ACTA PHYSICA

ACADEMIAE SCIENTIARUM HUNGARICAE

TOMUS XVII FASCICULI 1—2

AKADÉMIAI KIADÓ, BUDAPEST
1964

ACTA PHYS. HUNG.

ACTA PHYSICA

A MAGYAR TUDOMÁNYOS AKADÉMIA FIZIKAI KÖZLEMÉNYEI

SZERKESZTŐSÉG ÉS KIADÓHIVATAL: BUDAPEST V., ALKOTMÁNY UTCA 21.

Az *Acta Physica* német, angol, francia és orosz nyelven közöl értekezéseket a fizika tárgyköréből.

Az *Acta Physica* változó terjedelmű füzetekben jelenik meg: több füzet alkot egy kötetet.

A közlésre szánt kéziratok a következő címre küldendők:

Acta Physica, Budapest 502, Postafiók 24.

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Az *Acta Physica* előfizetési ára kötetenként belföldre 80 forint, külföldre 110 forint. Megrendelhető a belföld számára az Akadémiai Kiadónál (Budapest V., Alkotmány utca 21. Bankszámla 05-915-111-46), a külföld számára pedig a „Kultúra" Könyv- és Hírlap Külkereskedelmi Vállalatnál (Budapest I., Fő u. 32. Bankszámla 43-790-057-181 sz.) vagy annak külföldi képviseleteinél és bizományosainál.

Die *Acta Physica* veröffentlichen Abhandlungen aus dem Bereiche der Physik in deutscher, englischer, französischer und russischer Sprache.

Die *Acta Physica* erscheinen in Heften wechselnden Umfanges. Mehrere Hefte bilden einen Band.

Die zur Veröffentlichung bestimmten Manuskripte sind an folgende Adresse zu richten:

Acta Physica, Budapest 502, Postafiók 24.

An die gleiche Anschrift ist auch jede für die Redaktion und den Verlag bestimmte Korrespondenz zu senden.

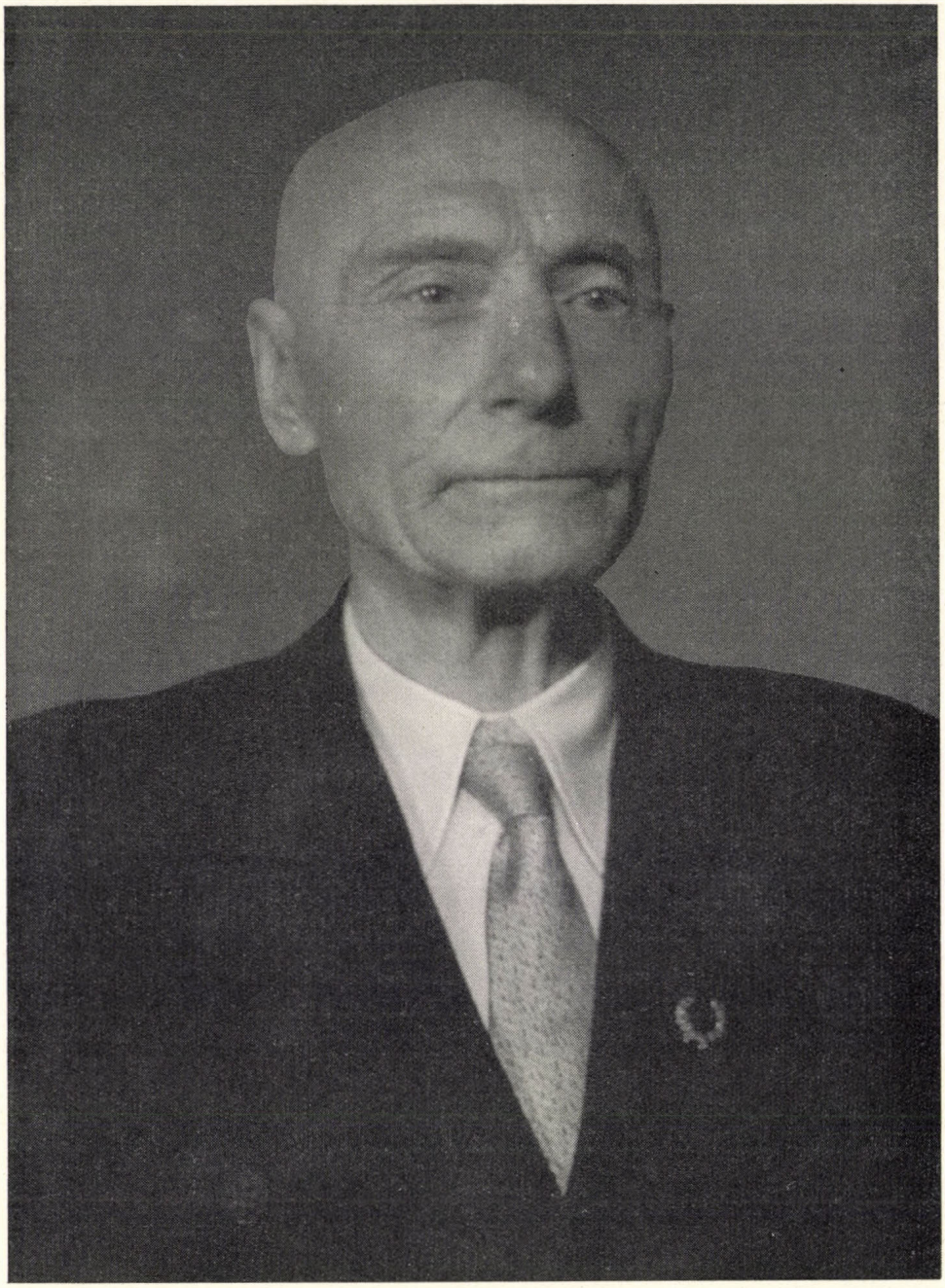
Abonnementspreis pro Band: 110 Forint. Bestellbar bei dem Buch- und Zeitungs-Aussenhandels-Unternehmen »Kultúra« (Budapest I., Fő u. 32. Bankkonto Nr. 43-790-057-181) oder bei seinen Auslandsvertretungen und Kommissionären.

ACTA PHYSICA

Tomus XVII

INDEX

CAROLO NOVOBÁTZKY
OCTOGENARIO



Prof. K. F. Novobátzky

K. F. NOVOBÁTZKY EIGHTY YEARS OLD

On the 3rd of March 1964 K. F. Novobátzky, Professor of Theoretical Physics, will celebrate his 80th birthday. On this occasion all Hungarian physicists as well as his friends and colleagues abroad express their sincere admiration and devotion and wish him further good health so that he may continue his fruitful work in the field of science.

K. F. Novobátzky was born in Temesvár in 1884. After the completion of his studies at the normal secondary school in Temesvár he continued his studies at the Department of Mathematics and Physics of the Faculty of Sciences of Budapest University, where at that time Roland Eötvös was the professor of experimental physics. After the end of the first world war, during which he served as an artillery officer, he taught for a while in the country, later in Budapest in secondary schools. After the end of the second world war in 1945 he was invited to occupy the vacant chair of the Department of Theoretical Physics of Budapest University. This post he has held ever since. In 1949 he has been elected a Member, and in 1958 Vice-President of the Hungarian Academy of Sciences.

The scientific activity of K. F. Novobátzky has always been closely related to the most interesting problems. As his particular field of research he chose two of the most modern subjects i.e. the physics of fields and the theory of relativity. In his scientific papers he always considers some fundamental problem, solving it in his characteristically simple and always original manner. When Novobátzky began his researches he had to rely completely on himself as in the theory of fields and relativity there was no tradition in Hungary.

In several publications Novobátzky dealt with the formulation of a unified field theory. After Albert Einstein had shown in the general theory of relativity that gravitation is closely connected with the geometrical structure of space, investigations attempting a geometrical interpretation of the electromagnetic field started.

The relevant papers of Novobátzky are characterized by the endeavour to avoid the introduction of the physically non-interpretable new dimensions and unobservable quantities. He has clearly seen that such investigations must not be restricted only to electromagnetic fields, but that they must now also be extended to the meson field accompanying the elementary particles.

In the early thirties research turned towards the quantum theory of electromagnetic radiation. Here the quantum mechanical treatment becomes rather complicated, because the field quantities are not independent of one another. It is known that the theory makes use of subsidiary conditions in order to eliminate the superfluous field components. Novobátzky rejected the subsidiary conditions which are physically meaningless and showed that the field equations themselves eliminate the superfluous components.

Dealing with the electromagnetic field Novobátzky also made important contributions. Based on the electromagnetic theory of light he gave a rigorous foundation to, and developed further, the Kirchhoff theory of light diffraction. The simple solution he gave to the old dispute in the electrodynamics of moving dielectrics, is characteristic of his grasp of the essential. The question of energy-momentum relations have been clarified by his use of the methods of the theory of relativity. By demonstrating the fact that electromagnetic stresses, which are present in vacuum, also accelerate mass, namely the inertial mass of the electromagnetic radiation, he was able to derive the equations of motion of radiant energy.

During the last ten years a great many papers have been published by the most eminent physicists concerning the interpretation of quantum theory. Novobátzky was among the first researchers to join these discussions. Making use of the variational principle he showed that one can arrive at the fundamental equation of quantum mechanics, the Schrödinger equation, in a purely mechanical way, without any recourse to optical or wave-theoretical analogies. The method proposed by Novobátzky touches not only questions of interpretation as is shown by the fact that he was able to carry out a consistent statistical treatment of the relativistic quantum theory.

Only some of the more important results of K. F. Novobátzky's scientific activities are reviewed here. The analysis and detailed enumeration of the whole of his work is not the task of this short introduction all the more so as we may hope for further contributions.

This introduction would not be complete if it did not remember K. F. Novobátzky, the teacher and pedagogue. Besides research work his long and active life

has been devoted to teaching. His pedagogic activity is highly esteemed by all his students.

His lectures at the University lead the students from classical physics to the most advanced problems of modern physics. He has written several text books and lecture-notes in order to improve university education. The introduction of several subjects in the curriculum of Hungarian university studies is due to him. He has gathered and educated a group of young and enthusiastic researchers; some of them now teach as professors and carry on the researches in the theory of fields and relativity initiated by him in Hungary.

This issue of Acta Physica Hungarica presenting the papers dedicated to him on his birthday by his students, coworkers, friends and colleagues is meant to serve as a tribute to Professor Novobátzky's services rendered to physics.

Budapest, 1964.

K. NAGY

THE EQUATIONS OF MOTION OF A RADIATING ELECTRON AND ITS LAGRANGIAN

By

L. INFELD

INSTITUTE FOR THEORETICAL PHYSICS, WARSAW UNIVERSITY, WARSAW, POLAND

(Received 23. X. 1963)

The equations of motion of an electron moving in an electromagnetic field are derived from a Lagrangian by taking into account the electron's own retarded field.

1. Introduction

We consider the motion of an electron in an external given field. If we restrict ourselves to the classical treatment and if we do not take into account the own field of the electron, then the equations of motion are simple. They can be easily derived from a Lagrangian. What happens, if we take into account the own field of an electron? Then, we have to distinguish between two possibilities. 1. The own field is a standing field, that is its potential is equal 1/2 advanced 1/2 retarded. Then the equations of motion are the same as if only the external field were present and there is no radiation. 2. The field of the electron can be derived from either an advanced or retarded potential. Then the equations of motion change considerably. We ask: could the more complicated equations of motion be derived from a Lagrangian?

We have asked before [1] a similar question concerning gravitational radiation and we found the Lagrangian giving us the radiative terms. The considerations here will be somewhat similar. Indeed, they can be still further generalized so as to include equations of motion for any field theory where radiation is present. Here, however, we shall restrict ourselves to the case of an electron in a given electromagnetic field. Since all our considerations will be Lorentz invariant we shall use both Latin and Greek indices, the Latin indices as summation indices, and the Greek as free indices.

2. General remarks on the relativistic Lagrangian

We assume the relativistic Lagrangian for a moving particle as given

$$L = L(x_\mu, x'_\mu, s). \tag{2.1}$$

Here

$$x'_\mu = \frac{dx_\mu}{ds}, \qquad dx_a\, dx_a = -ds^2 = \sum_0^3 (dx_a)^2, \tag{2.2}$$

$$x_0 = ict\,, \quad c = \text{velocity of light} = 1.$$

When we vary the integral

$$\int_{s_1}^{s_2} L\, ds \tag{2.3}$$

with respect to x_μ, where δx_μ vanishes at the end of the eigentime interval, we must not forget that the x'_μ are not arbitrary. Indeed, they fulfil the condition

$$x'_a\, x'_a + 1 = 0\,. \tag{2.4}$$

To obtain a proper Lagrangian in which the x'_μ can be treated as independent from each other we must modify L by adding to it the left side of the last equation multiplied by a factor, say $\lambda/2$,

$$L^* = L + \frac{\lambda}{2}(x'_e\, x'_e + 1) \tag{2.5}$$

and change (2.3) into

$$\delta \int_{s_1}^{s_2} L^*\, ds = 0\,, \tag{2.6}$$

where the variation has to be performed with respect to x_μ and λ. Thus our equations of motion are:

$$L_{,\mu} - (L_{,\mu'})' - (\lambda\, x'_\mu)' = 0\,, \tag{2.7}$$

where

$$L_{,\mu} = \frac{\partial L}{\partial x_\mu}: \qquad L_{,\mu'} = \frac{\partial L}{\partial x'_\mu}\,. \tag{2.7a}$$

Equations (2.4) and (2.7) determine x_μ and λ. Multiplying (2.7) by x_μ we find [2]:

$$x'_a\, L_{,a} - x'_a (L_{,a'})' + \lambda' = 0\,. \tag{2.8}$$

If L does not depend explicitly on s we can integrate (2.8) and obtain

$$\lambda = x'_a\, L_{,a'} = L + \text{constant} \tag{2.9}$$

which relation reminds us of the energy integral in classical mechanics.

As the most trivial case we take $L = 0$. Then $\lambda = \text{const} = m$ and the equations (2.7) become simply

$$m\, x''_\mu = 0 \tag{2.10}$$

which is the law of inertia. We, therefore, identify λ with the rest-mass of the particle. It is constant, if no external field is acting. In the more general case, if $\lambda = m$ depends on eigentime, $\lambda' = m'$ is the change in rest-mass and $\lambda' x'_\mu$ the change in the energy momentum.

3. The Lagrangian for a moving electron

We assume

$$L = e A_a x'_a . \tag{3.1}$$

Here e is the charge of an electron and A the four-potential of the electromagnetic field. Then calculating λ according to (2.8) we find

$$\lambda' = 0 \tag{3.2}$$

and according to our previous interpretation of λ

$$\lambda = m = \text{constant}. \tag{3.3}$$

For our equations of motion we find, because A_a is a function of x_a:

$$m x''_\nu = e(A_{a,\nu} - A_{\nu,a}) x'_a \tag{3.4}$$

or

$$m x''_\nu = e f_{\nu a} x'_a ; \quad f_{\nu\varrho} = A_{\varrho,\nu} - A_{\nu,\varrho} , \tag{3.5}$$

which are the well-known equations of motion for a charged particle if radiation is not taken into account. Indeed, in this case the change of the mechanical energy momentum vanishes. From the law of conservation of the moments we obtain the radiation vector P

$$P'_\nu = \lambda' x'_\nu = 0 . \tag{3.6}$$

We now ask: how to change the Lagrangian in order to get the proper equations of motion if the own field of the electron with the, say, retarded potential is taken into account. The equations of motion in this case, as it is well known, are:

$$m x''_\nu = e f_{\nu a} x'_a + \frac{2}{3} e^2 x'''_\nu - \frac{2}{3} e^2 x''_a x''_a x'_\nu . \tag{3.7}$$

Or, because

$$(x'_a x''_a)' = 0 = x''_a x''_a + x'_a x'''_a , \tag{3.8}$$

we can write (3.7) in the form

$$m x''_\nu = e f_{\nu a} x_a + \frac{2}{3} e^2 x'''_\nu + \frac{2}{3} e^2 x'_a x'''_a x'_\nu \tag{3.9}$$

which is entirely equivalent to (3.7).

The added two last expressions in (3.7) and (3.9) on the righthand side are small (we assume) compared with the first. Let us denote the solution of (3.5) by $\mathring{x}_\nu$, $\mathring{x}'_\nu$, $\mathring{x}''_\nu$. Then it is immaterial whether in the last two expressions we introduce $\mathring{x}$, $\mathring{x}'$, $\mathring{x}''$ or x, x', x''.

Therefore we can write instead of x''' in (3.7) $\mathring{x}'''$:

$$\mathring{x}'''_\nu = \frac{e}{m}(\mathring{f}_{\nu a})' x'_a + \frac{e}{m} f_{\nu a} \mathring{x}''_a = \frac{e}{m} \mathring{f}_{\nu a,k} \mathring{x}'_a \mathring{x}'_k + \frac{e}{m} f_{\nu a} \mathring{x}''_a \tag{3.10}$$

from which follows

$$m x''_\nu = e f_{\nu a} x'_a + \frac{2}{3} \frac{e^3}{m} \mathring{f}_{\nu a,b} \mathring{x}'_a \mathring{x}'_b + \frac{2}{3} \frac{e^3}{m} \mathring{f}_{\nu a} \mathring{x}''_a - \frac{2}{3} e^2 x''_a x''_a x'_\nu. \tag{3.11}$$

It is again immaterial whether in the last three expressions in (3.11) we put $\mathring{x}$, $\mathring{x}'$, $\mathring{x}''$ or x, x', x''. Equation (3.11) has the advantage over (3.7) and (3.8) that it does not contain higher derivatives with respect to s than the second. Thus our problem is to find a Lagrangian giving us (3.11) when in the last three expressions the $\mathring{x}$, $\mathring{x}'$, $\mathring{x}''$ and x, x', x'' can be interchanged.

4. The Lagrangian leading to (3.11)

We saw in Section 3 that

$$L = e A_a x'_a \tag{4.1}$$

was the right Lagrangian when no radiation was present. In the case of radiation the coordinates will be somehow in the neighbourhood of the coordinates, when radiation is not present. We assume therefore for the case of retarded potential

$$L(x + \varepsilon \alpha). \tag{4.2}$$

Here the α are *known* functions of s, to be determined at the end of our calculation and ε is a small constant. Neglecting consistently all expressions proportional to ε^2 we may write instead of (4.2)

$$L(x) + \varepsilon \alpha_a L_{,a} + \varepsilon \alpha'_a L_{,a'} = e A_a x'_a + \varepsilon e A_{b,a} x'_b \alpha_a + \varepsilon \alpha'_a e A_a. \tag{4.3}$$

We subtract and add

$$\alpha_a A'_a = \alpha_a A_{a,b} x'_b, \tag{4.4}$$

then we have

$$\mathscr{L}(x) = L(x + \varepsilon \alpha) = e A_a x'_a + e \varepsilon f_{ab} \alpha_a x'_b + (\alpha_a A_a)'. \tag{4.5}$$

Assuming that the δa, and their derivatives vanish at the end of the eigentime interval (s_1, s_2), this is equivalent to

$$\mathscr{L}(x) = eA_a x'_a + \varepsilon e f_{ab} a_a x'_b = L + \Delta L; \quad \Delta L = e\varepsilon a_a f_{ab} x'_b. \tag{4.6}$$

We find the equations of motion by calculating

$$(\Delta L)_{,\nu} - [(\Delta L)_{,\nu'}]' = \varepsilon a_a e f_{ab,\nu} x'_b - \varepsilon (a_a e f_{a\nu})' =$$

$$= \varepsilon a_a [e f_{ab,\nu} + e f_{\nu a,b}] x'_b - \varepsilon a'_a e f_{a\nu} = \varepsilon a_a e f_{\nu b,a} x'_b + \varepsilon a'_a e f_{\nu a}. \tag{4.7}$$

We find, because of (2.8):

$$\lambda' = m' = \varepsilon a'_a e f_{ab} x'_b. \tag{4.8}$$

Therefore our equations of motion are

$$m x''_\nu = e f_{\nu a} x'_a + \varepsilon a_a e f_{\nu b,a} x'_b + \varepsilon a'_a e f_{\nu a} - \varepsilon a'_a e f_{ab} x'_b x'_\nu. \tag{4.9}$$

We see that (4.9) and (3.11) are identical up to the order of ε^2 if we replace εa by

$$\varepsilon a_\mu = \frac{2}{3} \frac{e^2}{m} \overset{\circ}{x}'_\mu. \tag{4.10}$$

Thus as our right Lagrangian we find:

$$\mathscr{L} = eA_a x_a + \frac{2}{3} \frac{e^3}{m} f_{ab} \overset{\circ}{x}'_a x'_b. \tag{4.11}$$

Here the $\overset{\circ}{x}'$ have to be treated as known functions. They are gained by solving (3.5). The Lagrangian (4.11) gives us the right equations of motion for a radiating particle. However, m, is not constant and m' is different from zero.

5. The Lagrangian leading to (3.9)

Since the equations of motion contain third derivatives with respect to eigentime we must consider a Lagrangian with second derivatives of x_μ and assume that δx as well as $\delta x'$ vanishes at the end of the interval (s_1, s_2) If L is a function of x_ν, x'_ν, x''_ν then the equations of motion are:

$$L_{,\nu} - (L_{,\nu'})' + (L_{,\nu''})'' - (\lambda x_\nu)' = 0. \tag{5.1}$$

Equation (4.6) suggest for

$$\mathscr{L}(x) = eAa x'_a + \varepsilon a_a x''_a m. \tag{5.2}$$

Therefore the equations of motion are

$$mx''_\nu = ef_{\nu a}\, x'_a + \varepsilon m a''_\nu + \varepsilon m a''_a\, x'_a\, x'_\nu. \tag{5.3}$$

Comparing this with (3.9) we find again:

$$ma''_\nu = \frac{2}{3}\, e^2\, \mathring{x}'''_\nu. \tag{5.4}$$

Therefore finally the equations of motion following from (5.3) and (5.4) are

$$mx''_\nu = ef_{\nu a}\, x'_a + \frac{2}{3}\, e^2\, \mathring{x}'''_\nu + \frac{2}{3}\, \mathring{x}'''_a\, x'_a\, x'_\nu. \tag{5.5}$$

λ' obtained by multiplication of (5.1) with x'_ν is

$$\lambda' = -\varepsilon m\, a''_a\, x'_a.$$

But there is one essential difference between (5.5) and (3.9). Equations (3.9) are ordinary differential equations of the third order while (5.5) is of the second order. Therefore some of the known difficulties connected with (3.9) are not present in (5.5) where $\mathring{x}'''$ is a well-known function defined through equations of motion, where radiation is absent [3].

