is logically not necessary: it would be necessary only if it would be made such by experience. But it is made only probable by experience.

Prof. Meissner: The discussion has shown what is the first thing to be done. All physical concepts will have to be revised, they will have to be reformulated, indeed, in such a way as to bring out any invariance with respect to the transformation of the relativity principle that may be present. Klein has in fact already pointed out in a lecture that one must extract from each concept that which can be maintained unchanged when one applies the remarkable transformation of space and time. Only then will one have [12] extracted epistemologically one of the main results. Even if the whole theory of relativity were to prove untenable, this would represent an extraordinary advance.

Prof. Einstein: The main thing now is to set up the most exact experiments possible in order to test the foundation. In the meantime, all this brooding is not going to take us far. Only those consequences can be of interest that lead to results that are, in principle, accessible to observation.

Prof. Meissner: You have brooded over this, and discovered the magnificent time concept. You found that it is not independent. This must be investigated for other concepts as well. You have shown that mass depends on the energy content, and you have made the concept of mass more precise. You did not carry out any physical investigations in the laboratory—you were brooding instead.

Prof. Einstein: The observations we made created an embarrassing predicament for us.

Prof. Meissner: Just think of non-Euclidean geometry. People thought they knew what an angle was, but they did not.

Dr. Lämmel: Regarding these speculations, the question is whether we are dealing with mathematical or physical considerations. Purely mathematical considerations cannot produce anything but premises, while physical considerations can open new paths. Hence I understand the statements made earlier by Prof. Einstein.

Doc. 19

Notes for a Lecture on Fluctuations

[10 February 1911]

[p. 1] [1]

Outline of a lecture in Leiden [2]

$S = \kappa \lg W$ Funct. determ by consid. two indep. systems. [3]

from osmot. pressure and gas pressure $\kappa = \frac{R}{N}$

Equilibrium of the particle.

G = /excess/weight $\Phi = Gz$

Gz to be supplied $W = -Gz$ suppl. entropy $= -\frac{Gz}{T}$

$W = e^{-\frac{\Phi}{\kappa T}}$ Probabil. that the particle is shifted upwards by at least z

$dW = -\frac{dW}{dz}dz = \frac{G}{\kappa T}e^{-\frac{\Phi}{\kappa T}}dz$ If many particles $NdW = dN$

Derivation of the law of Brownian motion. [4]

1) Gravitational flow $<d> n \cdot \frac{G}{6\pi\eta P}dt$

2) Brownian motion vertical path Δ in time τ

$$\frac{1}{2}n_{z-\Delta}\Delta - \frac{1}{2}n_{(z+\Delta)}\Delta$$

$$= -\frac{1}{2}\frac{\partial n}{\partial z}\Delta^2\cdot 2 = -\frac{\partial n}{\partial z}\Delta^2$$

In unit time $-\frac{\partial n}{\partial z}\frac{\Delta^2}{\tau}$

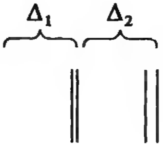
$$-n\frac{G}{6\pi\eta P} - \frac{\partial n}{\partial z}\frac{\Delta^2}{\tau} = 0$$

$$\frac{\Delta^2}{\tau} = -\frac{G}{6\pi\eta D\frac{d\lg n}{dz}} \quad\Bigg|\quad \frac{d\lg n}{dz} = -\frac{G}{\kappa T}$$

[5] $$\boxed{\frac{\Delta^2}{\tau} = \frac{\kappa T}{6\pi\eta P}}$$

$dw = \varphi(\Delta, t)\,d\Delta$

symmetrical in Δ



$$\int_{\Delta_1=-\infty}^{\Delta_1=+\infty}\varphi(\Delta_1, t_1)\,d\Delta_1\cdot\varphi(\Delta_2 t_2)\,d\Delta$$

$$\Delta_1 + \Delta_2 = \Delta$$

$$= \varphi(\Delta, t)\Delta\Delta$$

More detailed discussion of the law.

[6] Radiation.

$$\rho = \frac{8\pi h\nu^3}{c^3}\frac{1}{e^{h\nu/\kappa T}-1} \sim \frac{8\pi h\nu^3}{c^3}e^{-h\nu/\kappa T}$$

[p. 2] $$\cancel{S = \int V dv \int \frac{\frac{\partial\rho}{\partial T}dT}{T} = +V dv\cdot\frac{8\pi h\nu^3}{c^2}\frac{de^{h\nu/\kappa T}}{(\quad)^2}}$$

$$\cancel{= V dv\left(\frac{\rho}{T} - \int_0^T\frac{\rho}{T^2}dT\right)}$$

$$\frac{\partial\rho}{\partial\left(\frac{1}{T}\right)}\left(-\frac{1}{T^2}\right) = \frac{\partial\rho}{\partial T}$$

$$s = \frac{8\pi h\nu^3}{c^3} \int \overset{d(1/T)}{\frac{\rho}{dT}} \frac{dT}{T} = \left(-\frac{1}{N}\right)$$

$$s = \int \frac{\rho' \, dT}{T} = \frac{\rho}{T} - \int \frac{\rho \, dT}{T^2} = \frac{\rho}{T} + \int \rho d\left(\frac{1}{T}\right)$$

$$\int \rho d\left(\frac{1}{T}\right) = \int f \frac{\kappa}{h\nu} \frac{e^{-} d \frac{h\nu}{\kappa T}}{1 - e^{-}} = \int \frac{f\kappa}{h\nu} \frac{d(-e^{-h\nu/\kappa T})}{1 - e^{-}}$$

$$= \frac{f\kappa}{h\nu} \lg(1 - e^{-h\nu/\kappa T})$$

$$S = sV \, d\nu = \frac{\rho V \, d\nu}{T} + \frac{8\pi\kappa\nu^2}{c^3} V \, d\nu \lg(1 - e^{-h\nu/\kappa T})$$

$$- \frac{8\pi\kappa\nu^3 V \, d\nu}{c^3} \lg \frac{\rho c^3}{8\pi h\nu^3}$$

$$\left\langle \frac{\rho}{f} \right\rangle \qquad 1 + \frac{f}{\rho} = e^{h\nu/\kappa T} \qquad \frac{h\nu}{\kappa T} = \lg\left(1 + \frac{f}{\rho}\right) \qquad \left| \frac{d\rho}{T} = ds \right.$$

$$d\rho \frac{\kappa}{h\nu} \lg\left(1 + \frac{f}{\rho}\right) = ds$$

$$\frac{\kappa\rho}{h\nu} \lg(\qquad) - \int \rho \frac{\kappa}{h\nu} \frac{-\frac{f}{\rho^2}}{1 + \frac{f}{\rho}} d\rho = s$$

$$+ \frac{\kappa}{h\nu} \int \frac{\frac{f}{\rho}}{1 + \frac{f}{\rho}} d\rho \left| \frac{\kappa}{h\nu} \int \frac{f}{f + \rho} d\rho \right.$$

$$\left| \frac{\kappa f}{h\nu} \lg(f + \rho) \right.$$

[p. 3]

$$\rho = \frac{8\pi h\nu^3}{c^3}\frac{1}{e^{h\nu/\kappa T} - 1} \qquad \int\frac{d\rho}{T} = \frac{\kappa}{h\nu}\int d\rho \lg\left(1 + \frac{f}{\rho}\right) \tag{2}$$

$$\frac{f}{\rho} = e^{\cdots} - 1$$

$$\left|\rho\lg\left(1 + \frac{f}{\rho}\right)\right| - \int\frac{\rho}{1 + \frac{f}{\rho}}\cdot - \frac{f}{\rho^2}d\rho$$

$$\lg\left(1 + \frac{f}{\rho}\right) = \frac{h\nu}{\kappa T}$$

$$+ \int\frac{f}{\rho + f}d\rho$$

$$\int\frac{d\rho}{T} = \frac{\kappa}{h\nu}\left|f\lg\rho + (f + \rho)\lg\left(1 + \frac{f}{\rho}\right)\right|$$

$$+ f\lg(\rho + f)$$

$$f\lg\rho - (f + \rho)\lg\rho + (f + \rho)\lg(f + \rho)$$

$$f\lg\rho + f\lg\left(1 + \frac{f}{\rho}\right)$$

$$\frac{\kappa}{h\nu}\{(f + \rho)\lg(f + \rho) - \rho\lg\rho - f\lg f\}$$

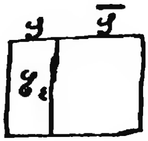
$$\rho \ll f$$

$$\int\frac{d\rho}{T} = \frac{\kappa}{h\nu}\rho\lg\frac{f}{\rho} \qquad S = \frac{\kappa}{h\nu}E\lg\frac{f}{\rho} = \frac{\kappa E}{h\nu}\lg\left(\frac{fV\,d\nu}{E}\right) =$$

$$S - S_0 = \frac{\kappa E}{h\nu}\lg\left(\frac{V}{V_0}\right) = \kappa\lg W$$

$$W = \left(\frac{V}{V_0}\right)^{E/h\nu}$$

Fluctuation of the Radiation Energy



$$S_{\text{tot}} = S(E + \varepsilon) - \bar{S}(\bar{E} - \varepsilon)$$

$$\frac{\partial S}{\partial E} = \frac{\partial\bar{S}}{\partial\bar{E}} \qquad S_{\text{tot}} = \frac{\partial^2 S}{\partial E^2}\frac{\varepsilon^2}{2} = \kappa\lg\frac{W}{W_0}$$

$$S = sV\,d\nu \qquad \frac{\partial^2 S}{\partial E^2} = \frac{1}{V\,d\nu}\frac{d^2 s}{d\rho^2}$$

$$E = \rho V\,d\nu$$

$$= -\frac{\kappa f}{h\nu V\,d\nu}\frac{1}{f + \rho}$$

$$\frac{ds}{d\rho} = \frac{\kappa}{h\nu}\{\lg(f + \rho) + \cancel{1} - \lg\rho - \cancel{1}$$

$$\frac{d^2 s}{d\rho^2} = \frac{\kappa}{h\nu}\left\{\frac{1}{f + \rho} - \frac{1}{\rho}\right\} = \frac{\kappa}{h\nu}\left(-\frac{f}{(f + \rho)\rho}\right)$$

$$W = W_0 e^{(\partial^2 S/\partial E^2)(\varepsilon^2/2\kappa)}\, d\varepsilon \qquad \frac{\int_{-\infty}^{-\infty} \varepsilon^2 e^{-\varepsilon^2/2a^2}\, d\varepsilon}{\int_{-\infty}^{\infty} e^{-\varepsilon^2/2a^2}\, d\varepsilon} = 2a^2 \frac{\int x^2 e^{-x^2}\, dx}{\int e^{-x^2}\, dx} = a^2 \qquad (3)$$ [p. 4]

$$\overline{\varepsilon^2} = \frac{\kappa}{-\dfrac{\partial^2 S}{\partial E^2}} = \frac{f+\rho}{f} \frac{h\nu V\, d\nu}{\langle\kappa\rangle}$$

$$\overline{\varepsilon^2} = \left(1 + \frac{\rho}{f}\right) h\nu\rho V\, d\nu = h\nu\rho V\, d\nu + \langle 8\pi h\nu^3\rangle \underbrace{\frac{c^3}{8\pi h\nu^3}}_{\nu^2} \overset{\rho^2}{\overbrace{h\nu}}\, V\, d\nu$$

$$\overline{\left(\frac{\varepsilon}{E}\right)^2} = \frac{h\nu}{E} + \langle E\rangle \underbrace{\frac{c^3}{8\pi\nu^2 V\, d\nu}}_{\text{wave theory}} \qquad \frac{l^3}{t^3} \cdot \frac{t^3}{V} \quad \text{correct}$$

Temperature fluctuation of a body in the radiation space. Brownian motion of a monochromatic mirror: [7]

Doc. 20

Statement on the Light Quantum Hypothesis

by A. Einstein

[IN: *Naturforschende Gesellschaft in Zürich. Sitzungsberichte* (1911): XVI. Published in vol. 4 of *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich* 56 (1911). Minutes of the meeting of 21 February 1911]

[1] It turns out that when Maxwell's theory of electricity and the molecular-kinetic approach are applied to some phenomena of light production and conversion, contradictions with observed facts, in particular those concerning "black-body radiation" and generation of cathode rays, come to light. These contradictions can be removed by introducing the working hypothesis that in the propagation of light, energy does not fill space in a continuous manner, but that, instead, it consists of a finite number of energy quanta localized at spatial points, which move without dividing and can be absorbed and [2] generated only as a whole. If these energy quanta impinge on a photoluminescent substance, then, according to the principle of the conservation of energy, the radiant energy emitted in an elementary process must be equal to or smaller than the incident radiant energy, and from the formula for the energy of a light quantum one arrives in a simple manner at Stokes' well-known frequency rule. When cathode rays are generated by illuminating solid bodies, the energy of the light quanta is converted into the kinetic energy of electrons, and only now do we realize that the quality of cathode radiation, i.e., the velocity of electrons, can be independent of the intensity of the exciting light, while the number of ejected electrons is proportional to the number of the light quanta. But from Planck's radiation formula it is to be inferred that, hand in hand with this, there has to come about a change regarding our conception of the molecular-kinetic mechanism for the transfer of energy to ions or electrons capable of oscillation (resonators), in that their energy can change only jumpwise by an integral multiple of precisely one quantum of light energy. If one also carries this mechanism over to the oscillations of the material molecules of a solid body that are due to the body's thermal motion, one arrives at a surprising elucidation of the change of specific (molecular) heats of solids with temperature, which until now had remained a total mystery.

(End of the session 10:30)

Doc. 21

Elementary Observations on Thermal Molecular Motion in Solids

by A. Einstein

[*Annalen der Physik* 35 (1911): 679–694]

I have shown in a previous paper[1] that a connection must exist between the law of radiation and the law of specific heats of solids (deviation from the Dulong-Petit law).[2] The investigations by Nernst and his students have now shown that, on the whole, specific heats indeed display the behavior deduced from the law of radiation, but that the true law of specific heats deviates systematically from the law established by theory. One of the first goals of this paper is to show that these deviations are due to the fact that the [2] oscillations of molecules are far from being *monochromatic* oscillations. The heat capacity of an atom of a solid is similar to that of a strongly damped oscillator in a radiation field and not like that of an oscillator that is only slightly damped. For that [3] reason specific heat decreases less rapidly toward zero with decreasing temperatures than the earlier theory would have it; the body behaves similarly to a mixture of resonators whose proper frequencies are distributed over a certain region. Further, it will be shown that Lindemann's formula, as well as my formula for the calculation of the proper frequencies ν of atoms, can be derived by dimensional arguments, with the latter also [4] yielding the order of magnitude of the numerical coefficients appearing in these formulas. Finally, it will be shown that the laws of heat conduction in crystallized insulators are not in accord with molecular mechanics, but that it is possible to derive the order of magnitude of the actually observable thermal conductivity by means of a dimensional argument, and thereby simultaneously to find out how the thermal conductivity of monatomic substances is probably related to their atomic weight, atomic volume, and proper frequency.

§ 1. On the Damping of Thermal Oscillations of Atoms

I showed in a recently published paper[3] that one arrives at approximately correct values for the proper frequencies of the thermal oscillations of atoms if one starts out from the following assumptions:

---

[1] A. Einstein, *Ann. d. Phys.* 22 (1907): 184. [1]

[2] Thermal motion in solids was conceived there as consisting in monochromatic oscillations of atoms. Cf. §2 of this paper. [5]

[3] A. Einstein, *Ann. d. Phys.* 34 (1911): 170.

1. The forces binding the atoms to their positions of rest are essentially identical with the elastic forces of mechanics.

2. The elastic forces operate only between immediately neighboring atoms.

To be sure, the theory is not completely determined by these two assumptions, for the elementary laws of interaction between immediately neighboring atoms can still be chosen freely to some extent. Also, it is not a priori clear how many molecules are to be viewed as "immediately neighboring." However, the specific choice of a pertinent hypothesis changes little in the results, so I will again stick with the simple assumptions I introduced in the above-mentioned paper. I will also use the same notation as there.

In the paper cited I imagined that each atom has 26 neighboring atoms with which it interacts elastically, and that all these atoms may be viewed as mathematically equivalent with respect to their elastic effect on the atom under consideration. The proper frequency was calculated in the following way. One thinks of the 26 neighboring atoms as being at rest, while only the atom under consideration oscillates; the latter then performs an undamped pendular oscillation, whose frequency one calculates (from the cubic compressibility). Actually, however, the 26 neighboring molecules are not at rest, but oscillate about their equilibrium position in a similar way as to the atom under consideration. Through their elastic connections with the atom considered, they influence the oscillations of the latter, so that its oscillation amplitudes in the coordinate directions are changing all the time, or—what comes to the same thing—the oscillation deviates from a monochromatic oscillation. Our first task is to estimate the magnitude of this deviation.

Let M be the molecule considered, whose oscillations in the x-direction we are investigating; let x be the momentary distance of the molecule from its rest position; if M_1' is a neighbor molecule of M that is in its rest position, but is at the moment at the distance $d + \xi_1$ from the rest position of M, then M_1' exerts on M a force of the magnitude $a(\xi_1 - x \cos \varphi_1)$ in the direction MM_1'. The X-component of this force is

$$a(\xi_1 - x \cos \varphi_1) \cos \varphi_1.$$

If m is the mass of M, then one obtains for M the equation of motion

$$m \frac{d^2x}{dt^2} = -x \cdot \sum a \cos^2 \varphi_1 + \sum a \, \xi_1 \cos \varphi_1,$$

where one has to sum over all of the 26 neighboring atoms.

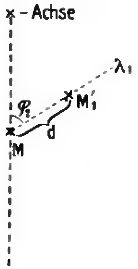


Fig. 1.

Now we calculate the energy transferred to the atom from the neighboring atoms during half an oscillation. We calculate as if the oscillation of the molecule considered as well as the neighboring molecules, proceeds sinusoidally during the half-oscillation period, i.e., we set

$$x = A \sin 2\pi\nu t,$$
$$\xi_1 = A_1' \sin (2\pi\nu t + \alpha_1).$$

. . . .

Multiplying the above equation by $(dx/dt)dt$ and integrating over the time indicated, we obtain the expression for the change of the energy,

$$\int d \left\{ m\frac{\dot{x}^2}{2} + \sum (a \cos^2 \varphi) \cdot \frac{x^2}{2} \right\} = \sum a \cos \varphi_1 \int \xi_1 \frac{dx}{dt} dt.$$

If we denote by Δ the total energy increase of the atom, and by η_1, η_2, etc., the amounts of energy transferred to the atom from the individual neighboring atoms during a half-oscillation period, we can write this equation in the form

$$\Delta = \Sigma\eta_n,$$

where we set

$$\eta_n = a \cos \varphi_n \int \xi_n \frac{dx}{dt} dt.$$

With the above conventions for x, $\xi_1 \dots$, we obtain

$$\eta_n = \frac{\pi}{2} a \cos \varphi_n \sin \alpha_n A A_n'.$$

From this it follows that the individual quantities η_n are as likely to be positive as negative, considering that the angles α_n take on each value with equal frequency and, indeed, independently of each other. For that reason we also have $\overline{\Delta} = 0$. Now we form the mean value $\overline{\Delta^2}$ as a measure of the energy change. Due to the indicated statistical property of η_1 etc., we have

$$\overline{\Delta^2} = \sum \overline{\eta_n^2}.$$

Since, as can easily be seen,

$$\overline{\sin^2 \alpha_n A^2 A_n'^2} = \frac{1}{2} \overline{A^2}^2,$$

we have

$$\overline{\eta_n^2} = \left(\frac{\pi}{2} a\right)^2 \cdot \frac{1}{2} \overline{A^2}^2 \cdot \cos^2 \varphi_n$$

and

$$\overline{\Delta^2} = \frac{\pi^2}{8} a^2 \overline{A^2}^2 \sum \cos^2 \varphi_n .$$

For an approximate calculation of this sum, we assume that two of the 26 atoms M' lie on the x-axis, 16 form an angle of close to 45° (or 135°) with the x-axis, and the remaining eight lie in the y-z plane. We then obtain $\Sigma \cos^2 \varphi_n = 10$, so that we get

$$\sqrt{\overline{\Delta^2}} = \sqrt{\frac{10}{8}}\, \pi a \overline{A^2}.$$

Now we compare this mean value for the atom's energy increase with the mean energy of the atom. The instantaneous value of the atom's potential energy is

$$a\frac{x^2}{2}\sum \cos^2 \varphi = a\frac{x^2}{2} \cdot 10.$$

The mean value of the potential energy is thus

$$5a\overline{x^2} = \frac{5}{2} a\overline{A^2}.$$

The mean value of the total energy E is then

$$\overline{E} = 5a\overline{A^2}.$$

The comparison of $\overline{E}$ with $\sqrt{\overline{\Delta^2}}$ shows *that the energy change during a half-oscillation period is of the same order of magnitude as the energy itself.*

Thus, in fact, not even for the time of a half-oscillation are the formulas for x, ξ_1, etc., from which we started out, approximately correct. However, this does not affect our result that the oscillation energy changes significantly during a half-oscillation.