6. The transition of equations without radiation to those with it

Let us go back to the considerations of Section 4. We took as one Lagrangian for motion without radiation

$$L(\mathring{x}, \mathring{x}') = \mathring{A}_a\, \mathring{x}'_a \tag{6.1}$$

and as the Lagrangian with radiation

$$L(x + \varepsilon a, x' + \varepsilon a') \approx eA_a\, x'_a + e\varepsilon a_a f_{ab}\, x'_b, \tag{6.2}$$

where

$$\varepsilon a_\nu = \frac{2}{3}\, \frac{e^2}{m}\, \mathring{x}'_\nu. \tag{6.3}$$

The small circles above x, x', A, L refer to the motion without radiation. The same functions and coordinates without the circles to the equations of motion with radiation. The eigentime s *is* the same in both these cases, since the change of

$$\mathring{x} \quad \text{to} \quad x + \varepsilon a \tag{6.4}$$

results in

$$\mathring{x}'_a \mathring{x}'_a \text{ into } (\mathring{x}'_a + \varepsilon a'_a)(\mathring{x}'_a + \varepsilon a'_a) = x'_a x'_a, \text{ since } x'_a a'_a \sim x'_a x''_a = 0. \quad (6.5)$$

Therefore the transition from a Lagrangian without radiation to that with one can be characterized by the following transition:

$$\mathring{x} \to x + a, \quad (6.6)$$

$$\mathring{\lambda}\, \mathring{x}'' = \mathring{m}\, \mathring{x}'' \to m x'' = \lambda x'',$$

$$0 = \mathring{\lambda}' \to m' = \lambda' = -\varepsilon m\, a''_a\, x'_a = \varepsilon m\, a'_a\, x''_a =$$

$$= \frac{2}{3} e^2 \mathring{x}''_a\, x''_a = \frac{2}{3} \frac{e^4}{m^2} (f_{ab}\, x'_b)^2.$$

Remembering the definition of L^* in (2.5) we have as the definition of the energy-momentum four-vector:

$$\mathring{P}_\mu = \mathring{L}^*_{,\mu'} = e \mathring{A}_\mu + \mathring{m}\, \mathring{x}'_\mu, \quad (6.7)$$

$$P_\mu = L^*_{,\mu'} = eA_\mu + \lambda x'_\mu - \frac{2}{3} \frac{e^3}{m} \mathring{x}'_a f_{\mu a} = eA_\mu + m x'_\mu - \frac{2}{3} e^2 x''_\mu.$$

Therefore

$$P'_\mu = eA'_\mu + \left(m x_\mu - \frac{2}{3} e^2 x'_\mu\right)'' + m' x'_\mu. \quad (6.8)$$

The last expression in (6.8) is responsible for the radiation. Therefore denoting the radiation four-vector by ΔP_μ we have

$$\Delta P'_\mu = m' x'_\mu = \frac{2}{3} \frac{e^4}{m^2} (f_{ab}\, x'_b)^2\, x'_\mu = \frac{2}{3} e^2 x''_a\, x''_a\, x'_\mu. \quad (6.9)$$

To interpret the result physically let us introduce c as the velocity of light. The transition

$$\mathring{x} \to x + \varepsilon a = x + \frac{2}{3} \frac{e^2}{\mathring{m} e^2} \mathring{x}', \quad x_0 = \text{ict}; \quad ds^2 = -dx_a\, dx_a. \quad (6.10)$$

We denote by

$$r = \frac{2}{3} \frac{e^2}{mc^2}, \quad \text{then} \quad \mathring{x} \to x + r\mathring{x}', \quad (6.11)$$

where r has the linear dimensions of the particle. Therefore, since

$$mc^2 \mathring{x}''_\nu = e \mathring{f}_{\nu a} \mathring{x}'_a, \quad (6.12)$$

we have

$$\Delta P'_\mu = \frac{2}{3} rmc\, x''_a x''_a x'_\mu = \frac{3}{2} \frac{1}{c} r^2 (f_{ab} x'_b)^2 x'_\mu \tag{6.13}$$

or in the usual three-dimensional notation, where the dot denotes the differentiation with respect to time and

$$f_{12},\ f_{13},\ f_{23} \to H_z,\ -H_y,\ H_x,\ \vec{v} \to \dot{x},\ \dot{y},\ \dot{z}, \tag{6.14}$$

$$f_{01},\ f_{02},\ f_{03} \to iE_x,\ iE_y,\ iE_z,$$

we have

$$c\left(1 - \frac{v^2}{c^2}\right)^{3/2} \Delta P'_0 = \frac{i}{e}\left(1 - \frac{v^2}{c^2}\right) E = \frac{3}{2} ir^2 \left\{\left[\bar{E} + \frac{1}{c}(\bar{v}\, x \bar{H})\right]^2 - \frac{1}{c^2}(\bar{v} \cdot \bar{E})^2\right\},$$

$$c\left(1 - \frac{v^2}{c^2}\right)^{3/2} \Delta \bar{P}' = \left(1 - \frac{v^2}{c^2}\right) \bar{P} = \frac{3}{2c} r^2 \left\{\left[\bar{E} + \frac{1}{c}(\bar{v}\, x \bar{H})\right]^2 - \frac{1}{c^2}(\bar{v} \cdot \bar{E})^2\right\} \vec{v}, \tag{6.15}$$

which is the accepted expression for radiation if by r we denote $2/3\ e^2/mc^2$ and by E the energy.

REFERENCES

1. L. Infeld, Bull. Acad. Polon. Sci., XI, 399, 1963.
2. L. Infeld, Planck's Festschrift "On Variational Principles in Relativistic Dynamics" VEB Verlag der Deutsch. Akad. der Wissenschaften, Berlin, p. 115, 1958.
3. C. J. Eliezer, Rev. Mod. Phys., "The Interactions of Electrons and Electromagnetic Field", **19**, 147, 1947.

УРАВНЕНИЕ ДВИЖЕНИЯ ИЗЛУЧАЮЩЕГО ЭЛЕКТРОНА И ЕГО ЛАГРАНЖИАН

Л. ИНФЕЛЬД

Р е з ю м е

При помощи функции Лагранжа выводится уравнение движения электрона, движущегося в электромагнитном поле, принимая во внимание собственное ретардированное поле электрона.

BEITRAG ZUR RELATIVISTISCHEN KINEMATIK EINES STARREN PUNKTSYSTEMS

Von

T. MÁTRAI

PHYSIKALISCHER LEHRSTUHL DER PÄDAGOGISCHEN HOCHSCHULE, EGER

(Vorgelegt von Z. Gyulai. — Eingegangen: 20. IX. 1963)

In einer vorläufigen Mitteilung [1] gab der Verfasser bereits eine auf M. BORN's Grundgedanken [2] beruhende Ableitung der Lorentz-invarianten Bedingungen für die erzwungene Bewegung eines sich in endlicher Entfernung voneinander befindenden starr verbundenen Punktepaares an. Es wurde hier untersucht, in welcher Weise die klassisch-kinematischen Eigenschaften des starren Körpers durch das Relativitätsprinzip modifiziert werden, wenn der starre Körper nicht als Medium, sondern als aus Paaren von Massenpunkten aufgebaut aufgefasst wird. Es gelang, auf elementare Weise zu zeigen, dass das aus allen möglichen Verbindungen zwischen vier Punkten bestehende starre System im allgemeinen sechs und nicht drei Freiheitsgrade besitzt, wie dies auf Grund der HERGLOTZschen Überlegungen [3] aus den sich auf Medien beziehenden BORNschen Bedingungen der erzwungenen Bewegung zu erwarten wäre. Mehr als vier Punkte jedoch können sich im allgemeinen nicht so bewegen, dass jede Verbindung zwischen ihnen starr bleibt, ausgenommen eben gerade den speziellen Fall der sogenannten BORNschen Bewegung. Es gelang weiter zu zeigen, dass in dem zuerst von BORN beschriebenen starren Medium, dessen Punkte relativistische Hyperbelbewegungen ausführen, notwendigerweise euklidische innere Raummetrik herrscht.

§ 1. Lorentz-invariante Bedingungen der erzwungenen Bewegung eines starr verbundenen Punktepaares, und die Anzahl der Freiheitsgrade eines relativistischen starr verbundenen Punktepaares

Zur Bestimmung des Masstensors mittels materieller Masstäbe und natürlicher Uhren muss man jene allgemeine kovarianten Bedingungen für die erzwungene Bewegung kennen, welchen die Endpunkte des Masstabes, d. h. einfach ein starr verbundenes Punktpaar, bei dem mit der Längenmessung verbundenen Transport genügen müssen. HERGLOTZ weist jedoch in dem obenerwähnten Paradoxon [3] darauf hin, dass die Bornschen Lorentz-invarianten Gleichungen höchstens hinreichende Bedingungen für die starre Bewegung des Mediums darstellen. Zur Bestimmung der zugleich notwendigen Bedingungen ist es am zweckmässigsten, die Untersuchungen zuerst auf das einfachste Punktsystem, nämlich das eines Punktepaares, zu beschränken.

BORN's Bedingungen für die Starrheit vereinfachen sich im Falle eines Punktepaares eindeutig folgendermassen: *Die Bewegung eines Punktes* P_2 $(\mathfrak{r}_2 = \mathfrak{r}_2(t))$ *kann in Bezug auf einen Punkt* P_1 $(\mathfrak{r}_1 = \mathfrak{r}_1(t))$ *als starr angesehen werden, wenn der von jedem beliebigen, auf der Weltlinie von* P_2 *gelegenen Punkt* $\mathfrak{r}_2 = \mathfrak{r}_2(t_2)$ *mit der (Koordinaten-) Zeit* t_2 *senkrecht auf die Weltlinie des Punktes*

P_1 gezogene Abstand s_{12} eine reelle nichtnegative Konstante ist, das heisst:

$$\frac{\mathrm{d}\mathfrak{r}_1(t_1)}{\mathrm{d}s_1}[\mathfrak{r}_1(t_1)-\mathfrak{r}_2(t_2)]-c^2\frac{\mathrm{d}t_1}{\mathrm{d}s_1}[t_1-t_2]=0, \tag{1,1}$$

$$[\mathfrak{r}_1(t_1)-\mathfrak{r}_2(t_2)]^2-c^2[t_1-t_2]^2=s_{12}^2. \tag{1,2}$$

In diesem Lorentz-invarianten Funktional-Gleichungssystem bedeutet s_1 den Parameter der Bogenlänge der Weltlinie des Punktes P_1 bei jenem Wert $t_1=t_1(t_2)$ der Koordinatenzeit, der durch die sog. *Orthogonalitätsbedingung* (1,1) einem gegebenen Wert von t_2 »zugeordnet« wird. Die Gleichung (1, 2) kann »*sphärische*« *Bedingung* genannt werden.

In der Relativitätstheorie bezieht sich also die Definition der Starrheit nicht auf die Raumlage zweier Punkte mit gleicher Koordinatenzeit.

Die Gleichung (1,1) kann bei Berücksichtigung der Bedingung $\dot{\mathfrak{r}}^2<c^2$ auch in folgender einfachen (obwohl scheinbar nicht invarianten) Form aufgeschrieben werden:

$$\frac{\mathrm{d}\mathfrak{r}_1(t_1)}{\mathrm{d}t_1}[\mathfrak{r}_1(t_1)-\mathfrak{r}_2(t_2)]-c^2[t_1-t_2]=0. \tag{1,3}$$

Wegen der Einwertigkeit der Funktion $\mathfrak{r}_j=\mathfrak{r}_j(t)$, $j=1,2$ kann $\mathrm{d}t_2/\mathrm{d}t_1$ nicht negativ sein. Mit dieser Bedingung lässt sich die Symmetrie der für j asymmetrisch erscheinenden Bedingungen (1,1) und (1,2) leicht zeigen. Bildet man die Ableitung der Gl. (1,2) nach t_1 und dividiert man die so erhaltene Gleichung durch $\mathrm{d}t_2/\mathrm{d}t_1$, so erhält man unter Berücksichtigung von (1,3) die folgende Gleichung:

$$\frac{\mathrm{d}\mathfrak{r}_2(t_2)}{\mathrm{d}t_2}[\mathfrak{r}_2(t_2)-\mathfrak{r}_1(t_1)]-c^2[t_2-t_1]=0. \tag{1,4}$$

Diese Gleichung zusammen mit der »sphärischen« Gleichung (1,2) drückt dagegen die auf den Punkt P_2 bezogene Starrheit des Punktes P_1 aus. Ist also P_2 in bezug auf P_1 starr, so ist auch umgekehrt P_1 in bezug auf P_2 starr, und da die Starrheit wechselseitig ist, kann man einfach von einem sich starr bewegenden Punktepaar sprechen.

Durch Subtraktion von (1,4) von (1,3) erhält man die symmetrische Gleichung

$$[\mathfrak{r}_1(t_1)-\mathfrak{r}_2(t_2)][\dot{\mathfrak{r}}_1(t_1)-\dot{\mathfrak{r}}_2(t_2)]=0, \tag{1,5}$$

die einfach die relativistische Verallgemeinerung der klassisch-kinematischen Gleichung der erzwungenen Bewegung eines starren Punktepaares P_1, P_2 darstellt.

So wie in der klassischen Kinematik besitzt auch relativistisch ein *starres Punktepaar* fünf Freiheitsgrade. Nimmt man nämlich fünf Raumkoordinaten des Punktpaares, u. zw. alle drei Raumkoordinaten des Punktes P_1, jedoch nur zwei des Punktes P_2 als Funktionen der (Koordinaten-) Zeit als gegeben an:

$$\begin{gathered} x_1 = x_1(t), \quad y_1 = y_1(t), \quad z_1 = z_1(t), \\ x_2 = x_2(t), \quad y_2 = y_2(t), \end{gathered} \tag{1,6}$$

so lässt sich die unbekannte Funktion $z_2 = z_2(t)$ aus dem Gleichungssystem (1,2), (1,3) mit Hilfe von (1,6) folgenderweise bestimmen. Mit Hilfe von (1,3) kann die $t = t_1$ »zugeordnete« Koordinate $z_2(t_2)$ folgendermassen ausgedrückt werden:

$$\begin{aligned} z_2(t_2) = z_1(t_1) + [\dot{x}_1(t_1)(x_1(t_1) - x_2(t_2)) + \dot{y}_1(t_1)(y_1(t_1) - y_2(t_2)) - \\ - c^2(t_1 - t_2)]\dot{z}_1(t_1). \end{aligned} \tag{1,7}$$

Setzt man diesen Ausdruck für $z_2(t_2)$ in (1,2) ein, so sind in der so erhaltenen Gleichung alle Koordinaten bereits bekannte Funktionen von t_1 bzw. von t_2. Setzt man die Existenz einer und nur einer reellen Wurzel t_1 voraus, so lässt sich aus dieser Gleichung wenigstens im Prinzip (falls notwendig auch mit Hilfe einer geeigneten Näherungsmethode) jener Wert der Koordinatenzeit t_1 des Punktes P_1 ermitteln, dem eben gerade der Wert t_2 der Koordinatenzeit des Punktes P_2 »entspricht«, also die Funktion

$$t_1 = t_1\{t_2\}. \tag{1,8}$$

Setzt man dies nun wieder in (1,7) ein, so erhält man $z_2 = z_2(t_2)$ bereits in der gewünschten Form. Dieses Verfahren* überzeugt gleichzeitig auch davon, dass zur Bestimmung von $z_2(t)$ die Angabe der fünf Koordinaten, die in (1,6) vorkommen, nicht nur hinreichend, sondern auch notwendig ist. Die obige Behauptung bezüglich der relativistischen Kinematik daher als bewiesen betrachtet werden kann.

Da für $\dot{\mathfrak{r}}_j = 0$ (sowohl für $j = 1$ als auch für $j = 2$) $t_1 = t_2$ ist, muss *die invariante Grösse s_{12} den Abstand d_{12} des Punktepaares voneinander bedeuten.* K. NOVOBÁTZKY wies auf eine äquivalente, besonders einfache (symmetrische) Definition der relativistisch starren Bewegung eines Punktepaares hin. *Seiner*

* Hier und im folgenden soll stets versucht werden, die Lösung der sich ergebenden Funktionalgleichungssysteme auf die Bestimmung der Lösung reeller Gleichungen zurückzuführen, wobei die Existenz von reellen und einfachen Wurzeln stets vorausgesetzt wird. Auf die Untersuchung der Existenzbedingungen solcher Wurzeln kann hier nicht eingegangen werden.

Definition nach ist ein Punktepaar immer dann als starr zu betrachten, wenn in jedem Augenblick ein Inertialsystem existiert, in dem die klassische Bedingung der Starrheit erfüllt ist, d. h. in dem der Abstand der zwei Punkten voneinander zeitlich konstant ist. (Ein solches Koordinatensystem ist z. B. ein System, in dem einer der Punkte des Punktepaares ruht.)

Im Hinblick auf weiter unten folgende Ausführungen sei hier bemerkt, dass die zehn Abstände d_{jk} zwischen fünf im euklidischen Raum ruhenden Punkten nicht unabhängig voneinander sind, sondern dass sie die folgende Determinantengleichung stets befriedigen:

$$\mathrm{Det.}(d_{1j}^2 + d_{1k}^2 - d_{jk}^2) \equiv 0, \quad j, k = 2, 3, 4, 5. \tag{1.9}$$

Dieser Satz kann auch umgekehrt werden: besteht (1,9) für beliebige fünf Punkte eines ruhenden Mediums, so kann in diesem Medium die Metrik nur euklidisch sein.

§ 2. Untersuchung einiger kinematischen Beispiele

I. *Bewegen sich zwei Punkte P_1, P_2 in Bezug aufeinander auf einer in einem Inertialsystem ruhenden Geraden mit gleicher und konstanter Geschwindigkeit $\mathfrak{v}$, so ist die Bewegung der beiden Punkte starr*, und ihre Raumvektoren befriedigen die Gl. (1,2) und (1,3). Während dieser Bewegung zeigt nun *das starre Punktepaar genau die Lorentz—Fitzgeraldsche Kontraktion*, also

$$s_{12}^2 = d_{12}^2/(1 - \mathfrak{v}^2/c^2), \tag{2,1}$$

wobei jetzt d_{12} den Abstand jener zwei Punkte P_1 und P_2 des Inertialsystems bedeutet, die zu derselben Koordinatenzeit gehören.

II. *Zwei Punkte, die sich mit konstanter und gleicher Beschleunigung kollinear bewegen, können relativ zueinander nicht starr sein.*

Bewegt sich nämlich der Punkt P_1 mit konstanter Beschleunigung a entlang der x-Achse gemäss der Bahngleichung $x_1 = at^2/2$, so kann die Bahn $x_2(t_2)$ und damit auch die Beschleunigung a_2 des sich auf derselben Achse bewegenden und in Bezug auf P_1 starren Punktes P_2 mit dem im Zusammenhang mit den Gln. (1,6)—(1,8) beschriebenen Verfahren auf elementare Weise berechnet werden. Die Rechnung ergibt

$$a_2 = a \cdot \frac{b + s_{12} \cdot a\,(1 - 2f - 2f^2)/c^2}{b + s_{12} \cdot a(1 - f)/c^2},$$

wobei

$$f = a^2t^2/c^2, \qquad b = \sqrt{1 - f}^{\,5}.$$

Laut diesem Ausdruck ist für keinen reellen Wert von t_1 $a_2 = a$, das bedeutet, der in Bezug auf den Punkt P_1 starre Punkt P_2 kann nicht die gleiche Beschleunigung haben wie P_1. Damit ist die obige Behauptung bewiesen.

Die obige Aussage führt weiter zu der folgenden: *Im Gegensatz zu der klassischen Kinematik, kann kein starres Koordinatensystem konstruiert werden, in dem jeder beliebige Punkt des Systems sich in Bezug auf das Inertialsystem mit konstanter und gleicher Beschleunigung bewegt.*

III. *Das starre Medium mit Hyperbelbewegung und seine innere Metrik*

Der Punkt P_1 bewege sich auf einer relativistischen Hyperbelbahn, d.h. er bewege sich z. B. entlang der x-Achse des Inertialsystems in der Weise, dass seine in dem augenblicklichen Koordinatensystem gemessene Beschleunigung stets den konstanten Wert a habe [4]. Von einer additiven Konstante abgesehen entspricht dieser Bedingung die folgende Bahngleichung:

$$x_1 = \sqrt{\frac{c^4}{a^2} + c^2 \cdot t^2}, \quad y_1 = z_1 = 0. \tag{2,2}$$

(Die Bezeichnung relativistische Hyperbelbewegung stammt bekanntlich daher, dass Gl. (2,2) in der (ict, x)-Ebene eine Welthyperbel darstellt, deren reelle Halbachse c^2/a und deren imaginäre Halbachse c/a ist.)

Berechnen wir nun gleichfalls unter Anwendung der im Zusammenhang mit den Gln. (1,7)—(1,8) angegebenen Methode die Bahngleichung des sich entlang der x-Achse bewegenden und in Bezug auf den Punkt P_1 starren Punktes P_2. Der Abstand zwischen den Punkten P_1 und P_2 sei s_{12}. Die elementare Rechnung ergibt

$$x_2 = \sqrt{\frac{c^4}{a^2}\left(1 \pm \frac{as_{12}}{c^2}\right)^2 + c^2 t^2}, \qquad y_2 = z_2 = 0. \tag{2,3}$$

Es zeigt sich in Übereinstimmung mit der Feststellung von M. BORN, dass auch Punkt P_2 eine Hyperbelbewegung ausführt, die reelle Halbachse der Hyperbel ist jedoch nicht c^2/a, sondern $c^2\left(1 \pm \frac{as_{12}}{c_2}\right)/a$. Die zwei Werte bedeuten, dass Punkt P_2 sich entweder vor oder hinter dem Punkt P_1 bewegen kann.

Es soll nun weiter untersucht werden, ob die Bewegung des sich in der (x, y)-Ebene befindenden Punktes P_3, dessen Koordinaten

$$x_3 = x_2, \quad y_3 = d, \quad z_3 = 0 \tag{2,4}$$

sind — wobei der Parameter d zeitlich konstant ist — relativ zu dem sich entlang der x-Achse bewegenden Punkt P_1 als starr zu betrachten ist.

Die linke Seite der für die Punkte P_1 und P_3 aufgeschriebenen sphärischen Gleichung (1,2) enthält, wenn man probeweise $t_3 = t_2$ setzt, im Ver-

gleich zu dem Fall der (kollinearen) Bewegung von P_1 und P_2 das zusätzliche Glied d^2. Folglich ist $s_{13}^2 = s_{12}^2 + d^2$ konstant und erfüllt die sphärische Gleichung (2,4). Ausserdem wird aber auch die Orthogonalitätsgleichung (1,4) befriedigt: Die Vermutung $t_3 = t_2$ bewährte sich ja bereits bei der Erfüllung der sphärischen Gleichung, und da ausserdem $\dot{y}_3 = \dot{z}_3 = 0$, so drückt die Orthogonalitätsgleichung eben gerade die Starrheit der Punkte P_1 und P_2 in Bezug aufeinander aus.

Die Parameter s_{12} und d können als die Zylinderkoordinaten derjenigen relativistischen Hyperbel um die x-Achse aufgefasst werden, auf der sich die Punkte des Mediums bewegen. Gleichzeitig ist auch offensichtlich, dass zwei beliebige Punkte des Mediums, also auch jedes aus beliebigen fünf Punkten gebildete Punktepaar, als starr zu betrachten ist.

Aus (2,2) und (2,4) ergibt sich ferner, dass für $t = 0$ die Geschwindigkeit aller fünf Punkte Null ist. Zu dieser Zeit müssen jedoch sämtliche zwischen den fünf Punkten bestehenden Eigenabstände s_{jk} mit den Raumabständen der im Inertialsystem ruhenden Punkte übereinstimmen. Für diese muss hingegen die Beziehung (1,6) bestehen, woraus folgt, dass die in der Zeit konstanten Eigenabstände s_{jk} auch (1,6) erfüllen müssen. Dies bedeutet, dass *in dem starren Körper, dessen Punkte relativistische Hyperbelbewegungen ausführen, auch euklidische innere Metrik herrscht.* Dieses Ergebnis ist deswegen bemerkenswert, da es bedeutet, dass *in einem beschleunigten starren Kasten die Geometrie durch die dort auftretende (Inertial-) Kraft nicht geändert wird.*

IV. *Das gleichförmig rotierende starre Medium und seine innere Metrik*

Untersuchen wir nun die Bewegung einer gleichförmig rotierenden Scheibe, d. h. die Bewegung eines Mediums, für das die Koordinaten irgendeines seiner Punkte durch die Gleichungen

$$\begin{aligned} x &= r\cos(\omega t + \alpha), \\ y &= r\sin(\omega t + \alpha), \\ z &= 0 \end{aligned} \tag{2,5}$$

gegeben sind. Die Parameter r und α können als ebene Polarkoordinaten angesehen werden. Sie sowie die Winkelgeschwindigkeit ω hängen hier von der Koordinatenzeit t nicht ab.

Die durch die Konstanten r_j, α_j $(j = 1, 2)$ charakterisierten beliebigen zwei Punkte der Scheibe sind starr. Setzt man nämlich die durch Gln. (2,5) gegebenen Koordinaten in (1,3) ein, so kann die so erhaltene Gleichung

$$\frac{\omega^2 r_1 r_2}{c^2}\sin[\omega(t_1 - t_2) + (\alpha_1 - \alpha_2)] = \omega(t_1 - t_2) \tag{2,6}$$

nur dann bestehen, falls $t_1 - t_2$ in der Zeit konstant ist. (Diese Konstante kann aber nur dann reell sein, wenn $\omega^2 r_1 r_2/c^2 \leqslant 1$.) Deshalb ist der sich aus (1,2) ergebende Ausdruck

$$s_{12}^2 = r_1^2 + r_2^2 - 2\, r_1 r_1 \cos[\omega(t_1 - t_2) + (a_1 - a_2)] - c^2(t_1 - t_2)^2 \qquad (2,7)$$

tatsächlich konstant, und sämtliche Bedingungen der Starrheit sind also befriedigt. Beliebige zwei Punkte einer gleichförmig rotierenden Kreisscheibe sind also in Bezug auf einander, d. h. auch in der Bornschen Interpretation starr.

Für $r_2 = 0$ und $r_1 \equiv r$ ist $s_{12} \equiv s = r$, daher gilt $s\omega = v$, d. h. *für eine gleichmässig rotierende Kreisscheibe besteht ein streng lineares »Geschwindigkeit-Abstandsgesetz«*, in Übereinstimmung mit dem Resultat von C. W. BERENDA [5] und N. ROSEN [6], jedoch im Gegensatz zu E. L. HILL [7]. Letzterer gab jedoch für die starre Bewegung eine von der obigen abweichende Definition.

Die Lorentz-invariante Definition der starren Bewegung eines Punktepaares bietet gleichzeitig eine einheitliche Methode auch zur Bestimmung der inneren Metrik eines im Bornschen Sinne starren Mediums. Diese Methode soll hier bei der Berechnung des zu den in Polarkoordinaten gegebenen Differentialen dr und da gehörenden (in dem betrachteten Falle zweidimensionalen ebenen) Bogenelements ds einer Scheibe angewandt werden. Für $a_1 - a_2 \equiv$ $\equiv da$ ergibt sich aus (2,6) $t_1 - t_2 \equiv dt = \omega r_1 r_2\, da/(c^2 - \omega^2 r_1 r_2)$. Setzt man diesen Ausdruck für dt in (2,7) ein und benutzt man für den Fall $r_1 - r_2 \equiv dr \to 0$ die Bezeichnung $r_1 = r$, so ergibt sich für $s_{12}^2 \equiv (ds)^2$ der folgende ganz offensichtlich nicht-pythagoreische Ausdruck

$$(ds)^2 = (dr)^2 + r^2(da)^2/(1 - \beta^2); \quad \beta = v/c, \qquad (2,8)$$

wiederum in vollkommener Übereinstimmung mit BERENDA und ROSEN, jedoch im Gegensatz zu A. S. EDDINGTON [8] und H. A. LORENTZ [9], die im Falle der Scheibe aus der allgemeinen Relativitätstheorie auf eine euklidische innere Metrik gefolgert haben.

§ 3. Die Bewegung einer durch ein starres Punktepaar bestimmten starren Geraden

Die Bewegung zweier nicht zusammenfallender Punkte P_1 und P_2 relativ zu einander sei starr, d. h. die Weltkoordinaten dieser Punkte sollen die Gln. (1,2) und (1,3) erfüllen. Die Koordinaten des Punktes P_3 seien durch die folgenden Gleichungen gegeben:

$$\begin{aligned} t_3 &= t_1 \pm (t_1 - t_2) s_{13}/s_{12}, \\ \mathfrak{r}_3(t_3) &= \mathfrak{r}_1(t_1) \pm [\mathfrak{r}_1(t_1) - \mathfrak{r}_2(t_2)] s_{13}/s_{12}, \end{aligned} \qquad (3,1)$$

wobei s_{13} ($\geqslant 0$) eine reelle Konstante bedeuten soll.

Unter diesen Bedingungen — wie dies auf elementare Weise ersichtlich ist — führen sowohl die Punkte P_1 und P_3 als auch die Punkte P_2 und P_3 relativ zu einander starre Bewegungen aus, weiterhin ist

$$s_{23}^2 = (s_{12} \pm s_{13})^2 , \tag{3,2}$$

Auf Grund dieser Eigenschaften kann gesagt werden, dass *der Punkt P_3 auf einer durch das Punktepaar P_1, P_2 bestimmten starren Geraden liegt.*

Gleichzeitig sieht man auch, dass zwischen den auf einer Geraden liegenden Punkten die »Zuordnung« der Koordinatenzeit transitiv ist.

Auf Grund der obigen Feststellungen ist es leicht einzusehen, dass die Bedingungen (1,2) und (1,3) die folgenden klassisch-kinematischen Sätze, von denen die ersten zwei gleichzeitig auch die mathematischen Vorbedingungen für die Längenmessung sind, unverändert lassen.

A) *Die beliebig gewählten Punkte P_3 und P_4 einer durch das starre Punktepaar P_1, P_2 bestimmten starren Geraden bilden ebenfalls ein starres Punktepaar.*

B) Umgekehrt: *Punkt P_1 ist dabei auch ein Punkt der starren Geraden, die durch das nicht zusammenfallende starre Punktepaar P_3, P_4 bestimmt wird.*

C) Ist (3,2) für die Punkte P_2 und P_3 — die in Bezug auf den Punkt P_1 starr sind — erfüllt, so gibt (3,1) die Bahn des Punktes P_3.

D) Sind die Geschwindigkeiten des Punktes P_1 und des in Bezug auf den Punkt P_1 starren Punktes P_2 gleichzeitig Null, so ist die Geschwindigkeit des Punktes P_3, der auf der durch P_1 und P_2 gebildeten starren Geraden liegt, zu der gleichen Koordinatenzeit auch Null.

§ 4. Die Bewegungsfreiheit eines in Bezug auf zwei Punkte vorgegebener Bahn starren dritten Punktes

Die Ortsvektoren $\mathfrak{r}_1 = \mathfrak{r}_1(t)$ und $\mathfrak{r}_2 = \mathfrak{r}_2(t)$ der Punkte P_1 und P_2 seien gegeben. Von dem Punkt P_3 — der relativ zu den zwei Punkten P_1 und P_2 starr ist — sei jedoch nur die Ortskoordinate $x_3 = x_3(t)$ bekannt. Es sollen nun die Bedingungen, unter denen die unbekannten Funktionen $y_3 = y_3(t)$ und $z_3 = z_3(t)$ bestimmt werden können, untersucht werden. Aus der Bedingung, dass die Punktepaare P_j, P_3 $(j = 1, 2)$ starr sind, ergibt sich

$$\begin{aligned} &\dot{\mathfrak{r}}_j(t_j)\{\mathfrak{r}_j(t_j) - \mathfrak{r}_3(t)\} - c^2\{t_j - t\} = 0, && \text{a)} \\ &\qquad\qquad j = 1,2. \\ &\{r_j(t_j) - \mathfrak{r}_3(t)\}^2 - c^2\{t_j - t\}^2 = s_{j3}^2 . && \text{b)} \end{aligned} \tag{4,1}$$

Ist z. B. $\mathfrak{r}_1(t) = \mathfrak{r}_2(t)$, so sind die Gleichungen des Gleichungssystems (4,1) für $j = 1$ mit denen für $j = 2$ identisch und $y_3(t)$ und $z_3(t)$ können aus dem Gleichungssystem nicht bestimmt werden. Schliesst man diesen Fall aus,

so können $y_3(t)$ und $z_3(t)$ mit Hilfe der Gleichungen b) des Gleichungssystems (4,1) im allgemeinen in der folgenden Form ausgedrückt werden:

$$y_3(t) \equiv y_3\{t, t_1, t_2\}; \quad z_3(t) \equiv z_3\{t, t_1, t_2\}\,. \tag{4,2}$$

Setzt man nun (4,2) in die Gleichungen a) von (4,1) ein, so stehen für die Unbekannten $t_1 = t_1(t)$ und $t_2 = t_2(t)$ gerade zwei Gleichungen zur Verfügung; bestimmt man aus diesen $t_1(t)$ und $t_2(t)$ und setzt die erhaltenen Werte in (4,2) ein, so erhält man y_3 und z_3 bereits in der gewünschten Form.

Da der reelle Wert von $s_{j3} \geqslant 0$ ist, so kommen — wie dies auf Grund von (4,1) leicht einzusehen ist — nur solche Werte der gesuchten Funktionen $y_3(t)$ und $z_3(t)$ in Betracht, für die

$$(s_{13} + s_{23})^2 \geqq \sigma_{12}^2 \equiv \{\mathfrak{r}_1(t_1) - \mathfrak{r}_2(t_2)\}^2 - c^2\{t_1 - t_2\}^2. \tag{4,3}$$

Stellt während der Bewegung zufälligerweise diese Beziehung beinahe eine Gleichheit dar, so muss auch σ_{12}^2 (> 0) in der Zeit konstant sein und wegen Punkt C) des vorhergehenden Paragraphen werden dann alle drei Koordinaten des Punktes P_3 durch die Koordinaten des starren Punktepaares P_1, P_2 eindeutig bestimmt.

Zusammenfassend kann man feststellen, dass *ein in Bezug auf zwei nicht zusammenfallende Punkte starrer dritter Punkt (mindestens und höchstens) einen Freiheitsgrad besitzt, angenommen, dass dieser Punkt sich nicht auf einer durch die zwei gegebenen Punkte bestimmten starren Geraden befindet.*

§ 5. Die kinematischen Eigenschaften dreier nicht auf einer Geraden liegenden und in Bezug auf einander starren Punkte (»starres Dreieck«)

Da die in § 1 definierte »Zuordnung« der Koordinatenzeiten zwar reflexiv, jedoch im allgemeinen nicht transitiv ist, muss man im Fall der gegenseitig starren Bewegung von mehr als zwei Punkten zur Bezeichnung der »zugeordneten« Kooordinatenzeiten auf die Benutzung don Doppelindizes übergehen. Es sei also t_{jk} jene Koordinatenzeit des Punktes P_j, die der Koordinatenzeit t_{kj} des in Bezug auf P_j starren Punktes P_k entspricht. Im allgemeinen ist $t_{jk} \neq t_{kj}$.

Die Glieder $j = k$ spielen in den Starrheitsgleichungen keine Rolle; von den im Falle $j, k = 1, 2, 3$ übrigbleibenden sechs Matrixelemten t_{jk} können nur drei willkürlich gewählt werden, da die weiteren drei Elemente durch die der dreifachen Paarbildung der drei Punkte entsprechenden Starrheitsbedingungen bestimmt werden.

I. *Man sieht, dass die Angabe von sechs Koordinaten eines nicht in eine Gerade entarteten starren Dreiecks zur Berechnung der übrigen Koordinaten hinreichend und gleichzeitig notwendig ist*

Von den neun Koordinaten eines starren Dreiecks können die sechs Koordinaten auf mehrere Weisen ausgewählt werden:

1. Es seien z. B. die Koordinaten $x_1 = x_1(t)$, $y_1 = y_1(t)$, $z_1 = z_1(t)$ des Punktes P_1 und die Koordinaten $x_2 = x_2(t)$, $y_2 = y_2(t)$ des Punktes P_2, dagegen lediglich die Koordinate $x_3 = x_3(t)$ des Punktes P_3 gegeben (dies ist der Fall »3 + 2 + 1«). Gemäss der im Zusammenhang mit den Gln. (1,7) — (1,9) angegebenen Methode sind diese Koordinaten notwendig und

hinreichend zur Bestimmung der Koordinate $z_2(t)$. Danach können mit Hilfe der im § 4 erwähnten Methode dann auch die weiteren Koordinaten, d. h. $y_3(t)$ und $z_3(t)$ bestimmt werden.

2. Werden jedoch je zwei Koordinaten der drei Punkte als gegeben angenommen, z. B. $x_j(t)$, $y_j(t)$, $j = 1, 2, 3$ (dies ist der Fall »2 + 2 + 2«), so lassen sich die fehlenden Koordinaten $z_j(t)$, $j = 1, 2, 3$ in folgenden Schritten berechnen:

Die Variablen $t \equiv t_{12} = t_{13} = t_{23}$ seien beliebig gewählt. Dann können die Gleichungen für die erzwungene starre Bewegung des Dreiecks auch folgendermassen geschrieben werden:

$$\dot{\mathfrak{r}}_1(t)\,\{\mathfrak{r}_2(t) - \mathfrak{r}_2(t_{21})\} - c^2\,\{t - t_{21}\} = 0, \qquad \text{a)}$$
$$\{\mathfrak{r}_1(t) - \mathfrak{r}_2(t_{21})\}^2 - c^2\,\{t - t_{21}\}^2 = s_{12}^2, \qquad \text{b)} \qquad (5{,}1)$$

$$\dot{\mathfrak{r}}_2(t)\,\{\mathfrak{r}_2(t) - \mathfrak{r}_3(t_{32})\} - c^2\,\{t - t_{32}\} = 0, \qquad \text{a)}$$
$$\{\mathfrak{r}_2(t) - \mathfrak{r}_3(t_{32})\}^2 - c^2\,\{t - t_{32}\}^2 = s_{23}^2, \qquad \text{b)} \qquad (5{,}2)$$

$$\dot{\mathfrak{r}}_1(t)\,\{\mathfrak{r}_1(t) - \mathfrak{r}_3(t_{31})\} - c^2\,\{t - t_{31}\} = 0, \qquad \text{a)}$$
$$\{\mathfrak{r}_1(t) - \mathfrak{r}_3(t_{31})\}^2 - c^2\,\{t - t_{31}\}^2 = s_{13}^2. \qquad \text{b)} \qquad (5{,}3)$$

Es sei wiederum vorausgesetzt, dass die drei Punkte nicht auf einer Geraden liegen.

In den weiteren Überlegungen wird auch die folgende aus den Gleichungen (5,1) ableitbare Gleichung gebraucht:

$$\dot{\mathfrak{r}}_2(t_{21})\,\{\mathfrak{r}_2(t_{21}) - \mathfrak{r}_1(t)\} - c^2\,\{t_{21} - t\} = 0. \qquad (5{,}4)$$

Zur Bestimmung der Funktionen $z_j = z_j(t)$ werden die mit b) bezeichneten Gleichungen der Reihe nach nach $z_2(t_{21})$, $z_3(t_{32})$ und $z_3(t_{31})$ aufgelöst. Die so gewonnenen Ausdrücke, in denen t_{21}, t_{32}, und t_{31} explizit und ausserdem nur bekannte Funktionen implizit vorkommen, werden in die entsprechenden mit a) bezeichneten Gleichungen eingesetzt. Die Wurzeln t_{21}, t_{32} und t_{31} des in dieser Weise aufgeschriebenen Gleichungssystems werden nun mit einer geeigneten Näherungsmethode bestimmt und die noch unbekannten Werte der Funktionen z_1, z_2, $\dot{z}_1$, $\dot{z}_2$ vorläufig als Parameter betrachtet. Als Ergebnis erhält man im allgemeinen die Funktionen in folgender Form:

$$t_{21} = t_{21}\,\{z_1(t),\ z_2(t),\ \dot{z}_1(t), \dot{z}_2(t),\ t\}\,,$$
$$t_{32} = t_{32}\,\{z_1(t), z_2(t), \dot{z}_1(t), \dot{z}_2(t), t\}\,, \qquad (5{,}5)$$
$$t_{31} = t_{31}\,\{z_1(t), z_2(t), \dot{z}_1(t), \dot{z}_2(t), t\}\,.$$

Die erhaltenen Funktionen werden nun in die entsprechenden mit b) bezeichneten Gleichungen eingesetzt und diese wiederum nach $z_2(t_{21})$, $z_3(t_{32})$, $z_3(t_{31})$ aufgelöst. Diese Ausdrücke samt den Funktionen (5,5) werden in die ihnen entsprechenden mit a) bezeichneten Gleichungen (5,1), (5,2) und (5,3), weiterhin in die Gleichung (5,4) eingesetzt. Keine der so erhaltenen Gleichungen enthält nun die Veränderlichen t_{21}, t_{32} und t_{31}, noch die ihnen entsprechenden Koordinatenwerte, sondern nur t und die dazugehörigen Koordinatenwerte, bzw. die Werte der Ableitungen. Aus der aus den Gln. (5,2) und (5,3) abgeleiteten Gleichung, sowie aus dem in Bezug auf die unbekannten Funktionen $z_1(t)$ und $z_2(t)$ gekoppelten Differentialgleichungssystem erster Ordnung können die Funktionen $z_1(t)$ und $z_2(t)$ bestimmt werden. Die Lösung enthält notwendigerweise zwei unbestimmte Integrationskonstanten C_1, C_2:

$$z_1 = z_1(t, C_1, C_2)\,,$$
$$z_2 = z_2(t, C_1, C_2). \qquad (5{,}6)$$

Zur Bestimmung der Integrationskonstanten setzt man nun Gl. (5,6) in die aus Gl. (5,1) und Gl. (5,4) in der beschriebenen Weise abgeleitete und bisher nicht benützte Gleichung ein. Diese zwei Gleichungen enthalten die noch unbekannte Funktion $z_3(t)$ nicht, so dass aus ihnen die Konstanten C_1 und C_2 berechnet werden können.

Die Bahnen der Punkte P_1 und P_2 sind nun bekannt, und es ist nur noch die unbekannte Funktion $z_3(t)$ zu bestimmen. Dies kann jedoch nach der Methode und unter den Bedingungen des vorhergehenden Paragraphen ohne Weiteres geschehen.

Aus dem Obigen geht also auch hier hervor, dass die Vorgabe von sechs (x, y) Koordinaten (Fall »2 + 2 + 2«) die Bestimmung der fehlenden übrigen Koordinaten ermöglicht. Dies ist daher eine hinreichende Bedingung. Sie ist aber gleichzeitig auch notwendig, da ja sämtliche (voneinander unabhängigen) Bedingungsgleichungen verwandt wurden und mit

weniger vorgegebenen Koordinaten die in dem Verfahren vorkommenden Gleichungen nicht hätten gelöst werden können.

3. In ähnlicher Weise folgert man auch, wenn von jedem Punkt wi derum zwei von der Zeit abhängige Koordinaten gegeben werden, die jedoch nicht einheitlich die x- und y-Koordinaten zu sein brauchen.

4. Der Fall, in dem sowohl von Punkt P_1 wie auch von Punkt P_2 alle drei Raumkoordinaten, von Punkt P_3 jedoch keine einzige Raumkoordinate gegeben ist, muss aus den vorliegenden Untersuchung auf Grund der im letzten Satz des § 4 enthaltenen Aussage ausgeschlossen werden.

Für ein starres Dreieck gilt also das Paradoxon von HERGLOTZ [3] nicht, nach dem die Bewegung eines starren Körpers bereits durch die Angabe dreier zeitabhängiger Koordinaten oder durch die Bewegung eines einzigen Punktes bestimmt wäre.

II. *Die Bewegungen der drei Ecken eines sich zur Zeit t in beliebiger Bewegung befindlichen starren Dreiecks können zu einer späteren, für alle drei Ecken gleichen Koordinatenzeit $\tau > t$ zum Stillstand gebracht werden*

Es ist bekannt, dass die BORNschen Bedingungen diese Forderung im allgemeinen nicht befriedigen können. Weitere interessante Eigenschaften der BORNschen starren Bewegung wurden in [10] eingehend untersucht.

EHRENFEST [11] hat z.B. gezeigt, dass eine im BORNschen Sinne starre gleichförmig rotierende Scheibe nicht so verlangsamt werden kann, dass sie dabei starr bleibt. Dieses Paradoxon soll nun auch für ein starres Dreieck untersucht werden.

Die zu beweisende Behauptung beruht darauf, dass die die Bewegung eines starren Dreiecks bestimmenden, also auch willkürlich veränderbaren sechs Raumkoordinaten $x_1(t)$, $y_1(t)$, $z_1(t)$; $x_2(t)$, $y_2(t)$ und $x_3(t)$ von irgendeinem beliebigen Zeitpunkt τ_0 an als Funktionen der Zeit so gewählt werden können, dass die Koordinaten sich zur Zeit τ_0 nicht sprunghaft ändern, und dass ihre Geschwindigkeiten zu einer Zeit τ verschwinden, zu der die folgende Determinante ungleich Null ist:

$$D \equiv \begin{vmatrix} y_3(\tau) - y_2(\tau) & z_3(\tau) - z_2(\tau) \\ y_3(\tau) - y_1(\tau) & z_3(\tau) - z_1(\tau) \end{vmatrix} \neq 0. \tag{5,7}$$

Sind von der Zeit τ_0 an die betrachteten Koordinaten durch Ausdrücke, die von zweitem Grade sind und je drei Konstanten enthalten, gegeben, so können zur Bestimmung der Konstanten auf Grund der obigen Voraussetzungen lineare Gleichungen aufgestellt werden, die in Bezug auf die in den sechs Koordinaten vorkommenden $3 \times 6 = 18$ Konstanten linear sind. Und zwar: sechs Gleichungen, die den kontinuierlichen Übergang der Geschwindigkeiten der Koordinaten zur Zeit $t = \tau_0$ berücksichtigen, weitere sechs Gleichungen, die das Verschwinden der Geschwindigkeiten der Koordinaten zur Zeit $t = \tau$ berücksichtigen und zuletzt drei sich aus der Beschränkung (5,1) ergebende Gleichungen. (Die letzten drei Gleichungen lassen sich z. B. in folgender Weise aufschreiben:

$$\text{falls } y_1(\tau) = y_2(\tau), \text{ so ist } z_1(\tau) = z_2(\tau) + d;$$

$$\text{falls } y_3(\tau) = y_2(\tau), \text{ so ist } y_3(\tau) = y_1(\tau) + d;$$

$$\text{falls } z_3(\tau) = z_2(\tau), \text{ so ist } z_3(\tau) = z_1(\tau) + d,$$

wo d eine geeignet zu wählende nichtverschwindende Konstante ist. Diese Gleichungen sind nämlich hinreichend für das Nichtverschwinden der Determinante D.)

Man sieht, dass zur Bestimmung der 18 Konstanten, die in den Ausdrücken zweiten Grades enthalten sind, 15 lineare Gleichungen zur Verfügung stehen, demnach können die Konstanten in Mächtigkeit eines Kontinuums widerspruchlos bestimmt werden.

Man hat nun nur noch zu beweisen, dass wenn die Geschwindigkeiten der Raumkoordinaten $x_1, y_1, z_1; x_2, y_2; x_3$ des starren Dreiecks $P_1P_2P_3$ zu einer Zeit τ verschwinden, zu der gleichzeitig $D \neq 0$ ist, auch die Geschwindigkeiten der weiteren Koordinaten zur selben Zeit τ verschwinden.

Bevor der Beweis für die obige Behauptung geführt wird, soll noch gezeigt werden, dass das Verschwinden zu ein und derselben Zeit τ von fünf willkürlich wählbaren Koordinatengeschwindigkeiten eines starren Punktepaares P_2, P_2 das Verschwinden der aus den fünf Koordinaten berechenbaren Geschwindigkeit der sechsten Koordinate (§ 4) nach sich zieht. Wegen $\dot{\mathfrak{r}}_1 = 0$ ist nämlich $t_{12} = t_{21} = \tau$ und daher hat man auf Grund der Symmetrieeigenschaft der Bedingungsgleichungen (1,4):

$$\dot{\mathfrak{r}}_2(\tau)\,\{\mathfrak{r}_2(\tau) - \mathfrak{r}_1(\tau)\} = 0\,.$$

Da hier $\dot{x}_2(\tau) = \dot{y}_2(\tau) = 0$, so folgt aus der obigen Gleichung $\dot{z}_2(\tau) = 0$. In ähnlicher Weise lässt sich der Satz über das gleichzeitige Stehenbleiben der Eckpunkte eines starren Dreiecks aus der Bedingungsgleichung (1,4) beweisen. Es verschwinde die Geschwindigkeit der gegebenen sechs Koordinaten $x_1(t), y_1(t), z_1(t); x_2(t), y_2(t)$ und $x_3(t)$ zur Zeit τ. Dann ist die Koordinatenzeit des Punktes $P_1 : t_{21} = \tau$ und auch die dem Punkt P_2 entsprechende Koordinatenzeit $t_{21} = \tau$, da $\dot{\mathfrak{r}}_1 = 0$ und die Starrheitsbedingung (1,4) symmetrisch ist. Laut obiger Überlegung ist daher

$$\dot{z}_2(\tau) = 0\,. \tag{5,8}$$

Beachtet man, dass wegen (5,8) gleichzeitig auch die der im Punkt P_2 gemessenen Koordinatenzeit $t_{23} = \tau$ entsprechende Koordinatenzeit im Punkt P_3: $t_{32} = \tau$ ist, und dass weiter auch die im Punkt P_1 gemessene Zeit $t_{13} = \tau$ der Zeit $t_{31} = \tau$ im Punkt P_3 entspricht, so kann man die Gleichung (1,4) der Starrheit der Punktepaare P_2, P_3 und P_1, P_3 zu der Zeit $t_{23} = \tau$ bzw. $t_{13} = \tau$ aufschreiben:

$$\begin{aligned} &\dot{y}_3(\tau)\{y_3(\tau) - y_2(\tau)\} + \dot{z}_3(\tau)\{z_3(\tau) - z_2(\tau)\} = 0\,,\\ &\dot{y}_3(\tau)\{y_3(\tau) - y_1(\tau)\} + \dot{z}_3(\tau)\{z_3(\tau) - z_1(\tau)\} = 0\,. \end{aligned} \tag{5,9}$$

Diese zwei Gleichungen stellen ein homogenes lineares Gleichungssystem für $\dot{y}_3$ und $\dot{z}_3$ dar, das eine von Null verschiedene Lösung nur dann hat, falls die Determinante D des Gleichungssystems Null ist. Dies ist jedoch wegen der aus dem Verschwinden der Geschwindigkeiten der sechs Koordinaten folgenden Bedingung (5,7) unmöglich.