§ 2. The Specific Heats of Simple Solids and the Theory of Radiation

Before asking ourselves what consequences the result just obtained has for the theory of specific heats, we must recall the train of thought that leads from the theory of radiation to the theory of specific heats. Planck has shown that, in a radiation field of density $\mathfrak{u}$ ($\mathfrak{u}d\nu$ = radiation energy of the frequency region $d\nu$ per unit volume), an oscillator slightly damped by emission assumes the mean energy

$$\overline{E} = \frac{c^3 \mathfrak{u}_0}{8 \pi \nu_0^2}$$

when c denotes the velocity of light in vacuum, ν_0 the proper frequency of the oscillator, and $\mathfrak{u}_0$ the radiation density for the frequency ν_0. [6]

Let the oscillator considered be an ion bound to an equilibrium position by quasi-elastic forces. Suppose the radiation space also contains gas molecules, which are in statistical (thermal) equilibrium with the radiation, and which may experience collisions with the ion constituting our oscillator. On the average, no energy may be transferred to the oscillator through these collisions; for, otherwise, the oscillator would disturb the thermodynamic equilibrium between the gas and the radiation. Hence one must conclude that the mean energy that the gas molecules alone would impart to our oscillator is exactly equal to the mean energy imparted to the oscillator by the radiation alone, which is to say that it is equal to $\bar{E}$. Further, since it is, in principle, irrelevant for the [7] molecular collisions whether the structure in question carries an electric charge or not, the above relation holds for every structure that oscillates approximately monochromatically. Its mean energy is related to the mean density $\mathfrak{u}$ of the radiation of the same frequency at the temperature considered. Hence, if one conceives of the atoms of solids as nearly monochromatically oscillating structures, then one obtains directly from the radiation formula the formula for specific heat, whose value should be $N(d\bar{E}/dT)$ per gram-molecule.

We see that this argument, the result of which, as we know, does not agree with the results of statistical mechanics, is independent of the quantum theory, as well as of any particular theory of radiation whatsoever. It is based solely on

1. the empirically established law of radiation,
2. Planck's analysis of resonators, which is based, in turn, on Maxwell's electrodynamics and mechanics,
3. the assumption that atomic oscillations are sinusoidal to a great degree of accuracy.

Regarding (2), it should be expressly noted that the oscillation equation for the oscillator employed by Planck cannot be derived rigorously without mechanics. For when solving problems of motion, electrodynamics makes use of the assumption that the sum of the electrodynamic and other forces acting on the framework of an electron is always zero, or—if one ascribes a ponderable mass to the structure—that the sum of the electrodynamic and other forces equals the mass times the acceleration. Thus, one has a priori a good reason to doubt the correctness of the result of Planck's analysis, seeing that the application of the fundamental postulates of our mechanics to rapid periodical processes leads to results that are in conflict with experience,[4] and that, therefore, the application of these fundamental postulates must raise doubts here too. Nevertheless,

---

[4] That is to say that our mechanics is not able to explain the small specific heats of solids at very low temperatures.

I believe that Planck's relation between u_0 and E should be retained, if for no other reason than because it has led to an approximately correct description of specific heats at low temperatures.

On the other hand, we have shown in the last section that assumption (3) cannot be supported. Atomic oscillations are not even approximately harmonic. The frequency region of an atom is so great that the change of the oscillation energy during a half-oscillation period is of the same order of magnitude as the oscillation energy. Thus, we must ascribe to each atom not a specific frequency, but rather a frequency range $\Delta\nu$ that is of the same order of magnitude as the frequency itself. To derive rigorously a formula for the specific heats of solids, one would have to carry out, for an atom of a solid, an analysis that is based on a mechanical model and is completely analogous to the analysis carried out by Planck for the infinitesimally damped oscillator. One would have to calculate the mean oscillation energy at which an atom, when provided with an electric charge, emits as much energy in a thermal radiation field as it absorbs.

[8] While I was laboring rather fruitlessly on this project, Nernst sent me the proofs of a paper[5] that contains a surprisingly useful tentative solution of the problem. He finds that the expression

$$\frac{3}{2}R\left(\frac{\left(\frac{\beta\nu}{T}\right)^2 e^{\frac{\beta\nu}{T}}}{\left(e^{\frac{\beta\nu}{T}}-1\right)^2}+\frac{\left(\frac{\beta\nu}{2T}\right)^2 e^{\frac{\beta\nu}{2T}}}{\left(e^{\frac{\beta\nu}{2T}}-1\right)^2}\right)$$

is an excellent representation of the temperature dependence of atomic heat. The fact
[10] that this expression shows a better agreement with experience than the one I chose
originally is easy to explain in the light of what has been said above. After all, one
obtains this expression by assuming that half of the time the atom performs quasi-
undamped sinusoidal oscillations with the frequency ν, and the other half of the time
with the frequency $\nu/2$. This is the manner in which the considerable deviation of the
[11] structure from monochromatic behavior finds its most primitive expression.

It is certainly not justified to consider ν as the proper frequency of the structure; instead, a value between ν and $\nu/2$ is to be taken as the mean proper frequency. Further, it should be noted that an exact coincidence between thermal and optical proper frequencies is out of the question, even if the proper frequencies of the different atoms of the compound in question closely coincide, for while the atom oscillates with respect to all neighboring atoms in thermal oscillations, it does so only with respect to the
[12] neighboring atoms with an opposite sign in optical oscillations.

---

[9] [5] W. Nernst and F. A. Lindemann, *Sitzungsber. d. preuss. Akad. d. Wiss.* 22 (1911): 65–90.

§ 3. Dimensional Argument Concerning Lindemann's Formula and My Formula for the Determination of the Proper Frequency

As we all know, dimensional arguments allow us chiefly to find general functional relations between physical quantities if all physical quantities occurring in the relation in question are known. For example, if we know that the oscillation period θ of a mathematical pendulum can only depend on the length of the pendulum l, on the acceleration of free fall g, on the mass of the pendulum m, and on no other quantity, then a simple dimensional argument will lead us to the conclusion that the relation must be given by the equation [13]

$$\theta = C \cdot \sqrt{\frac{l}{g}},$$

where C is a dimensionless number. But as we know, there is still something more that can be inferred from the dimensional argument, even though not in a completely rigorous way. Namely, dimensional numerical factors (as the factor C in this instance), the magnitude of which can only be deduced by means of a more or less detailed mathematical theory, are generally on the order of magnitude one. To be sure, this cannot be strictly required, because why should it not be possible for a numerical factor $(12\,\pi)^3$ to appear in a mathematical-physical analysis? But such cases are unquestionably rare. Suppose, therefore, that we had measured the oscillation period θ and the pendulum length l of an individual mathematical pendulum and that the above formula had yielded us 10^{10} as the value of the constant C; in that case we would already look upon our formula with justified suspicion. Conversely, our trust would grow if we found from our experimental data that C is, say, 6.3; our basic assumption that the relation sought contains only the quantities θ, l, and g, but no other quantities, would gain in probability in our eyes.

Let us now seek to determine the proper frequency ν of an atom of a solid by means of a dimensional argument. The simplest possibility would evidently be that the oscillation mechanism is determined by the following quantities:

1. The mass m of an atom (dimension m);
2. The distance d between two neighboring atoms (dimension l);
3. The forces with which the neighboring atoms oppose a change in their distance from each other. These forces also manifest themselves in elastic deformations; their magnitude is measured by the compressibility coefficient κ (dimension lt^2/m).

The only expression for ν that consists of these three quantities and has the right dimension is

$$\nu = C\sqrt{\frac{d}{m\kappa}},$$

where C is again a dimensionless numerical factor. Substituting the molecular volume v for d ($d = \sqrt[3]{v/N}$) and the so-called atomic weight M for m ($M = Nm$), one obtains

[14] $$\nu = CN^{\frac{1}{3}}v^{\frac{1}{6}}M^{-\frac{1}{2}}\kappa^{-\frac{1}{2}} = C \cdot 1.9 \cdot 10^7 M^{-\frac{1}{3}}\rho^{-\frac{1}{6}}\kappa^{-\frac{1}{2}},$$

where ρ denotes the density.

The formula I found by means of a molecular-kinetic argument,

[15] $$\lambda = 1.08 \cdot 10^3 M^{\frac{1}{3}}\rho^{\frac{1}{6}}\kappa^{\frac{1}{2}},$$

or

$$\nu = 2.8 \cdot 10^7 M^{-\frac{1}{3}}\rho^{-\frac{1}{6}}\kappa^{-\frac{1}{2}},$$

agrees with this formula with a factor C whose order of magnitude is one. The numerical factor obtained from my earlier argument is in satisfactory agreement with experiment.[6] Thus, the value for copper is

[16] $$\nu = 5.7 \cdot 10^{12},$$

when calculated from the compressibility by means of my formula, and

[17] $$\nu = 6.6 \cdot 10^{12}$$

when calculated from specific heat using the formula of Nernst, discussed in §2. However, this value of ν is not to be conceived as the "true proper frequency." We only know about the latter that it lies between Nernst's ν and half of this value. In the absence of an exact theory, the most logical thing to do is to view $\frac{\nu + \nu/2}{2}$ as the "true proper frequency," from which value one obtains, for copper, according to Nernst

$$\nu = 5.0 \cdot 10^{12},$$

which is in close agreement with the value calculated from the compressibility.

Let us now turn to Lindemann's formula.[7] We assume again that, above all else, the mass of an atom and the distance d between two adjacent atoms influence the proper frequency. Besides that, we assume the existence of a law of corresponding states for the solid state, whose degree of accuracy is sufficiently good for our present purposes. The

---

[6] Regarding the degree of approximation with which this formula holds, cf. the last paragraph of this section.

[18] [7] F. Lindemann, *Physik. Zeitschr.* 11 (1910): 609.

behavior of the substance, and thus also the proper frequency, will then completely be determined with the addition of a further characteristic quantity of the substance that is not determined by the two quantities mentioned above. As this third quantity we choose the melting point T_s. Of course, the latter cannot immediately be applied in the dimensional argument, because it cannot be measured directly in the C.G.S. system. Instead of T_s, we therefore choose the energy quantity $\tau = RT_s/N$ as the measure of temperature. τ is one-third of the energy that an atom possesses at the melting point according to the kinetic theory of heat (R = gas constant, N = number of atoms in a gram-atom). The dimensional argument yields immediately

$$\nu = C \cdot \sqrt{\frac{\tau}{md^2}} = C \cdot R^{\frac{1}{2}} N^{\frac{1}{3}} \sqrt{\frac{T_s}{Mv^{2/3}}} = C \cdot 0.77 \cdot 10^{12} \sqrt{\frac{T_s}{Mv^{2/3}}} \, .$$

The Lindemann formula reads

$$\nu = 2.12 \cdot 10^{12} \sqrt{\frac{T_s}{Mv^{2/3}}} \, .$$ [19]

Thus, the dimensionless constant C is here also of the order of magnitude one.

The investigations by Nernst and his students[8] show that, even though it is based on a very daring assumption, this formula yields a surprisingly good agreement with the ν values determined from the specific heat. From this it seems to follow that the law of corresponding states holds in remarkably good approximation for simple bodies in the solid and liquid states. It even appears that Lindemann's formula holds much better than my formula, which rests on a less daring assumption. This is all the more remarkable [21] because my formula also, of course, can be deduced from the law of corresponding states. If both formulas, mine as well as Lindemann's, are correct, then it follows from the division of the two formulas that $M/\rho T_s \kappa$ must be independent of the nature of the substance; in fact, this relation can also be deduced directly from the law of corresponding states. However, if one uses Grüneisen's values[9] for the compressibility of metals, one obtains values for this quantity that fluctuate roughly between $6 \cdot 10^{-15}$ and $15 \cdot 10^{-15}$! In view of the fact that the law of corresponding states holds up so well in the case of Lindemann's formula, this is quite peculiar. Might it not be possible that systematic [23] errors still lie hidden in all determinations of the cubic compressibility of metals? Compression under equal pressure from all sides has not yet been applied for the purpose of measurement, probably because of the considerable experimental difficulties

---

[8] Cf. especially W. Nernst, *Sitzungsber. d. preuss. Akad. d. Wiss.* 13 (1911): 311. [20]

[9] E. Grüneisen, *Ann. d. Phys.* 25 (1900): 848. [22]

involved. It is possible that such measurements, with deformation without angular deformation, would lead to values of κ significantly different from those obtained by the [24] measurements conducted thus far. At least this seems highly probable from the theoretical point of view.

§ 4. Remarks about the Thermal Conductivity of Insulators

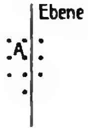


Fig. 2.

The result obtained in §1 makes it seem justified to attempt an approximate calculation of the thermal conductivity of nonmetallically conducting solids. If ε is the mean kinetic energy of an atom, then, according to §1, the atom releases on the average the quantity of energy $\alpha \cdot \varepsilon$ to the surrounding atoms during half of an oscillation period, α being a coefficient of the order of magnitude one, but smaller than one. If we imagine that the atoms are arranged in a lattice, and consider an atom A, which lies right next to an imaginary plane that does not intersect any molecule, then the atom A will, on the average, send the energy

$$\alpha \cdot \varepsilon \frac{9}{26}$$

across the plane during the time of half an oscillation, and thus the energy

$$\alpha\, \varepsilon \cdot \frac{9}{26} \cdot 2\nu$$

in unit time. If d is the smallest distance between neighboring atoms, then there are $(1/d)^2$ atoms per unit surface area abutting one side of the plane, and together they send the energy

$$\alpha \cdot \frac{9}{13} \nu \cdot \frac{1}{d^2} \varepsilon$$

per unit surface area in one direction (the direction of increasing x) across the unit surface area of the plane. Since the molecules on the other side of the layer are sending the quantity of energy

$$-\alpha \frac{9}{13} \nu \frac{1}{d^2} \left(\varepsilon + \frac{d\varepsilon}{dx} \cdot d \right)$$

per unit time across the unit surface area in the negative x-direction, the total flow of energy will be

$$-\alpha \cdot \frac{9}{13} \nu \cdot \frac{1}{d} \frac{d\varepsilon}{dx}.$$

If we make use of $d = (v/N)^{1/3}$ and denote by W the heat content of one gram-atom at the temperature T, we obtain the expression

$$-\alpha \frac{9}{13} \nu v^{-1/3} N^{-2/3} \frac{dW}{dT} \frac{dT}{dx},$$

and, hence, for the coefficient of thermal conductivity k

$$k = \alpha \cdot \frac{9}{13} \nu v^{-1/3} N^{-2/3} \frac{dW}{dT}.$$

If W is measured in calories, one obtains k in the customary units (cal/cm · sec · deg). If the substance obeys the Dulong-Petit law in the temperature range considered, then, because

$$\frac{dW}{dT} = \frac{3R}{\text{heat equivalent}} = \frac{3 \cdot 8.3 \cdot 10^7}{4.2 \cdot 10^7} \sim 6,$$

we can, perhaps, set

$$k = \alpha \cdot 4N^{-2/3} \nu v^{-1/3}.$$

We first apply this formula to KCl, which, according to Nernst, behaves with regard to its specific heat like a substance composed only of identical atoms. Taking for ν the [25] value $3.5 \cdot 10^{12}$, obtained by Nernst from the specific heat curve, we get [26]

$$k = \alpha \cdot 4 \cdot \left(6.3 \cdot 10^{23}\right)^{-2/3} \cdot 3.5 \cdot 10^{12} \cdot \left(\frac{74.4}{2 \cdot 2}\right)^{-1/3} = \alpha \cdot 0.0007,$$ [27]

whereas experiment at ordinary temperature[10] yields about

$$k = 0.016.$$

Thus, the thermal conductivity is much greater than was to be expected from our argument. But this is not all. According to our formula,[11] within the validity range of the Dulong-Petit law k should be independent of the temperature. According to Eucken's results, however, the actual behavior of crystalline nonconductors is entirely different; κ varies approximately as $1/T$. From this we must conclude that mechanics is [29] not capable of explaining the thermal conductivity of nonconductors.[12] It should be added that the assumption of a quantized distribution of energy also does not contribute anything to the explanation of Eucken's results. [30]

---

[10] Cf. A. Eucken, *Ann. d. Phys.* 34 (1911): 217. [28]

[11] Or according to a quite obvious argument by analogy.

[12] It must be noted that this also makes the arguments in §§1 and 2 questionable.

Eucken's important result, that the thermal conductivity of crystalline insulators is nearly proportional to $1/T$, can be used as a basis for a very interesting dimensional argument. We define the "thermal conductivity in natural units" k_{nat} by the equation

$$\text{Heat flow per unit surface area per second} = -k_{\mathrm{nat}} \frac{d\tau}{dx},$$

where heat flow should be thought of as expressed in absolute units, and $\tau = RT/N$. k_{nat} is a quantity to be measured in the C.G.S. system and its dimension is $[l^{-1}t^{-1}]$. In the case of a monatomic solid insulator, this quantity can depend on the following quantities:

d (distance between adjacent atoms; dimension l),
m (mass of an atom; dimension m),
ν (frequency of the atom; dimension t^{-1}),
τ (measure of the temperature; dimension $m^1 l^2 t^{-2}$).

If we assume that k_{nat} does not depend on any additional quantities, then the dimensional argument shows that k_{nat} can be expressed by an equation of the form

$$k_{\mathrm{nat}} = C \cdot d^{-1} \nu^1 \varphi\left(\frac{m^1 d^2 \nu^2}{\tau^1}\right),$$

where C denotes again a constant of the order of magnitude one, and φ an a priori arbitrary function, which, however, according to the mechanistic model, would have to be a constant if quasi-elastic forces between atoms are assumed. But according to Eucken's results, we have to set φ approximately proportional to its argument in order for k_{nat} to be inversely proportional to the measure of absolute temperature τ. We thus obtain

$$k_{\mathrm{nat}} = C\, m^1\, d^1\, \nu^3\, \tau^{-1},$$

where C denotes another constant of the order of magnitude one. If, instead of k_{nat}, we introduce k again, while using calories to measure the heat flow, and degrees Celsius to measure the drop in the temperature, and if we replace m, d, τ by their expressions in M, v, T, we obtain

$$k = \frac{1}{4.2 \cdot 10^7} \cdot \frac{R}{N} \cdot C \cdot \frac{M}{N} \cdot \left(\frac{v}{N}\right)^{1/3} \cdot \nu^3 \cdot \frac{N}{RT} = C \frac{N^{-4/3}}{4.2 \cdot 10^7} \frac{m v^{1/3} \nu^3}{T}.$$

This equation expresses the relation between the thermal conductivity, the atomic weight, the atomic volume, and the proper frequency. This formula yields for KCl

[31] $$k_{273} = C \cdot 0.007.$$

[32] Experiment yields $k_{273} = 0.0166$, so that C is really of the order of magnitude one. We must view this as confirming the assumptions that underlie our dimensional argument.

Experiments will have to decide whether C is independent to some extent of the nature of the substance; it will be the task of theory to modify molecular mechanics in such a way that it will yield the law of specific heats as well as the seemingly so simple law of thermal conductivity.

Prague, May 1911. (Received on 4 May 1911)

NOTE ADDED IN PROOF.

To make clearer the last sections of §2, let me add the following. If $\varphi(\nu/\nu_0)$ denotes a function that is to be conceived as the temporal frequency of the instantaneous frequency ν, and $\phi(\nu_0/T)$ the specific heat of the monochromatic structure of frequency ν_0, then the specific heat of the nonmonochromatic structure can be expressed by the formula

$$\tau = \int_{x=0}^{x=\infty} \phi\left(\frac{\nu_0 x}{T}\right) \varphi(x)\,dx.$$

One arrives at Nernst's formula if the function $\varphi(x)$ is given values different from zero for the arguments 1 and 1/2 only.

Doc. 22

[1] On the Ehrenfest Paradox. Comment on V. Varičak's Paper

by A. Einstein

[*Physikalische Zeitschrift* 12 (1911): 509–510]

Recently V. Varičak published in this journal some comments[1] that should not go unanswered because they may cause confusion.

The author unjustifiably perceived a difference between Lorentz's conception and [3] mine with *regard to the physical facts*. The question of whether the Lorentz contraction does or does not exist *in reality* is misleading. It does not exist "in reality" inasmuch as it does not exist for a moving observer; but it does exist "in reality," i.e., in such a way that, in principle, it could be detected by physical means, for a noncomoving observer. This is just what Ehrenfest made clear in such an elegant way.

We obtain the shape of a body moving relative to the system K with respect to K by finding the points of K with which the material points of the moving body coincide *at a specific time* t of K. Since the concept of simultaneity with respect to K that is being used in this determination is completely defined, i.e., is defined in such a way that, on the basis of this definition, the simultaneity can, in principle, be established by experiment, the Lorentz contraction as well is observable in principle.

Perhaps Mr. Varičak might admit—and thus in a way retract his assertion—that the Lorentz contraction is a "subjective phenomenon." But perhaps he might cling to the view that the Lorentz contraction has its roots solely in the arbitrary stipulations about [4] the "manner of our clock regulation and length measurement." The following thought experiment shows to what extent this view cannot be maintained.