Dadurch ist die am Anfang dieses Paragraphen gemachte Behauptung bewiesen. Aus der Behauptung folgt auch gleichzeitig, dass das starre Dreieck in einem Inertialsystem für eine beliebige Dauer zum Stillstand gebracht werden kann, ohne dass er dabei seine Starrheit auch nur für einen Augenblick verlieren würde.

§ 6. Zwei wichtige kinematische Eigenschaften eines aus vier Punkten bestehenden starren Punktsystems (starres »Tetraeder«)

I. Berechnung der Bahn eines in Bezug auf drei Punkte P_1, P_2, P_3 starren vierten Punktes P_4, wenn die Bahn der drei Punkte gegeben ist. Um Missverständnissen vorzubeugen, sei erwähnt, dass die gegenseitige Starrheit der drei Punkte (P_1, P_2, P_3) vorerst nicht vorausgesetzt wird. Mit der in § 5 eingeführten Bezeichnungsweise kann man die Bedingungen für die Starrheit der Punktepaare P_i, P_4 $(i = 1, 2, 3)$ folgendermassen aufschreiben: Wählt man die Zeiten t_{41}, t_{42}, t_{43} gleich, d.h. $t = t_{41} = t_{42} = t_{43}$, so hat man

$$\begin{aligned} &\dot{\mathfrak{r}}_1(t_{14})\{\mathfrak{r}_1(t_{14}) - \mathfrak{r}_4(t)\} - c^2\{t_{14} - t\} = 0\,, && \text{a)}\\ &\{\mathfrak{r}_1(t_{14}) - \mathfrak{r}_4(t)\}^2 - c^2\{t_{14} - t\}^2 = s_{14}^2\,, && \text{b)} \end{aligned} \tag{6,1}$$

$$\begin{aligned} &\dot{\mathfrak{r}}_2(t_{24})\{\mathfrak{r}_2(t_{24}) - \mathfrak{r}_4(t)\} - c^2\{t_{24} - t\} = 0\,, && \text{a)}\\ &\{\mathfrak{r}_2(t_{24}) - \mathfrak{r}_4(t)\}^2 - c^2\{t_{24} - t\}^2 = s_{24}^2\,, && \text{b)} \end{aligned} \tag{6,2}$$

$$\begin{aligned} &\dot{\mathfrak{r}}_3(t_{34})\{\mathfrak{r}_3(t_{34}) - \mathfrak{r}_4(t)\} - c^2\{t_{34} - t\} = 0\,, && \text{a)}\\ &\{\mathfrak{r}_3(t_{34}) - \mathfrak{r}_4(t)\}^2 - c^2\{t_{34} - t\}^2 = s_{34}^2\,. && \text{b)} \end{aligned} \tag{6,3}$$

In diesen Gleichungen seien die Vektorfunktionen $\mathfrak{r}_i = \mathfrak{r}_i(t)$, $i = 1, 2, 3$ als gegeben angenommen, die Funktion $\mathfrak{r}_4 = \mathfrak{r}_4(t)$ sei hingegen die gesuchte, unbekannte Funktion. Zu ihrer Bestimmung lösen wir die mit a) bezeichneten Gleichungen, die ein lineares Gleichungssystem darstellen, der Reihe nach für $x_4(t)$, $y_4(t)$ und $z_4(t)$ auf. Die Funktion $\mathfrak{r}_4(t)$ ergibt sich dann im allgemeinen in folgender Form:

$$\mathfrak{r}_4 = \mathfrak{r}_4(t_{14}, t_{24}, t_{34}, t). \tag{6,4}$$

Die erhaltenen Ausdrücke werden nun in die mit b) bezeichneten Gleichungen eingesetzt, die jetzt für die unbekannten Grössen t_{14}, t_{24}, t_{34} ein nur bekannte Funktionen enthaltendes Gleichungssystem bilden. Wird dieses mit Hilfe einer geeigneten Näherungsmethode gelöst, so ergeben sich die unbekannten Grössen t_{i4} $(i = 1, 2, 3)$ als Funktionen von t:

$$t_{i4} = t_{i4}(t), \quad i = 1, 2, 3.$$

Setzt man diese nun bekannten Funktionen in die Gl. (6,4) ein, so gelangt man zu dem gesuchten Ausdruck für $\mathfrak{r}_4$: $\mathfrak{r}_4 = \mathfrak{r}_4(t)$.

Auf Grund der beschriebenen Methode können die mathematischen Voraussetzungen für die Bestimmung von $\mathfrak{r}_4$ auch festgestellt werden. Ohne in Einzelheiten einzugehen, soll hier nur darauf hingewiesen werden, dass sich die Bahn des Punktes P_4 sicherlich nicht bestimmen lässt, wenn von den Punkten P_1, P_2, P_3 zwei auf die Dauer zusammenfallen. In diesem Fall werden nämlich zwei von den Gleichungen (6,1), (6,2), (6,3) identisch.

Aus dem Obigen geht hervor, dass jede einzelne irgendeinem Punkt aufgelegte neue, unabhängige Starrheitsbedingung den Freiheitsgrad des Punktes — wie das ja auch in der klassischen Kinematik der Fall ist — um einen Grad verringert.

Es sei nun der spezielle Fall betrachtet, in dem die Punkte P_1, P_2, P_3 gegenseitig starr sind, und in dem der Punkt P_3 sich auf der gemäss Gleichung (3,1) bestimmten starren Geraden befindet.

In der klassischen kinematik lässt sich in diesem Falle die Bahn des Punktes P_4 aus der der Punkte P_1, P_2, P_3 nicht bestimmen. Auch in der relativistischen Kinematik tritt eine Schwierigkeit auf, deren Ursache einfach darin liegt, dass die die Starrheit der Punktpaare P_j, P_4 $(j = 1, 2, 3)$ ausdrückenden Gleichungen (6,1—3) mit den Gleichungen, laut denen die Eckpunkte P_1, P_2, P_3 des Dreiecks in eine Gerade fallen, im Widerspruch stehen können.

Im allgemeinen gibt es keinen Punkt P_4, der in Bezug auf jede einzelnen der auf einer starren Geraden liegenden Punkte P_1, P_2, P_3, die gegenseitig starr sind, starr sein könnte. Gäbe es nämlich einen solchen Punkt P_4, so könnten im Sinne der obigen Feststellungen (§ 5 II und § 6 II) in einem Inertialsystem alle vier Punkte auf die Dauer zum Stillstand gebracht werden, ohne dass dabei in der gegenseitigen Starrheit von irgendwelchen zwei Punkten, d. h. in dem Abschnittswert s_{jk} auch nur für einen Augenblick eine Änderung eintreten würde. Dies würde jedoch bedeuten, dass zwischen den sechs Entfernungen (Abschnittswerten), die durch die auf eine Gerade fallenden Punkte P_1, P_2, P_3 sowie durch den Punkt P_4 bestimmt werden, dieselbe Beziehung besteht, wie zwischen den Abständen der im Inertialsystem ruhenden Punkte. Dies ist jedoch unmöglich, da laut der in § 2 Punkt IV bewiesenen Feststellung, die euklidische Massbestimmung für eine sich gleichförmig drehende Ebene nicht gültig ist.

Aus dieser elementaren Überlegung folgt gleichzeitig, dass *in der relativistischen Kinematik* im allgemeinen *nicht von einer starren Achse gesprochen werden kann.*

Dies ist der Grund dafür,warum in dem am Anfang, unter I, stehenden Satze ausgeschlossen werden musste, dass die Punkte auf einer Geraden liegen können und weiter, warum die am Anfang des folgenden Punktes, unter II, stehende Aussage auf ausserhalb der Ebene der Punkte P_1, P_2, P_3 liegende Punkte P_4 zu beschränken ist.

II. Gibt es einen solchen Wert der Koordinatenzeit, bei dem irgendeiner der Punkte eines nicht in eine Gerade entarteten starren Dreiecks $P_1P_2P_3$ stehen bleibt, so bleibt zu der gleichen Zeit auch der ausserhalb der Ebene des Dreiecks liegende und mit den zur Ruhe gekommenen Punkten starr verbundene Punkt P_4 stehen. (Der durch die Punkte P_1, P_2, P_3, P_4 gebildete starre Körper kann starres Tetraeder genannt werden.)

Vor dem Beweis der obigen Behauptung soll noch auf folgenden Satz hingewiesen werden. Es seien P_1, P_2, P_3 Punkte der in § 3 erwähnten sich starr bewegenden Gerade. Bleiben P_1 und P_2 bei dem gleichen Wert der Koordinatenzeit stehen, so bleibt zu derselben Zeit auch P_3 stehen. (Die Richtigkeit dieses kinematischen Satzes kann aus den nach der Zeit abgeleiteten nur Raumkoordinaten enthaltenden Gliedern des Gleichungssystems (6,3) gefolgert werden, falls man berücksichtigt, dass der im Punkte P_1 und auch im Punkte P_2 gemessenen Koordinatenzeit τ im Punkte P_3 ebenfalls die Zeit τ und nur die Zeit τ entspricht.)

Zum Beweis des am Anfang des gegenwärtigen Punktes (II) stehenden Satzes bedenke man, dass wegen der gemachten Voraussage

$$\dot{\mathfrak{r}}_j(\tau) = 0\,, \quad j = 1, 2, 3.\,, \tag{6,5}$$

also entspricht im Punkt P_4 gemessenen Zeit τ in allen Punkten P_j $(j = 1, 2, 3)$ auch die Zeit τ. Damit muss wegen der geforderten Starrheit des Punktes P_4 in Bezug auf die Punkte P_j $(j = 1, 2, 3)$ die folgende symmetrische Gleichung (1,5) erfüllt werden, in der nach der in § 5 gegebenen Bezeichnung $t_{4j} = \tau$ zu setzen ist:

$$\{\dot{\mathfrak{r}}_j(\tau) - \dot{\mathfrak{r}}_4(\tau)\}\{\mathfrak{r}_j(\tau) - \mathfrak{r}_4(\tau)\} = 0, \quad j = 1, 2, 3.$$

Wegen der Bedingung (6,5) kann für die drei Komponenten von $\dot{\mathfrak{r}}_4(\tau)$ folgendes homogene Gleichungssystem aufgeschrieben werden:

$$\dot{\mathfrak{r}}_4(\tau)\{\mathfrak{r}_j(\tau) - \mathfrak{r}_4(\tau)\} = 0\,, \quad j = 1, 2, 3\,.$$

Dieses Gleichungssystem hat nur dann eine von Null verschiedene Lösung, falls die Determinante D des Gleichungssystems gleich Null ist. In unserem Fall ist aber:

$$D \equiv \left| \begin{array}{c} \mathfrak{r}_j(\tau) - \mathfrak{r}_4(\tau) \\ j = 1, 2, 3 \end{array} \right| \neq 0.$$

Die Forderung dieser Ungleichheit drückt aus, dass der Punkt P_4 weder einer der im Inertialsystem ruhenden Punkte ist, noch sich in der durch die Punkte P_j $(j = 1, 2, 3)$ bestimmten Ebene befindet. Da dies ja gerade die Voraussetzung war, ist gewiss, dass in diesem Falle $\dot{\mathfrak{r}}_4(\tau)$ nur Null sein kann, d.h. dass Punkt P_4 zur Zeit τ ebenfalls ruht.

Ist $D = 0$, so befindet sich P_4 zur Zeit τ in der durch die nicht in eine Gerade fallenden Punkte P_1, P_2, P_3, P_4 zur Zeit τ bestimmten Ruhebene. In diesem Fall kann der oben angeführte Satz in ähnlicher einfacher Weise unter der Beschränkung, dass irgendeiner der Punkte P_1, P_2, P_3 von der Zeit τ ab dauernd in Ruhe bleibt, eingesehen werden.

§ 7. Eine von der klassischen abweichende Eigenschaft des starren Bitetraeders

Zwei nicht gemeinsame Eckpunkte von zwei Tetraedern mit gemeinsamer Grundfläche, d. h. Bitetraeder, können in Bezug zu einander im allgemeinen nicht starr sein.

Laut § 6 können nämlich sämtliche Eckpunkte eines Bitetraeders im Inertialsystem dauernd zur Ruhe gebracht werden, d.h. alle fünf Eckpunkte eines Bitetraeders können auf ein im euklidischen Raum ruhendes kongruentes Bitetraeder gelegt werden, d.h. darauf abgebildet werden. Falls, im Gegensatz zu der obigen Behauptung, zwei nicht gemeinsamen Eckpunkte des sich starr bewegenden Bitetraeders in Bezug aufeinander ebenfalls starr wären, so würde zwischen den zu den zehn Punktepaaren gehörenden zeitlich konstanten, invarianten Raumabständen d_{jk} während der Bewegung dieselbe die Raummetrik bestimmende Beziehung (1,9) bestehen, die für das im euklidischen Raum ruhende kongruente Bitetraeder gültig ist. Es wurde jedoch bereits im

§ 2 IV gezeigt, dass für einen sich bewegenden starren Körper auch eine nicht euklidische Metrik gelten kann. Damit ist die Behauptung bewiesen, da man sonst zu einem Widerspruch geführt würde.

Da für die Bornsche starre Bewegung eine solche Beschränkung per def. nicht bestehen kann, ist es klar, dass

a) *die Bornschen Bedingungen für ein aus mindestens fünf Punkten bestehendes Punktsystem eine spezielle (3 Freiheitsgrade besitzende) Bewegung (Bewegungsgruppe) auswählen,*

b) *die starre Bewegung eines aus mehr als vier Punkten bestehenden Punktsystems (6 Freiheitsgrade), bei der alle Paare nach der hier gegebenen Definition (1. 1—2) starr sind, im allgemeinen nicht verwirklicht werden kann.*

Herrn Prof. G. Marx sei für sein anspornendes Interesse der Dank des Verfassers ausgesprochen.

LITERATUR

1. T. Mátrai, Nature, **172**, 858, 1953.
2. M. Born, Ann. d. Phys., **30**, 1, 1909.
3. G. Herglotz, Ann. d. Phys., **31**, 393, 1910.
4. K. Novobátzky, Die Relativitätstheorie, S. 34. Tankönyvkiadó, Budapest, 1951. (ungarisch).
5. C. W. Berenda, Phys. Rev., **62**, 280, 1942.
6. N. Rosen, Phys. Rev., **71**, 54, 1947.
7. E. L. Hill, Phys. Rev., **65**, 488, 1946.
8. A. S. Eddington, Relativitätstheorie in mathematischer Behandlung, Springer, Berlin, 1923.
9. H. A. Lorentz, Nature, **106**, 795, 1921.
10. G. Salzmann und A. H. Taub, Phys. Rev., **95**, 1659, 1954.
11. P. Ehrenfest, Phys. Zeitschr., **10**, 918, 1909.

О НЕКОТОРЫХ РЕЛЯТИВИСТСКИХ КИНЕМАТИЧЕСКИХ СВОЙСТВАХ ЖЕСТКОЙ ТОЧЕЧНОЙ СИСТЕМЫ

Т. МАТРАИ

Резюме

В предыдущей статье [1] сообщилось, что из основых соображений М. Борна [2] были выведены мною Лоренц-инвариантные принудительные условия для жесткого принужденного движения пар точек конечного расстояния. Твердое тело рассматривалось не как среда, а предполагалось построенным из пар точек. В такой концепции исследовалось, каким образом видоизменяются классические кинематические свойства в специальной теории относительности. Элементарным путем удалось показать, что жесткая система четырех точек в любом спаривании обладает не тремя, а шестью степенями свободы, как это можно ожидать на основе соображений Г. Герглоца [3], которые базируются на принудительных условиях Борна, касающихся сред. Однако, движение точек, число которых больше четырёх, вообще не может осуществляться так, что любое спаривание оставалось жестким, за исключением специального случая, так называемого движения типа Борна. Между прочим, удалось доказать, что в жесткой среде релятивистского гиперболического движения, описанной первый раз Борном, по необходимости господствует внутренняя пространственная метрика Евклида.

MACH'S PRINCIPLE AND GENERAL RELATIVITY

By

F. KÁROLYHÁZY

INSTITUTE FOR THEORETICAL PHYSICS, ROLAND EÖTVÖS UNIVERSITY, BUDAPEST

(Presented by Z. Gyulai. — Received 20. IX. 1963)

It is argued, after a short review of the problem, that MACH's principle, in an interpretation near to MACH's original ideas, is not incorporated in general relativity. On the other hand, one can formulate a selection principle on the basis of which it is possible to divide the solutions of gravitational equations into two classes. Those of the first class do, those of the second do not contradict the reasonable content of MACH's requirements. Thus, these latter solutions may perhaps remain permissible even in the future, when a fuller knowledge proves the others to be an unpermissible extrapolation beyond the domain of validity of EINSTEIN's equations. We formulate a selection principle, which seems to be more general and flexible than previous ones.

§ 1

The existence of numerous contradictory approaches to "MACH's principle" in general relativity in itself proves that here we are faced with an unsolved problem. Part of the confusion stems from the fact that MACH's criticism of the basic notions of Newtonian mechanics preceded the birth of special relativity, that is to say, the recognition of the fusion of space and time into four-space, and further that MACH did not formulate his own requirements sufficiently rigorously. Therefore, it may be justifiable to present a short historical review.

The very generality in stating the axioms of dynamics on the one hand and the law of gravitation on the other led NEWTON to an extraordinary difficulty in the choice of a reference system. In relation to what are the bodies accelerated, if a force acts upon them? To find a general answer, valid beyond the individual features of single bodies, NEWTON established (or rather confirmed) the notion of absolute space. This, in a certain sense, has to be regarded as a return to the static world picture of antiquity (expressed by PTOLEMY). In this picture the natural state of bodies is to be at rest in a particular position. The statement that the rigid arrangement of bodies is the natural one leads immediately to the notion of "mere places" filling up side by side in empty space. Time plays an inactive role, running uniformly and at the same rate at all places. Motion appears as something disturbing, as seen in the famous apories of ZENON, etc.

It was this "grained" character of empty space, which is physically unobservable (there is no possibility of identifying a definite point in space at a later time), or rather the fact that it has no other physical manifestation than to make bodies resist acceleration against which E. MACH protested We reproduce here his criticism in a somewhat more modern form. Let K and K' be two Cartesian systems of reference, one of which is accelerated relative to the other. Events will be labelled by the coordinates x, y, z, t in K and x', y', z', t' in K'. The connection between then may be something like

$$\begin{aligned} x' &= x - \frac{1}{2}at^2, & x &= x' + \frac{1}{2}at'^2, \\ y' &= y, & y &= y', \\ z' &= z, & z &= z', \\ t' &= t, & t &= t'. \end{aligned} \tag{1}$$

(This corresponds to the simplest choice of acceleration and direction of axes.) The equality $t = t'$ is natural, since we have absolute time. The role of the frames K and K' is symmetrical in the above formulae. If we now ask for the spatial distance between two simultaneous events, or for the time interval between two arbitrary events (in accordance with the concept of absolute time), we have for K and K', respectively,

$$d\sigma^2 = dx^2 + dy^2 + dz^2 = dx'^2 + dy'^2 + dz'^2 \qquad dt^2 = dt'^2. \tag{2}$$

That is to say, K and K' also show equal intrinsic geometric properties which means that they are completely equivalent. We have no reason to state that the one is at absolute rest and that the other suffers an absolute acceleration. If undisturbed test bodies distinguish one of the systems of reference, let us say K, by moving in it uniformly along straight lines so that we need forces to deflect them, then this must have some other detectable reason. MACH was convinced that the explanation of all inertial properties can be found in the distant masses of the Universe. He attempted to construct (without success) a new mechanics in which in the absence of distant masses test bodies would not show any tendency in their motion and would not make a choice between K and K'. Or, to put it in another way, if, assuming the existence of distant masses, with the help of the test bodies K is shown to be an inertial frame, then a consequence of this should be that the distant masses, as a whole, do not suffer any acceleration in K. Conversely, if we need forces in K' to keep a test body in the origo, then the whole of the distant masses should be accelerated in K'. All these requirements are usually illustrated in connection with the well-known example of the rotating bucket. To summarize the properties of such mechanics: the different coordinate frames, exhibiting Euclidean character

of space, are equivalent. But motions and inertial properties are perfectly relative. We are allowed, for instance, to imagine the Universe rotating as a rigid whole in (Euclidean) space without fear of any disaster through centrifugal effects.

§ 2

The above criticism of the absolute, by pointing out the role of distant masses and leading to the concept of the relativity of motion, may, however, in spite of its clearness at first sight, be objected to in several respects. Comparing the coordinate systems K and K' we treated them as completely abstract reference frames, and we considered the motion of distance and spatial structure in a similar manner. But in physics distance can only be defined by tools made of matter and this involves dynamics. One might very well imagine that for the material construction of K we would have a simple prescription whereas the construction of K' in accordance with equ. (1) could be carried out only in a complicated and clumsy fashion. This asymmetry would then correspond to the different behaviour of freely moving test particles with respect to K and K'. There is nothing to be objected to when it is supposed that dynamical laws take a particularly simple form in a frame or in a class of frames which can be constructed particularly easily as they obey these same dynamical laws. (It may be noted that L. LANGE, a contemporary of E. MACH, was able to give an essentially correct definition of inertial frames, without reference to the distant masses of the Universe.) If in one such frame, distinguished by the simple method of its construction, the physical laws show some suitable symmetry, this then indicates that a whole set of equally simple frames of reference exist.

The fact is that in classical physics no such consistency exists between dynamical laws in general on the one hand, and preferred systems of reference (and space structure defined by them) on the other, (NEWTON's axioms assuming constant mass are invariant with respect to GALILEI transformations; MAXWELL's equations are not), and this circumstance again may justify MACH's attempt to introduce distant masses. The GALILEI group does not indicate any limitation in relative velocities, nor does it reject from the beginning the concept of arbitrarily elastic measuring rods with arbitrarily small mass. Thus the lack of consistency explains to a certain extent the demand for a description of nature such as was outlined at the end of the foregoing section. On the other hand, MAXWELL's equations — in their final consequences — have led to special relativity.

The recognition that space and time do not represent two separate "beds" for the flow of physical events, but that they are amalgamated to a four-space of space-time establishes full consistency between the structure of space-time

and the dynamical laws. Let us disregard gravitation for the moment. Then the physically well-defined motion which is independent of the choice of coordinates leads to the absolute straight (timelike) lines of the four-manifold of events. The separate $d\sigma^2$ and dt^2 in (2) are then replaced by the absolute four-distance ds^2 between two events, measured by physical clocks, and in contrast to the symmetry of equs. (2), K and K' no longer appear geometrically equivalent, ds^2 being represented by the simple expression $ds^2 = dx^2 + dy^2 + dz^2 - c^2dt^2$ in K, but by a more complicated one in K'. Nor are they equivalent physically. The material constituents of K are free-moving, whereas K' may turn out to be not directly realizable at all. The paths of other free-moving test bodies appear uncurved in K, but curved in K'. The relevant invariance group of the laws of nature, as expressed in the simplest possible frames, is the Lorentz-group, which indicates the corresponding set of equivalent simple frames. At this stage the protest against the absolute and the reference to the role of the distant masses appears to be out of date. Epistemologically we could very well resign ourselves once for all to the Minkowskian structure of space-time as given by MINKOWSKI.

§ 3

The justification for MACH's demands has been recognized by the general theory of relativity, which has revived them in a necessarily modified form. EINSTEIN's theory was the first to succeed in accounting for the simple empirical fact that all bodies fall at the same rate. Both in a laboratory floating in space far from all masses, and in another passing by some celestial body and influenced by nothing else than the gravitation of the latter, the state of weightlessness prevails equally. The most concise rendering of what is common to both is to say that in both cases motion is free. We have now connected the notion of free motion with the straight lines of space-time. This immediately leads to the concept of geodetics in a curved space-time, the curvature depending on the distribution of matter. The local connection between curvature and energy-stress density is given by the gravitational equations

$$R_{ik} - \frac{1}{2} g_{ik} R = - \varkappa T_{ik}. \tag{3}$$

Once experience disproves the prejudice that space-time is independent of matter the demand for space-time to be completely determined by matter rightly arises anew, a demand which is similar e.g. to the tacit assumption of classical physics that all electromagnetic fields originate (or perhaps terminate) in moving charges or magnetic dipoles etc. This new principle, stated

by EINSTEIN and called by him MACH's principle, is the legitimate successor to MACH's original requirements. If they could be fulfilled within the framework of general relativity, this would mean that geodetics and so the motion of test particles would be uniquely determined by the masses. But that, it must be emphasized, would not mean that "motion is relative". We see this most easily in the following way. Suppose we deflect a particle from its geodetic path. To do this, we need some external force. One would perhaps wish to fix a coordinate system to this particle, accelerated in the direction of the external force to find that in this system the particle is at rest, and that the distant masses are accelerated in the opposite direction. However, this desire retains the notion of rigid-type accelerated frames, described e.g. by eqs. (1) and attached to the Newtonian concepts of space and time. It fails because it is not uniquely and reasonably possible even in the case of a Minkowskian plane to extend a small fragment of a coordinate system—accelerated locally with respect to a local inertial frame — to the entire space-time manifold. If we give the test body short pulses and make it vibrate, we can choose a periodically deformed reference frame, in which the test particle and the distant masses as well remain at rest. So long as the lack of absolute simultaneity is disregarded this choice may be declared unreasonable, no justification for the rejection of such reference frames can be found in a general space-time. The traditional claim for the relativity of motion has been confused with the absolute space-time of general relativity (see [1] for details) as for example in the famous calculation of H. THIRRING and J. LENSE, the well-known result of which was hailed as a partial success and was praised by EINSTEIN himself. However, all considerations of this type are misleading. A real approach to the concept of the relativity of motion would need the following. For a model, in which a small test body is rotating "in the centre" with respect to distant masses at rest in a nearly flat space — and such a model does exist — we ought first of all to find another model, as a counterpart, in which there are big masses rotating with respect to a common centre arbitrarily distant, again in a nearly flat universe. Such a model, however, cannot exist without contradicting the field equations.

Thus the demand that the structure of absolute space-time be exclusively determined by the matter in the Universe, must have a deeper content, different from that for the relativity of motion. But, surprisingly, it is very hard to formulate this precisely. It certainly will mean something like this: The distribution of matter, that is, the distribution of energy flow and density at a given time t uniquely determines the curvature, etc. of the whole space-time, in the past as well as in the future. But what does the term "at a given time" mean? It can only mean a space-like "hypersurface" in the curved space-time. Prescription of the distribution of matter on such a hypersurface automatically involves the prescription of the hypersurface itself with all

its geometrical properties, something which we wished to deduce from the "given distribution" of matter. Of course, as a consequence of the requirement that the gravitational equation (3) should be satisfied, the freedom in the choice of the structure of the 3-dimensional hypersurface and in the choice of the distribution of matter mutually restrict each other, in other words, initial values must necessarily be consistent. This mutual restriction, together with the remaining freedom shows the "space-time aspect" and "matter aspect" of this initial value problem to be of equal intrinsic importance. It is true that the information obtained by a given hypersurface for the structure of the four-manifold is not so large, as that to be obtained from the intrinsic geometrical properties of the hypersurface. We can choose a very complicated hypersurface in a smooth four-manifold. This implies that the real freedom in "originating" a four-geometry from a three-surface is smaller than the apparent freedom of the three-surface. Nevertheless, there is a certain amount of freedom left, as is most easily seen by considering two infinitesimally near neighbouring three-surfaces. The best approach to an interpretation of the MACH—EINSTEIN requirement as a demand for the possibility of defining a unique initial value problem, seems to have been that of WHEELER [2]. Starting from the assumption that for a closed universe the specification of the three-geometry on a three-surface and its time rate of change uniquely determines the extrinsic curvature and, therefore, the self-consistent initial value data by which the entire four-geometry is completely determined, he arrives at the following conclusion, which at the same time is the best possible formulation of MACH's principle: for a closed universe the specification of sufficiently regular three-dimensional geometry at two instants immediately succeeding each other and that of the density and flow of energy determines the geometry of space-time — past, present and future — and thereby the inertial properties of every infinitesimal test particle. The most remarkable feature of this conjecture is that it automatically postulates a closed universe in order to avoid certain ambiguities in the position of the two three-surfaces relative to each other, that is, in embedding them into the entire four-geometry. Namely, in the case of open space-like surfaces we should be faced with the problem of finding appropriate boundary conditions to fix the extrinsic curvature at infinity, and this would by no means be governed by the distribution of matter. This approach is unsatisfactory because, apart from the lack of rigorous proofs, even though it excludes certain models of space-time in which the presence of matter is strikingly incidental, one cannot appreciate that matter plays the most important role in the remaining models, allowed by the above mentioned interpretation formulated in terms of an initial value problem.

It seems obvious, therefore, that the equations (3) permit the existence of a completely independent, individual space-time. As a matter of fact, this

is shown most simply by the fact that the perfectly empty Minkowskian space-time is a solution of (3). For a long time the vague hope subsisted that the independent aspect of space-time is only apparent and that space-time does not possess any real degree of freedom. Namely, it had been conjectured that the only regular and complete solution of (3), with an everywhere vanishing energy-stress tensor, is the flat space-time of MINKOWSKI. Clearly this would mean, that the variability of the structure of space-time in the solution of (3) is a consequence of the matter present.* This expected consequence of (3) has even been called MACH's principle in the literature [3], but this hope failed. Now many solutions are known (see for example [4]) which in spite of their non-vanishing Riemann-tensor describe model-universes completely empty and yet complete and without singularities.

§ 4

Of course, one could reason as follows: if space-time is unlikely per se, to lose its own degrees of freedom, we can force it to do so by the requirement of suitable boundary conditions. In fact this would be merely an evasion of the problem, not its solution. The undesirable release of space-time structure from the influence of the distribution of matter would then be replaced by the conceptual independence of the boundary conditions. Nobody would agree, for example, that in an asymptotically flat universe embodying a finite amount of mass, the geodetics are predominantly governed by matter. With, or without boundary conditions, the following conclusion is inescapable. Equations (3) deny the possibility of a consistent picture about space-time and matter, as sketched at the end of § 2, but they do not premise the monism of matter. This profound epistemological difficulty may become a starting point for further development of the theory. It suggests that the gravitational equations (3) are not exact laws having general validity, but are approximations, working well under circumstances actually observed in the Universe (e.g. nearly uniform distribution of matter in the visible region, etc.) but failing for other conceivable distributions of matter. This would then remove the necessity of excluding solutions like empty universes, which obviously contradict the spirit of the MACH—EINSTEIN principle from eq. (3) as the application of (3) to extreme distributions of matter would then be an impermissible extrapolation. All these considerations lead to two important consequences.

* We will not touch upon the case in which the so-called cosmological term is added to the left-hand side of (3) because at present there is no reason to introduce this term and besides, it would be superfluous in our considerations.

1. It is justifiable to seek *a)* for more direct theoretical indications (more direct than the above considerations) concerning the limited validity of equations (3), *b)* for experiments, detecting eventual gravitational effects of distant masses which would go beyond the boundaries of interpretation in terms of Riemann geometry.

2. It is desirable to have a selection principle, which limits the applicability of equation (3), since we have seen that in (3) the simple concept of Riemannian space-time may, at a later stage of deeper understanding of the connection between space-time and matter, prove to be an erroneous extrapolation.

As to 1*a*, it seems that we can have theoretical indications in the sense mentioned. Quite apart from the well-known apparent contraditions involved in any attempt to apply EINSTEIN's theory to microphysics, it is conceivable, that difficulties are encountered even in the field of macrophysics. We mention the problem of the "falling star", investigated on the ground of reasonable models by WHEELER [2]. It has long been known that under any reasonable assumption relating to the connection between density and pressure, the accumulation of a sufficiently large amount of mass to form a star would necessarily lead as a result of equations (3) to a breakdown of metric in the centre (not on the surface) of the star. Of course, in explaining the results of this, in principle we can always discount, in the form of a suitable radiation, the dangerous energies which would make the star explode. To insist upon the validity of the concept of Riemannian space-time and equations (3) even in regions filled with matter in such an extreme state would impose almost incredible restrictions on the behaviour of matter.

As to 1*b* we observe first that a number of speculative theories on gravitation, deviating to a greater or lesser extent from equations (3) have been enunciated without any supporting experimental result. Some of them (see [5] for example and for further references) make use of the fact [6] that we can treat gravitation as a long range force in ordinary flat space, corresponding to a zero mass boson filled in universal interaction with all matter. The assumption of the existence of other zero mass boson fields means indeed, that the unique attraction of interpreting the theory in terms of Riemannian geometry ceases. But the experiments performed so far in this direction (for example that of HUGHES and DREVER [5] planned to investigate the spatial anisotropy of inertia, or that of K. TURNER [7] in an attempt to find an effect of the motion of the Earth against the whole of the galaxies) have yielded only negative results. It should not be surprising if in the near future the increasing accuracy of measurements should only lead to a strengthening of the reliability of the famous tests of general relativity.

In view of these tests a fairly large domain of validity for EINSTEIN's equations can be expected. This turns our attention to point 2, that is to the

problem of distinction between those solutions of equations (3) which can, and those which cannot, be regarded as governed predominantly by matter. It is clear that such a distinction must be something like this: "It is not allowed to deal with solutions of equations (3), which correspond to a universe very different from ours." Indeed, many versions of the desired selection principle postulate a completely homogeneous, isotropic cosmological background, sometimes even with everywhere constant curvature. In such models even the definition of a special class of parameter-networks may turn out to be possible, on purely mathematical grounds, making use of the high degree of uniformity, which we may regard physically to correspond to uniformly-accelerated coordinate systems. In this manner we can partly force out the "relativity of motion" (for the very artificial case of permanently constant acceleration).

In our opinion, similar approaches exaggerate the restriction on the mutual freedom of space-time and matter following from the equations (3), and greatly hinder the theory from developing its full efficiency.

We propose the following criterion for the selection of cosmological solutions of the equations (3), which are dominated by matter: A physical solution S, that is a complete regular space-time with an everywhere acceptable (non-negative) mass density, etc. obeying equations (3) is permissible, if it is impossible to give a set $S(p)$ of other physical solutions, depending continously on the parameter p, and a set $\Sigma(p)$ of space-time domains, where $\Sigma(p)$ is a suitably chosen domain of $S(p)$, in such a manner, that $S(p_0) = \Sigma(p_0) = S$ and for some other value p_1 of p $\lim_{p \to p_1} \Sigma(p) = S_F$, where S_F describes a completely empty but otherwise complete and regular space-time. In short, matter dominates a solution if it is "essentially different" from an empty solution in the above sense, that is, if it cannot be deformed through a series of other solutions into a matter-free one. The notion of continuity, as well as the meaning of the limit, occurring in the above definition, ought to — and could easily — be explained rigorously. But it is not worth while to do this, because of the lack of formulae for the general solutions of equations (3) at present. The method has been applied to spherically symmetric cosmological models, with the restriction, however, that $S(p)$ also should be spherically symmetric. Here the meaning of the above notions is clear. As a matter of fact, most of these models have singular points in space-time (corresponding to the beginning of expansion or to the end of the contraction), probably due to the strict spherical symmetry. But our criterion can be made effective also in these cases, admitting the analogous singularities also in $S(p)$. It has been found, that closed models are allowed, open ones are not. This is in remarkable agreement with the conclusion of WHEELER. On the other side, if we compare our criterion with WHEELER's attempt to regard the MACH—EINSTEIN principle

as a requirement for a satisfactory initial-value problem, we see, that according to our definition the prevalence of matter does not mean the imposition of freedom and variability of the structure of space-time but rather it simply means that in this type of solution the existence of space-time presupposes the existence of matter.

REFERENCES

1. V. Fock, The Theory of Space Time and Gravitation, Pergamon, Press, 1959, p. 369.
2. J. A. Wheeler, Lectures at L'École d'Été de Physique, Les Houches, 1963.
3. F. A. E. Pirani, Helv. Phys. Acta. Suppl., IV., 199, 1956.
4. I. Ozsváth and E. Schücking, An anti-Mach metric, Recent developments in General Relativity, Pergamon Press, 1962.
5. C. Brans and R. H. Dicke, Phys. Rev., **124**, 925, 1961.
6. W. E. Thirring, Annals of Phys, **16**, 96, 1961.
7. R. H. Dicke, Experimental relativity, Lectures at L'École d'Été de Physique, Les Houches, 1963.

ПРИНЦИП МАХА В ОБЩЕЙ ТЕОРИИ ОТНОСИТЕЛЬНОСТИ

Ф. КАРОЛЬХАЗИ

Резюме

После краткого обзора темы показывается, что принцип Маха в интерпретации, близкой к оригинальным соображениям Маха, не является следствием общей теории относительности. Однако можно формулировать такой принцип отбора, который разделяет решения уравнений гравитационного поля на две группы. Члены одной группы противоречат, члены другой группы не противоречат рациональному содержанию требования Маха. Таким образом, может быть, что последние решения останутся разрешимыми и в будущем, когда на основе более глубокого понимания связи пространства-времени с веществом выяснится о других, что они являются экстраполяциями, выходящими за пределы применимости уравнений Эйнштейна. Формулируется принцип отбора, который кажется более общим и гибким, чем остальные.

CLEBSCH-TRANSFORMATIONEN IN DER ALLGEMEINEN RELATIVITÄTSTHEORIE

Von

HANS-GEORG SCHÖPF

INSTITUT FÜR THEORETISCHE PHYSIK DER ERNST–MORITZ–ARNDT-UNIVERSITÄT, GREIFSWALD, DDR

(Eingegangen: 20. IX. 1963)

Es werden CLEBSCH-Transformationen für relativistische ideale Flüssigkeiten hergeleitet. Die erhaltenen Resultate werden für den Fall verallgemeinert, in dem die Flüssigkeit mit einem elektromagnetischen Feld in Wechselwirkung steht. Insbesondere wird ein Variationsprinzip für Dielektrika angegeben.

§ 1. Einleitung

Die Arbeiten von A. CLEBSCH über die Integration der hydrodynamischen Gleichungen liegen über ein Jahrhundert zurück. 1857 veröffentlichte er eine erste Arbeit [1], die sich auf eine ideale, inkompressible und stationäre Flüssigkeit bezog. In einer zweiten Arbeit [2] wird die Bedingung der Stationarität fallengelassen. Die Bedingung der Inkompressibilität kann durch die der Barotropie ersetzt werden. Vereinfachte Darstellungen findet man in dem Lehrbuch von H. LAMB [3], sowie in dem Handbuchartikel von J. SERRIN [4].

Man kann den Gedankengang dieser Überlegungen in vier Schritte gliedern. Erstens stellt man fest, dass ein dreidimensionales kovariantes Vektorfeld v_i die Darstellung

$$v_i = \psi_{,i} + \alpha\mu_{,i} \tag{1.1}$$

mit Skalaren ψ, α, μ gestatten wird.
(In unserer Arbeit nehmen lateinische Indizes die Werte 1, 2, 3, griechische die Werte 1, 2, 3, 4 an. Ein Komma bedeutet die gewöhnliche partielle Ableitung nach den Koordinaten, während durch ein Semikolon die kovariante Ableitung bezeichnet werden wird.)

Um zu unterstreichen, dass die Darstellung (1.1) kovariant ist und nichts mit der Metrik des Raumes zu tun hat, wollen wir auf folgendes hinweisen: Sei $\mathfrak{P}^i$ eine kontravariante orientierte Vektordichte, welche divergenzfrei ist,

$$\mathfrak{P}^i_{,i} = 0\,, \tag{1.2}$$

so wird die allgemeinste, zwei willkürliche Funktionen α und μ enthaltende Lösung von (1.2)

$$\mathfrak{P}^i = \varepsilon^{ijk}\,\alpha_{,j}\,\mu_{,k} \tag{1.3}$$

lauten, wobei ε^{ijk} das Levi-Civita-Symbol ist, welches als orientierte dreistufige Tensordichte aufgefasst werden kann. Ist nun $\mathfrak{P}^i$ die Rotation eines Vektors v_j

$$\mathfrak{P}^i = \varepsilon^{ijk} v_{j,k} \tag{1.4}$$

womit (1.2) identisch erfüllt ist, so kann durch Integration auf (1.1) geschlossen werden.

In der nichtrelativistischen Hydrodynamik wird nun zweitens v_i mit dem Geschwindigkeitsvektor identifiziert. Die obigen Skalare müssen dann die Zeit als Parameter enthalten. Sie können so gewählt werden, dass vermöge der Bewegungsgleichungen die substantiellen zeitlichen Ableitungen von α und μ verschwinden:

$$\frac{d\alpha}{dt} = 0\,, \qquad \frac{d\mu}{dt} = 0\,. \tag{1.5}$$

Zum Beweis dessen bedient man sich am besten der Lagrangeschen Betrachtungsweise.

Drittens können mit Hilfe von (1.1) und (1.5) die Bewegungsgleichungen integriert werden. Viertens kann ein Variationsprinzip für die Theorie angegeben werden. Es findet sich bereits in den Arbeiten von CLEBSCH. Eine neuere Arbeit zu diesem Punkt stammt von H. ITO [5], der den entsprechenden kanonischen Formalismus entwickelt.

In derselben Arbeit wird ausserdem die Voraussetzung der Barotropie durch thermodynamische Betrachtungen etwas abgeschwächt. Es ist dann der geometrische Parameter μ durch die spezifische Entropie η zu ersetzen, während an die Stelle von (1.5) die Relationen

$$\frac{d\alpha}{dt} = T\,, \qquad \frac{d\eta}{dt} = 0 \tag{1.6}$$

treten, wobei T die Temperatur ist. Die letzte Beziehung ist insofern plausibel, als bei Vorhandensein von dissipativen Effekten die Existenz eines Variationsprinzips nicht zu erwarten ist.

Die Benutzung thermodynamischer Relationen scheint für die Entwicklung eines relativistischen Analogons der hier referierten Theorie von vornherein notwendig zu sein. Das bestätigt auch ein Artikel von Z. KOBA [6], in welchem ein Spezialfall der von uns darzulegenden Theorie kurz behandelt wird. Im § 2 formen wir daher die relativistischen Bewegungsgleichungen mittels thermodynamischer Relationen für unsere Zwecke um.

Sodann leiten wir in § 3 CLEBSCH-Transformationen für eine ideale relativistische Flüssigkeit her. Dabei benutzen wir zunächst eine dreidimen-

sionale Betrachtungsweise, indem wir ein mitbewegtes Koordinatensystem verwenden. Diese Koordinaten können als Lagrangesche Koordinaten angesehen werden, so dass ein enger Zusammenhang mit der nichtrelativistischen Verfahrensweise besteht.

Auf Fragen der Integration der Bewegungsgleichungen gehen wir nicht ein, sondern wenden uns in § 4 dem Variationsprinzip zu. Bei diesem werden neben anderen Grössen die kontravarianten Komponenten der Vierergeschwindigkeit u^μ, der metrische Tensor $g_{\mu\nu}$ sowie ψ, α, η unabhängig voneinander variiert. Diese Methode unterscheidet sich von einer anderen, der »Variation der Weltlinien«, über die wir an anderer Stelle berichtet haben [7]. Sie sieht ausser der Variation der Weltlinien (und eventuell einer thermodynamischen Grösse α) nur die von $g_{\mu\nu}$ vor. Daher wird hier unter der letztgenannten Variation u^μ nicht konstant zu halten sein. Während das aus den CLEBSCH-Transformationen fliessende Variationsprinzip auf ideale Flüssigkeiten beschränkt ist, kann die Methode der Variation der Weltlinien für allgemeine konservative mechanische Kontinua angewandt werden.

In den letzten beiden Paragraphen verallgemeinern wir unsere Ergebnisse für den Fall, dass eine Kopplung der Flüssigkeit mit einem elektromagnetischen Feld besteht. Dabei wird insbesondere eine Bestätigung und Rechtfertigung der Untersuchungen von K. F. NOVOBÁTZKY [8] erfolgen.

§ 2. Eine Umformung der relativistischen Bewegungsgleichungen

Die Bewegungsgleichungen für eine relativistische ideale Flüssigkeit,

$$T^\nu_{\mu;\nu} = 0\,, \tag{2.1}$$

$$T^\nu_\mu = (w + p)u^\nu u_\mu + \delta^\nu_\mu p\,, \tag{2.2}$$

formen wir mit Hilfe einiger thermodynamischer Grössen und Beziehungen um. Bezüglich ihrer näheren Erläuterung für die relativistische Theorie verweisen wir auf eine frühere Arbeit (l. c. [7]).

Zunächst konzipieren wir eine Dichte ϱ, welche der Kontinuitätsgleichung

$$(\varrho u^\nu)_{,\nu} = 0 \tag{2.3}$$

genügt.

$$\frac{1}{\varrho} \equiv v \tag{2.4}$$

ist das relativistische Analogon zum spezifischen Volumen. Die spezifische

Energie Φ genügt den Relationen

$$\varrho \Phi = w , \tag{2.5}$$

$$\frac{\partial \Phi}{\partial v} = - p , \tag{2.6}$$

$$\frac{\partial \Phi}{\partial \eta} = T . \tag{2.7}$$

Dabei ist η die spezifische Entropie und T die Temperatur. Schliesslich verwenden wir die spezifische Enthalpie

$$h = \Phi + pv , \tag{2.8}$$

für welche sich gemäss (2.6), (2.7), (2.4)

$$h_{,\mu} = \frac{p_{,\mu}}{\varrho} + T\eta_{,\mu} \tag{2.9}$$

ergibt. ϱ, Φ, η und T werden als skalare Grössen angesehen. Es folgt sodann

$$T^{\nu}_{\mu} = \varrho u^{\nu} h u_{\mu} + \delta^{\nu}_{\mu} p , \tag{2.10}$$

und mit (2.3) und

$$\frac{D(\ldots)}{ds} \equiv (\ldots)_{;\nu} u^{\nu} \tag{2.11}$$

die Beziehung

$$\varrho \frac{D u_{\mu}}{ds} + p_{,\mu} = 0 . \tag{2.12}$$

Unter Verwendung von (2.9) erhalten wir schliesslich

$$\frac{D u_{\mu}}{ds} = - h_{,\mu} + T\eta_{,\mu} = \tag{2.13}$$

$$= h \frac{D u_{\mu}}{ds} + u_{\mu} \frac{Dh}{ds} .$$

Überschieben wir diesen Ausdruck mit u^{μ}, so erkennen wir, dass

$$\frac{D\eta}{ds} = 0 \tag{2.14}$$

gilt. Das Fehlen einer Entropieerzeugung ist also eine Folge unserer Voraussetzungen für das diskutierte Medium.

Mit Hilfe des spezifischen Gibbsschen Potentials

$$\zeta = h - T\eta$$

kann (2.13) in der Form

$$\frac{D\zeta u_\mu}{ds} + \eta \frac{DTu_\mu}{ds} = -\zeta_{,\mu} - \eta T_{,\mu}$$

geschrieben werden. Die Theorie von KOBA [6] basiert auf der speziellen Annahme, dass diese Beziehung in

$$\frac{D\zeta u_\mu}{ds} = -\zeta_{,\mu}, \qquad \frac{DTu_\mu}{ds} = -T_{,\mu}$$

aufspaltet.

(2.13) kann mit der Abkürzung

$$U_\varrho \equiv h u_\varrho \tag{2.15}$$

auch in die Gestalt

$$(U_{\mu,\varrho} - U_{\varrho,\mu}) u^\varrho = T\eta_{,\mu} \tag{2.16}$$

gebracht werden, indem man nämlich beachtet, dass

$$\begin{aligned} -h_{,\mu} &= -h_{;\mu} = h_{;\mu}\, u_\varrho\, u^\varrho = \\ &= h_{;\mu}\, u_\varrho\, u^\varrho + h u_{\varrho;\mu}\, u^\varrho = \\ &= (h u_\varrho)_{;\mu}\, u^\varrho = U_{\varrho;\mu}\, u^\varrho \end{aligned}$$

und

$$U_{\mu;\varrho} - U_{\varrho;\mu} = U_{\mu,\varrho} - U_{\varrho,\mu}$$

ist.