Consider two equally long rods (when compared at rest) $A'B'$ and $A''B''$, which can slide along the X-axis of a nonaccelerated coordinate system in the same direction as and parallel to the X-axis. Let $A'B'$ and $A''B''$ glide past each other with an arbitrarily large, constant velocity, with $A'B'$ moving in the positive, and $A''B''$ in the negative direction of the X-axis. Let the endpoints A' and A'' meet at a point A^* on the X-axis, while the endpoints B' and B'' meet at a point B^*. According to the theory of relativity, the distance A^*B^* will then be smaller than the length of either of the two rods $A'B'$ and $A''B''$, which fact can be established with the aid of one of the rods, by laying it along the stretch A^*B^* while it is in the state of rest.

Prague, May 1911. (Received on 18 May 1911)

---

[2] [1] This jour. 12 (1911): 169.

Doc. 23

On the Influence of Gravitation on the Propagation of Light

by A. Einstein

[Annalen der Physik 35 (1911): 898–908]

In a paper published three years ago,[1] I already tried to answer the question as to whether the propagation of light is influenced by gravitation. I now return to this topic because my former treatment of the subject does not satisfy me, but, even more importantly, because I have now come to realize that one of the most important consequences of that analysis is accessible to experimental test. In particular, it turns out that, according to the theory I am going to set forth, rays of light passing near the sun experience a deflection by its gravitational field, so that a fixed star appearing near the sun displays an apparent increase of its angular distance from the latter, which amounts to almost one second of arc.

In the course of carrying through the analysis, further results regarding gravitation were obtained. However, since the presentation of the argument in its entirety would be [2] rather difficult to follow, I shall present in what follows only a few quite elementary considerations on the basis of which one can easily orient himself regarding the assumptions and the line of reasoning of the theory. Even if their theoretical basis is correct, the relationships here derived are valid only in first approximation.

§ 1. A Hypothesis Concerning the Physical Nature of the Gravitational Field [3]

In a homogeneous gravitational field (acceleration due to gravity, γ) let there be a coordinate system at rest K, which is oriented in such a way that the lines of force of the gravitational field run in the direction of the negative z-axis. In a space free of gravitational fields, let there be another coordinate system K' that moves with a uniform acceleration (acceleration γ) in the direction of its positive z-axis. So as not to complicate the analysis unnecessarily, we will disregard the theory of relativity for the time being, and consider, instead, the two systems according to conventional kinematics, and the motions occurring in them according to customary mechanics.

Material points not subjected to actions of other material points move relative to K as well as relative to K' according to the equations

---

[1] A. Einstein, *Jahrb. f. Radioakt. u. Elektronik* IV.4. [1]

$$\frac{d^2x_v}{dt^2} = 0, \quad \frac{d^2y_v}{dt^2} = 0, \quad \frac{d^2z_v}{dt^2} = -\gamma.$$

For the accelerated system K', this follows directly from Galileo's principle, but for the system K at rest in a homogeneous gravitational field, this follows from the experience that all bodies undergo the same, constant, acceleration in such a field. This experience of the identical falling of all bodies in the gravitational field is one of the most universal experiences that the observation of nature has yielded to us; nevertheless, this law has not been granted a place in the foundations of our physical edifice.

But we arrive at a very satisfactory interpretation of the empirical law if we assume that the systems K and K' are, physically, perfectly equivalent, i.e., if we assume that the system K could likewise be conceived as occurring in a space free of a gravitational field; but in that case, we must consider K as uniformly accelerated. Given this conception, one can no more speak of the *absolute acceleration* of the reference system than one can speak of a system's *absolute velocity* in the ordinary theory of relativity.[2] With this conception, the equal falling of all bodies in a gravitational field is self-evident.

As long as we confine ourselves to purely mechanical processes within the range of validity of Newton's mechanics, we can be sure of the equivalence of the systems K and K'. However, for our conception to acquire deeper significance, the systems K and K' must be equivalent with respect to all physical processes, i.e., the natural laws with respect to K must coincide completely with those with respect to K'. If we accept this assumption, we obtain a principle that possesses great heuristic significance, provided that it is really correct. For through a theoretical analysis of processes taking place relative to a uniformly accelerating reference system, we obtain information about the course of processes taking place in a homogeneous gravitational field.[3] In what follows, I shall first show that from the point of view of the ordinary theory of relativity our hypothesis has considerable probability.

§ 2. *On the Gravitation of Energy*

The theory of relativity has shown that the inertial mass of a body increases with its energy content; if the energy increase is E, then the increase in the inertial mass is E/c^2, where c denotes the velocity of light. But is there also an increase in gravitational mass corresponding to this increase in inertial mass? If not, then a body would fall with different accelerations in the same gravitational field, depending on its energy content. [4]

---

[2] Of course, one cannot replace an *arbitrary* gravitational field by a state of motion of the system without a gravitational field, just as one cannot transform to rest all the points of an arbitrarily moving medium by means of a relativistic transformation.

[3] It will be shown in a subsequent paper that the gravitational field considered here is homogeneous only to first approximation.

The very satisfying result of the relativity theory, according to which the principle of the conservation of mass merges into the principle of the conservation of energy, would not be possible to maintain, because the old formulation of the principle of the conservation of mass would indeed have to be abandoned for the *inertial* mass, but maintained for the gravitational mass.

This must be considered very unlikely. On the other hand, the ordinary theory of relativity does not provide us with any argument from which we could conclude that the weight of a body depends on its energy content. But we will show that the gravitation of energy is a necessary consequence of our hypothesis of the equivalence of the systems K and K'.

Consider two material systems S_1 and S_2 which are equipped with measuring instruments and situated on the z-axis of K at a distance h from each other,[4] in such a way that the gravitational potential in S_2 is greater by $\gamma \cdot h$ than that in S_1. Suppose that S_2 has sent off a certain amount of energy E toward S_1 in the form of radiation. Let the energies in S_1 and S_2 be measured with sets of apparatus that are completely identical when brought to the *same* place in the system z and there compared with each other. Nothing can be asserted a priori about the process of this energy transfer, because we do not know how the gravitational field influences the radiation and the measuring instruments in S_1 and S_2.

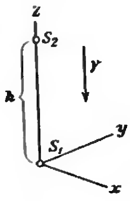


Fig. 1.

But in accordance with our assumption of equivalence of K and K', we can replace the system K, which is situated in a homogeneous gravitational field, by the gravitation-free system K', which moves with uniform acceleration in the direction of the positive z-axis, and to whose z-axis the material systems S_1 and S_2 are rigidly bound.

We will evaluate the process of energy transfer by radiation from S_2 to S_1 from a nonaccelerated system K_0. At the moment when the radiation energy E_2 has been emitted from S_2 toward S_1, the velocity of K' with respect to K_0 will be zero. The radiation will arrive at S_1 after a time h/c has elapsed (to a first approximation). But at that moment the velocity of S_1 with respect to K_0 will be $\gamma \cdot h/c = v$. Hence, according to the ordinary theory of relativity, the radiation arriving in S_1 will not possess the energy E_2, but the greater energy E_1, which is related to E_2, to a first approximation, by the equation[5]

$$E_1 = E_2\left(1 + \frac{v}{c}\right) = E_2\left(1 + \frac{\gamma h}{c^2}\right). \tag{1}$$

[5]

---

[4] S_1 and S_2 are considered infinitely small compared with h.

[5] A. Einstein, *Ann. d. Phys.* 17 (1905): 913, 914.

According to our assumption, exactly the same relation will hold in the case where this same process takes place in the system K that is not accelerated but is provided with a gravitational field. In that case we can replace γh by the potential Φ of the gravitation vector in S_2, if the arbitrary constant of Φ in S_1 is set equal to zero. We thus have

$$E_1 = E_2 + \frac{E_2}{c^2}\Phi. \tag{1a}$$

This equation expresses the energy principle for the process under consideration. The energy E_1 arriving in S_1 is greater than the energy E_2 (measured by the same kinds of instruments), which was emitted in S_2, by the potential energy of the mass E_2/c^2 in the gravitational field. Thus, for the energy principle to be satisfied, a potential energy of gravitation corresponding to the (gravitational) mass E/c^2 must be ascribed to the energy E before its emission at S_2. Our assumption of the equivalence of K and K' thus removes the difficulty mentioned at the beginning of this section, which the ordinary theory of relativity leaves unresolved.

The meaning of this result becomes especially clear upon consideration of the following cyclic process:

1. Energy E (measured at S_2) is sent in the form of radiation from S_2 to S_1, where, according to the result we have just obtained, the energy $E(1 + \gamma h/c^2)$ is absorbed (as measured at S_1).
2. A body W of mass M is lowered from S_2 to S_1, in which process an amount of work $M\gamma h$ is released.
3. The energy E is transferred from S_1 to the body W while W is in S_1. This changes the gravitational mass M such that its new value will be M'.
4. W is lifted back to S_2, which requires the application of work $M'\gamma h$.
5. E is transferred from W back to S_2.

The only effect of this cyclic process is that S_1 has undergone an energy increase of $E(\gamma h/c^2)$ and that the quantity of energy

$$M'\gamma h - M\gamma h$$

has been conveyed to the system in the form of mechanical work. According to the energy principle, we must then have

$$E\frac{\gamma h}{c^2} = M'\gamma h - M\gamma h$$

or

$$M' - M = \frac{E}{c^2}. \tag{1b}$$

The increase in *gravitational* mass is thus equal to E/c^2, thus equaling the increase in *inertial* mass obtained from the theory of relativity.

This result follows even more directly from the equivalence of the systems K and K', according to which the *gravitational* mass with respect to K is perfectly equal to the *inertial* mass with respect to K'; hence, energy must possess a *gravitational* mass that is equal to its *inertial* mass. If a mass M_0 is suspended from a spring balance in the system K', the balance will indicate the apparent weight $M_0\gamma$ because of the inertia of M_0. If the energy quantity E is transferred to M_0, the spring balance will indicate $\left(M_o + \frac{E}{c^2}\right)\gamma$, in accordance with the principle of the inertia of energy. According to our basic assumption, exactly the same thing must happen if the experiment is repeated in the system K, i.e., in the gravitational field.

§ 3. *Time and the Velocity of Light in the Gravitational Field*

If the radiation emitted in S_2 toward S_1 in the uniformly accelerated system K' had the frequency ν_2 with respect to a clock located at S_2, then upon its arrival at S_1, its frequency with respect to an identically constituted clock located at S_1 will no longer be ν_2, but a larger frequency ν_1, such that, to a first approximation,

$$\nu_1 = \nu_2\left(1 + \frac{\gamma h}{c^2}\right). \tag{2}$$

For if we again introduce the nonaccelerated reference system K_0, relative to which K' has no velocity at the time the light is emitted, then the velocity of S_1 with respect to K_0 will be $\gamma(h/c)$ at the time the radiation arrives at S_1, and from this we immediately obtain the relation given above with the help of Doppler's principle.

According to our assumption of the equivalence of the systems K and K', this equation also holds for the coordinate system K, which is at rest and is endowed with a uniform gravitational field, if the radiation transfer described above takes place in it. Thus, it follows that a ray of light emitted at a given gravitational potential in S_2, and possessing at its emission the frequency ν_2—compared with a clock located at S_2—will possess a different frequency ν_1 at its arrival at S_1, if this frequency is measured by an identically constituted clock located at S_1. We substitute for γh the gravitational potential Φ of S_2, with respect to S_1 as the zero point, and assume that our relation, which was derived for the *homogeneous* gravitational field, holds for otherwise constituted fields as well; we have then

[6] (2a)
$$\nu_1 = \nu_2\left(1 + \frac{\Phi}{c^2}\right).$$

This result (valid to a first approximation according to our derivation) allows, to begin with, the following application: Let ν_0 be the frequency of an elementary light source, measured by a clock U that is read at the same location. This frequency is then independent of the location at which both the light source and the clock are set up. We shall imagine that both are set up on the surface of the sun (this is where our system S_2 is located). A part of the light there emitted reaches earth (S_1), where we measure the frequency ν of the arriving light by means of a clock U, with exactly the same constitution as the clock mentioned above. According to (2a), we will then have

$$\nu = \nu_o\left(1 + \frac{\Phi}{c^2}\right),$$

where Φ is the (negative) gravitational potential difference between the solar surface and the earth. Thus, according to our conception, the spectral lines of solar light must be shifted somewhat toward red as compared with the corresponding spectral lines of terrestrial light sources, which has the relative shift amounting to

$$\frac{\nu_2 - \nu}{\nu_o} = \frac{-\Phi}{c^2} = 2 \cdot 10^{-6}.$$

If the conditions under which the solar lines are generated were known exactly, this shift would be accessible to measurement. However, since additional factors (pressure, temperature) influence the position of the center of density of the spectral lines, it is difficult to establish whether the influence of the gravitational potential that has been
[7] derived above really exists.[6]

At first glance, equations (2) and (2a) seem to assert something absurd. If the transmission of light from S_2 to S_1 is continuous, then how can the number of periods arriving per second at S_1 be different from that emitted at S_2? But the answer is simple. We cannot simply consider ν_2 and ν_1 as frequencies (numbers of periods per second) because we have not yet defined a time in the system K. ν_2 denotes the number of periods referred to the time unit of the clock U at S_2, and ν_1 the number of periods referred to the time unit of the identically constituted clock U at S_1. There

---

[8] [6] L. F. Jewell (*Journ. de phys.* 6 [1897]: 84) and especially Ch. Fabry and H. Boisson (*Compt. rend.*
[9] 148 [1909]: 688–690) did actually establish such shifts of fine spectral lines toward the red end of the spectrum of the order of magnitude calculated above, but they attributed them to an effect of the pressure in the absorbing layer.

is nothing that compels us to assume that the clocks U, which are situated in different gravitational potentials, must be conceived as going at the same rate. On the contrary, we must surely define the time in K in such a manner that the number of wave crests and troughs between S_2 and S_1 be independent of the absolute value of the time, because the process under consideration is stationary by its nature. If we did not satisfy this condition, we would arrive at a definition of time upon whose application time would enter explicitly the laws of nature, which would surely be unnatural and inexpedient. Thus, the clocks in S_1 and S_2 do not both give the "time" correctly. If we measure the time at S_1 with the clock U, *then we must measure the time at* S_2 *with a clock that runs* $1 + \Phi/c^2$ *times slower than the clock* U *when compared with the latter at one and the same location.* Because, when measured with such a clock, the frequency of the light ray considered above is, at its emission at S_2,

$$\nu_2\left(1 + \frac{\Phi}{c^2}\right),$$

and is thus, according to (2a), equal to the frequency ν_1 of the same ray of light on its arrival at S_1. [10]

From this follows a consequence of fundamental significance for this theory. Namely, if the velocity of light is measured at different places in the accelerated, gravitation-free system K' by means of identically constituted clocks U, the values obtained are the same everywhere. According to our basic assumption, the same holds also for K. But, according to what has just been said, we must use clocks of unlike constitution to measure time at points of different gravitational potential. To measure time at a point whose gravitational potential is Φ relative to the coordinate origin, we must employ a clock which, when moved to the coordinate origin, runs $(1 + \Phi/c^2)$ times slower than the clock with which time is measured at the coordinate origin. If c_0 denotes the velocity of light at the coordinate origin, then the velocity of light c at a point with a gravitational potential Φ will be given by the relation

$$c = c_0\left(1 + \frac{\Phi}{c^2}\right). \tag{3}$$

The principle of the constancy of the velocity of light does not hold in this theory in the formulation in which it is normally used as the basis of the ordinary theory of relativity.

§ 4. Bending of Light Rays in the Gravitational Field

From the proposition just proved, that the velocity of light in the gravitational field is a function of place, one can easily deduce, via Huygens' principle, that light rays propagated across a gravitational field must undergo deflection. For let ε be a plane of equal phase of a plane light wave at time t, and P_1 and P_2 two points in this plane a unit distance apart from each other. Let P_1 and P_2 lie in the plane of the paper, which is chosen in such a way that, when taken along the normal to the plane, the derivative of Φ, and thus also of c, vanishes. We obtain the corresponding plane of equal phase—or, rather, its intersection with the plane of the paper—at the time $t + dt$ by drawing circles with radii $c_1\,dt$ and $c_2\,dt$ around the points P_1 and P_2 and plotting the tangent to these circles, where c_1 and c_2 denote the velocities of light at P_1 and P_2, respectively. The angle of deflection of the light ray on the path cdt is then

$$\frac{(c_1 - c_2)dt}{1} = -\frac{\partial c}{\partial n'}dt,$$

if we take the angle of deflection as positive when the ray of light bends in the direction of increasing n'.

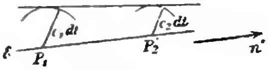


Fig. 2.

Thus, the angle of deflection per unit path length of the light ray will then be

$$-\frac{1}{c}\frac{\partial c}{\partial n'},$$

or, according to (3),

$$-\frac{1}{c^2}\frac{\partial \Phi}{\partial n'}.$$

Finally, we obtain for the deflection α which the light ray undergoes in the direction n' on any arbitrary path (s) the expression:

$$\alpha = -\frac{1}{c^2}\int \frac{\partial \Phi}{\partial n'}ds. \tag{4}$$

We could have obtained the same result by directly considering the propagation of a light ray in the uniformly accelerated system K' and transferring the result to the system K, and from there to the case of an arbitrarily constituted gravitational field.

According to equation (4), a ray of light traveling past a celestial body undergoes a deflection in the direction of decreasing gravitational potential, and thus, in the direction toward the celestial body, the magnitude of the deflection being

$$\alpha = \frac{1}{c^2} \int_{\vartheta=-\frac{\pi}{2}}^{\vartheta=+\frac{\pi}{2}} \frac{kM}{r^2} \cos\vartheta \cdot ds = 2\frac{kM}{c^2\Delta},$$

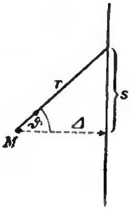


Fig. 3.

where k denotes the gravitation constant, M the mass of the celestial body, and Δ the distance of the ray of light from the center of the celestial body. *Accordingly, a ray of light traveling past the sun would undergo a deflection amounting to* $4 \cdot 10^{-6}$ *= 0.83 seconds of arc.* This [11] is the amount by which the angular distance of the star from the center of the sun seems to be increased owing to the bending of the ray. Since the fixed stars in the portions of the sky that are adjacent to the sun become visible during total solar eclipses, it is possible to compare this consequence of the theory with experience. In the case of the planet Jupiter, the displacement to be expected comes to about 1/100 of the amount indicated. It is greatly to be desired that astronomers [12] take up the question broached here, even if the considerations here presented may appear insufficiently substantiated or even adventurous. Because apart from any theory, we must ask ourselves whether an influence of gravitational fields on the propagation of light can be detected with currently available instruments.

Prague, June 1911. (Received on 21 June 1911)

Doc. 24

Excerpts of Discussions Following Lectures Delivered at 83rd Meeting of the Gesellschaft Deutscher Naturforscher und Ärzte, 25 and 27 September 1911

[*Physikalische Zeitschrift* 12 (1911): 978, 1068–1069, and 1084]

Stark: The speaker said that an electron resonates without damping in response to incident monochromatic radiation until the time a quantum of action, or, to use the Planck-Einstein expression, an energy quantum has accumulated in it. I would like to ask whether a calculation has been performed regarding the length of time the resonance needs for the accumulation of a light quantum. This question comes to mind because it was precisely Lorentz who performed detailed calculations on this point and who pointed out that, even in the case of perfect undamped resonance, the lengths of time one would have to assume in order to explain the maximum kinetic energy in the photoelectric effect are so great that they cannot be reconciled to within an order of magnitude with the actually observed situation regarding the intensity and maximum kinetic energy of the emitted ions.

Sommerfeld: Of course, the magnitudes that Lorentz found for the time length τ are here valid for monochromatic light within the same order of magnitude. It can be shown that our results are not confined to monochromatic light, and this must be shown because monochromatic light does not exist. If one uses natural light as the incident light, and the laws found for monochromatic light remain valid up to the dispersion mentioned, then the difficulty regarding the time length is eliminated.

Stark: But the light that was used in spectral resolution was practically monochromatic. Ladenburg worked with the lines of the mercury arc. Their frequency range is very narrow.

Have you performed the calculation for a specific frequency range? For example, for a spectral line with a width of 1 ångstrom.

Sommerfeld: Yes, for one spectral line.

Stark: The consistency of such a result with Lorentz's calculation would be amazing indeed. I assume that a detailed report of the calculation is to follow?

Einstein: Would the time required for a complete act of absorption according to the [1] theory presented not be long enough to be accessible to observation? In this case, the experimental investigation would be of enormous interest.

Sommerfeld: I believe that it would be possible to compare the theoretical [2] accumulation time with experiment.

Stark: Can the resonance theory presented be extended to the emission of cathode rays to the Roentgen-induced rays? This would be desirable because the phenomenon in question is completely analogous to the photoelectric effect.

Sommerfeld: I hope so.

Koenigsberger: In several metallic elements, the specific heat at constant volume exceeds the value 6 at high temperatures. The specific heat of metals cannot be represented by Einstein's original formula either, though the formula holds much better for an insulator such as diamond. Perhaps the free electrons in metals do have to be taken into account after all. It seems to me therefore that the behavior of specific heat supports the quantum theory more in a qualitative than in a quantitative way.

Rubens: But Nernst's experiments on specific heats at low temperatures can be completely represented by Einstein's formula if one does not stop at a single oscillation but assumes many proper frequencies instead, just as Einstein's formula also presupposes.