§ 3. Clebsch-Transformationen für relativistische Flüssigkeiten

Wir wählen zunächst ein mitbewegtes Koordinatensystem $a^\varkappa$, in welchem die Weltlinien der Flüssigkeit durch $a^k = \text{const}$, $a^4 = \text{variabel}$, beschrieben werden. Grössen in diesem System kennzeichnen wir durch einen Querstrich. Es gilt (1. c. [7])

$$\bar{u}^\alpha = \frac{\delta_4^\alpha}{\sqrt{-\bar{g}_{44}}}, \qquad \bar{u}_\beta = \frac{\bar{g}_{\beta}}{\sqrt{-\bar{g}_{44}}}, \tag{3.1}$$

insbesondere

$$\bar{u}_4 = -1/\bar{u}_4 = -\sqrt{-\bar{g}_{44}}. \tag{3.1a}$$

Die Gleichung (2.16) schreibt sich hier

$$\begin{aligned} \bar{T}\,\bar{\eta}_{,r} &= \bar{u}^4 (\bar{U}_{r,4} - \bar{U}_{4,r}) = \\ &= \bar{u}^4 \bar{U}_{r,4} + \bar{h}_{,r} + \bar{h}\,\bar{u}_4\,\bar{u}^4{}_{,r}. \end{aligned} \tag{3.2}$$

Wir fassen ferner eine bestimmte Anfangsmannigfaltigkeit ins Auge, $a^4 = a_0^4$, und bezeichnen auf ihr erklärte Grössen mit einem zweifachen Querstrich. Den dreidimensionalen kovarianten Vektor $\overline{\overline{U}}_r$ denken wir uns gemäss (1.1) dargestellt:

$$\overline{\overline{U}}_r = \overline{\overline{\psi}}_{,r} + \overline{\overline{\alpha}}\,\overline{\overline{\mu}}_{,r}. \tag{3.3}$$

Hierbei wird also a_0^4 als Parameter eingehen. Wir differenzieren nach diesem Parameter und erhalten nach leichter Umformung:

$$\overline{\overline{u}}^4\,\overline{\overline{U}}_{r,4} = (\overline{\overline{\psi}}_{,4}\,\overline{\overline{u}}^4)_{,r} - \overline{\overline{\psi}}_{,4}\,\overline{\overline{u}}^4{}_{,r} + \\ + \overline{\overline{\alpha}}_{,4}\,\overline{\overline{u}}^4\,\overline{\overline{\mu}}_{,r} + \overline{\overline{\alpha}}\,\overline{\overline{\mu}}_{,r4}\,\overline{\overline{u}}^4. \tag{3.4}$$

Gehen wir hiermit in (3.2) für $a^4 = a_0^4$ ein, so sehen wir, dass die Feldgleichungen auf der Anfangsmannigfaltigkeit erfüllt sind, wenn wir auf ihr folgende Wahl treffen:

$$\overline{\overline{\mu}} = \overline{\overline{\eta}} = \overline{\eta}. \tag{3.5}$$

(Es ist nämlich nach (2.14) $\overline{\eta} = \overline{\eta}\ (a^1, a^2, a^3) = \overline{\overline{\eta}}$.)

$$\overline{\overline{\alpha}}_{,4}\,u^4 = \overline{\overline{T}}, \tag{3.6}$$

$$\overline{\overline{\psi}}_{,4}\,\overline{\overline{u}}^4 = -\,\overline{\overline{h}}, \tag{3.7}$$

woraus

$$\overline{\overline{\psi}}_{,4}\,\overline{\overline{h}}\,\overline{\overline{u}}_4 = \overline{\overline{U}}_4 \tag{3.8}$$

folgt.

Nunmehr können wir schliessen, dass diese Darstellung für alle a^4 benutzt werden kann. Es ist nämlich nach (3.2), (3.1a)

$$\frac{\partial \overline{U}_r}{\partial a^4} = -\,\overline{\eta}_{,r}\,\overline{T}\,\overline{u}_4 + (\overline{h}\,\overline{u}_4)_{,r} = \\ = \overline{\eta}_{,r}\sqrt{-\bar{g}_{44}}\,\overline{T} - (\sqrt{-\bar{g}_{44}}\,\bar{h})_{,r},$$

woraus durch Integration

$$\overline{U}_r = \overline{\overline{U}}_r - \overline{\eta}_{,r}\int_{a_0^4}^{a^4} \overline{T}\,\overline{u}_4\,da^4 + \\ + \frac{\partial}{\partial a^r}\left(\int_{a_0^4}^{a^4} \bar{h}\,\bar{u}_4\,da^4\right) = \tag{3.9}$$

$$= \frac{\partial}{\partial a_r}\left(\overline{\overline{\psi}} - \int_{s(a_0^4)}^{s(a^4)} \bar{h}\,ds\right) + \overline{\eta}_{,r}\left(\overline{\alpha} + \int_{a\,(a_0^4)}^{s(a^4)} \overline{T}\,ds\right) \equiv \overline{\psi}_{,r} + \overline{\alpha}\,\overline{\eta}_{,r}$$

folgt. Da ausserdem eine Beziehung der Form (3.8) für beliebige a^4 gilt, kann

$$\overline{U}_\alpha = \overline{\psi}_{,\alpha} + \overline{\alpha}\,\overline{\eta}_{,\alpha}$$

und wegen der Kovarianz dieser Gleichung allgemein

$$U_\alpha = \psi_{,\alpha} + \alpha\eta_{,\alpha} \tag{3.10}$$

geschrieben werden, wobei

$$\frac{D\psi}{ds} = -h, \qquad \frac{D\alpha}{ds} = T \tag{3.11}$$

ist.

§ 4. Das Variationsprinzip

Die Grundgleichungen der hier diskutierten Theorie, einschliesslich der Einsteinschen Feldgleichungen lassen sich nun aus folgendem Variationsprinzip herleiten:

$$\delta \int \sqrt{-g}\left(\varrho K - p + \frac{1}{2\varkappa} R\right) d^4x = 0. \tag{4.1}$$

Hierbei ist R der Riemannsche Krümmungsskalar, $\varkappa$ die Einsteinsche Gravitationskonstante und

$$K \equiv -\frac{h}{2} g_{\mu\nu} u^\mu u^\nu + u^\mu (\psi_{,\mu} + \alpha\eta_{,\mu}) + \frac{h}{2}. \tag{4.2}$$

Unabhängig voneinander variiert werden die Grössen $g_{\mu\nu}$, u^μ, ψ, α, η und p. Die Beziehung $u_\mu u^\mu = -1$ wird nicht vorausgesetzt, sondern soll aus dem Variationsprinzip folgen. ϱ und h werden als Funktionen von p und η angesehen, wobei wir nur die Gültigkeit der thermodynamischen Relationen

$$\frac{\partial h}{\partial p} = \frac{1}{\varrho}, \qquad \frac{\partial h}{\partial \eta} = T \tag{4.3}$$

voraussetzen [s. (2,9)].

Die Variation bezüglich ψ ergibt die Kontinuitätsgleichung

$$(\varrho u^\mu)_{;\mu} = 0. \tag{4.4}$$

Variieren wir bezüglich α, sok ommt

$$\eta_{,\mu} u^\mu = \frac{D\eta}{ds} = 0. \tag{4.5}$$

Die Variation nach u^μ liefert

$$-hu_\mu + \psi,_\mu + \alpha\eta,_\mu = 0. \tag{4.6}$$

Hiermit kann für K

$$K = \frac{1}{2} h(u_\mu u^\mu + 1) \tag{4.7}$$

geschrieben werden. Wir variieren jetzt bezüglich p und erhalten zunächst

$$\frac{\partial \varrho}{\partial p} K - \varrho \frac{\partial h}{\partial p}\left(\frac{1}{2} u_\mu u^\mu - \frac{1}{2}\right) - 1 = 0. \tag{4.8}$$

Unter Verwendung von (4.3) und (4.7) ergibt sich

$$(u_\mu u^\mu + 1)\left(\frac{\partial \varrho}{\partial p} \cdot \frac{h}{2} - \frac{1}{2}\right) = 0,$$

woraus wir auf

$$u_\mu u^\mu = -1 \tag{4.9}$$

schliessen, da der zweite Faktor nicht allgemein verschwinden wird. Bei Erfüllung der Feldgleichungen ist also nach (4.7)

$$K = 0. \tag{4.10}$$

Schliesslich variieren wir bezüglich η. Das Ergebnis ist

$$\frac{\partial \varrho}{\partial \eta} K - \frac{1}{2} \varrho \frac{\partial h}{\partial \eta}(u_\mu u^\mu - 1) - (\varrho u^\mu \alpha)_{;\mu} = 0.$$

Bei Beachtung von (4.10), (4.9), (4.3) und (4.4) resultiert

$$T = u^\mu \alpha,_\mu = \frac{D\alpha}{ds}. \tag{4.11}$$

Um noch die Variation nach den $g_{\mu\nu}$ auszuführen, beachten wir die Beziehungen

$$\frac{\delta\sqrt{-g}}{\delta g_{\mu\nu}} = \frac{1}{2}\sqrt{-g}\, g^{\mu\nu}, \tag{4.12}$$

$$\frac{\delta\sqrt{-g}\, R}{\delta g_{\mu\nu}} = -\sqrt{-g}\left(R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R\right).$$

Es ergibt sich dann

$$0 = -\frac{\sqrt{-g}}{2}\left\{(-\varrho K + p) g^{\mu\nu} + \varrho h u^\mu u^\nu + \frac{1}{\varkappa}\left(R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R\right)\right\},$$

d. h. wegen (4.10) erscheinen die Einsteinschen Feldgleichungen

$$R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R = - \varkappa (\varrho h u^{\mu} u^{\nu} + p g^{\mu\nu}). \tag{4.13}$$

§ 5. Elektrisch geladene Flüssigkeit

Wir verallgemeinern die Theorie jetzt für den Fall, dass die Flüssigkeit elektrisch geladen ist, und anstelle von (2.1) die Beziehung

$$T^{\nu}_{\mu;\nu} = \varrho^{el.} F_{\mu\varrho} u^{\varrho} \tag{5.1}$$

mit dem elektromagnetischen Feldstärkentensor $F_{\mu\varrho}$ und der elektrischen Ladungsdichte $\varrho^{el.}$ gilt. Behandeln wir die linke Seite wie in § 2, so werden wir anstatt auf (2.13) auf

$$\frac{D h u_{\mu}}{ds} = - h_{,\mu} + T \eta_{,\mu} + \frac{\varrho^{el.}}{\varrho} F_{\mu\varrho} u^{\varrho} \tag{5.2}$$

geführt, woraus wegen der Antisymmetrie von $F_{\mu\varrho}$ wieder $D\eta/ds = 0$ folgt. Wir setzen

$$\frac{\varrho^{el.}}{\varrho} \equiv \frac{e}{m}, \tag{5.3}$$

und gehen vermöge

$$F_{\mu\varrho} = A_{\varrho,\mu} - A_{\mu,\varrho} \tag{5.4}$$

zum Viererpotential A_{μ} über. Aus (5.2) ergibt sich dann als Analogon zu (2.16):

$$(U_{\mu,\varrho} - U_{\varrho,\mu}) u^{\varrho} = \frac{e}{m} (A_{\varrho,\mu} - A_{\mu,\varrho}) u^{\varrho} + T \eta_{,\mu}. \tag{5.5}$$

Es gelten die Maxwellschen Gleichungen

$$F^{\mu\nu}{}_{;\nu} = \varrho^{el.} u^{\mu}, \tag{5.6}$$

woraus die Kontinuitätsgleichung für die Ladung

$$(\varrho^{el.} u^{\mu})_{;\mu} = 0 \tag{5.7}$$

folgt, und daraus wiederum zusammen mit (2.3)

$$\left(\frac{e}{m}\right)_{,\mu} u^{\mu} = \frac{D e/m}{ds} = 0. \tag{5.8}$$

Für unsere Zwecke ist es nützlich, die Eichung

$$A_\mu u^\mu = 0 \tag{5.9}$$

zu wählen, d. h. $A_4 = 0$. Das ist immer erreichbar. Ist nämlich $\bar{A}_4^+$ zunächst von Null verschieden, so führe man mit

$$B = \int_{a_0^4}^{a^4} \bar{A}_4^+ \, da^4 = B(a^1, a^2, a^3, a^4)$$

die Umeichung

$$\bar{A}_\mu = \bar{A}_\mu^+ - \frac{\partial B}{\partial a^\mu}$$

durch, womit das Gewünschte erreicht ist.

(5.8) und (5.9) gestatten es nun, in (5.5) auf der rechten Seite e/m unter die Differentiation zu ziehen. Mit

$$V_\mu \equiv h u_\mu + \frac{e}{m} A_\mu \tag{5.10}$$

kann nun

$$(V_{\mu,\varrho} - V_{\varrho,\mu}) u^\varrho = T\eta_{,\mu} \tag{5.11}$$

geschrieben werden in völliger Analogie zu (2.16).

Die Überlegungen des § 3 können sodann wiederholt werden und führen zu dem Schluss, dass die Darstellung

$$V_\mu = \psi_{,\mu} + \alpha\eta_{,\mu} \tag{5.12}$$

mit unveränderter Bedeutung von ψ und α (s.(3.11)) möglich ist.

Die erweiterte Theorie besitzt folgendes Variationsprinzip:

$$\delta \int \sqrt{-g} \left\{ \varrho M - p + \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2\varkappa} R \right\} d^4x = 0, \tag{5.13}$$

wobei

$$M \equiv -\left(\frac{h}{2} g_{\mu\nu} u^\nu + \frac{e}{m} A_\mu \right) u^\mu + u^\mu (\psi_{,\mu} + \alpha\eta_{,\mu}) + \frac{h}{2} \tag{5.14}$$

ist. Ausser den bei (4.2) angegebenen Grössen werden jetzt auch noch e/m und die A_μ variiert. Diese Variationen liefern die Eichgleichung (5.9) und, wie bekannt ist, die Maxwellschen Gleichungen (5.6), wenn wir (5.3) als Definitionsgleichung für $\varrho^{el.}$ benutzen. Bei der Variation von ψ und α wirkt sich das in M im Vergleich zu K auftretende Zusatzglied $e/m \, A_\mu u^\mu$ nicht aus, so dass dieselben Ergebnisse (4.4) und (4.5) resultieren. Die Variation bezüglich

u^μ ergibt jetzt

$$-hu_\mu - \frac{e}{m} A_\mu + \psi_{,\mu} + \alpha\eta_{,\mu} = 0, \tag{5.15}$$

wie es sein soll. [s. (5,12), (5,10).] Ist (5.15) und (5.9) erfüllt, so kommt in Analogie zu (4.7)

$$M = \frac{1}{2} h(u_\mu u^\mu + 1),$$

was zur Folge hat, dass auch die Variationen bezüglich p und η dieselben Resultate wie in § 4 zeitigen, nämlich (4.9) und (4.11). Ebenso wird bei der Variation nach den $g_{\mu\nu}$ zusätzlich nur der Term $\sqrt{-g}\, F_{\mu\nu} F^{\mu\nu}$ zum Ergebnis beitragen. Dieser Beitrag ist bekanntlich

$$\frac{1}{4} \frac{\delta\sqrt{-g}\, F_{\alpha\beta} F^{\alpha\beta}}{\delta g_{\mu\nu}} = -\frac{1}{2}\sqrt{-g}\left(F^\mu_\varrho F^{\nu\varrho} - \frac{1}{4} g^{\mu\nu} F_{\alpha\beta} F^{\alpha\beta}\right).$$

Daher lauten die Einsteinschen Feldgleichungen jetzt:

$$-\varkappa\left(\varrho h u^\mu u^\nu + p g^{\mu\nu} + F^\mu_\varrho F^{\nu\varrho} - \frac{1}{4} g^{\mu\nu} F_{\alpha\beta} F^{\alpha\beta}\right) = R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R. \tag{5.16}$$

Wir wollen zur Probe bestätigen, dass aus den Gleichungen, die aus dem Variationsprinzip fliessen, tatsächlich die Ausgangsgleichungen (5.2) reproduzierbar sind. Zu diesem Zwecke verweisen wir auf nachfolgende Rechnung, auf deren Kommentierung wir verzichten können:

$$\begin{aligned}
\frac{D(hu_\mu + e/m\, A_\mu)}{ds} &= \psi_{;\mu\alpha} u^\alpha + \alpha_{,\alpha} u^\alpha \eta_{,\mu} + \alpha\eta_{;\mu\alpha} u^\alpha = \\
&= -h_{,\mu} + T\eta_{,\mu} - (\psi_{,\alpha} + \alpha\eta_{,\alpha}) u^\alpha{}_{;\mu} = \\
&= -h_{,\mu} + T\eta_{,\mu} - \frac{e}{m} A_\alpha u^\alpha{}_{;\mu} - hu_\alpha u^\alpha{}_{;\mu} = \\
&= -h_{,\mu} + T\eta_{,\mu} + \frac{e}{m} A_{\alpha;\mu} u^\alpha, \\
\frac{Dhu_\mu}{ds} &= -h_{,\mu} + T\eta_{,\mu} + \frac{e}{m}(A_{\alpha,\mu} - A_{\mu,\alpha}) u^\alpha.
\end{aligned}$$

Es sei noch darauf hingewiesen, dass das nichtrelativistische Analogon dieser Theorie der geladenen Flüssigkeit in einer Arbeit F. MÖGLICH [9] zu finden ist.

§ 6. Variationsprinzip für Dielektrika

Die in der Einleitung erwähnten Untersuchungen NOVOBÁTZKYS [8] basieren auf folgender Überlegung: Die kovarianten Materialgleichungen eines isotropen Dielektrikums lauten:

$$G^{\mu\nu} = \frac{1}{\mu}\left\{F^{\mu\nu} + k(u^{\mu} F^{\nu} - u^{\nu} F^{\mu})\right\}. \tag{6.1}$$

Hierbei ist μ die Permeabilität. (Eine Verwechslung mit der am Anfang dieser Arbeit benutzten Grösse μ ist nicht zu befürchten.) Ferner ist

$$F^{\nu} \equiv F^{\nu\varrho}\, u_{\varrho}, \tag{6.2}$$

$$k \equiv \varepsilon\mu - 1\,, \tag{6.3}$$

mit ε = Dielektrizitätskonstante. Die Maxwellschen Gleichungen sind — neben (5.4) — in Abwesenheit von Ladungen und Strömen

$$G^{\mu\nu}{}_{;\nu} = 0. \tag{6.4}$$

Lassen sich nun diese Gleichungen aus einem Variationsprinzip in der Form

$$\frac{\delta\sqrt{-g}\, L^{el.}}{\delta A_{\mu}} = 0 \tag{6.5}$$

herleiten, so sollte der elektromagnetische Energie-Impulstensor aus

$$\frac{\delta\sqrt{-g}\, L^{el.}}{\delta g_{\mu\nu}} = -\frac{1}{2}\sqrt{-g}\, T^{el.\,\mu\nu} \tag{6.6}$$

abzulesen sein.

Gegen diese Argumentation liesse sich einwenden [10], dass eine derartige Bestimmung des Energie-Impulstensors eigentlich nur für abgeschlossene Systeme vom Standpunkt der allgemeinen Relativitätstheorie zu rechtfertigen ist. In unserem Fall ist das elektromagnetische Feld aber nicht abgeschlossen. Es befindet sich vielmehr in Wechselwirkung mit der phänomenologischen Materie. Man sollte daher ein Variationsprinzip angeben, aus welchem nicht nur die Maxwellschen Gleichungen (6.4), sondern durch Variation von Grössen, welche für die phänomenologische Materie massgebend sind, auch deren Bewegungsgleichungen

$$T^{\nu}_{\mu;\nu} = f_{\mu} \tag{6.7}$$

fliessen. Die Variationsableitung der gesamten Lagrangedichte nach $g_{\mu\nu}$ sollte dann nach dem bekannten Schema den (divergenzfreien) Energie-Impuls-

tensor des abgeschlossenen, aus elektromagnetischem Feld und phänomenologischer Materie bestehenden Systems ergeben.

Zur Konstruktion einer solchen Lagrangedichte müsste eigentlich die ponderomotorische Kraft f_μ bekannt sein, während man sie doch anderseits erst vermöge

$$T^{el.\ \nu}_{\mu;\nu} = -f_\mu \tag{6.8}$$

zu bestimmen wünscht.

Nun wird aber, wie NOVOBÁTZKY l. c. [8] gezeigt hat, (6.5) mit (6.4) identisch, wenn man

$$L^{el.} = \frac{1}{2\mu}\left(\frac{1}{2} F_{\alpha\beta} F^{\alpha\beta} - k F_\alpha F^\alpha\right) \tag{6.9}$$

wählt. Die gesamte Lagrangedichte für das abgeschlossene System kann also ausser (6.9) nur solche von A_μ abhängigen Terme Z enthalten, deren Variationsableitungen nach A_μ modulo derjenigen Feldgleichungen verschwinden, welche aus der Variation der sich auf die phänomenologische Materie beziehenden Feldgrössen fliessen. Die von A_μ freien Terme müssen offenbar mit der Lagrangedichte für die kräftefreie phänomenologische Materie identisch sein.

Ist nun die phänomenologische Materie eine ideale Flüssigkeit, wie wir sie in dieser Arbeit behandelt haben, so dürfen wir die Gültigkeit der Feldgleichungen (4.4), (4.5) und (4.9) annehmen, während (4.6) in noch unbekannter Weise modifiziert werden muss. Das soll nun so geschehen, indem in dem Ansatz

$$L = \varrho K - p + L^{el.} + Z \tag{6.10}$$

Z so gewählt wird, dass (4.4), (4.5), (4.9) und (6.4) sicher aus dem Variationsprinzip folgen. Und das wiederum ist nun mit

$$Z \equiv -\frac{k}{2\mu}(1 + g_{\mu\nu} u^\mu u^\nu) F_\alpha F^\alpha , \tag{6.11}$$

d. h.

$$L^{el.} + Z = \frac{1}{2\mu}\left\{\frac{1}{2} F_{\alpha\beta} F^{\alpha\beta} - k(2 + u_\mu u^\mu) F_\alpha F^\alpha\right\} \tag{6.12}$$

der Fall.

Wie in § 4 ergeben die Variationen nach ψ und α die Beziehungen (4.4) und (4.5). Anstelle von (4.6) ergibt sich aber

$$\varrho(-h u_\mu + \psi_{,\mu} + \alpha\eta_{,\mu}) - \frac{k}{\mu}(F^\alpha F_\alpha u_\mu + F_{\alpha\mu} F^\alpha [2 + u_\lambda u^\lambda]) = 0 . \tag{6.13}$$

Für K gilt somit nach (4.2) bei Erfüllung von (6.13)

$$K = (u_\mu u^\mu + 1)\left(\frac{1}{2} h + \frac{2k}{\varrho\mu} F_\alpha F^\alpha\right) . \tag{6.14}$$

Hiermit und mit Hilfe der unverändert gültigen Formel (4.8) kann aber wie in § 4 auf (4.9) geschlossen werden. Da nun bei Erfüllung der Feldgleichungen $K = 0$, $Z = 0$ ist, ergeben die noch ausstehenden Variationen nach η und A_μ die Beziehungen (4.11) und (6.4). Die Variation von (6.12) nach $g_{\mu\nu}$ führt nun gemäss

$$\frac{\delta\sqrt{-g}\,(L^{el.} + Z)}{\delta g_{\mu\nu}} = -\frac{\sqrt{-g}}{2}\,T_A{}^{\mu\nu} \tag{6.15}$$

zu dem Abrahamschen Ausdruck für den elektromagnetischen Energie-Impulstensor:

$$T_A{}^{\mu\nu} = \frac{1}{\mu}\left\{F^{\mu\lambda}F^{\nu}{}_{\lambda} - \frac{g^{\mu\nu}}{4}F_{\alpha\beta}F^{\alpha\beta} + k\left(F_\lambda F^\lambda u^\mu u^\nu - F^\mu F^\nu + \frac{1}{2}g^{\mu\nu}F_\lambda F^\lambda\right)\right\}, \tag{6.16}$$

wie bereits NOVOBÁTZKY l. c. [8], gezeigt hat. Die Einsteinschen Feldgleichungen lauten daher:

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = -\varkappa(\varrho h u^\mu u^\nu + p g^{\mu\nu} + T_A{}^{\mu\nu}) \,. \tag{6.17}$$

NOVOBÁTZKY ist in seiner Arbeit von (6.6) zu (6.15) übergegangen, da (6.6) zu einem nicht akzeptablen elektromagnetischen Energie-Impulstensor führt. Das ist freilich nur dann der Fall, wenn man das kontravariante u^λ bei der Variation nach $g_{\mu\nu}$ konstant hält. Verwendet man die in der Einleitung erwähnte Methode der Variation der Weltlinien, bei der $\delta u^\lambda/\delta g_{\mu\nu} \neq 0$ ist, so erhält man aus (6.6) für $T^{el.\mu\nu}$ ebenfalls $T_A{}^{\mu\nu}$. (1. c. [10]). Anderseits entspricht die voneinander unabhängige Variation von u^λ und $g_{\mu\nu}$ genau dem Variationsprinzip für ideale Flüssigkeiten, wie es sich aus den CLEBSCH-Transformationen ergibt. In diesem Fall aber wird die Annahme des NOVOBÁTZKYschen Zusatzterms Z gerade dadurch gerechtfertigt, dass er in der Lagrangefunktion L (6.15) für das abgeschlossene System zur Erfüllung der notwendig zu fordernden Relation $u_\mu u^\mu = -1$ führt.

Um die Konsistenz unserer Theorie zu erhärten, wollenwir noch die Bewegungsgleichungen für die Flüssigkeit ausrechnen. Nach (6.13), (4.9), (4.5), (4.11), sowie der dann aus (6.13) durch Überschieben mit n^μ unverändert erscheinenden Formel $\psi_{,\mu}u^\mu = -h$ folgt:

$$\begin{aligned}\frac{Dhu_\mu}{ds} = &- h_{,\mu} + T\eta_{,\mu} - (\psi_{,\nu} + \alpha\eta_{,\nu})u^\nu{}_{;\mu} + \\ &+ \frac{k}{\mu\varrho^2}\varrho_{,\nu}u^\nu(F^\alpha F_\alpha u_\mu + F_{\alpha\mu}F^\alpha) - \\ &- \frac{k}{\mu\varrho}(F^\alpha F_\alpha u_\mu + F_{\alpha\mu}F^\alpha)_{;\mu}u^\nu .\end{aligned}$$

Da gemäss (4.4)

$$\varrho_{,\nu} u^{\nu} = -\varrho u^{\nu}{}_{;\nu}$$

und gemäss (6.13), (4.9)

$$-(\psi_{,\nu} + \alpha\eta_{,\nu}) u^{\nu}{}_{;\mu} = -\frac{k}{\mu\varrho} F_{\alpha\nu} F^{\alpha} u^{\nu}{}_{;\mu}$$

ist, haben wir

$$\frac{Dhu_{\mu}}{ds} = -h_{,\mu} + T\eta_{,\mu} - \frac{k}{\mu\varrho} F_{\alpha\nu} F^{\alpha} u^{\nu}{}_{;\mu} -$$

$$- \frac{k}{\mu\varrho} (F^{\alpha} F_{\alpha} u_{\mu} u^{\nu} + F_{\alpha\mu} F^{\alpha} u^{\nu})_{;\nu},$$

und schliesslich nach (2.9) und (4.4)

$$(\varrho h u^{\nu} u_{\mu} + \delta^{\nu}_{\mu} p)_{;\nu} = -\frac{k}{\mu} F_{\alpha\nu} F^{\alpha} u^{\nu}{}_{;\mu} - \frac{k}{\mu} (F^{\alpha} F_{\alpha} u_{\mu} u^{\nu} + F_{\alpha\mu} F^{\alpha} u^{\nu})_{;\nu}, \qquad (6{,}18)$$

wobei die rechst stehende Kraftdichte in der Tat die durch die negative Divergenz des Abraham-Tensors definierte ist.

LITERATUR

1. A. Clebsch, J. reine angew. Math., **54**, 293, 1857.
2. A. Clebsch, J. reine angew. Math., **56**, 1, 1859.
3. H. Lamb, Hydrodynamics, 6. edit. Cambridge, 1932; §167.
4. J. Serrin, Mathematical Principles of Classical Fluid Mechanics, in Flügges Handbuch der Physik, VIII/1, Springer, 1959; § 29.
5. H. Ito, Prog. Theor. Phys., **9**, 117, 1953.
6. Z. Koba, Prog. Theor. Phys., **14**, 488, 1955.
7. H.—G. Schöpf, Ann. Physik, Leipzig, im Druck.
8. K. F. Novobátzky, Hung. Acta Phys., **1**, 25, 1949.
9. F. Möglich, Zur Hydrodynamik eines Elektronengases, Sitzungsber. DAW, Berlin, 1957.
10. H.—G. Schöpf, Ann. Physik, Leipzig, **9**, 301, 1962.

ПРЕОБРАЗОВАНИЯ КЛЕБША В ОБЩЕЙ ТЕОРИИ ОТНОСИТЕЛЬНОСТИ

Г. Г. ШЕПФ

Резюме

Устанавливаются преобразования Клебша для простой жидкости в общей теории относительности. Полученные результаты распространяются для жидкости, взаимодействующей с электромагнитным полем. Главным образом истолкуется вариационный принцип диэлектриков.

EINIGE BEMERKUNGEN ZUR THEORIE DER SPINOREN UND BISPINOREN IN DER GEKRÜMMTEN RAUM-ZEIT

Von

ERNST SCHMUTZER

THEORETISCH-PHYSIKALISCHES INSTITUT DER UNIVERSITÄT JENA, JENA, DDR

(Eingegangen 20. IX. 1963)

Das Postulat der Lorentzinvarianz der Diracschen Gleichung führt zu zwei Transformationsversionen: 1. Konstanz der Dirac-Matrizen und daher das wohlbekannte Transformationsgesetz für Spinoren. 2. Auffassung der Dirac-Matrizen als Tensoren und Interpretation der Wellenfunktion als einer Art von Invarianten. In der vorliegenden Arbeit wird die Sachlage für die gekrümmte Raum-Zeit untersucht. Zu diesem Zweck wird die Theorie der Spinoren und der Bispinoren in der gekrümmten Raum-Zeit kurz dargestellt.

§ 1. Einführung

Gemäss dem allgemeinen Relativitätsprinzip darf es zur Beschreibung eines physikalischen Sachverhaltes kein Verbot bestimmter Koordinaten geben, d. h. die physikalischen Grundlagen einer Theorie müssen allgemein-kovariant formulierbar sein. Aus diesem Grund ist man seit Jahrzehnten bemüht, dem Spinorkalkül eine allgemein-kovariante Form zu geben, denn die Existenz von Spin-Elektronen in der gekrümmten Welt ist eine Realität, die beschrieben werden muss. Dennoch ist die Ausarbeitung der Theorie, die von FOCK, IWANENKO, INFELD, VAN DER WAERDEN, SCHRÖDINGER, BARGMANN begonnen wurde und mit der sich in den letzten Jahren ausserordentlich viele Physiker wieder beschäftigt haben, noch keineswegs zu einem allgemeinen Abschluss gekommen. Wir verzichten hier auf die Zitierung all dieser Arbeiten, sondern verweisen lediglich auf einige eigene Arbeiten zu diesem Problemkreis, in welchen der Leser weitere Hinweise findet [1]. Ein allgemein-kovarianter Spinorapparat ist vor allem auch für die Theorie der Elementarteilchen von grosser Bedeutung, die bekanntlich auf der Grundlage der Spinoren aufzubauen ist.

§ 2. Skizzierung des kovarianten Spinorkalküls in der gekrümmten Raum-Zeit

Analog zu dem Transformationsgesetz

$$T_{m'} = A^{m}_{m'} T_{m} \qquad \left(A^{m}_{m'} = \frac{\partial x^{m}}{\partial x^{m'}}\right) \tag{1}$$

für Tensoren, wird für Spinoren das Transformationsgesetz

$$\chi_{M'} = A^{M}_{M'}\,\chi_{M} \tag{2}$$

gefordert (kleine lateinische Indizes beziehen sich auf den Tensorraum und laufen von 1 bis 4; grosse lateinische Indizes beziehen sich auf den Spinorraum und laufen von 1 bis 2; Punktierung von Spinorindizes bedeutet komplexe Konjugation). Dabei soll $A^{M}_{M'}$ den Transformationseffekt beinhalten, der durch die Koordinatentransformation $(A^{m}_{m'})$ induziert wird. Eine Aussage über die Struktur von $A^{M}_{M'}$ können wir vorerst nicht machen.

Das Analogon zum metrischen Tensor g_{mn} der Tensortheorie ist der in diesem Fall antisymmetrische metrische Spinor $\gamma_{MN} = -\gamma_{NM}$, mit dem die Spinorindizes wie folgt bewegt werden:

$$\chi^{M} = \gamma^{MN}\,\chi_{N}\,. \tag{3}$$

Die kovariante Ableitung wird durch

$$\chi_{A;l} = \chi_{A,l} - \left\{{B \atop Al}\right\}\chi_{B} \tag{4}$$

definiert, wobei von den Spinoraffinitäten $\left\{{B \atop Al}\right\}$ Überschiebungsinvarianz gefordert wird. Der Ausdruck

$$S_{Al\,B} = \frac{1}{2}\left(\{Al,B\} - \{Bl,A\}\right) \tag{5}$$

heisst Torsions-Spintensor. Für die kovariante Ableitung des metrischen Spinors postulieren wir

$$\gamma_{DB;l} = -\,i\Phi_{l}\,\gamma_{DB}\,, \tag{6}$$

wobei Φ_l ein reeller Tensor ist, der mit dem elektromagnetischen Viererpotential verknüpft ist.

Von den verallgemeinerten Pauli-Matrizen, die metrische Spintensoren heissen, wird Hermitizität

$$\sigma_{k\dot{A}B} = \sigma_{kB\dot{A}} \tag{7}$$

und Verschwinden der kovarianten Ableitung gefordert:

$$\sigma_{k\dot{A}B;l} = 0\,. \tag{8}$$

Die gesamte Algebra der metrischen Spintensoren lässt sich auf dem Axiom

$$\sigma^{\dot{B}A}_{m}\,\sigma_{n\dot{B}C} = g_{mn}\,\gamma^{A}{}_{C} \pm \frac{i}{2}\,\varepsilon_{mnrs}\,\sigma^{r\dot{B}A}\,\sigma^{s}_{\dot{B}C} \tag{9}$$

aubfauen. Insbesondere folgt daraus sofort die gut bekannte Vertauschungsregel

$$\sigma_m^{\dot{B}A}\,\sigma_{n\dot{B}C} + \sigma_n^{\dot{B}A}\,\sigma_{m\dot{B}C} = 2g_{mn}\,\gamma_C^A \tag{10}$$

(ε_{mnrs} ist der Levi-Civitasche Pseudotensor bei der Signatur $+\,+\,+\,-$).

Aus der eben kurz skizzierten Axiomatik lässt sich ein expliziter Ausdruck für die Spinoraffinitäten herleiten, nämlich

$$\left\{{F \atop Bk}\right\} = -\frac{1}{4}\,\sigma^{j\dot{A}F}\,\sigma_{j\dot{A}B,k} + \frac{1}{4}\,\sigma^{j\dot{A}F}\,\sigma_{l\dot{A}B}\left\{{l \atop jk}\right\} - \frac{1}{2}\,\gamma_B{}^F\left[\hat{\Gamma}_{,k} - i(\Phi_k + \Theta_{,k})\right], \tag{11}$$

wobei $\hat{\Gamma}$ und Θ aus den Relationen

$$\hat{\Gamma} = \ln\sqrt{\gamma_{1\dot{2}}\,\gamma_{12}}\,, \qquad \gamma_{12} = \sqrt{\gamma_{1\dot{2}}\,\gamma_{12}}\,e^{i\theta} \tag{12}$$

zu entnehmen sind.

§ 3. Skizzierung des kovarianten Bispinorkalküls in der gekrümmten Raum-Zeit

Ein Bispinor wird in der folgenden Weise aus zwei Spinoren aufgebaut:

$$\Psi = \begin{pmatrix}\chi^A \\ \varphi_{\dot{A}}\end{pmatrix}. \tag{13}$$

Durch hermitische Konjugation entsteht

$$\Psi^+ = (\chi^{\dot{A}}\,\varphi_A)\,. \tag{14}$$

Sofort erkennt man, dass die Bildung $\psi^+\,\psi$ keine Kovariante ergibt. Dagegen ist

$$\overline{\Psi}\,\Psi = \varphi_A\,\chi^A + \varphi_{\dot{A}}\,\chi^{\dot{A}} \tag{15}$$

eine Invariante, wobei der adjungierte Bispinor durch

$$\overline{\Psi} = \Psi^+\beta \tag{16}$$

mit

$$\beta = \begin{pmatrix} 0 & \gamma_{\dot{A}}{}^{\dot{B}} \\ \gamma_B{}^A & 0 \end{pmatrix} \tag{17}$$

zu bilden ist. Auf diese Weise wird der Übergang von einem Bispinor zu seinem adjungierten eine kovariante Operation, im Gegensatz zur geläufigen Definition $\overline{\psi} = \psi^+\,\gamma_4$, die γ_4 auszeichnet, was relativistisch unverständlich ist und

ausserdem noch zu manchen Schwierigkeiten Anlass gibt. Damit wird das γ_4-Dilemma beseitigt. In Minkowski-Koordinaten lässt sich durch Zufall durch die Identifizierung $\beta = \gamma_4$ die Kovarianz retten. Eine Folge dieser Identifizierung ist aber dort die Forderung der Hermitizität für die Diracschen Matrizen, während hier die verallgemeinerten Dirac-Matrizen, genannt metrische Bispintensoren und definiert durch

$$\gamma^m = -i \begin{pmatrix} 0 & \sigma^{mA\dot{B}} \\ \sigma^m_{\dot{A}B} & 0 \end{pmatrix}, \tag{18}$$

nicht hermitisch gewählt werden müssen. Von β und γ^m wird Verschwinden der kovarianten Ableitung, also

$$\beta_{;l} = 0, \qquad \gamma^m{}_{;l} = 0 \tag{19}$$

gefordert.

Durch (9) ist die Algebra der metrischen Bispintensoren vorbestimmt:

$$\gamma_m \gamma_n = g_{mn} + \frac{i}{2} \varepsilon_{mn}{}^{rs} \gamma_r \gamma_s \gamma \,, \tag{20}$$

wobei

$$\gamma = \frac{1}{4!\, i} \varepsilon_{nmkl} \gamma^n \gamma^m \gamma^k \gamma^l \tag{21}$$

die kovariante Formulierung des γ_5 darstellt. Eine Folge von (20) ist die gut bekannte Vertauschungsregel

$$\gamma_m \gamma_n + \gamma_n \gamma_m = 2 g_{mn} \,. \tag{22}$$

Die Formel (4) zieht für die kovariante Ableitung eines Bispinors die Beziehung

$$\Psi_{;l} = \Psi_{,l} + \Gamma_l \Psi \tag{23}$$

nach sich, wobei die Bispinoraffinität die Struktur

$$\Gamma_l = \begin{pmatrix} \{^A_{Bl}\} & 0 \\ 0 & -\{^{\dot{B}}_{\dot{A}l}\} \end{pmatrix} \tag{24}$$

erhält.

Der aus der Gleichung

$$\chi_{A;n;k} - \chi_{A;k;n} = \chi_c R^c_{Ank} \tag{25}$$

definierte Krümmungs-Spintensor

$$R^c_{Ank} = \{^C_{Ak}\}_{,n} - \{^C_{An}\}_{,k} + \{^C_{Dn}\}\{^D_{Ak}\} - \{^C_{Dk}\}\{^D_{An}\} \tag{26}$$

und der aus der Gleichung

$$\Psi_{;n;k} - \Psi_{;k;n} = \Gamma_{kn}\,\Psi \tag{27}$$

definierte Krümmungs-Bispintensor

$$\Gamma_{kn} = \Gamma_{n,k} - \Gamma_{k,n} + [\Gamma_k, \Gamma_n] \tag{28}$$

sind durch die Beziehung

$$\Gamma_{kn} = \begin{pmatrix} R^{A}{}_{Bkn} & 0 \\ 0 & R^{\dot{B}}{}_{\dot{A}kn} \end{pmatrix} \tag{29}$$

miteinander verknüpft. In diesem Zusammenhang ist die Relation

$$\text{Spur}\,\Gamma_{kn} = 2i\,\Phi_{kn} = 2i\,(\Phi_{n,k} - \Phi_{k,n}) \tag{30}$$

interessant.

§ 4. Die Transformationsproblematik bei Koordinatentransformationen für Spinoren und Bispinoren

Die kovariante Form der Dirac-Gleichung lautet:

$$\gamma^l\,\Psi_{;l} + \frac{m_0\,c}{\hbar}\,\Psi = 0\,. \tag{31}$$

Spezialisiert man auf Minkowski-Koordinaten und vergleicht man mit der bekannten Gestalt der Dirac-Gleichung, so erkennt man, dass die kovariante Ableitung in die Eichableitung entartet. Insbesondere ergibt sich die Relation

$$\Phi_l = -\frac{2e}{\hbar c}\,A_l\,. \tag{32}$$

für den Zusammenhang zwischen dem früher eingeführten Tensor Φ_l und dem elektromagnetischen Viererpotential A_l.

Die Kovarianz der Dirac-Gleichung (31) gegenüber Koordinatentransformationen ist gesichert, wenn folgendermassen transformiert wird:

$$\Psi' = \mathfrak{S}\,\Psi, \quad \gamma'^{l'} = \mathfrak{S}_{\gamma}{}^{m}\,\gamma^{-1}\,A^{l'}_{m}\,. \tag{33}$$

Dabei bedeutet der Strich am Tensorindex den Transformationseffekt bezüglich des Tensorindex und der Strich am Symbol den Transformationseffekt bezüglich der bei Bispinoren in Matrizenschreibweise verschluckten Spinorindizes.

Die Bispinor-Transformationsmatrix $\mathfrak{S}$ ist mit den in (2) eingeführten Spinor-Transformationskoeffizienten durch die Formel

$$\mathfrak{S} = \begin{pmatrix} \Lambda_B^{A'} & 0 \\ 0 & \Lambda_{\dot{A}'}^{\dot{B}} \end{pmatrix} \tag{34}$$

verknüpft.

SOMMERFELD [2] schreibt, dass es für Minkowski-Koordinaten zwei Versionen zur Herstellung der Lorentz-Invarianz der Dirac-Gleichung gibt:

1. Transformation der Dirac-Matrizen wie Tensoren und Behandlung der Bispinoren wie Invarianten:

$$\gamma^{l'} = \Lambda_m^{l'} \gamma^m, \qquad \mathfrak{S} = 1 \qquad (\Lambda_B^{A'} = \delta_B^A); \tag{35}$$

2. Konstanthaltung der Dirac-Matrizen und damit eine implizite Definition des Transformationsgesetzes für Bispinoren:

$$\gamma'^{l'} = \gamma^l, \qquad \gamma^l = \mathfrak{S}_{\gamma'}^{\ m} \mathfrak{S}^{-1} \Lambda_m^{l'}. \tag{36}$$

Die übliche Auffassung der Spinoren ist in der historischen Entwicklung der Konzeption 2 gefolgt. Erst kürzlich konnte LANDHÄUSER [3] eine geschlossene Form für die Bispinortransformation bei allgemeinen endlichen Lorentz-Transformationen angeben. Der Grund dafür ist vor allem der gruppentheoretischen Untersuchung der Lorentz-Gruppe und deren Darstellungen zuzuschreiben. Es zeigte sich, dass die Lorentzgruppe zwei grundsätzlich verschiedene Darstellungen, nämlich die Tensor- und die Spinordarstellungen besitzt.

Im folgenden gehen wir der Situation bei beliebigen Koordinaten etwas genauer nach:

Wir betrachten die infinitesimale Koordinaten-transformation

$$x^{m'} = x^m + \xi^m, \tag{37}$$

woraus für die Transformationskoeffizienten bei Tensortransformationen

$$\Lambda_n^{m'} = \delta_n^m + \xi^{m}{}_{,n} \tag{38}$$

folgt. Im Falle beliebiger Koordinatentransformationen haben wir es bei der Auffindung von Transformationsgesetzen für geometrische Objekte mit der 16-parametrigen reellen linearen Gruppe (38), deren Darstellungen wir zu suchen haben, zu tun. Die einfachste nichttriviale Darstellung ist die Tensordarstellung (Selbstdarstellung)

$$D(\xi^{m}{}_{,n}) = 1 + I_m^n \xi^{m}{}_{,n}, \tag{39}$$

wobei die I_m^n die infinitesimalen Gruppenelemente sind, die die Form

$$I_m^n = \begin{pmatrix} \delta_m^1 \delta_1^n \, . \, \delta_m^1 \delta_4^n \\ \vdots \qquad \vdots \\ \delta_m^4 \delta_1^n \, . \, \delta_m^4 \delta_4^n \end{pmatrix} \tag{40}$$

besitzen. Mit ihrer Hilfe gelingt die Bestimmung der Strukturparameter der untersuchten Gruppe:

$$C_{nsr}^{mtl} = \delta_s^m \delta_r^t \delta_n^l - \delta_r^m \delta_s^l \delta_n^i \, , \tag{41}$$

so dass die Lie-Cartansche Integrabilitätsbedingung die Gestalt

$$[I_t^s, I_l^r] = I_t^r \delta_l^s - I_l^s \delta_t^r \tag{42}$$

annimmt. Existieren neben den Tensoren noch andere geometrische Objekte U mit dem Transformationsgesetz

$$U' = S(\xi^m{}_{,n}) \, U \, , \tag{43}$$

wobei für infinitesimale Transformationen

$$S(\xi^m{}_{,n}) = 1 + S_m^n \, \xi^m{}_{,n} \tag{44}$$

gilt, so müssen die Matrizen S_m^n der Lie-Cartanschen Integrabilitätsbedingung

$$[S_t^s, S_l^r] = S_t^r \delta_l^s - S_l^s \delta_t^r \tag{45}$$

genügen.

Denken wir speziell an die Möglichkeit der Existenz von Bispinoren, so haben wir für die Grössen S_m^n 4-4-reihige Matrizen anzusetzen. Nun ist jede 4-4-reihige Matrix als Linearkombination der 16 unabhängigen Matrizen

$$1, \quad \gamma_m, \quad \gamma_m \gamma, \quad \mathfrak{B}_{ts} = \frac{1}{4} [\gamma_t, \gamma_s], \quad \gamma \tag{46}$$

darstellbar, so dass der allgemeinste Ansatz durch

$$S_t^s = A_t^s \, 1 + M_t^{sm} \gamma_m + N_t^{sm} \gamma_m \gamma + K_t^{smn} \mathfrak{B}_{mn} + L_t^s \gamma \tag{47}$$

mit den freien Koeffizienten A_t^s, M_t^{sm}, N_t^{sm}, $K_t^{smn} = - K_t^{snm}$, L_t^s gegeben ist. Gehen wir mit diesem Ansatz in (45) ein, so ergeben sich durch Koeffizientenvergleich nach längerer Rechnung die Resultate

$$A_t^s = A \delta_t^s \, , \tag{48}$$

$$L_t^s = L\delta_t^s\,, \tag{49}$$

$$M_t^{sk} = \pm N_t^{sk}\,, \tag{50}$$

$$\begin{aligned} N_l^{sm}\,\delta_t^r - N_t^{rm}\,\delta_l^s \pm 2L(N_t^{sm}\,\delta_l^r - N_l^{rm}\,\delta_t^s) + \\ + 2g_{ij}(K_t^{smj}\,N_l^{ri} - K_l^{rmj}\,N_t^{si}) = 0\,, \end{aligned} \tag{51}$$

$$K_l^{smn}\,\delta_t^r - K_t^{rmn}\,\delta_l^s + 2g_{ij}(K_t^{sin}\,K_l^{rjm} - K_t^{sim}\,K_l^{rjn}) = 0\,. \tag{52}$$

Ohne Schwierigkeiten erkennen wir, dass neben der Tensordarstellung (40) auch

$$S_m^n = \delta_m^n\,\mathfrak{A} \qquad (\mathfrak{A} \text{ beliebige quadratische Matrix}) \tag{53}$$

eine Darstellung ist. Jedoch gelingt es nicht, Dasrtellungen zu finden, die man als Verallgemeinerungen der Spinordarstellungen bei Verwendung von Minkowski-Koordinaten interpretieren könnte. Das liegt ganz im Sinne der Erkenntnis, dass allgemeine Transformationsgruppen weniger Darstellungen als spezielle Transformationsgruppen besitzen.