*****

[1]

Einstein: I would like to ask whether the speaker does not think it possible that the conductivity of pure metals becomes infinite as the temperature approaches the absolute zero. After all, Kamerlingh Onnes has found that even the slightest impurities have a very great effect, and that the purest metals have an extremely small resistance. Were [2] there, by any chance, different samples of aluminum available, and were they tested for this?

Nernst: I believe, too, that the high value for aluminum is due to impurities. After all, pure aluminum is currently very difficult to produce. Certainly, one could assume the value zero for the resistance of absolutely pure metals, but I believe that it will have a finite characteristic value for each metal.

Sommerfeld: How does the quantum theory envision the influence of small amounts of impurities?

*****

Einstein: We are indeed facing a certain difficulty, because we do not know how to understand luminescence radiation. The more a system deviates from the state of thermodynamic equilibrium, the more blurred become the differences between thermal radiation and luminescence radiation, because the concept of temperature loses its meaning. In the case of a mercury lamp, we cannot say what the temperature in the [1] lamp is. Surely, there will exist a certain temperature with respect to the molecules undergoing translatory motion, but not so with respect to the ions. In this sense the radiation of the mercury lamp is surely luminescence radiation, since at the temperature a termometer inside the lamp would indicate, the lamp would not emit radiation without current. But it does not seem out of the question that such a radiation might be emitted without a current at a higher temperature, i.e., that the radiation of the mercury lamp is a kind of radiation that is at certain temperatures essentially identical to thermal radiation. [2]

Rubens: But, from the absorption experiments with mercury vapor one can surely draw the conclusion that the long-wave radiation is most likely to originate from charged ions rather than from neutral molecules. This would support the view advanced by Mr. Lindemann. If a gas can be ionized through a mere increase in the temperature, which, incidentally, is not yet quite certain, then it would hardly be possible to draw a sharp line between luminescence and thermal radiation, at least in the high temperature range, on the basis of the old definitions.

Doc. 25

Discussion Remarks Following Lectures Delivered at First Solvay Congress

[30 October–3 November 1911]

III. LORENTZ

Lorentz's lecture (*Lorentz 1912*) discusses several ways of studying the applicability of the law of the equipartition of energy to heat radiation, one of which is related to the work by Einstein and Hopf (see *Einstein and Hopf 1910b* [Doc. 8]). Following this approach, which was outlined by Einstein in 1909 (see *Einstein 1909b* [Vol. 2, Doc. 56], p. 190), Lorentz assumes a frictionlike force acting on an electron moving in the radiation field, and he inserts the velocity change in the time τ due to this force into an expression for the fluctuations of the velocity of the electron (see *Lorentz 1912*, pp. 35–39). From his expression for these fluctuations, he attempts to determine the mean energy of the electron but obtains unsatisfactory results. Einstein's first comment refers to an objection raised by Planck against the separability of oscillatory and linear motion assumed by Lorentz (see *Lorentz et al. 1912*, pp. 46–47, and *Lorentz et al. 1914*, pp. 39–40), and his second comment refers to two alternative proposals to Lorentz's procedure, one suggested by Planck in his discussion remark, the other by Langevin (see *Lorentz et al. 1912*, pp. 42–44, and *Lorentz et al. 1914*, pp. 36–37). Contrary to Lorentz, Planck and Langevin in their respective comments describe the motion of the electron by means of ordinary differential equations instead of an expression for fluctuations.

No. 22 (*Lorentz et al. 1914*, p. 40; *Lorentz et al. 1912*, p. 47)

The smaller the radiation density, the more completely can the oscillatory motion of the electron that is caused by the momentary influence of the radiation be separated from its translational motion.

No. 25 (*Lorentz et al. 1914*, p. 40; *Lorentz et al. 1912*, pp. 47–48). The last sentence in the following text reads in the published version: "For this reason neither the consideration of Mr. Langevin nor that of Mr. Planck solves the problem, in my opinion."

The <consideration> differential equation neglects those terms by virtue of which the mean translational motion of the electron (independent of the momentary radiation field) can change. Mathematically, this manifests itself in the circumstance that an additive constant in v remains undetermined. For this reason the consideration does not solve the problem in my opinion.

IV. PLANCK

In his lecture Planck examined various ways of accounting for the spectral distribution of black-body radiation. In one approach, which corresponds to his earlier derivation of his formula for black-body radiation by methods of statistical physics, he determined the probability of a given macroscopic state by counting the combinatorial possibilities for realizing this state in terms of microscopic configurations (Boltzmann's "complexions"); see *Planck 1914*, pp. 86–87. In his first comment on Planck's lecture, Einstein summarized the critique of this approach, which he had earlier presented in *Einstein 1909b* (Vol. 2, Doc. 56), pp. 187–188. In line with his earlier analysis and in contrast to Einstein, Planck applied the quantum hypothesis as well as statistical methods only to matter that interacts with radiation and not directly to radiation itself. In the discussion this controversial point was first taken up by Jeans, and subsequently commented upon by Einstein in his second remark, referring to Lorentz's analysis of radiation (*Lorentz 1912*). In his lecture, Planck also presented his second attempt at a theory explaining the black-body radiation formula (for a historical discussion, see *Kuhn 1978*, pp. 235ff). According to Planck's "second theory," the quantum hypothesis plays a role only for the emission of radiation, while Maxwell's equations are supposed to be valid for absorption as well as for radiation in matter-free space. In his last remark during the discussion, Einstein argues that it is not possible to introduce any form of the quantum hypothesis for the emission by an oscillator, but he upholds classical electrodynamics in the space surrounding it. His reference to Planck's original theory is probably a reference to Planck's attempts at an analysis of black-body radiation prior to the introduction of the quantum hypothesis (see *Planck 1900a*).

No. 51 (*Planck et al. 1914*, p. 95; *Planck et al. 1912*, p. 115)

1) What I find strange about the way Mr. Planck applies Boltzmann's equation is that he introduces a state probability W without giving this quantity a physical definition. If one proceeds in such a way, then, to begin with, Boltzmann's equation does not have any physical meaning. The circumstance that W is equated to the number of complexions belonging to a state does not change anything here; for there is no indication of what is supposed to be meant by the statement that two complexions are equally probable. Even if it were possible to define the complexions in such a manner that the S obtained from Boltzmann's equation agrees with experience, it seems to me that with this conception of Boltzmann's principle it is not possible to draw <any> conclusions about the admissibility of any fundamental theory whatsoever on the basis of the empirically known thermodynamic properties of a system.

No.[53] (*Planck et al. 1914a*, p. 98; *Planck et al. 1912*, p. 119)

2) Objections have often been raised against the application of statistical methods to radiation. But I do not see any reason why these methods should be excluded here (cf. Lorentz's report, §6–§13).

3) <Omit!>

No. 100 (*Planck et al. 1914a*, p. 106; *Planck et al. 1912*, p. 129)

4) If an oscillator is to emit radiation in a manner different from that assumed in Mr. Planck's original theory, then this means a renunciation of the validity of Maxwell's equations in the vicinity of the oscillator. For according to Maxwell's equations, the quasi-static field of the oscillating dipole necessarily results in the release of energy in the form of spherical waves.

V. KNUDSEN

Knudsen had reviewed the available evidence in favor of the kinetic theory of gases, emphasizing the good agreement between theory and experiment in the limiting case that the mutual interaction between the molecules of a gas is small in comparison to the interaction between the gas and its container. In the first comment during the discussion of Knudsen's contribution, Nernst claimed that Maxwell's law of the distribution of molecular velocities might have to be changed because the quantum hypothesis implies a change of the law of molecular collisions. (For the implications of the quantum hypothesis for molecular collisions, see Einstein's lecture *Einstein 1914* [Doc. 26], p.352.) In his response to Nernst's comment, Einstein shows himself convinced of the validity of the Maxwell distribution and hence of the theorem of the equipartition of energy, at least for the linear motion of gas molecules, a conviction that also underlies his contemporary studies of radiation in interaction with a gas (see *Einstein and Hopf 1910b* [Doc. 8]). If the mean length of the path of a molecule is small, however, Einstein argues that the validity of the equipartition theorem is no longer assured. Einstein's first comment is followed by a remark by Warburg on the Krakatoa eruption of 1883, which showed that the motion of dust particles in the higher atmosphere deviates from Stokes's law. The discussion thus turned to the problem of small spheres suspended in a medium. This problem, touched upon in Knudsen's talk, was at that time particularly important because of its role in Millikan's oil drop experiments on the value of the elementary charge, and quickly became the focus of the discussion (see *Holton 1978* for a historical study of Millikan's experiments). Perrin and Brillouin suggested possible deformations of spherical droplets in a medium as the cause for the deviation of their motion from Stokes's law. In his second remark during this discussion Einstein refuted the conjecture that thermodynamic fluctuations could give rise to such deformations by arguing that the work to produce these deformations exceeded the energy transferred to the drops by collisions; see *Einstein 1907b* (Vol. 2, Doc. 39).

No. 114 (*Knudsen et al. 1914*, p. 121; *Knudsen et al. 1912*, p. 147)

5) Even though it is certain that our mechanics fails with regard to the oscillatory thermal motions of atoms and molecules, it can<not> hardly be doubted that Maxwell's distribution law is valid for the translational motion of gas molecules involving sufficiently large free paths. For Maxwell's law assumes only the momentum and energy conservation laws for individual collisions; these will certainly remain valid even if our mechanics does not hold during individual collisions. However, Maxwell's law presumably does not hold when, at a given temperature, the free path length becomes much too small. For in that case the molecule describes a zigzag line, that is, a kind of oscillatory motion, for which, according to our present knowledge, the law of the equipartition of energy does not hold.

No. 127 (*Knudsen et al. 1914*, p. 123; *Knudsen et al. 1912*, p. 150)

6) A noteworthy deformation of small droplets owing to irregular thermal motion must be ruled out because of the considerable capillary forces. Only such deviations from thermodynamic equilibrium take place whose average value is such that the mechanical work necessary, according to thermodynamics, for producing the deviations is equal to $\frac{RT}{2N}$, i.e., to a third of the average kinetic energy of a monatomic gas molecule.

VI. PERRIN

Perrin's lecture (*Perrin 1912*) is an exhaustive review of the experimental evidence in favor of the existence of atoms. Although much of it concerns the work of his own group on Brownian motion and related matters, Perrin also mentioned studies of critical opalescence by Smoluchowski, Keesom, and Einstein (see the editorial note, "Einstein on Critical Opalescence"), and experiments on the "atom of electricity." Einstein's first discussion remark refers to Keesom's derivation of a formula for light scattering by critical opalescence, first published in a footnote to *Kamerlingh Onnes and Keesom 1908b*, pp. 621–622. This remark does not appear in the published version of the discussion, probably because Einstein changed his opinion on the significance of Keesom's contribution (see Einstein to W. H. Julius, 18 December 1911). Einstein's second discussion remark concerns the experimental evidence for the existence of a natural unit of electric charge. In contrast to Millikan's results, which seemed to demonstrate conclusively the existence of such a unit, Ehrenhaft's experiments on ultramicroscopic silver particles suggested that there is no lower limit on electric charge. (See *Holton 1978* for a historical discussion.) How to explain Ehrenhaft's apparent "subelectronic" charges constituted a puzzle. In his remark, Einstein claims that the puzzle was solved by his colleague Edmund Weiss in Prague. Weiss found that contrary to Ehrenhaft's claims, Stokes's law does not apply to these small silver particles (see *Weiss 1911*, p. 631). Weiss evaluated the coefficient of mobility for each individual particle in his experiments, and found that this coefficient differed from particle to particle. His conclusion, repeated here by Einstein, was that Ehrenhaft's charge determinations were not valid. For Einstein's role in Weiss's experiments, see Einstein to Heinrich Zangger, 7 April 1911. Einstein had communicated Weiss's results to Perrin, who referred to them in his report (see *Perrin 1912*, p. 234). Millikan later discussed Weiss's experiments and emphasized that in the sequel of the experiments by Weiss and Przibram, the scientific world "ceased to concern itself with the idea of a sub-electron" (see *Millikan 1917*, p. 163, and also p. 153).

No. 137 (*Perrin et al 1914*, p. 251). The first remark does not appear in the German or in the French printed versions, presumably indicating Einstein's intervention.

9) It should be noted that Mr. Keesom was the first to derive the opalescence formula for homogeneous substances, something which he did in a very elegant way.

Further, I would like to remind you that Mr. Weiss from Prague was able to show why Ehrenhaft had been led to such small values of ε. He investigated silver particles in the air, and determined their motility from their Brownian motion and their charge from their velocity in the electrical field, which showed good agreement with the other determinations of ε. It turned out that no connection exists between the velocity of fall in the gravitational field and the motility <of different particles>, from which it follows that the particles must be of very irregular shape. Ehrenhaft's determinations of ε are thus illusory, because it will not do to draw conclusions about the mass of the particles from the velocity of fall.

VII. NERNST

Einstein's first comment on Nernst's lecture (*Nernst 1914*) refers to the difficulties of generalizing the quantum hypothesis to more than one dimension. Einstein confronted an objection raised by Lorentz against Nernst's decomposition of a classical three-dimensional oscillation into three circular components. Nernst not only used this decomposition to infer the equality of kinetic and potential energy for each of the circular components, but also to attempt to make the different roles of kinetic and potential energy in his understanding of the quantum hypothesis plausible. According to Lorentz, however, Nernst's decomposition of an elliptic oscillation into three mutually perpendicular circular oscillations is "artificial" and does not correspond to a decomposition of the energy into three additive components.

No. 149 *(Nernst et al. 1914,* pp. 235–236; *Nernst et al. 1912,* p. 239)

7) It has been pointed out several times that the application of the quantum hypothesis to structures with more than one degree of freedom meets with difficulties of a <conceptual> formal kind, irrespective of whether one views the quanta as energy quanta or as indivisible elementary regions of the q-p-manifold. If one modifies the equation for the mean energy $\overline{E}$ of a *three*-dimensional oscillator that is yielded by statistical mechanics,

$$\overline{E} = \frac{\int E^3 e^{-\frac{E}{\kappa T}} dE}{\int E^2 e^{-\frac{E}{\kappa T}} dE}$$

by introducing sums instead of integrals, which one does by giving to E in sequence the values $0, h\nu, 2h\nu$ etc., one does not thus arrive at three times of the energy of the linear Planck oscillator. Thus, the quantum theory in its current form leads to contradictions as soon as one seeks to apply it to structures with several degrees of freedom.

In his second comment, Einstein attempted to explain the temperature independence of what he interpreted as the damping of ionic oscillations within a crystal, referring to observations of residual rays reported in the preceding comment by Rubens. Einstein's comment is related to an extended controversy among himself, Rubens, and Nernst about the interpretation of the experiments on residual rays performed by Rubens and his group. Nernst in his lecture and Rubens in his comment argued that the results of these experiments are in conflict with Einstein's interpretation of the Nernst-Lindemann formula for specific heats as being the consequence of a strong damping of the elementary oscillators constituting the solid body (see *Einstein 1911g* [Doc. 21], p. 679). Rubens argued that the results of his measurements can be interpreted by assuming two proper frequencies of the solid body, an interpretation that Einstein did not accept.

No. 156 *(Nernst et al. 1914,* p. 238; *Nernst et al. 1912,* pp. 295–296)

8) The fact that the damping of optically discernible ionic oscillations is independent of the temperature had to be expected based on conventional mechanics. For if one assumes that, in the solid state, atoms are bound to each other by elastic forces, then according to mechanics the equations of motion become linear homogeneous differential equations, so that from one solution of the latter one obtains another one by merely multiplying the amplitudes by a constant, without otherwise having to change the time functions. From this it follows that the degree to which individual oscillating structures deviate from monochromatic behavior does not depend on the temperature. — It is odd

that this one inference from mechanics seems to be right, whereas thermal conduction seems not to be amenable to any mechanical interpretation.

Einstein's third comment is a response to a remark by Rutherford on Nernst's lecture. In it Rutherford inquired about the possibility of explaining the decreasing specific heats of solids for lower temperatures by assuming that a "polymerization" takes place within the solid. In a comment preceding Einstein's, Nernst excluded this possibility by arguing that chemical transformations are unlikely to take place at such temperatures. A proposal similar to Rutherford's had been made earlier by Lorentz and discussed by Einstein (see H. A. Lorentz to Einstein, 6 May 1909, and Einstein to H. A. Lorentz, 23 May 1909). In a footnote to the published text of his discussion remark, Einstein added further arguments against Rutherford's proposal: "The specific inductivity would have to approach unity if the temperature decreases to absolute zero. According to this hypothesis, the ultraviolet proper oscillations should not, for ordinary temperatures, exert an influence on the index of refraction or on the specific inductivity." After the Solvay Congress, the polymerization hypothesis was explored by a number of researchers (see, e.g. *Duclaux 1912b* and *Benedicks 1913*), but eventually rejected for reasons such as the ones mentioned by Einstein in his comment (see *Verhandlungen 1914*, p.371)

No. 171 (*Nernst et al. 1914*, p. 239; *Nernst et al. 1912*, pp. 296–297)

10) It is absolutely impossible to explain the decrease of specific heats at low temperatures by assuming rigid bonds between the atoms (reduction in the degrees of freedom). For according to this assumption, solid bodies would have to lose their elastic deformability as they approach absolute zero (the compressibility would have to vanish for $T = 0$), and the infrared proper frequencies would have to become less and less optically discernible, neither of which is true.

The first of the following two comments by Einstein follows a longer explanation by Kamerlingh Onnes, while his second comment responds to a suggestion made by Lindemann; both Kamerlingh Onnes and Lindemannn argued in favor of Rubens's interpretation of the Nernst-Lindemann formula (see the editorial note to Einstein's comment 8 on Nernst's lecture). Kamerlingh Onnes agreed with Rubens that the two frequencies appearing in this formula correspond to two different oscillations of the solid body. But whereas Rubens attempted to identify these oscillations as those of the neutral molecule and the electrically charged atoms, repectively, Kamerlingh Onnes argued that in a molecular system longitudinal and transverse oscillations exist which could have diffeent frequencies because of the way in which spatially extended atoms interact via parts of their surfaces. Lindemann, on the other hand, attempted to explain the existence of two different frequencies by assuming the interatomic forces to be directed so that, for example, oscillations along the diagonal and oscillations along one of the sides of a cubic lattice would have different frequencies. For a modern discussion of the role of the modes of oscillation of a lattice, see, e.g., *Born and Huang 1954*.

No. 177 (*Nernst et al. 1914*, p. 241; *Nernst et al. 1912*, p. 291)

11) The formula of Nernst and Lindemann undoubtedly represents a significant step forward. But we should beware, in my opinion, of seeing in it more than an empirical formula. It was clear a priori that atoms of solid bodies cannot behave thermodynamically exactly like infinitely weakly damped radiation resonators; in my opinion, the incompletely monochromatic character of atomic oscillations is the reason why experience deviates

from theory. A more careful investigation must show whether this conception will hold up.

The following discussion remark is transcribed from *Nernst et al. 1914*, p. 241. See also *Nernst et al. 1912*, p. 300. A manuscript version does not exist.

If the forces that cause the oscillations are proportional to the distance from the equilibrium position, then it follows from the symmetry of the cubic system that a material point cannot possess two frequencies, at least not as long as one adheres to the laws of mechanics.

In a comment following Einstein's previous remark, Poincaré brought the subject of the behavior of gases at low temperatures into the discussion. In the course of the ensuing exchange among Nernst, Poincaré, Rutherford, Kamerlingh Onnes, Einstein and Langevin, Nernst related this behavior to the rotational motion of the molecules and mentioned the difficulties of applying the "quantum theory" to this motion. In his Solvay lecture, Einstein criticized Nernst's theoretical treatment of the rotational motion of molecules, and made a remark similar to the comment printed below; see *Einstein 1914* (Doc. 26), pp. 350–351.

No. 181 (*Nernst et al. 1914*, p. 242; *Nernst et al. 1912*, p. 301)

12) The optical <and energetical> investigation of the optical properties of gases with a diatomic molecule with an electric moment is in fact of the greatest importance, because from the relation between the coefficient of emission and the frequency (or the temperature, if the frequency is given) one can obtain directly (using electrodynamics, to be sure) the statistical law of rotational motion.

In §6 of his lecture, Nernst claimed that his heat theorem (the third law of thermodynamics) can be derived from the quantum theory of specific heats. This claim gave rise to an extended controversy between Einstein and Nernst on the status of the heat theorem, starting with the discussion remark printed below. The conflict resurfaced during the second Solvay conference where it led to a lengthy discussion following *Grüneisen 1921* (see *Grüneisen et al. 1921*, pp. 290–301).

No. 186 (*Nernst et al 1914*, p. 243; *Nernst et al. 1912*, p. 302)

13) I would like to remark here that, as far as I can see, Nernst's heat theorem cannot be inferred from the vanishing of the specific heat in the vicinity of absolute zero, even though its validity is made much more plausible by this. For the question is whether, in a sufficiently close proximity to abs. zero, a system can be brought reversibly & isothermally from a state A to a state B *without the addition of heat*. This could not be inferred from the weakness of the molecular agitation if the transition from A to B could only be produced by using this minimal residue of thermal agitation; in that case it would be absolutely impossible to transfer the system from state A to state B at absolute zero. Nernst's theorem amounts to the assumption (that is quite plausible, to

be sure) that a transition from A to B in a way that is *purely* static from the viewpoint of molecular mechanics is always possible in principle.

In his lecture, Nernst only briefly mentioned the problem of heat conduction (on p. 231). See *Einstein 1911g* (Doc. 21), §4, and *Einstein 1914* (Doc. 26), p. 341, for a more detailed account of Einstein's contemporary thoughts on this problem.

Nos. [191]–192 (*Nernst et al. 1914*, p. 244; *Nernst et al. 1912*, p. 303)

14) The considerable thermal conductivity of insulators can neither be explained by the conventional mechanical theory nor with the auxiliary concept of energy quanta. According to both conceptions, during the time of half an oscillation, the oscillation energy bound to an atom should not propagate farther than up to the immediately neighboring atom, and consecutive energy transfers of this kind should be conceived as mutually independent processes. But on the basis of these assumptions one arrives at values for thermal conductivity that are far too small. It appears, accordingly, that at low temperatures thermal motion does not possess the character of complete disorder.

VIII. SOMMERFELD

In the introductory section of his lecture, Sommerfeld introduced his version of the quantum hypothesis, which he considered to be compatible with classical electrodynamics (see *Sommerfeld 1914*, p. 294) in the form of the principle that in "every purely molecular process," p. 254, the quantity of action

$$\int_0^\tau H dt = \frac{h}{2\pi}$$

is exchanged where τ is the duration of the process, H the Lagrangian, and h Planck's constant (for a historical study of Sommerfeld's work, see, e.g., *Hermann 1971*, pp. 103–123). In the context of his talk, Sommerfeld restricted the notion of a purely molecular process to the interaction between an electron and an atom (*Sommerfeld 1914*, p.254), but demonstrated the relativistic invariance of the action integral for the case of a single mass point, which is taken up by Einstein in his comment. The function $L - U$ mentioned by Einstein is the Lagrangian written in terms of the kinetic energy L and potential energy U.

No. 197 (*Sommerfeld et al. 1914*, p. 301; *Sommerfeld et al. 1912*, p. 373)

15) It seems to me that Sommerfeld's interpretation of the physical meaning of Planck's constant h involves a difficulty in that the function $L - U$ for a freely movable particle can hardly be set equal to zero, so that the existence of a freely movable mass point decomposes, so to speak, into quanta of action, which it does in a manner that depends on the state (of velocity) of the coordinate system.

The following comments by Einstein refer to Sommerfeld's analysis of X rays generated by the impact of electrons on an obstacle. In a letter to Besso, Einstein had earlier stressed his view that Sommerfeld had postulated his hypothesis on the role of collision times in this process without any theory (see Einstein to

Michael Besso, 21 October 1911). The essential points of Einstein's discussion remarks also appear in §4 of his Solvay lecture, *Einstein 1914* (Doc. 26). Einstein's second comment is a response to the following objection by Lorentz: "Mr. Einstein decomposes an arbitrary motion of a particle into a Fourier series, every term of which has a certain frequency ν. Did I understand correctly that, according to his understanding, there will be a radiation corresponding to some term, if the $h\nu$ characterizing this term is smaller than the total quantity of the available energy?" (*Sommerfeld et al. 1914*, p. 308; *Sommerfeld et al. 1912*, p. 382). Einstein's last comment is a response to Planck, who had suggested that the quantum hypothesis should apply only to monochromatic radiation and not to γ- and X-rays because the measured energy of these rays exceeds the energy obtained by dividing the quantum of action by the impulse time of the radiation.

No. 215 (*Sommerfeld et al. 1914*, pp. 307–308; *Sommerfeld et al. 1912*, pp. 381–382)

16) Sommerfeld's important result, which yields the energy emitted as X-ray energy during the collision of an electron with an obstacle, can also be derived in another way. I mention this in order that the satisfactory agreement between the theoretical formula and experience not be viewed as a direct confirmation of the underlying equation $\int(L - U)dt = \frac{h}{4\pi}$.

During a sudden collision, the electron emits energy in such a way that the quantity of energy

$$\frac{1}{3\pi}\frac{\varepsilon^2}{c^3}v^2 d\nu$$

is emitted from the frequency range $d\nu$. (ε = electrostatically measured charge, c = velocity of light, v = velocity of the electron.) <The loss of velocity in the collision is neglected.> It is assumed here that the electron is at rest after the collision. In order to obtain the total emitted energy, one would have to integrate this expression between $\nu = 0$ and $\nu = \infty$, which would lead to an infinitely large emission. But if one assumes that the electron cannot emit a ν that is greater than the one that corresponds to its kinetic energy L according to the quantum conception, then the upper limit of the frequency of the emitted radiation is given by the equation

$$L = h\nu,$$

so that the indicated integration yields for the emitted energy

$$\frac{1}{3\pi}\frac{\varepsilon^2}{hc^3}v^2 L,$$

essentially in agreement with Sommerfeld's result.

No. 224 (*Sommerfeld et al. 1914*, pp. 308–310; *Sommerfeld et al. 1912*, pp. 382–383)

17) The objection touches upon a sensitive point of the conception. According to the quantum theory in its original formulation, the way it is applied in the consideration just presented, one would have to imagine that always only *one* quantum of a specific frequency at a time is emitted in a collision, so that the result of our integration would be correct only as an average value over many collisions. But this conception is artificial; rather, the consideration reveals clearly a weak point of the conception characterized by monochromatic energy quanta.

19) According to Sommerfeld's conception, the frequencies $\nu > \frac{L}{h}$ are not emitted in a collision of an electron because the collision is not sudden. According to this conception, the higher terms of the Fourier expansion do not appear in the emitted field because they already do not appear in the Fourier expansion of the accelerations occurring in the collision. This conception has the great advantage that one can adhere to Maxwell's equations when calculating the emitted field. Unfortunately, however, this conception also brings with it a serious difficulty, which must not be left unmentioned.

If a radiation space contains a gas with electrically charged atoms, then these atoms emit and absorb radiation energy when they collide, and it would have to be possible to derive the radiation formula by a statistical investigation of such a system. The fact that one arrives here at Rayleigh's formula, if one starts off from classical mechanics and Maxwell's electrodynamics, can be taken as indisputably proved. To reach agreement with experience, one has to modify the theoretical foundations in such a way that, at a given temperature of the gas, the quotient $\frac{\text{emission coefficient}}{\text{absorption coefficient}}$ for the gas becomes exceedingly small for large ν. Thus, for large ν, the emission coefficient must become exceedingly small relative to the absorption coeff. Is this likely to be achieved by Sommerfeld's collision law?

Basically, the latter amounts to the assumption that the higher terms are missing in the Fourier expansion of the collision acceleration of the individual charged mass points. From this it follows directly that the corresponding terms are missing in the emission. But it seems that the absence of these terms in the Fourier expansion of the collision acceleration also excludes an *absorption* of those frequencies, so that in essence the above quotient should not be at all influenced by Sommerfeld's hypothesis.

No. 233 (*Sommerfeld et al. 1914*, p. 310; *Sommerfeld et al. 1912*, p. 384)

18) It may nevertheless be difficult to maintain the view that radiation of a given frequency can be emitted only in quanta of magnitude $h\nu$ by monochromatic oscillators, but in arbitrarily small portions by colliding electrons.

In §4 of his lecture, Sommerfeld presented an explanation of the photoelectric effect which he had developed in collaboration with Debye and earlier sketched in a lecture *(Sommerfeld 1911b)*, in the discussion of which Einstein had participated; see *Sommerfeld et al. 1911* (Doc. 24). For evidence of further exchanges between Sommerfeld and Einstein on this topic, see *Sommerfeld 1914*, p. 257. Since Sommerfeld's explanation is based on a resonance effect between the incident radiation and an atom, the photoelectric effect should, as he acknowledged in his lecture (see *Sommerfeld 1914*, p. 284; *Sommerfeld 1912*, p. 355), be more susceptible to material properties such as the damping of atomic oscillations than it should be according to the explanation given in *Einstein 1905* (Vol. 2, Doc. 14), §8. For a historical overview of alternative explanations of the photoelectric effect, see *Stuewer 1970* and *Wheaton 1978*.

No. 242 *(Sommerfeld et al. 1914*, p. 315; *Sommerfeld et al. 1912*, p. 390)

[2]0) According to Sommerfeld's theory of the photoelectric effect, the number of electrons emitted per unit of time can be proportional to the intensity of light only if one totally rules out a damping of the oscillatory motion.

IX. LANGEVIN

In his lecture, Langevin reviewed the kinetic theory of magnetism and in particular the work of Pierre Weiss. Following a suggestion made earlier by Gans, Langevin attempted, in the last part of his lecture, to use Sommerfeld's principle (see above) in a speculative construction of a molecular model of magnetization. He assumed that an electron circulates around a center that attracts it by a force characterized by a power law, and he obtained in this way a rough agreement with experiment. In the discussion following Weiss's presentation of his theory on an earlier occasion, Gans mentioned that Einstein had recently suggested to him an explanation of the units of magnetism by quantizing rotations (see *Weiss 1911*; this has been noted in *Kuhn 1978*, p. 312, fn. 40). In a letter written to Lorentz soon after the Solvay meeting, Einstein remarked: "The case of electrons in a magnetic field already mentioned in Brussels is interesting, but not as much as I thought in Brussels. . . . In any case, the thing seems to show that mechanics already ceases to hold in the case of the electron moving in a magnetic field" (Einstein to H. A. Lorentz, 23 November 1911).

This discussion remark is transcribed from *Langevin et al. 1914*, p. 328. See also *Langevin et al. 1912*, p. 405. A manuscript version does not exist.

It would be interesting to apply the above calculation to a single electron in the magnetic field.

Doc. 26

On the Present State of the Problem of Specific Heats

by A. Einstein

[Eucken, Arnold, ed., *Die Theorie der Strahlung und der Quanten. Verhandlungen auf einer von E. Solvay einberufenen Zusammenkunft (30. Oktober bis 3. November 1911), mit einem Anhange über die Entwicklung der Quantentheorie vom Herbst 1911 bis Sommer 1913.* Halle a. S.: Knapp, 1914. (Abhandlungen der Deutschen Bunsen Gesellschaft für angewandte physikalische Chemie, vol. 3, no. 7), pp. 330–352]

[1] *§1. The Connection between Specific Heats and the Radiation Formula*

It was in the domain of specific heats that the kinetic theory of heat achieved one of its earliest and finest successes in that it permitted the exact calculation of the specific heat of a monatomic gas from the equation of state. It is now, again, in the domain of specific heats that the inadequacy of molecular mechanics has come to light.

According to molecular mechanics, the mean kinetic energy of an atom not bound rigidly to other atoms is in general $\frac{3}{2}\frac{RT}{N}$, if one lets R denote the gas constant, T the absolute temperature, and N the number of molecules in a gram-molecule. From this it follows directly that the specific heat of an ideal monatomic gas at constant volume is $\frac{3}{2}R$, or 2.97 calories, per gram-molecule, which is in very good agreement with experience. If the atom does not move freely but is bound in an equilibrium position, then it possesses not only the mean kinetic energy mentioned above, but, in addition, also a potential energy; we must assume this to be the case for solid bodies. For the arrangement of atoms to be stable, the potential energy corresponding to the displacement of an atom from its equilibrium position must be positive. Further, since the mean distance from the equilibrium position must increase with the thermal agitation, i.e., with the temperature, this potential energy must always correspond to a *positive* component of the specific heat. Thus, according to our molecular mechanics, the atomic heat of a solid body must always be greater than 2.97. As we know, in the case where the forces binding the atom to its equilibrium position are proportional to the displacement, the theory yields the value of $2 \cdot 2.97 = 5.94$ for the atomic heat. It has been known for a long time that for most of the solid elements the atomic heats possess values that do not deviate substantially from 6 at ordinary temperatures (Dulong-Petit law). But it also has been known for a long time that there are elements with smaller atomic heats. Thus, already in 1875, H.F. Weber found that the value of the atomic heat of diamond at
[2] -50°C is roughly the value 0.76, far smaller than that permitted by molecular mechan-

ics. This one result already shows that molecular mechanics cannot yield correct specific heats for solid bodies—at least not at low temperatures. Further, the laws of dispersion led to the conclusion that instead of consisting of only *one* material point, the atom may possess electrically charged material points (polarization electrons) that move independently of the atom as a whole and which—statistical mechanics notwithstanding—make no contribution to the specific heat. [3]

We were not in the position to relate these inconsistencies of the theory to other physical properties of matter until a few years ago, when Planck's investigations on thermal radiation quite unexpectedly shed new light on this area.[1] Though we have not yet come to the point where we need to supplant classical mechanics with a mechanics that would be able to yield correct results for fast thermal oscillations as well, still we have found the law from which the deviations from the Dulong-Petit law follow, and we learned that these deviations are related by law to other physical properties of the substances. In what follows, I shall outline the train of reasoning in Planck's investigations in a manner that will bring out clearly the connection with our problem.

It is possible to arrive at a theory of the law of cavity radiation at thermal equilibrium (the law of black-body radiation) by doing a theoretical analysis to determine the density and composition at which the radiation is in statistical equilibrium with an ideal gas, given the presence of structures that make an energy exchange between the radiation and the gas possible. One such structure is a material point bound to a point in space by forces proportional to its displacement from this point (oscillator); we shall assume that the material point of the oscillator is provided with an electric charge. Let thermal radiation, an ideal gas, and oscillators of the kind indicated be enclosed in a volume bounded by perfectly reflecting walls. By virtue of their electric charges, the oscillators must emit radiation and continually receive new momentum from the radiation field. On the other hand, the material point of the individual oscillator collides with gas molecules and in this way exchanges energy with the gas. The oscillators thus bring about an energy exchange between the gas and the radiation, and the energy distribution of the system in the state of statistical equilibrium is completely determined by the total energy, if we assume that oscillators of all frequencies are present.

In an investigation based on Maxwell's electrodynamics and on the mechanical equations for the motion of the material point of the oscillator, Planck has now shown that—assuming that only oscillator and radiation are present, but not the gas—the following relation exists between the mean kinetic energy $\overline{E_\nu}$ of an oscillator of frequency ν, and the radiation density u_ν[2]

[4]

---

[1] M. Planck, *Vorl. über d. Theorie der Wärmestrahlung*, pp. 104–166.

[2] We assume here an oscillator with three degrees of freedom.

[5] (1)
$$\bar{E}_\nu = \frac{3}{8}\frac{c^3 u_\nu}{\pi\nu^2}.$$

On the other hand, statistical mechanics implies the following: If the volume contains only a gas and oscillators (without charge), there is a relation between the temperature T and the mean energy $\bar{E}_\nu$ of the oscillator of the form

(2)
$$\bar{E}_\nu = \frac{3RT}{N}.$$

But if the oscillators interact simultaneously with the radiation and the gas, as we must assume in our analysis, then equations (1) and (2) must be satisfied simultaneously if they hold individually in the special cases discussed; for if one of these equations were not satisfied, this would have to result in a transport of energy, whether between radiation and resonators, or between the gas and resonators.

Eliminating $\overline{E_\nu}$ from both equations, we find as the condition for equilibrium between radiation and gas the equation

$$u_\nu = \frac{8\pi R}{c^3 N}\nu^2 T.$$

This is the only radiation equation that is simultaneously in agreement with our mechanics and our electrodynamics. However, it is now generally recognized that this equation does not correspond to reality. For this equation allows the integral $\int_0^\infty u_\nu d\nu$ to become infinite, and this would make a thermal equilibrium between radiation and matter impossible in the case when the heat content of the latter is different from zero, whereas it can be considered as experimentally proven that a statistical equilibrium at finite radiation density does exist in reality.

Faced with this failure of our theories to conform to reality, Planck proceeds in the following fashion: He rejects (2), and thereby a foundation in mechanics, but keeps (1), even though mechanics has been applied in the derivation of (1) as well. He obtains his theory of radiation by replacing (2) by a relation in whose derivation he introduced, for the first time, the quantum hypothesis. However, for what follows, we need neither (2) nor a corresponding relation, but only equation (1). The latter tells us how large the mean energy of an oscillator must be for it to emit on average as much radiation as it absorbs. Even if we abandon (2), we must adhere to the proposition that (1) is valid not only when the oscillator is influenced by radiation alone, but also when molecules of a gas having the same temperature collide with the oscillator. Because if these molecules were to alter the mean energy of the oscillator, then more radiation would be emitted by the oscillator than it absorbs, or vice versa. Equation (1) also remains valid when the energy variations in the resonator are preponderantly determined by the interaction between the oscillator and the gas; it is certainly therefore also valid in the total absence of an interaction with radiation, for example, when the oscillators have no charge whatsoever. The equation is also valid if the body interacting with the oscillator is not an ideal gas but any other kind of body, as long as the oscillator vibrates approximately monochromatically.

Thus, if the function of ν and T obtained in the investigations of black-body [6] radiation is substituted for the radiation density u_ν in (1), we arrive at the mean thermal energy of an approximately monochromatically vibrating structure as a function of ν and T. Starting out from Planck's radiation formula as the formula confirmed to the highest degree of approximation, we obtain from equation (1)

$$\overline{E_\nu} = \frac{3h\nu}{e^{\frac{h\nu}{kT}} - 1}, \tag{3}$$

where $k = \frac{R}{N}$ and h is the second constant of Planck's formula ($6.55 \cdot 10^{-27}$). If we assume that one gram-atom of a solid element consists of N such approximately monochromatic oscillators, we obtain its atomic heat c by differentiating with respect to T and multiplying by N, where we put $h/k = \beta$:

$$c = 3R\frac{e^{\frac{\beta\nu}{T}}\left(\frac{\beta\nu}{T}\right)^2}{\left(e^{\frac{\beta\nu}{T}} - 1\right)^2}. \tag{4}$$

The accompanying Fig. 22, taken from a paper by Nernst,[3] shows the extent to which this formula yields correct values for the specific heats of solid elements at low temperatures in the figure. In the figure, the experimentally obtained curves are drawn in thick lines, and the theoretical ones in thin lines; appended to the latter are the corresponding values of $\beta\nu$.

Even though systematic differences between the observed and the theoretical values do exist, the agreement is nevertheless astonishing, if one takes into account that each individual curve is completely determined by a single parameter ν, namely, the proper frequency of the atom of the element in question. Thus, the retaining of equation (1), which, according to the foregoing, did not seem completely justified from a purely theoretical point of view, has been completely justified by experiment. One thing needs

---

[3] *Zeitschr. f. Elektrochemie* 17 (1911): 274. [7]

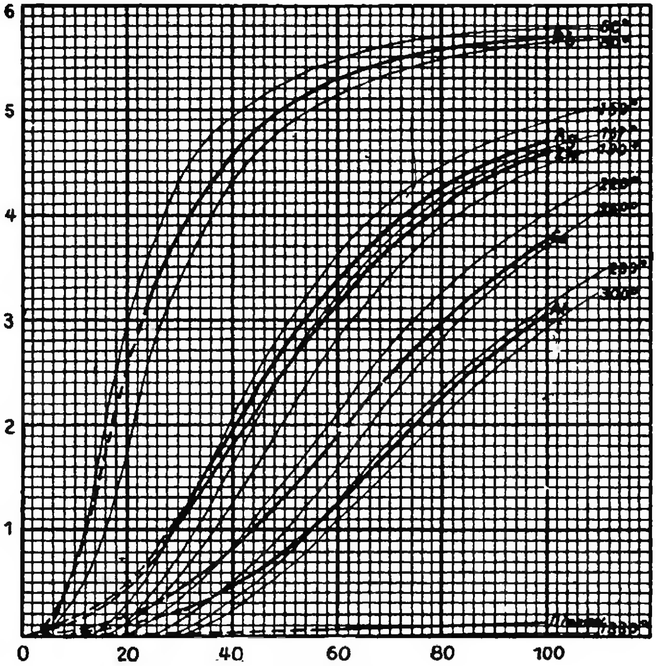


Fig. 22

to be especially emphasized: *In no way should it be inferred from the empirical confirmation of formula (1) that the quantum hypothesis is correct.* In general, nothing about mechanics can be concluded from the confirmation of (1) that could not be derived from the radiation formula and equation (2).

But what is the source of the systematic discrepancies between the observed and the theoretical curves? Why is it that, with decreasing temperatures, the specific heat approaches zero less rapidly than the theory would lead us to expect? To get what I see as the correct answer to this question, we must try to delve deeper into the mechanism of the thermal oscillations of atoms. Madelung,[4] and then, independently of him, Sutherland[5] discovered the following: In binary salts (e.g., KCl), the frequency of elastic waves (as calculated from elasticity constants) at which the wavelength attains the order of magnitude of the intermolecular distance, is of the same order of magnitude as the infrared frequencies of these bodies (as calculated from the residual radiation). This fact suggests that those atomic interaction forces that determine the infrared proper

---

[8] [4] E. Madelung, *Nachrichten d. königl. Ges. d. W. z. Göttingen, Mat.-Phys. Kl.* 20 (1909): 100–106.

[9] [5] W. Sutherland, *Phil. Mag.* (6) 20 (1910): 657.

frequencies or, more generally, the oscillations of the atoms about their equilibrium positions, are essentially identical to the forces opposing the deformation of solids. Motivated by this idea, Madelung[6] and I[7] have tried to make an approximate calculation of these proper frequencies from the elastic constants, with Madelung turning his attention to optical proper frequencies of simple compounds, while I turned my attention to those proper frequencies that determine the specific heat. The following is probably [12] the most primitive model on which one can base the calculation. Starting out from a representation in which the atoms are arranged in a cubic lattice, one comes up with a picture in which each atom has 26 neighboring atoms, all of which are located at approximately the same distance d from it. Let each change Δ of this distance d be opposed by a force $a\Delta$, where the constant a determines the degree of rigidity of this model body. The compressibility k of this model body, as well as the proper frequency of the atom ν, can then be expressed as a function of a. We then obtain the latter frequency by keeping the 26 neighboring atoms in their rest position, while the atom under consideration is supposed to oscillate. Eliminating the auxiliary variable a from these two relations, we obtain the following relation between ν and k:

$$\frac{c}{\nu} = \lambda = 1.08 \cdot 10^3 M^{\frac{1}{3}} \rho^{\frac{1}{6}} k^{\frac{1}{2}}, \tag{5}$$

where c denotes the velocity of light in vacuum, λ the wavelength in vacuum that corresponds to ν, M the gram-atomic weight, and ρ the density.

Using this formula, I obtained for silver $\lambda \cdot 10^4 = 73$, whereas Nernst obtained $\lambda \cdot 10^4 = 90$ from the specific heat. Since this good agreement in the order of magnitude is hardly a matter of chance, the essential identity of the forces determining the degree of rigidity and those determining the thermal proper frequency can be considered firmly established. Naturally, such a formula can give only a rough approximation, because it does not take into account the individual properties of the substance (e.g., the crystal structure), which do not occur in the formula.

The degree of approximation with which formula (5) is able to represent the actual situation depends, ultimately, on the extent to which a particular body is characterized, if at all, by the distance d between neighboring atoms, the mass of the individual atoms, and the compressibility. Insofar as this is the case, then in place of the compressibility, for example, one can posit some other fundamental characteristic as the defining quantity of the substance, and derive an expression for the proper frequency by dimensional [13]

---

[6] E. Madelung, *Physik. Zeitschr.* 11 (1910): 898. [10]

[7] A. Einstein, *Ann. d. Phys.* (4) 34 (1911): 120. [11]

analysis. Lindemann[8] chose the melting temperature T_s as the third defining quantity and so obtained the formula

$$\nu = 2.12 \cdot 10^{12} \sqrt{\frac{T_s}{Mv^{\frac{2}{3}}}}, \tag{6}$$

in which the numerical factor is determined empirically, and in which T_s denotes the melting temperature, v the atomic volume, and M the (gram) atomic weight.

Thus far, this formula has shown an unexpectedly close agreement with the facts. The [15] following table has been taken from the previously cited paper by Nernst:

Element	$\nu \cdot 10^{-12}$ from specific heat	$\nu \cdot 10^{-12}$ from Lindemann's formula
Pb	1.44	1.4
Ag	3.3	3.3
Zn	3.6	3.3
Cu	4.93	5.1
Al	5.96	5.8
I	1.5	1.4

Now we ask again: Why does the observed temperature dependence of specific heat deviate from the theoretically determined dependence? In my opinion, the cause for this deviation must be sought in the fact that the thermal oscillations of the atoms deviate markedly from monochromatic oscillations, and therefore do not actually have a definite frequency but rather a range of frequencies.[9] Above we mentioned the calculation of [17] ν from elastic forces; in this calculation we introduced the simplifying assumption that the atoms surrounding the oscillating atom under consideration are kept fixed in place. In actual fact, however, they do oscillate as well, and continually influence the motions of the atom considered. I will not seek to go into a more detailed examination of the actual atomic motion, but will only use an intuitive special case to demonstrate that a definite frequency is out of the question. If we picture two adjacent atoms oscillating along the line connecting them, while all other atoms stay fixed, then it is obvious that these atoms must have a greater frequency when they oscillate in opposite directions (i.e.,

---

[14] [8] *Physik. Zeitschr.* 11 (1910): 609.

[9] On this question there is no consensus whatsoever. Thus Nernst, who rescued all the results pertaining to this question from their theoretical limbo, does not share my opinion (cf., e.g., [16] *Sitzungsberichte d. Berl. Akad.* (1911), part XXII).

so that their displacements have opposite signs at all times) than when they move in the same direction, since elastic forces act between the two in the first, but not in the second case. Hence, it must be assumed that the body behaves as a mixture of oscillators of different frequencies. Now, Nernst and Lindemann found that one takes sufficiently good account of the existing experimental evidence if one assumes that the substance behaves as a mixture of oscillators, with half of them having the frequency ν, and the other half the frequency ν/2. To this assumption there corresponds the formula

$$c = \frac{3}{2}R\left[\frac{\left(\frac{\beta\nu}{T}\right)^2 e^{\frac{\beta\nu}{T}}}{\left(c^{\frac{\beta\nu}{T}} - 1\right)^2} + \frac{\left(\frac{\beta\nu}{2T}\right)^2 c^{\frac{\beta\nu}{2T}}}{\left(e^{\frac{\beta\nu}{2T}} - 1\right)^2}\right]. \tag{4a}$$

However, in accordance with what I have said before, I do not believe that we are dealing here with a theoretical formula. The only way to obtain an exact formula from (4) would be to sum over infinitely many values of ν. But with this formula Nernst and Lindemann made a very valuable advance in that they obtained a better agreement with experience without having to introduce a new constant characterizing the particular substance.[10]

Naturally, equation (4) or (4a) also makes it possible to represent the specific heat of compounds in the solid state. All one has to do is to set up an expression of the form (4a) for each kind of atom, and add up these expressions. Compounds usually display infrared proper frequencies that show up as optical absorption bands in the infrared region and as corresponding regions of metallic reflection. As Drude has shown, these infrared proper frequencies correspond to oscillations of charged ponderable atoms. [18] These are, therefore, oscillations of the same structures and under the influence of the same forces as those we have just studied. The only difference is that, in contrast to the forces mediating thermal interactions, the forces that set the atoms in motion when the body is irradiated show some degree of orderliness in space, so that the oscillation phases of identicallly charged adjacent atoms are not independent of each other. Hence, it cannot be stated without reservation that the optical proper frequencies are identical with the thermal frequencies; but, in any case, they are not likely to deviate too much from the latter.

This consequence of the theory also proves to be correct. According to Nernst, the molecular heats of KCl and NaCl can be satisfactorily explained based on the assumption that in each of these substances the metal atom and the halogen atom have the same proper frequency. The comparison of the proper frequency, as calculated from the [19]

---

[10] An exact investigation of the specific heat of solid binary compounds consisting of a very heavy and of a light atom might well be instructive, because the light atoms probably perform oscillations approximating the monochromatic oscillations assumed by the theory.

the specific heat, with the infrared absorption maximum proves, as the numbers taken from the paper by Nernst show,

[20]

$\beta\nu$ from specific heat	$\beta\nu$ from residual rays
218	203 and 232
287	309 and 265

that the expected agreement really exists, and that it is very good. Further theoretical and experimental studies of this relation between the thermal and optical behavior of insulators are most likely to yield very interesting results. In particular, it is to be hoped that we will learn something about the nature of the absorption of radiation, since in the area of infrared oscillations we seem to be on the verge of understanding not only the optical, but also the thermal aspects of the phenomenon. An understanding of the temperature dependence of absorbtivity process would be of especially great interest. [21]

The important advances described above should not by any means deceive us about the fact that we are completely in the dark regarding the laws of periodic atomic motions and, in general, regarding the mechanical laws for the case where relatively small velocities appear alongside large time derivatives of velocity. This becomes obvious if we try to apply to structures with other kinds of motion the kind of argument that led us to [22] the temperature dependence of the mean energy of sinusoidally oscillating structures. This problem always leads us to try to find the mean energy that the structure (endowed with electric charges) assumes in a black-body radiation field. But it is not possible to solve this problem without recourse to mechanics, the same mechanics that has been irrefutably proved to be invalid! As things now stand, we should consider it a pure stroke of luck that Planck's arguments yield correctly—or seem to do so—equation (1), on which the theory of specific heat is based. In fact, totally analogous reasoning yields erroneous results in other cases.

Specifically, if we imagine that a resonator—say, a monatomic molecule with an ultraviolet proper frequency—moves freely in a radiation-filled space, we can calculate the mean kinetic energy of the translatory motion acquired by this structure by investigating the oscillations and the forces that the radiation produces in the structure.[11] In this case, this mean kinetic energy must be equal to that derived for a gas molecule from the kinetic theory of gases. However, when it is based on the empirically established radiation law (say, Planck's formula), the indicated argument yields values [24] that are much too small for the kinetic energy of translatory motion. We thus see that

---

[23] [11] A. Einstein and S. Hopf, *Ann. d. Phys.* (4) 33 (1910): 1105.

we should greet with skepticism each new application of the method of deducing the thermal properties of matter from the radiation formula; for in each such application we must base ourselves on a mechanics that is most certainly lacking in general validity, and on an electrodynamics that probably cannot be upheld.

This fundamental misgiving notwithstanding, one should try to apply this method to the rotary motion of a rigid diatomic molecule about an axis perpendicular to the line connecting the atoms. One would have to assume that the atoms have opposite electric charges, and one would have to restrict oneself to considering the rotation about an axis fixed in space.

I tried to solve this problem, but did not succeed because of mathematical difficulties. [25] The solution would give us a clue as to how low the temperature is at which the ratio of the specific heats should be expected to drop under the value $^7/_5$.[12]

§ 2. *Theoretical Remarks on the Quantum Hypothesis*

We now turn to the highly important but, unfortunately, mainly unsolved question: How is mechanics to be reformulated so that it does justice to the radiation formula as well as the thermal properties of matter? The most important thing we know in this respect is already contained in Planck's fundamental paper on the radiation formula,[13] and it consists in the following: One arrives at a formula for the mean energy of the oscillator as a function of temperature that is consistent with past experience if one assumes that the oscillator can take on only such energy values that are integer multiples of $h\nu$ $(0 \cdot h\nu, 1 \cdot h\nu, 2 \cdot h\nu$, etc.).

According to statistical mechanics, the probability dW that the energy of a (one- [28] dimensional) oscillator at temperature T lies between E and $E + dE$ is given by

$$dW = \text{const.}\, e^{-\frac{E}{kT}}\, dE.$$

According to the above hypothesis, and building heavily on this result, one would have to put for those values of energy E that are multiples of $h\nu$

$$W = \text{const.}\, e^{-\frac{E}{kT}},$$

but for all other energy values, $W = 0$. For the mean energy of the oscillator we get $\bar{E} = \Sigma EW$, or, since we must have $\Sigma W = 1$,

---

[12] Nernst followed another course in solving this problem (*Z. f. Elektroch.* 17 [1911]: 270). We will return to this in §4. [26]

[13] M. Planck, *Ann. d. Phys.* 1 (1900): 69. [27]

$$\bar{E} = \frac{\Sigma EW}{\Sigma W} = \frac{0e^{-\frac{0}{kT}} + h\nu e^{-\frac{h\nu}{kT}} + 2h\nu e^{-\frac{2h\nu}{kT}} + \dots}{e^{-\frac{0}{kT}} + e^{-\frac{h\nu}{kT}} + e^{-\frac{2h\nu}{kT}} + \dots} = \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1}.$$

This is the expression found by Planck, which according to his theory has to replace formula (2), and which, together with (1), leads to Planck's radiation formula.

Simple as this hypothesis is, and simple as it is to arrive at Planck's formula with its aid, its contents strike us as counterintuitive and outlandish on closer inspection. Let us consider a diamond atom at 73° (abs): What can be said about the oscillation of the
[29] atom on the basis of Planck's hypothesis? If, with Nernst, we set $\nu = 27.3 \cdot 10^{12}$, we obtain from the oscillator formula[14]

$$\frac{\bar{E}}{h\nu} = e^{-18.6}.$$

However, the mean energy of the oscillator $\bar{E}$ is a vanishingly small fraction (about 10^{-8}) of the energy quantum $h\nu$. Only one of 10^8 atoms oscillates at any given moment, while the other atoms are completely at rest. No matter how firm one's conviction that our current mechanics is not applicable to such motions, such a picture strikes one as
[30] extremely strange.

I would like to make one additional remark. According to Eucken,[15] at low temperatures diamond does not conduct heat much more poorly than copper; the temperature dependence of thermal conductivity is here not very large anyway. Let us try to picture this from the standpoint of the quantum theory. To do this, we must form a picture of the way the quanta move. Since they are very far apart at very low temperatures, they will probably move independently of one another. Further, if it is possible to speak of a sinusoidal motion of an atom, a quantum must be bound to the atom during at least half of the period of oscillation. But when the quantum transfers to another atom, it certainly would have to transfer to one of the neighboring atoms, and always do so according to the rules of chance. I am not going to carry out the simple calculation that can be performed on this basis, but will only note that the heat flow must be proportional to the spatial derivative of the quantum density, so that for low temperatures

[32] $$\text{heat flux} \sim -\frac{d}{dx}\left(e^{\frac{h\nu}{kT}}\right) \sim -\frac{1}{T^2} e^{-\frac{h\nu}{kT}} \frac{dT}{dx},$$

---

[14] For the sake of a clean theoretical interpretation, I have based my calculations on the original formula rather than on Nernst's improved formula; incidentally, this is permitted since we are only dealing with a rough estimate.

[31] [15] *Phys. Zeitschr.* 12 (1911): 1005.

hence $$\text{thermal conductivity} \sim \frac{1}{T^2} e^{-\frac{h\nu}{kT}}.$$

Thus, in contrast to Eucken's finding, thermal conductivity would have to tend exponentially to zero at low temperatures.[16] In order to avoid this conclusion, one must introduce quite implausible assumptions about the motion of quanta. One sees that it will be difficult to bring the quantum theory in its simplest form into conformity with experience in a satisfactory fashion.

In this situation, it is a good idea to reverse direction, and to try to deduce the statistical properties of thermal processes from the now empirically well-known thermal behavior of bodies. We do so by relying upon Boltzmann's general theorem about the relation between the statistical probability and the entropy of states

$$S = k \log W + \text{const.}$$

Boltzmann's theorem gives directly the statistical probability W of the individual states that an isolated system can assume, if the entropy S is known.

We apply the theorem to a solid body of thermal capacity c that is in (thermal) [34] contact with a reservoir of infinitely large heat capacity whose temperature is T. Suppose the body possesses the energy E when in ideal thermal equilibrium. But its instantaneous energy will deviate from E by a mostly very small quantity ε, and so too its instantaneous temperature, which we will denote by $T + \tau$; this is an inevitable consequence of the disorderliness of thermal motion. The entropy corresponding to a specific value of ε or τ is obtained from the equation

$$dS = \frac{c d\tau}{T + \tau} - \frac{c d\tau}{T},$$

and thus, choosing an appropriate integration constant and neglecting the terms of higher than second order in τ,

---

[16] When I carry out the calculation indicated, I obtain for the upper limit of thermal conductivity

$$\frac{9}{13} v^{-\frac{1}{3}} N^{-\frac{2}{3}} \nu c.$$ [33]

The values yielded by this formula are furthermore much too small in comparison with experience; this result is also obtained without the quantum hypothesis.

$$S = -\frac{c\tau^2}{2T^2} = -\frac{\varepsilon^2}{2cT^2}.$$

From this we obtain, with the aid of Boltzmann's theorem,

$$W = \text{const.}\, e^{-\frac{\varepsilon^2}{2kcT^2}}.$$

Hence, the mean square $\overline{\varepsilon^2}$ of the deviation of the energy from the mean value $\overline{E}$ is

$$\overline{\varepsilon^2} = kcT^2.$$

This equation is completely general. We now apply it to an ideal, chemically simple solid consisting of n gram-atoms and having the frequency ν. For this solid we must put

$$c = 3 \cdot n\ R\ \frac{\left(\frac{h\nu}{kT}\right)^2 e^{\frac{h\nu}{kT}}}{\left(e^{\frac{h\nu}{kT}} - 1\right)^2}.$$

Substituting this expression in the preceding equation, and eliminating T with the aid of the relation

$$E = 3nN\, \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1},$$

we obtain the simple relation

$$\overline{\left(\frac{\varepsilon}{E}\right)^2} = \frac{h\nu}{E} + \frac{1}{3Nn} = \frac{1}{Z_q} + \frac{1}{Z_f},$$

where $Z_q = \frac{E}{h\nu}$ denotes the average number of Planck's "quanta" found in the body, and $Z_f = 3nN$ the total number of degrees of freedom of all the atoms of the system taken together.

One sees from this equation that the system's relative energy fluctuations, which are produced by the irregular thermal motion, result from two completely different causes, corresponding to the two right-hand terms. The relative fluctuation corresponding to the second term, which is the only fluctuation according to our mechanics,[17] results from the fact that the number of degrees of freedom of the body is finite; it is independent of the magnitude of the energy content. But the relative fluctuation corresponding to the first term has nothing to do with how many degrees of freedom the body has. This fluctuation depends solely on the proper frequency and the magnitude of the mean energy, and vanishes when this energy is very large. The magnitude of this fluctuation shows an exact agreement with the quantum hypothesis, according to which energy consists of quanta of magnitude $h\nu$, which change their location independently of each other; indeed, neglecting the second term, the equation can be written in the form

$$\sqrt{\overline{\left(\frac{\varepsilon}{\bar{E}}\right)^2}} = \frac{1}{\sqrt{Z_q}}.$$

But earlier we have seen that this conception may be difficult to reconcile with the empirical findings on heat conduction. This formula also clearly shows that the fluctuation corresponding to this term has nothing to do with the individual atom, or at least not with its size. This fluctuation might arise on account of the fact that, neglecting the carriers of energy, the less energy there is to be distributed, the smaller the manifold of energy distribution possibilities will become. The degree of orderliness of molecular motion that obtains at low total energy must be similar to that which obtains when only few degrees of freedom are present. Perhaps the flaw in the current quantum theory is mainly to be sought in the circumstance that this restriction on the possible states has been conceived as a property of the individual degree of freedom. But the essence of the quantum theory seems to remain valid all the same; if E becomes of the order of magnitude of $h\nu$, the relative energy fluctuation becomes of the order of magnitude 1, i.e., the fluctuation of the energy is of the same order of magnitude as the energy, or, the total energy is alternately present and not present; it behaves, in essence, as something with limited divisibility. But nevertheless, bounded energy quanta of definite magnitude need not necessarily exist.

The question now arises: Does the fluctuation equation just derived exhaust the thermodynamic content of Planck's radiation formula or of Planck's equation for the oscillator (3)? It can easily be seen that this is indeed the case. For, if in accordance with our inference from Boltzmann's theorem, we substitute

---

[17] This can easily be derived from the equation

$$dW = \text{const.}\, e^{-\frac{E}{KT}} dE_1 dE_2 \ldots dE_{3n},$$

where the indices refer to the individual degrees of freedom.

$$\overline{\varepsilon^2} = kcT^2 = kT^2\frac{dE}{dT}$$

for $\overline{\epsilon^2}$ in the fluctuation equation, we get (3) by integration. Thus, a mechanics that would lead to the equation we derived for the energy fluctuations of an ideal solid would also have to lead to Planck's oscillator formula.

Let us now address the question, to what extent are we forced to ascribe a special quantitative structure (in a broader sense) to radiation as well? I have investigated this question in several different ways, and have always arrived at the same results. [35]

We consider again a body K of heat capacity c, which is in a state of continuous heat exchange with an environment U of infinite heat capacity and of temperature T. Due to the irregularity of elementary thermal processes, the energy of K fluctuates around its mean value E, so that, in general, it deviates from the latter by a variable difference ε. As above, we deduce from Boltzmann's principle that the mean value of this fluctuation is given by the equation

$$\overline{\varepsilon^2} = kcT^2.$$

We assume now that the heat exchange between U and K is effected solely by thermal radiation. Assume that the surface of K is completely reflecting except for a section of surface f, which absorbs completely (is black) in the frequency region $d\nu$ and reflects completely otherwise. The surface f constantly receives radiation from U and sends radiation back to U. The radiation energy emitted from f during a given time will be greater or smaller than that absorbed by f, depending on whether the temperature of K is greater or smaller than T; the temperature of K tends therefore to approach the value T. The continuous fluctuations of the temperature or energy of K that follow directly from Boltzmann's principle result from irregular fluctuations of the radiation process over time; these must be of the magnitude that results precisely in the temperature fluctuations of K, and thus they can be calculated.

An important property of the fluctuation of the radiation emitted by f and of that absorbed by f can be deduced without calculation: namely, the property that these two fluctuations must, on the average, be equal. This is evident in the special case where the surface f stands directly opposite another surface of the envelope, f', at a very small distance; for in this case it is evident that the radiation emitted by f' fluctuates according to the same law as the radiation emitted by f, and that the radiation emitted by f' is identical to the radiation absorbed by f. But if the envelope U is positioned arbitrarily, the fluctuation of the energy absorbed by f cannot be different from that in the case we just considered; for the radiation emitted by f fluctuates independently of the position of U, and the total effect of the two fluctuations (the fluctuation of the energy of K) is also independent of the position of U. Thus, our assertion has been proved. The same argument also implies that the fluctuation in the radiation crossing, in the manner specified, a surface located somewhere in a space containing thermal radiation is equal to the fluctuation in emissions from a same-sized boundary surface of a black body.

If we denote by s the radiation energy emitted or absorbed on the average during a specified time interval t by the surface f at the temperature T, then s is a function of the temperature related to u_ν by the equation

$$s = \frac{1}{4} L u_\nu f d\nu t$$

(L = velocity of light in vacuum).

But the energy emitted or absorbed in an arbitrarily chosen time interval t will deviate from s by σ_e and σ_a, respectively, where σ_e and σ_a take positive and negative values with equal probability (equally often). Let the time t be chosen so large that σ_e and σ_a will be small compared with s, but nevertheless so small that the deviation τ of the temperature of the body K from its mean value changes only by a small fraction of its value during t.

If ε is the deviation of the energy of the body K from its mean value E at an arbitrary instant, then in the next time interval t, ε will change because of absorption by the quantity of energy

$$s_T + \sigma_a$$

and because of emission by the quantity

$$-\left(s_{T+\frac{\varepsilon}{c}} + \sigma_e\right),$$

where with sufficiently good accuracy,

$$s_{T+\frac{\varepsilon}{c}} = s_T + \frac{ds}{dT} \cdot \frac{\varepsilon}{c}.$$

Thus, after the lapse of time t, the deviation ε of the energy from its mean value will be

$$\varepsilon - \frac{ds}{dT} \frac{\varepsilon}{c} + \sigma_a - \sigma_e.$$

Since the mean square of ε is constant over time, we must have

$$\overline{\left(\varepsilon - \frac{ds}{dT} \frac{\varepsilon}{c} + \sigma_a - \sigma_e\right)^2} = \overline{\varepsilon^2}.$$

If we take into account that

$$\left(\frac{ds}{dT}\right)^2 \frac{\epsilon^2}{c^2}$$

can be neglected because it is proportional to t^2, and that, further,

$$\overline{\varepsilon\sigma_a} = \overline{\varepsilon\sigma_a} = 0,$$

just as

$$\overline{\varepsilon\sigma_e} = 0$$

and

$$\overline{\sigma_a\sigma_e} = 0,$$

then we get, if we also set

$$\overline{\sigma_a^2} = \overline{\sigma_e^2} = \overline{\sigma^2}$$

(the equality of these two quantities has been proved above),

$$\overline{\sigma^2} = \frac{ds}{dT}\,\frac{\overline{\varepsilon^2}}{c}.$$

If we insert herein the value of $\overline{\varepsilon^2}$ derived from Boltzmann's theorem, we obtain

$$\overline{\sigma^2} = kT^2\frac{ds}{dT}.$$

Thus, the fluctuations of thermal radiation prove to be independent of the heat capacity of the body K, which is how it should be. If we express s by means of u, using the relation given above, replace u using Planck's radiation formula, then differentiate, and finally eliminate T again by reintroducing the quantity s in place of T, we obtain

$$\overline{\left(\frac{\sigma}{s}\right)^2} = \frac{h\nu}{s} + \frac{c^2}{2\pi\nu^2 f d\nu t}.$$

This equation gives the expression for the mean relative fluctuation of the radiation energy that crosses f in *one* direction during time t, and indeed—as we have seen—both in the case where f is located in the immediate vicinity of a black wall and in the case where f is located at a great distance from the walls enclosing the volume.

Here as well, the square of the relative fluctuation consists of two parts, indicating two mutually independent causes of the fluctuations. The second term can easily be explained and accurately calculated from the wave theory. The fluctuation of the radiation energy crossing a surface f during time t, which corresponds to this term, is caused by the fact that, among the infinite number of plane ray bundles out of which one can compose the radiation crossing the surface f, those with almost identical directions and frequencies (and states of polarization) interfere with each other, i.e., depending on their phase angles, either predominantly reinforce or weaken each other in the time-space region under consideration. Since these phase angles of the different bundles must be totally independent of each other if the walls of the space are infinitely distant, a probability analysis yields exactly the mean value of these fluctuations. I have ascertained through calculations that this result coincides with the second term of our formula. Besides, one can see even without calculation that this relative fluctuation based on interference must be independent of the amplitude of the whole process, i.e., of s, and also that this fluctuation gets smaller the smaller the wavelength (thus, the greater ν) is, and the greater the time, space, and frequency region are over which the energy s is distributed.

However, wave optics cannot possibly explain the first term of our expression for the fluctuation. The latter corresponds to an unevenness in the distribution of the radiation energy, which is the more significant the smaller the quantity of the energy s involved. The conception that the radiation energy is distributed in localized quanta of the magnitude $h\nu$ does lead to a fluctuation of this kind. However, it seems absolutely impossible to explain the phenomena of light diffraction and interference on the basis of this conception. We stand here before an unsolved puzzle, just as in the study of thermal motion in a solid. At any rate, it seems to emerge from this analysis that our electrodynamics can no more be brought into agreement with the facts than can our mechanics.

On the other hand, this disappointing result compels us to subject the foundations of the argument just discussed to critical scrutiny. The most natural way out would be to assume that Boltzmann's theorem is in need of correction, that the formula for the mean energy fluctuation $(\overline{\varepsilon^2})$ is not correct. Such a modification would not be of any help. Because, for small values of ν at a given temperature, the theory yields fluctuations $\overline{\sigma^2}$ in agreement with the wave theory; and this agreement would disappear if the formula for $\overline{\varepsilon^2}$ were to be changed.

Further, one could suppose that $\overline{\varepsilon^2}$ may depend on the mechanism mediating the heat exchange between K and the environment. If this were the case, Boltzmann's conception of the nature of irreversible processes would be wrong in principle, because the "probability of the state" would depend on things on which, according to experience, entropy does not depend (the mode of thermal interaction between K and the environment).

Further, one could surmise that the heat absorbed by K when the latter is irradiated is not exactly equal to the radiation incident on K, so that the fluctuations of the heat taken up by K are not equal to the fluctuations of the radiations in the given wavelength range that are incident on the surface f. Such a conception does not necessarily amount to an actual violation of the energy law, because it is possible to assume that the hypothesized difference between the two quantities of energy is going to accumulate. Of course, one then faces the task of picturing the mechanism of such an accumulation, just as, analogously, one is faced with the task of picturing the immense disorderliness in the spatial distribution of the radiation energy. If we reject this accumulation hypothesis as well, then we must resolve to abandon the energy law in its present form, and conceive it as a law that can claim statistical validity only, in analogy with the conclusions from the second law of thermodynamics.[18] Who would have the audacity to give a categorical answer to these questions? I only intended to show here how fundamental and deep-rooted the difficulties are in which the radiation formula enmeshes us even if we view it as a purely empirical given.

§ 3. *The Quantum Hypothesis and the General Character of the Related Experiments*

The positive results produced by the investigations described in the last section can be summarized as follows: When a body absorbs or emits thermal energy by a quasi-periodical mechanism, the statistical properties of the mechanism are such as they would be if the energy were propagated in whole quanta of the magnitude $h\nu$. Though we have little insight into the details of the mechanism by which nature produces this property of these processes, we must expect all the same that the disappearance of such an energy of a periodic character is accompanied by the generation of packets of energy in the form of discrete quanta of magnitude $h\nu$, and second, that energy in discrete quanta of magnitude $h\nu$ must be available, so that energy of a periodic character in the frequency region ν may be produced. In particular, if a radiation in the frequency range $\Delta\nu$ is capable of producing a certain type of effect, e.g., a certain photochemical reaction, at

---

[18] In addition to what has been said in the text, let me point out that the formula for the energy fluctuations $\overline{\varepsilon^2}$ can also be applied to a radiation-filled space that is bounded by light-scattering, nonabsorbing walls and that can exchange radiation of the frequency range $d\nu$ with some body. Of course, one would again arrive at a fluctuation formula of similar construction. In this case the accumulation hypothesis is inconceivable, so that only the choice between the $h\nu$-structure and the abandoning of the strict validity of the energy principle seems to remain.

a particular density of the acting radiation, then it must induce the same effect at any lesser radiation density as well, no matter how small this density may be. [36]

These consequences seem to be confirmed in every particular, in which regard it must be noted that according to our conventional theoretical conceptions we would have expected a totally different behavior. One would have thought that a certain minimal density of electromagnetic oscillatory energy is necessary to induce, for example, the photochemical decomposition of a molecule; the electromagnetic vibration brought about in a molecule at a smaller radiation density should not be capable of causing the molecule's breakdown. On the other hand, our prevailing conceptions cannot explain why radiation of higher frequency should produce elementary processes of greater energy than radiation of lower frequency. In brief, we neither understand the specific effect of frequency nor the lack of a specific effect of intensity. Further, it has been pointed out in many a discussion that according to our theoretical conceptions it is inconceivable that light, and even more so Roentgen rays and γ-rays, no matter how low their intensities, should be able to accelerate electrons with such violence that they would fly out from bodies with their well-known high velocities. In the photoelectric effect, in particular, the kinetic energy of the ejected electrons is of the same order of magnitude as the product $h\nu$ of the acting radiation, and it even turns out that this kinetic energy increases approximately as $h\nu$ and ν in ranges devoid of resonance effects. In the face of this [37] experience, one cannot easily close one's mind to the conception (especially if one keeps in mind the great fluctuations in the conductivity of air that is irradiated with γ-rays) that energy appears in the form of large quanta in the course of absorption of radiation, and [38] also that the formation of secondary energy is in no way spatially and temporally somewhat uniform. These discontinuities, which we find so off-putting in Planck's theory, [39] seem really to exist in nature.

The difficulties which stand in the way of formulating a satisfactory theory of these fundamental processes seem insurmountable at this time. From where does an electron in a piece of metal that is struck by Roentgen rays take the great kinetic energy we are seeing in secondary cathode rays? After all, the field of the Roentgen rays impinges on all of the metal; why does only a small portion of electrons attain the velocity of those cathode rays? How is it that the absorbed energy shows up only in relatively exceedingly few places? What distinguishes these places from other places? These and many other questions are being asked in vain. [40]

An interesting question is whether absorption has the character of a random event also when viewed from the standpoint of the absorbed radiation. This amounts to the question as to whether two coherent ray bundles remain completely coherent if each of them is weakened to the same fraction of its value by absorption. No doubt, everybody presumes that the coherence will be completely maintained; still, it would be nice if we knew this for sure. [41]

And here is another question to which an experimental answer would be desirable: It is generally assumed that the high velocities displayed by electrons emerging from bodies irradiated by ultraviolet light or Roentgen rays are produced by one single elementary act. But in actual fact, we do not have any proof for that. It would be conceivable a priori that the electrons attain these high velocities little by little, through collisions with many irradiated molecules. If this were the case, we would be able to reduce the velocity of emergence by reducing the thickness of the effectively irradiated layer. Also, in that case—especially in irradiation with weak Roentgen rays—there would [42] have to pass some possibly measurable time from the beginning of the irradiation to the formation of the secondary rays. These kinds of experiments, if they were to turn out positive, could provide irrefutable proof that those high electron velocities are not due to a quantized distribution of radiation energy.

Finally, it would be of great importance to check with the greatest possible precision whether the secondary effects that arise in the absorption of radiation are really [43] absolutely independent of the *intensity* of the exciting radiation. Currently one must hold that the temperature of a bundle of rays of low intensity and high frequency depends only weakly on the intensity. Thus if the *temperature* of the radiation bundle (with or without the influence of the phase spread in the bundle) were to determine, for example, the velocity distribution of the electrons in the photoelectric effect, then a *slight*, but still measurable dependence of this velocity distribution on the radiation intensity would also come into evidence.

§ 4. The Rotation of Gas Molecules

Sommerfeld's Hypothesis.[19]

Two other important attempts to relate Planck's constant h to mechanical properties of elementary structures are known. First, using an approximate argument, Nernst [45] attempted to determine the rotational energy of gas molecules as a function of the temperature. Second, Sommerfeld calculated the electromagnetic radiation emitted in the stopping of cathode-ray electrons, as well as in the acceleration of β-ray particles, on the basis of the hypothesis $L\tau = h$, where L denotes the kinetic energy of the particle, τ its collision time, and h the Planck's constant. We will show to what extent these two [46] things can be derived from the radiation formula without recourse to special hypotheses. But we will have to content ourselves with rough approximations.

If, with Nernst, we assume for the sake of simplicity that all molecules of the diatomic gas being considered have a definite angular frequency ν, which is the same for

---

[44] [19] A. Sommerfeld, "Über die Struktur der γ-Strahlen." *Sitz.-Ber. d. Königl. Bayerischen Akad. d. Wiss. Phys. Klasse* (1911): 1–60.

all molecules, then the relation between the rotational energy E, the frequency, and the temperature will not differ substantially from the corresponding relation for the linear oscillator. We will have approximately

$$E = \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1}.$$

If we denote by I the moment of inertia with respect to an axis through the center of gravity of the molecule and perpendicular to the line connecting its atoms, then we must assume, in accordance with mechanics,

$$E = \frac{1}{2} I (2\pi\nu)^2.$$

These two equations contain the relation between E and T we have been looking for; all that remains is to eliminate ν.[20] Nernst and Lindemann have already pointed [47] out[21] that it would be of exceedingly great interest to investigate the infrared absorption of diatomic gases whose molecules, probably as in HCl, possess an electric moment. In such cases one could use Kirchhoff's law to find from absorption coefficients the emission coefficients for the different frequencies, and from these the number of molecules momentarily present with a particular angular velocity—the statistical law of rotational motion. Of course, a part of absorption phenomena would have to be ascribed to the relative oscillation of the two molecules [atoms].

Let us now turn to Sommerfeld's hypothesis concerning elementary collisions.

One of the areas left unaffected by molecular mechanics is the kinetic theory of monatomic gases, since in this case the mechanism of collisions is immaterial. But we can learn something about the latter from the radiation formula, using a procedure that is completely analogous to that adopted for the oscillator; unfortunately, for the time being we must do without an exact theory in this case as well.

As in §1, assume that thermal radiation and a monatomic gas are in thermal equilibrium in some enclosure. In this case, however, the possibility of thermal interaction between gas and radiation shall be brought about by endowing individual gas molecules with an electric charge. If these molecules collide with other molecules or with the wall, they will emit and absorb radiation. Let us assume that these collisions are so infrequent that each collision can be considered by itself, as an isolated event. The radiation emitted in a collision is easy to determine with the aid of Maxwell's theory, if the velocity of the emitting atom is given as a function of time.

---

[20] Instead of the second of these relations, Nernst assumed the relation $\beta\nu = a\sqrt{T}$ But this relation could only be satisfied if the specific heat were independent of the temperature.

[21] *Zeitschr. f. Elektroch.* 17 (1911): 826. [48]

According to Kirchhoff's law,

$$u_\nu = \frac{8\pi}{c}\frac{\varepsilon_\nu}{\alpha_\nu},$$

where ε_ν denotes the emission coefficient and α_ν the absorption coefficient of a medium. If ν is kept constant, u_ν will be practically zero up to a certain temperature, after which it will rapidly increase. Since α_ν remains finite, what we said about u_ν applies also to ε_ν. According to the formula of Wien or Planck, the condition for u_ν or ε_ν to become different from zero is

$$\frac{h\nu}{kT} < Z,$$

where Z is some number of the order of magnitude 1. Since, up to a negligible factor, kT equals the mean energy E of translatory motion of the gas molecules, this condition can also be written in the form

[49] $$h\nu < ZE.$$

Thus, if E is the energy of their translatory motion, charged gas molecules must collide in such a way that no frequencies inconsistent with this equation will be emitted.

If the collisions were *sudden,* the equation would be violated, according to Maxwell's theory, because even the greatest frequencies would have to occur in the radiation emitted in the collision. Thus, sudden collisions cannot exist; the collision must take place gradually, in such a way that frequencies greater than ν will not be produced. It can easily be demonstrated that the duration τ of collisions satisfying this condition is of the order of magnitude of

$$\frac{1}{\nu_{\max}}.$$

Accordingly, the above relation can also be written in the form $h = E\tau \times$ a number of the order of magnitude 1.

This is Sommerfeld's hypothesis, which permits the correct calculation, at least in order of magnitude, of the fraction of the energy of the cathode radiation that is converted to Roentgen radiation.

Thus, to derive Sommerfeld's hypothesis from the radiation equation, one need only assume, essentially, that the electron energy determines the emission. If this line of reasoning corresponds to reality, then a charged elementary structure, e.g., an electron, loses only a very small fraction of its kinetic energy in a collision electron velocities like those occurring in the photoelectric effect (without resonance) or if not-too-fast cathode rays are involved. If one looks upon the acceleration of electrons by radiation as the reverse of such emission processes, one will incline to the view that acceleration of this kind must also proceed in many stages. As already noted, we would then expect that under otherwise identical conditions, e.g., in the photoelectric effect, electrons emerge with lower velocities from very thin irradiated effective layers than from thicker layers.

Doc. 27

"Discussion" Following Lecture Version of "The Present State of the Problem of Specific Heats"

Einstein et al.

[*Abhandlungen der Deutschen Bunsen Gesellschaft für angewandte physikalische Chemie* 3 (1914): 353-364]

Einstein: We probably all agree that the so-called quantum theory of today is, indeed, a helpful tool but that it is not a theory in the usual sense of the word, at any rate not a theory that could be developed in a coherent form at the present time. On the other hand, it has also turned out that classical mechanics, which finds its expression in the equations of Lagrange and Hamilton, can no longer be viewed as a useful schema for the theoretical representation of all of physical phenomena (cf. especially the report of H. A. Lorentz). [1]

This raises the question of which general laws of physics we can still expect to be valid in the domain with which we are concerned. To begin with, we will all agree that the energy principle is to be retained. [2]

In my opinion, another principle whose validity we must maintain unconditionally is Boltzmann's definition of entropy through probability. It is to this principle that we owe the faint glimmer of theoretical light we now see shed over the question of states of statistical equilibrium in processes of an oscillatory character. But there is still the greatest diversity of opinion as regards the content and domain of validity of this principle. I will therefore first present in brief my view about this matter. [3]

If we have an isolated physical system of given energy, this system can still assume the most diverse states, which are characterized by a number of quantities that are observable in principle (e.g., volumes, concentrations, energies of parts of the system, etc). All of these states compatible with the given energy value shall be denoted by Z_1, Z_2 .... Z_l. According to thermodynamics, if the system is brought into one of these states (Z_a), it will pass successively through specific states Z_b, Z_c, tending toward a final state Z_g, the state of thermodynamic equilibrium, in which it remains indefinitely. But we know from the statistical theory of heat on the one hand, and from our experience with Brownian motion on the other, that this conception is only a more or less rough, approximate description of the *average* behavior of a system. In reality, the character of irreversibility ascribed to the phenomena in this description is only apparent; nor does the system remain in the state of thermodynamic equilibrium. Instead, in the course of time the system assumes again and again, in a perpetual alternation, all the states Z_1 .... Z_l without exception.

The apparently unidirectional succession of states from a state Z_a onwards, and the final apparent persistence in a state of thermodynamic equilibrium Z_g, Boltzmann attributes to the fact that in the overwhelming majority of cases a state Z_a is succeeded by a more probable state Z_b. From among all the states Z_b, $Z_{b'}$, $Z_{b''}$, etc., to which Z_a can pass in the very short time τ, the state Z_b will practically always appear, because it possesses an enormously greater probability than the state Z_a and all of the other states $Z_{b'}$, $Z_{b''}$, etc. Thus, the apparently unidirectional succession of states actually consists in states of ever greater probability following successively upon each other.

But such an argument gains some measure of persuasive power only when one has made clear what is to be understood by the "probability" of a state. If a system left to itself passes in an endless succession through the states $Z_1 \ldots Z_l$ (in the most varied sequences), each state will possess a definite temporal frequency.

There will be a fraction τ_1 of a very long time T, during which the system will be in the state Z_1; if, for large T, $\frac{\tau_1}{T}$ tends toward a limiting value, then we call this the probability W_1 of the first state, etc. Thus, the probability W of a state is conceived as the latter's temporal frequency in a system left to itself an infinitely long time. From this point of view, it is noteworthy that in the overwhelming majority of cases, if one starts out from a specific initial state, there will exist a neighboring state which the system—if left to itself an infinitely long time—assumes more often than it will do other states. But if we forgo such a physical definition of W, the statement that in the overwhelming majority of cases the system passes from one state to a state of greater probability is a statement devoid of meaning, or—if one has set W equal to an arbitrarily chosen mathematical expression—is an arbitrary assertion.

If W is defined in the manner indicated, then it follows from the very definition that a system left to itself in an arbitrary state (and isolated from without) must assume, in the majority of cases, successive states of ever greater probabilities, and from this it follows that W and the entropy S are connected by Boltzmann's equation

$$S = k \log W + \text{const.}$$

This follows from the circumstance that the probability W—insofar as the character of a unidirectional flow of events is maintained at all—must always grow with time, and that there cannot be a function independent of S that has this property at the same time as S. That the connection between S and W is exactly the one given in Boltzmann's equation follows from the relations

$$S_{\text{total}} = \sum S, \qquad W_{\text{total}} = \Pi(W),$$

which hold for the entropy and probability of states of systems composed of a number of subsystems.

If one defines W in the manner indicated, as temporal frequency, then Boltzmann's equation contains right away a physical statement. The equation contains a relation between quantities that are observable in principle, i.e., the equation is either correct or incorrect. Boltzmann's equation is usually applied in the following way: one starts out from a specific elementary theory (e.g., molecular mechanics), determines the probability of a state theoretically, and calculates the entropy from this by means of Boltzmann's equation in order to learn, finally, the thermodynamic properties of the system. But one can also proceed in the reverse direction: from the empirically ascertained thermal behavior of the system, one determines the entropy values of the individual states, and from these one calculates the probabilities of the states by means of Boltzmann's equation.

To illustrate this way of applying Boltzmann's principle, let me use the following [4] example: Suppose a cylindrical vessel contains a liquid, in which there shall be suspended a particle whose weight exceeds by P the liquid displaced by it. According to thermodynamics, the particle should sink to the bottom and remain there. From the point of view of the kinetic theory of heat, in an incessant fluctuation the particle will change its height above the bottom in irregular succession, without ever coming to rest. To lift the particle to the height z above the bottom, one has to perform the work Pz. In order for the energy of the system not to change in this process, one must simultaneously withdraw from the system an amount of heat that is equivalent to this work, so that the entropy of the system as a function of the height z of the particle is expressed by

$$S = \text{const.} - \frac{Pz}{T}.$$

Using Boltzmann's equation, one obtains from this the probability W that the particle will be found at the height z at an arbitrary instant of time:

$$W = Ce^{-\frac{Pz}{kT}}$$

[5] This is the law that Perrin actually obtained from his observations. It is clear that this relation expresses the state of affairs established by Perrin only if the probability W has been defined in the manner indicated above.

This simple example also provides a beautiful illustration of Boltzmann's conception of an irreversible process. Namely, if P is not far too small, then the exponent $\frac{Pz}{kT}$ will be of considerable magnitude for fairly large z because of the smallness of the constant k $(= \frac{R}{N})$; thus, W will be small and will decrease very rapidly with increasing z. If one raises the particle to a certain height above the bottom and then leaves it alone, in the overwhelming majority of cases the particle will sink to the bottom in an almost perpendicular line and with almost constant velocity (an irreversible process in the sense of thermodynamics). Yet, on the other hand we know that the particle can rise on its own, even though very rarely, to any height above the bottom of the vessel.

Lorentz: Mr. Einstein speaks about the probability of a given height z of the particle. But to be strict, the probability of finding the particle between z and dz must be expressed by Wdz. This difference is not without significance because it brings with it a difficulty. Instead of z, one can just as well choose any function of these variables, e.g., $z' = z^2$ as the coordinate. Then one would have to introduce a probability W', defined as follows:

$$W'dz' = Wdz,$$

or

$$W' = \frac{W}{2z}.$$

This would lead to the value $S' = k \log W'$ for the entropy, which differs from $S = k \log W$ by a *variable* quantity $k \log 2z$. But this is impermissible.

Einstein: Strictly speaking, one cannot actually speak of the probability that the particle (or its center of gravity) is to be found at the height z, but only of the probability that it is to be found in the height interval between z and $z + dz$. [6]

But by no means does this fact imply that Boltzmann's equation $S = k \log W$ cannot be strictly valid. For it is also easy to see that a remark can be made with regard to entropy that would correspond precisely to the remark Mr. Lorentz just made with regard to probability: strictly speaking, one cannot talk of the entropy of a given *state*, but only of the entropy of a state region.

To show this in a fairly simple case, imagine a cylindrical vessel that contains a liquid as before; in it, a particle is suspended, and the particle's variable height above the bottom is again denoted by z. To make it really simple, I also imagine that the weight of the particle is exactly compensated by its buoyancy. Now we seek the entropy of the state characterized by the fact that the center of gravity of the particle is found at a given height z. To find the entropy value of this state, it must be realized in a reversible way, which is possible to do in the following manner. We picture two sieves that are impermeable for the particle; the one sieve shall initially be at the height $z = 0$, and the second at $z = l$. We push these sieves infinitely slowly from both sides toward a specific height $z = z_0$. When this process has been completed, the particle is found at $z = z_0$. In this process we must perform mechanical work in order to overcome the osmotic pressure of the particle. If we brought the sieves to the distance δ from each other, this work is equal to $+\frac{RT}{N} \lg \frac{l}{\delta}$. For the particle to remain at the height $z = z_0$, S must be brought to the value zero, that is, a logarithmically infinitely great quantity of work must [7] [8] be performed. It can readily be seen, furthermore, that the entropy has the value $-\frac{\text{work}}{T}$, so that we have to set

$$S = \text{const.} + \frac{R}{N} \lg \delta .$$

Thus, S too becomes infinite with vanishing δ. The entropy belonging to the interval dz is therefore

$$S = \text{const.} + \frac{R}{N} \lg dz$$

On the other hand, the probability W for the interval dz is

$$W = \text{const.}\, dz.$$

Thus, independently of the choice of the interval dz, Boltzmann's equation

$$S = \frac{R}{N} \lg W + \text{const.}$$

is, in fact, satisfied here. It follows with great probability that Boltzmann's equation is strictly valid if S and W refer to the same state region.

Poincaré: When defining probability, the choice as to which differential to employ as a factor is not arbitrary; one must take an element of the phase space.

[9] *Lorentz*: Mr. Einstein does not follow Gibbs's method; he simply speaks about the probability of a specific value of the coordinate z.

Einstein: What is characteristic of this standpoint is that one uses the (temporal) probability of a state defined purely *phenomenologically*. In this way one gains the advantage of not having to base the analysis on any specific elementary theory (e.g., statistical mechanics).

Poincaré: In every theory that one introduces in the place of ordinary mechanics, instead of the element of the phase space one must use an *invariant* element as the differential.

Wien: In my opinion, one can set up a relation between entropy and probability for radiation only if one goes back to the emitting atoms.

Einstein: An analysis analogous to the one just brought forth for the case of a [10] suspended particle can be applied to radiation enclosed in a cavity. Imagine a chest of a total volume V that has perfectly reflecting or perfectly white inner walls and that encloses a radiation energy E, the frequency of which is close to ν. A partitioning wall that is also perfectly reflecting or white and that has a hole in it divides the interior of the chest into two parts, whose volumes are V_1 and V_2. The radiation will ordinarily be distributed over the volumes V_1 and V_2 in such a way that the ratio of the energy shares, E_1 of volume V_1 and E_2 of volume V_2, will be the same as the ratio of the two volumes. But because of the irregularities of the radiation process, all other distributions compatible with the given value E of the total energy will also occur. To each of the distributions (E_1, E_2) belongs a probability W. But to each of the distributions there belongs also a definite value of the entropy S. S and W must be connected by Boltzmann's equation. Since the entropy of any distribution of this kind can be found from the radiation law, one obtains the statistical probability W of any distribution from Boltzmann's equation. If the radiation is so attenuated that it falls within the range of validity of Wien's radiation law, then it turns out that the statistical distribution law is constituted as if the radiation consisted of pointlike structures, each of which possesses the energy $h\nu$. Specifically, for the probability that the total energy E will be localized in the subvolume V_1, one obtains the expression

$$W = \left(\frac{V_1}{V}\right)^{\frac{e}{h\nu}}.$$

This result is so interesting because it cannot be brought into agreement with the wave theory of radiation. This can be seen without calculation from the following argument by analogy. [11]

Let there be given a distribution of the radiation for a specific value E_0 of the total energy. If we now imagine all the electric and magnetic field components multipled by a constant α, we obtain a new vector field satisfying Maxwell's equations that has the same frequency range and degree of randomness as the original field. All energy densities in the latter field are exactly α^2 greater than those in the original field. From this it follows directly that, in the latter case, the energy distribution $\alpha^2 E_1$, $\alpha^2 E_2$ occurs exactly as probably, i.e., as frequently, in the second field, as the energy distribution E_1, E_2 does in the original field. From this it follows that according to the wave theory in its present form the frequency (probability) of a specific distribution ratio $\frac{E_1}{E_2}$ must be independent of the value of the total energy E. But this contradicts the expression for W that we derived from the entropy of radiation by means of Boltzmann's equation.

The quantum hypothesis is a provisional attempt to interpret the expression for the statistical probability W of the radiation. By conceiving radiation as consisting of small complexes of energy $h\nu$, one found an intuitive interpretation of the probability law for low-intensity radiation. I emphasize the provisional character of this auxiliary idea, which does not seem compatible with the experimentally verified conclusions of the wave theory. But since it follows, in my opinion, from such considerations that the localizations of energy in the radiation field implied by our current electromagnetic theory do not correspond to reality in the case of low-intensity radiation, we must also admit, in addition to Maxwell's electrodynamics, which is indispensable to us, a hypothesis such as that of quanta in some form or other.

Planck: I too hold firmly, and for all cases, to the relation

$$S = k \log W + \text{const.}$$

as the general expression of the principle that the second law of thermodynamics is basically a probability law. This is why the entropy of a state always immediately yields its probability as well. But on the other hand, I do not believe that there exists a completely general definition of probability that is also usable outside of classical dynamics and that permits the probability of a totally arbitrary state to be calculated on the basis of temporal (or spatial) fluctuations alone, without taking into account the mutually independent elementary regions of equal probability. Especially as viewed from the standpoint of the quantum hypothesis, there seem to exist states whose nature is too complicated to preserve the simple relationship between probability and fluctuations to which the consideration of elementary regions leads.

As concerns, in particular, thermal radiation in a vacuum, I am of the opinion that its entropy (or probability) cannot really be derived from the energy fluctuations of free radiation alone; rather, one must either go back to the emitting substance, from which [12] the radiation originates, or consider the absorption (cf. my report, p. 84). Otherwise it is not possible to recognize behind the complex event the equiprobable elementary events that are producing it.

Lorentz: All the same, it seems to me that one can always talk of a probability that the energy content in one of the halves of volume considered lies in the interval between ξ and $\xi + d\xi$. This probability could be measured by the time interval during which this energy distribution is actually in existence. Now if one assumes, on the one hand, that a particular energy distribution deviating from the uniform energy distribution possesses a certain probability, and, on the other hand, that this implies a perfectly definite value of the entropy, then I do not see why one should not apply Boltzmann's theorem.

Langevin: If one can define a probability as well as an entropy for radiation, then it seems difficult to bypass Boltzmann's general relation between these two quantities. If we consider a system consisting of matter and ether, then the probability of any configuration is equal to the product of the probability of the state of the matter and that of the ether, taken individually; the total entropy is equal to the sum of the individual entropies, and by virtue of an argument presented by Mr. Planck in his report, there must therefore exist a proportionality between the entropy and the logarithm of the probability; the proportionality factor for the matter as well as for the ether is Boltzmann's coefficient.

Poincaré: This is exactly what the definitions of both the entropy and the probability are based on.

[13] *Lorentz*: The first term $\frac{h\nu}{E}$ in Mr. Einstein's formula seems in fact to be totally incompatible with Maxwell's equations and with the prevailing views about electromagnetic processes. This can be recognized from Mr. Einstein's manner of representation as well as from the following reasoning: Let P denote a diathermic disk situated in a space filled with black-body radiation. We now consider the energy of the rays

emanating from the disk in a given direction and contained in an instant t in a limited volume v. This energy E derives from the energy quantities E_1 and E_2, which were present, at a somewhat earlier instant t', in two volumes v_1 and v_2, both of which are equal to v and lie on both sides of the disk, the one on the same side as v, the other on the opposite side. If one denotes the common average value of E, E_1, and E_2 by E_0, and the deviations from this value by α, α_1, and α_2, and neglects the fluctuations in the volume v caused by the interference of the reflected and transmitted rays, one obtains $\overline{\alpha_1^2} = \overline{\alpha_2^2}$; the same value would have to be found for α^2.

Meanwhile, we have (r denotes the coefficient of reflection)

$$E = rE_1 + (1 - r)\, E_2,$$
$$\alpha = r\alpha_1 + (1 - r)\, \alpha_2,$$
$$\overline{\alpha^2} = [r^2 + (1 - r)^2]\,\overline{\alpha_1^2},$$

where the last-mentioned value is smaller than α_1^2. This result derives from our having tacitly assumed that for a given frequency and a given angle of incidence, the fraction of the radiation reflected will always be the same.

Nernst: Could one not demonstrate the temperature fluctuations by measuring the electrical resistance at very low temperatures?

Wien: Perhaps it might be possible to avoid the difficulties with fluctuations by assuming an accumulation of energy in the atoms that does not lead to an immediate rise in the temperature. Such processes might also occur in thermal conduction.

Einstein: First of all, this hypothesis is of no use whatever in explaining the law for the distribution of radiation between two communicating spaces that follows from Boltzmann's principle. Furthermore, it is obvious that this hypothesis cannot be applied to ideal monatomic gases; but the body denoted by K can consist of exactly such gases, without changing what was essential in the last argument.

Langevin: Like Mr. Planck, I too think that the conditions are not the same when a body is, at one time, very close to the wall of a cavity and when it is, another time, very far from it. In the latter case, the fluctuations in emission and absorption on the surface of the wall and on the surface of the small body are independent of each other; the probability of both events is therefore a product of the individual probabilities. But when the surfaces are very close to each other, the intervening medium cannot absorb any energy, the fluctuations are not independent of each other, and statistical considerations can no longer be applied in the customary way.

Kamerlingh Onnes: Based on Nernst's conception, but in another way, Mr. Einstein has calculated that, at 0°, the molecular heat of hydrogen at constant pressure is to be expected to deviate from that of a diatomic gas by 4%. In this connection, I would like [14]
to revert to the remark on the specific heat of hydrogen made at the lecture by Mr. [15]
Nernst. The calculation mentioned there showed that from 14°K and on, hydrogen would

display a distinct deviation in the direction toward the value for monatomic gases, and this induced Mr. Keesom and me to embark upon the experimental investigation. Let it here be noted that this investigation looked promising, for on the basis of the same calculation, a deviation could be expected even at 0°C, a deviation already indicated by Pier's experimental results, which were mentioned by Mr. Nernst. According to a more accurate calculation, which, however, faithfully followed Nernst's way of doing it, the deviation would amount to about 3% of the molecular heat at constant volume. Pier's result gives about 4%.

Lorentz: Perhaps it might be of interest to mention the result at which one arrives if one applies the idea of the energy element to a rigid sphere able to rotate about a diameter.

If ν is the number of rotations per second, then the energy is equal to $q\nu^2$, where q is a constant. The hypothesis that this energy must be a multiple of $h\nu$ leads to the following formulas, in which n represents an integer:

$$q\nu^2 = nh\nu \qquad \nu = n\frac{h}{q} \qquad q\nu^2 = n^2\frac{h^2}{q}.$$

Hence the sphere would have to be able to rotate only with definite velocities that form an arithmetic progression; the possible energy values would therefore have to vary as the squares of integer numbers.

But then, this remark is of no great significance. When applying the hypothesis of energy elements, one can confine oneself to systems for which a definite frequency, determined by the nature of the process in question, is given in advance.

[16] *Poincaré*: Mr. Nernst cites a formula in which ν is proportional to $\sqrt{T}$.

Einstein: This formula contradicts the final result obtained by Nernst himself, and would therefore have to be changed.

Poincaré: At a given temperature, ν will be distributed according to a certain law; what result would one obtain for specific heat if all values of ν were taken into account corresponding to their relative frequency?

Hasenöhrl: Nernst's oscillator model, in which a light-weight atom revolves around [17] a much heavier one at a constant distance (*Zeitschr. f. Elektrochem.* 17 [1911]: 825), does not have a definite proper oscillation; but if one calculates its energy under the assumption of specific elementary regions in the phase space, one obtains an expression of the form[1]

---

[1] This formula is easy to derive. The energy is completely kinetic and has the value

$$9E = C_1(\dot{\vartheta}^2 + \sin^2\vartheta\dot{\varphi}^2) = C\left(p_1^2 + \frac{1}{\sin^2\vartheta}p_2^2\right)$$

$$\frac{c}{e^{\frac{c'}{T}} - 1},$$

in which c and c' depend solely on the moment of inertia, and no value of the oscillation frequency ν appears. To see whether this can be brought into accord with Planck's radiation formula, it would also be necessary to investigate the relation between the energy of the resonator and that of radiation, which is probably not so simple in this case as in that of Planck's resonator. The calculation of this relation seems to meet with very great mathematical difficulties.

(ϑ and φ denote spheric coordinates; $p_1 = \frac{\partial E}{\partial \dot{\vartheta}}, p_2 = \frac{\partial E}{\partial \dot{\varphi}}$; C_1 and C are constants). If one uses the mode of expression of Gibbs's statistical mechanics, one obtains

$$e^{-\frac{\Psi}{\Theta}} = \int_0^{\pi} d\vartheta \int_0^{2\pi} d\varphi \int \int_{-\infty}^{+\infty} dp_1 dp_2 e^{-\frac{C}{\Theta}\left(p_1^2 + \frac{1}{\sin^2\vartheta} p_2^2\right)} = \frac{4\pi^2}{C}\Theta,$$

from which follows

$$\overline{E} = -\Theta^2 \frac{d}{d\Theta}\left(\frac{\Psi}{\Theta}\right) = \Theta.$$

Now we introduce the phase volume

$$V = \int_0^{2\pi} d\varphi \int_0^{\pi} d\vartheta \int\int dp_1 dp_2,$$

where p_1 and p_2 are to be taken between the limits 0 and $C\left(p_1^2 + \frac{1}{\sin^2\vartheta} p_2^2\right) = E$.

A simple calculation gives

$$V = \frac{4\pi^2}{C} E, \qquad E = \frac{C}{4\pi^2} V,$$

$$e^{-\frac{\Psi}{\Theta}} = \int_0^{\infty} e^{-\frac{1}{\Theta}\frac{C}{4\pi^2} V}\, dV = \frac{4\pi^2}{C}\Theta.$$

Langevin: It seems to me, as Mr. Planck's argument on the basis of his hypothesis regarding the elements of the phase space shows, that the introduction of energy elements is admissible only if the system possesses a definite frequency that is independent of the accumulated energy. The situation is completely different in the case of rotation; here the period depends just on the kinetic energy, potential energy not being present at all. It seems to me therefore gratuitous to apply the hypothesis of energy quanta to the rotation of molecules.

Lindemann: The assumption that a diatomic gas molecule rotating with frequency ν can take up only quanta of the magnitude $h\nu$ is probably inadmissible. For if this were the case, a gas molecule that has been heated up from absolute zero would have to gain the frequency ν_1 from the first collision it experiences. Since it could then take up only an integral multiple of $h\nu_1$, its frequency would be $\nu_1\sqrt{1+n_1}$ after the second collision, $\nu_1\sqrt{1+n_1}\,\sqrt{1+n_2}$ after the third one, etc.

It is highly improbable that this molecule will interact at some time with another molecule that has an angular momentum of opposite direction but exactly the same magnitude. Thus, the velocities of rotation would finally be so great that they could not be exchanged at all, i.e., the atomic heat would be $\frac{3}{2}R$.

The introduction of quanta is not gratuitous at all, but absolutely necessary, and one must probably adhere to the formula $\frac{h\nu}{e^{\frac{h\nu}{kT}}-1} = (2\pi\nu)^2 I$, or one similar to it, for

---

According to the theory of quanta, instead of the integral one would have to write the sum

$$e^{-\frac{\Psi}{\Theta}} = \sum_{x=0}^{x=\infty} h e^{-\frac{1}{\Theta}\frac{C}{4\pi^2}xh} = \frac{h}{1-e^{-\frac{1}{\Theta}\frac{Ch}{4\pi^2}}}.$$

This yields

$$\overline{E} = -\Theta^2\frac{d}{d\Theta}\left(\frac{\Psi}{\Theta}\right) = \frac{Ch}{4\pi^2}\,\frac{1}{e^{\frac{1}{\Theta}\frac{Ch}{4\pi^2}}-1}.$$

The quantity h used here is not identical with that of Mr. Planck; here it has the dimensions of the square of action (Hasenöhrl).

otherwise one comes into conflict with the radiation laws; but it seems hardly possible to derive this formula by means of the usual conceptions of the theory of quanta.

Lorentz: I remember a conversation I had some time ago with Mr. Einstein. We [18] were talking about a simple pendulum that can be shortened by taking hold of the thread with two fingers and gliding them along it. If at the beginning the pendulum possessed an energy element corresponding exactly to its oscillation period, then its energy at the end of the experiment must obviously be smaller than that of an energy element that corresponds to the new frequency.

Einstein: If one changes the length of the pendulum infinitely slowly in a continous manner, then the oscillation energy remains equal to $h\nu$ if it was $h\nu$ in the beginning; the energy of oscillation varies as ν. The same applies to an oscillatory electrical circuit without resistance and to free radiation. [19]

Lorentz: This result is most noteworthy and removes this difficulty. In general, the hypothesis of energy quanta leads to interesting problems in all those cases in which the frequency can be changed arbitrarily.

Warburg: One can increase the frequency of an oscillating simple pendulum without expending work by letting, as in Galileo's experiment, a point of the thread fall against a fixed rod when the pendulum is in the equilibrium position, and then fixing this point during the ascent of the pendulum bob.