Wir haben bei unseren Untersuchungen den Weg über die Lie-Cartan-Bedingungen gewählt, weil auf diese Weise die Problematik auf ein algebraisches Gleichungssystem reduziert wird, ohne dass besonders abstrakte gruppentheoretische Hilfsmittel erforderlich werden. Ausserdem erschien es uns interessant, die Sachlage einmal von dieser Seite zu beleuchten. In der gruppentheoretischen Literatur findet man den allgemeinen Beweis, dass die allgemeine lineare reelle Gruppe nur Tensordarstellungen besitzt. Dabei werden ziemlich schwerige gruppentheoretische Methoden benutzt, wovon sich der Leser z. B. beim Studium des Lehrbuches von BOERNER [4] leicht überzeugen kann.

Wir haben also die grundsätzliche Erkenntnis gewonnen, dass es keine Verallgemeinerung der Spinortransformationen in Minkowski-Koordinaten gibt, somit der Spinorbegriff in Minkowski-Koordinaten nicht auf beliebige Koordinaten verallgemeinerungsfähig ist. Die historisch entstandene Bevorzugung der früher angeführten Version 2 hat also gewissermassen in eine Sackgasse geführt. Deshalb bleibt kein anderer Ausweg, als die Version 1 als die eigentliche, einer kontinuierlichen Weiterentwicklung zugängliche *Konzeption von dem Charakter dieser neuen geometrischen Objekte*, die man Spinoren genannt hat, zu verfolgen, sofern man nicht grundsätzlich die übliche Plattform verlässt und etwa die Tetradentheorie als physikalische Basis nimmt [5]. Dennoch sind die Spinoren und Bispinoren nicht als Invarianten im üblichen Sinne, also als Tensoren O. Stufe oder Spinoren O. Stufe aufzufassen, denn für geometrische Objekte O. Stufe fällt die kovariante Ableitung mit der partiellen Ableitung entsprechend der gewählten Axiomatik zusammen. Hingegen ver-

halten sich die Spinoren in dieser Hinsicht wesentlich anders, wie aus (4) oder (23) hervorgeht. Würde man nämlich Übereinstimmung der kovarianten und der partiellen Ableitung für Spinoren fordern, so hätte man die Spinoraffinitäten Null zu setzen. Das ergäbe aber dann nach (11) eine nicht verständliche Differentialgleichung für die metrischen Spintensoren. Somit spielen die Spinoren, wenn man sich auf den von uns als Ausweg dargelegten Standpunkt stellt, eine Zwischenrolle zwischen Invarianten (Transformationsverhalten) und höherstufigen geometrischen Objekten (Übertragungsverhalten).

LITERATUR

1. E. Schmutzer, Z. Naturf., **15***a*, 355, 1960; **16***a*, 825, 1961; **17***a*, 685, 707, 1962 (die vorliegende Arbeit ist in einigen Punkten gegenüber diesen Arbeiten modifiziert).
2. A. Sommerfeld, Atombau und Spektrallinien, Vieweg u. Sohn, Braunschweig 1953, Bd. II.
3. E. Landhäuser, Z. Physik, **162**, 438, 1961.
4. H. Boerner, Darstellungen von Gruppen, Springer-Verlag, Berlin, 1955.
5. C. Pellegrini und J. Plebanski, Mat. Fys. Skr. Dan. Vid. Selsk, **2**, 39, 1962.

НЕКОТОРЫЕ ЗАМЕЧАНИЯ К ТЕОРИИ СПИНОРОВ И БИСПИНОРОВ В ИСКРИВЛЕННОМ ПРОСТРАНСТВЕ-ВРЕМЕНИ

Э. ШМУЦЕР

Резюме

Предположение о лоренц-инвариантности уравнения Дирака приводит к двум трансформационным вариантам: 1. Постоянство матрицы Дирака и отсюда хорошо известный закон преобразования для спиноров. 2. Трактовка матрицы Дирака как тензора и интерпретация волновой функции как вида инвариантности. В данной работе исследуется положение вещей по отношению искривлённости пространства-времени. Для этой цели коротко рассматривается теория спиноров и биспиноров в искривлённом пространстве-времени.

SOME NEW RESULTS IN THE SPECTROSCOPIC INVESTIGATION OF DIATOMIC MOLECULES

By

I. Kovács

DEPARTMENT OF ATOMIC PHYSICS, POLYTECHNICAL UNIVERSITY, BUDAPEST

(Received 20. IX. 1963)

The anomalous multiplet splittings observed in the fine structure of the multiplet molecule states occurring most frequently in the spectra of molecules are explained. A number of theoretical and experimental results are compared and excellent agreement is found throughout.

Well-known formulae concerning the fine structure of the multiplet terms of diatomic molecules have long been established based on the approximative solution of the wave equation of the molecule and on perturbation calculations. These formulae describe — for the most part in agreement with observation — the position of energy levels taken as a function of the rotational quantum number. There are, however, a few molecules for which the above-mentioned formulae do not agree with the experimental results, that is to say, larger and smaller deviations can be observed. For the past few years part of our theoretical research work has been directed towards finding a theoretical explanation for these deviations. In other words, the author has tried to discover those special causes that give rise to these deviations, and which render it possible to explain these deviations quantitatively as well, in agreement with experimental results. In this paper we should like to discuss these questions.

Before coming to the main points of our discussion, however, we deem it necessary to give a short introduction concerning the bases of the molecular spectra in order to make the following more comprehensible [1]. As is well-known, in diatomic molecules the electrons (like the atoms) can move only in definite orbits, and the states corresponding to them are classified analogously — though in a slightly different manner — to those of the atoms. In the atoms, in accordance with the value of the resultant orbital angular moment of electrons, S, P, D . . . terms are distinguished according to whether this value is 0, 1, 2 . . ., respectively. Since in diatomic molecules, in contrast to the central field of atoms, a field of axial symmetry directed towards the molecular axis prevails, a specific direction — arising in the case of atoms only in the presence of an external magnetic field — is present in the case of molecules even without such a field. Consequently, the resultant orbital angular moment of electrons in the direction of the molecular axis can only have quantized

values. Depending on whether this quantum number is 0, 1, 2 . . . we distinguish Σ, Π, Δ, . . . electron states in the molecule. Again, depending on whether the value of the resultant spin moment is 0, 1/2, 1 . . ., we speak of singlet, doublet, and triplet . . . states, respectively. However, to the energy values of these electron states energies due to the vibration of nuclei are always added, and these are usually smaller by one order of magnitude, and always assume quantized values. In addition there are energy values, belonging to every single vibrational state, which are quantized and smaller by one order of magnitude again, due to the rotation of the nuclei round the axis which passes through the centre of gravity and which is perpendicular to the molecular axis. Thus, in the molecule there are many (sometimes hundreds) energy levels corresponding to a single energy level of the atom. Since the spectrum lines arise (at least within the visible and ultra-violet range) through the transitions between the vibrational and rotational levels of two different electron states, it is easy to understand that one atomic spectral line is replaced here by a whole, so-called band system. A knowledge of the rotational fine structure of multiplet terms is of great importance for the investigation of the fine structure of bands produced by the transitions between several rotational levels of two vibrational levels of two electron states; in other words, it is highly important to know the dependence of the energy values on the rotational quantum number. Knowing this, it is possible to predict the structure of the resulting band, what branches will occur within the bands, what their position will be, etc., that is to say, these facts afford an important support for analysis.

The formulae were derived some 25 to 30 years ago from the wave equation of the molecule. For, by the omission of certain terms, the wave equation may be separated into three parts: a part containing the electron co-ordinates, one containing the rotational co-ordinates, and one the vibrational co-ordinate (internuclear distance). Energy values calculated from the wave equation in this way, however, did not correspond closely with experimental results, and in order to obtain results in better agreement with experimental ones, the terms neglected in the separation of the wave equation had to be taken into account by means of the perturbation calculation, particularly those terms which refer to the mutual perturbation of the neighbouring components. Considering all these terms, formulae were derived for the rotational energies, which in most cases agreed with experimental results. Thus for multiplet Σ terms the following simple formulae have been given [2], [3], [4], [5], [6]:

$$^{2S+1}F_i(N) = BN(N+1) + {}^{2S+1}f_i(N), \tag{1}$$

where $2S + 1$ means the multiplicity of the respective Σ term, thus $i = 1, 2, \ldots 2S + 1$, B is the rotational constant, and N the rotational quantum number for the case when the resultant spin momentum of electrons — either

owing to the lack of orbital angular moment falling in the direction of the molecular axis, or to a strong rotation — is coupled not to the molecular axis but to the angular moment of the rotation. In ${}^{2S+1}f_i(N)$ apart from the rotational quantum number the constant ε of the spin-spin interaction and the constant γ of the interaction between rotation and spin, respectively, also appear.

A deviation appears, however, from this formula in the ground state of HgH, on the so-called $X^2\Sigma^+$ term observed as early as 1928 by HULTHÉN [7], at the $B^2\Sigma$ term of YO observed by UHLER and ÅKERLIND in 1961 [8], at the $A^3\Sigma_u^+$ and $B'^3\Sigma_u^-$ terms of the N_2 molecule observed by CARROLL in 1952 [9] and 1960 [10], and at the ${}^7\Sigma$ term of MnH observed by NEVIN in 1942 [11]. The basic idea governing the theory enunciated to account for these deviations is this: already at first sight the finer effects can obviously be interpreted by also taking into account those terms neglected at the separation of the wave equation, which describe not only the mutual perturbations between the neighbouring components of the same multiplet, but which also include the influence of more distant terms. As these terms usually lie far enough apart, their interaction may be taken into account by means of the perturbation calculation of nondegenerate systems. As is well-known, the following general formula results:

$$F_i'(N) = F_i(N) + \sum_k \frac{|H_{ik}|^2}{h\nu_{ik}}, \tag{2}$$

where H_{ik} is the perturbation matrix element, and $h\nu_{ik}$ represents the distance between the two terms perturbing each other.

Upon investigating these interactions in detail we are led to the surprising result that the influence of the Σ and Π terms of the same and different multiplicity on the Σ terms results only in the change of the values of the constants in (1), but the structure of the formulae remains unchanged [2], [13], [14], [12]. These constants, however, cannot be theoretically calculated beforehand, unless the wave equation of the molecule is solved exactly. Using both theory and experiment the values of these constants are calculated from the observed spectra, and from these such characteristic data of the molecular structure are inferred (i.e. nuclear distance, dissociation energy, vibration and rotation of nuclei, motion of the electrons etc.) as cannot possibly be obtained in a purely theoretical way owing to mathematical difficulties. Thus, since the experimental physicist determines these constants from the spectra, he is not able to observe these perturbations, the structure of formulae being unchanged. This accounts for the fact that the formulae derived formerly in a simpler way were in most cases in agreement with the observation.

The experimental examples mentioned already, however, imply perturbations that bring about changes even in the structure of formulae. What

is then the reason for such perturbations? This can be answered if we critically examine the foregoing, especially formula (2). Obviously this formula is valid only in the case when the second term on the right side is small and, within this, the term distances in the denominator do not change considerably with the rotational quantum number. It can be shown by closer examination that this denominator must be written more rigorously in the first approximation as:

$$h\nu_{ik} = h\nu + (B_i - B_k)\, N(N+1), \tag{3}$$

where $h\nu$ now means the constant distance of the vibrational levels within one band, whereas B_i and B_k are the rotational constants of the perturbed and perturbing level. A closer examination of this relationship shows that with a change in the rotational quantum number deviations of 4—5 per cent, or even more in certain cases, arise. Consequently, by the development in series of the reciprocal value of (3) we obtain terms containing — although a fixed distance of the vibrational levels appears in the denominator — correction terms as well, inversely proportional to the square of this distance. Among these correction terms there are some whose dependence on the rotational quantum number differs from that of the terms in the original formulae. Thus, terms occur which do not fit in to any of the terms of the old formulae, in other words, apart from the change of the values of the constants the structure of the formulae undergoes a change as well. The point we want to make is this: if the distance of two close lying interacting states changes rapidly with increasing rotational quantum number, the multiplet splitting will be different from what would be expected on the basis of the usual multiplet formulae. These theoretically calculated changes show, even quantitatively, a complete agreement with experimental results [15], [12].

The constants occurring in the formulae, can be brought into relation with the molecular constants determined from other measurements and with the distance of the perturbing term (or to be more precise, the constants are inversely proportional to the square of the latter). Thus, if we know the term or terms causing the perturbation (and this can always be established on the basis of the term scheme of the molecule), we can make estimates concerning the order of magnitude of these constants also based on the theory. Accordingly, in case of the experimental examples examined the constants obtained theoretically are of the same order of magnitude as the values determined experimentally. In addition, we are given a simple explanation as to why the absolute value of constants increases with the vibrational quantum number in those cases where these deviations have been observed on a number of consecutive vibrational states of the same electron term. Both for HgH and N_2 the perturbing term lies higher than the perturbed one. For HgH this cannot be otherwise as the perturbed term is the ground state of the HgH molecule,

whereas for N_2, too, the perturbed term is the lowest-lying known triplet term of the N_2 molecule. In this case, therefore, the distance between the perturbing and perturbed term gradually decreases with the increasing vibrational quantum number, and, since the square of this decreasing distance occurs in the denominators of the constants mentioned, it is clear that their values will increase. The degree of the increase corresponds to the decrease involved in the increase of the vibrational quanta. Since other terms are occasionally near, it may be assumed that such deviations occur elsewhere as well, and in all probability they do occur in several other cases, apart from those mentioned. This sufficiently clarifies the anomalies found for the Σ terms.

Now let us consider the multiplet splittings of the Π and Δ terms which apart from the Σ terms occur most frequently in molecular spectra. In connection with the rotational fine structure of the multiplet terms (up to the $^4\Pi$ terms) there are also formulae which take into account the mutual perturbations of a more complicated structure between the neighbouring components, but, for the most part, are in close agreement with the experimental data [16], [17], [18]. For $^2\Pi$ terms these formulae agreed in every case with the experiment, but for triplet and quartet terms deviations appeared here, too, in a few cases. These deviations have a different character from those observed for Σ terms, and the main point is that the middle component or components, taken as the function of the rotational quantum number, lie elsewhere than they should according to the triplet or quartet formulae. This can be found at one of the $^3\Pi$ terms of the PH [19] and NH molecule [20], respectively, at the $^4\Pi$ term of the O_2^+ molecule [21] and at the $^3\Delta$ term of the CO molecule [30].

The question was raised: what is the cause of these deviations? At first sight one quite obviously thinks of perturbation, that is, the influence of more distant lying terms. Since, however, it is only the middle component or components that are displaced, those perturbations may be taken into consideration that affect — in the first approximation — only these components. Thus, the perturbations of $^1\Pi$ terms in the case of $^3\Pi$ terms, the perturbations of $^2\Pi$ terms in the case of $^4\Pi$ terms and those of $^1\Delta$ terms in case of $^3\Delta$ terms had to be examined. Perturbations of this kind can be discussed by taking into account the spin-orbit interaction due to the electronic part of the wave equation of the molecule. This question was treated in two of our earlier works [22]. By taking into account these intercombination perturbations we succeeded in interpreting the deviations observed at the multiplet splittings in full agreement with the experiments [23].

Later on, however, it appeared that the observed deviations may be interpreted also by a mechanism quite different from the previous interpretation in the following way: in the case of multiplet Σ terms the multiplet splitting is brought about by the influence of the spin-spin interaction as well as the interaction between rotation and spin. Thus, after neglecting these there

would be no multiplet splitting at the Σ terms, whereas the spin-orbit interaction for Π terms brings about a more or less marked multiplet splitting even in a rotationless state. In the case of Π and Δ terms the determination of the fine structure of the multiplets means that we have to investigate how the multiplet splitting brought about by the spin-orbit interaction is modified owing to rotation while the resultant spin moment becomes gradually decoupled from the molecular axis towards the angular moment of the more and more increasing rotation. Essentially, this is what is involved when taking into account the mutual perturbation of the neighbouring components.

Since the formulae obtained are for most cases in agreement with the observation, the usual spin-spin interaction in the case of Σ terms as well as the interaction between rotation and spin is not usually taken into account here. In the case of the experimental disagreements mentioned before, however, it is rather surprising that the doublet formulae have been and still are in agreement with the experimental data. As is well known, doublet terms come into being when there is only one electron outside the closed shell. In such cases, therefore, when spin-spin interaction is out of the question the formulae agree with the observation, whereas in the case of higher multiplicity, that is, when more electrons are found outside the closed shell and an interaction between their spins is possible, there are cases when the formulae are not in agreement with the observation. This suggested the investigation of the results which would be obtained by taking into account the spin-spin interaction for Π and Δ terms of higher multiplicity [24]. Surprisingly it turned out that by taking into account the spin-spin interaction, exactly the same result is obtained as from the intercombination perturbations transmitted by the spin-orbit interaction; in other words, the experimental deviations can be interpreted equally well by the spin-spin interaction, that is, by a mechanism completely different from the earlier one. Thus we cannot but assume that the experimental deviations are brought about by the joint influence of these two interactions [31].

On the basis of the above considerations, in the case of triplet and quartet terms the following formulae result for the perturbed values of terms:

$$^{2S+1}F_i'(J) = {}^{2S+1}F_i(J) + \beta \sum_{k=2}^{2S} S_{ik}^2 \quad (i = 1,\ 2,\ldots\ 2S+1),$$

where J means the rotational quantum number in the case when the resultant spin moment is still quantized to the orbital angular moment pointing in the direction of the molecular axis, that is, for slower rotation. The first term on the right side signifies the term values calculated on the basis of the triplet and quartet formula, respectively, the constant is the constant of the spin-spin interaction and that of the matrix element of the perturbation with the

$^1\Pi, ^2\Pi$ and $^1\Delta$ terms, respectively, as well as the distance of the perturbing term. The S_{ik}-s mean the transformation matrix elements. These matrix elements are the coefficients of those linear combinations by which the perturbed eigenfunctions — obtained by taking into account the neglected terms — can be produced from the eigenfunctions resulting from the separation of the wave equation [25], [32]. These depend, especially in the case of quartet terms, in a very complicated manner on the rotational quantum number, on the rotational and multiplet splitting constant, therefore their calculation for each single quantum number is laborious and lengthy [26]. If on the basis of the above formulae the differences between the perturbed and unperturbed multiplet splittings between the neighbouring components are calculated and compared with those of the multiplet splittings experimentally observed and calculated by the old theory, the results obtained show an excellent quantitative agreement [27], [28], [23], [24], [29]. (Data for the $^3\Delta$ term of the CO molecule will be published.)

These deviations have been observed so far for two vibrational states of the PH molecule, for one vibrational state of the NH ($^3\Pi$), for seven vibrational states of the O_2^+ molecule ($^4\Pi$) and for two vibrational states of the CO molecule ($^3\Delta$), respectively.

In this paper our intention has been to convey only the main ideas of the theory; more information on the details will be found in the literature quoted.

REFERENCES

1. G. Herzberg, Molecular Spectra and Molecular Structure, I. Spectra of Diatomic Molecules, Van Nostrand Co, New-York, 1951.
2. $^2\Sigma$; J. H. Van Vleck, Phys. Rev., **33**, 497, 1929.
3. $^3\Sigma$; H. A. Kramers, ZS. f. Phys., **53**, 427, 1929; R. Schlapp, Phys. Rev., **39**, 806, 1932.
4. $^4\Sigma$: A. Budó, ZS. f. Phys., **105**, 73, 1937.
5. $^5\Sigma$, $^6\Sigma$; K. S. Rao, Ind. Journ. of Phys., **26**, 47, 1952.
6. $^7\Sigma$; T. E. Nevin, Proc. Roy. Ir. Acad., **50A**, 123, 1945.
7. E. Hulthén, ZS. f. Phys., **50**, 319, 1928.
8. U. Uhler, A. Akerlind, Ark. f. Fys., **19**, 1, 1961.
9. P. K. Carroll, Proc. Roy. Ir. Acad., **54A**, 369, 1952.
10. P. K. Carroll, H. E. Rubalcava, Proc. Phys. Soc., **76**, 342, 1960.
11. T. E. Nevin, Proc. Roy. Ir. Acad., **48**, 1, 1942.
12. I. Kovács, Proc. Roy. Ir. Acad., **60A**, 15, 1959.
13. M. H. Hebb, Phys. Rev., **49**, 610, 1936.
14. A. Budó, I. Kovács, Hung. Acta Phys., **1**, 1, 1948.
15. I. Kovács, Acta Phys. Hung., **15**, 1, 1962; **15**, 337, 1963.
16. $^2\Pi$; E. Hill, J. H. Van Vleck, Phys. Rev., **32**, 250, 1928.
17. $^3\Pi$; A. Budó, ZS. f. Phys., **96**, 219, 1935.
18. $^4\Pi$; W. H. Brandt, Phys. Rev., **50**, 778, 1936.
19. M. Ishaque, R. W. B. Pearse, Proc. Roy. Soc., **156A**, 221, 1936; **173A**, 265, 1939; F. Legay, Can. Journ. Phys., **38**, 797, 1960.
20. R. N. Dixon, Can. Journ. Phys., **37**, 1171, 1959.
21. T. E. Nevin, Phil. Trans. Roy. Soc. (London), **237**, 471, 1938.
22. I. Kovács, Can. Journ. Phys., **36**, 309, 329, 1958.

23. A. Budó, I. Kovács, Acta Phys. Hung., **4,** 273, 1954; Journ. Chem. Phys., **23,** 751, 1955.
24. I. Kovács, Acta Phys. Hung., **10,** 255, 1959.
25. A. Budó, ZS. f. Phys., **105,** 579, 1937.
26. A. Budó, I. Kovács, Phys. ZS., **45,** 122, 1944.
27. I. Kovács, Acta Phys. Hung., **12,** 67, 1960.
28. I. Kovács, Acta Phys. Hung., **13,** 303, 1961.
29. I. Kovács, S. Weniger, Journ. de Phys., **23,** 377, 1962.
30. P. K. Carroll, Journ. Chem. Phys., **36,** 2861, 1962.
31. I. Kovács, Phys. Rev., **128,** 663, 1962.
32. I. Kovács, Can. Journ. Phys., **38,** 955, 1960.

НЕКОТОРЫЕ НОВЫЕ РЕЗУЛЬТАТЫ В СПЕКТРОСКОПИЧЕСКОМ ИССЛЕДОВАНИИ ДВУХАТОМНЫХ МОЛЕКУЛ

И. КОВАЧ

Р е з ю м е

В работе дается интерпретация аномального мультиплетного расщепления, наблюдаемого в тонкой структуре мультиплетных молекулярных состояний, часто встречающихся в спектре молекул. Теоретические результаты сравниваются с рядом экспериментальных данных, они отлично согласуются.

MAGNETOHYDRODYNAMIC SHOCK WAVES

By

J. SZABÓ

INSTITUTE FOR THEORETICAL PHYSICS, ROLAND EÖTVÖS UNIVERSITY, BUDAPEST*

(Presented by Z. Gyulai. — Received 20. IX. 1963)

Starting from the fundamental equations of magnetohydrodynamics we derive the conservation laws for the case of an ideal plasma. Based on these laws the fundamental equations — that govern the behaviour of magnetohydrodynamic (mhd) shock waves along an arbitrary continuous surface of discontinuity — are obtained. It is shown that the mhd shock waves — similar to those of gas dynamics — are compression waves. Finally the properties of different kinds of mhd shock waves are discussed.

I. Introduction

It is well known that in gases shock waves may be generated by supersonic flows. The shock waves are such surfaces, along which the flow velocity and the thermodynamic parameters of the gas suffer a jump. In reality the shock waves are not discontinuity surfaces but are such regions where the hydro- and thermodynamical quantities change very quickly. One can estimate the thickness of the shock from the equations of gas dynamics. In most cases this thickness is not greater than the mean free path of the particles of the gas. Therefore one may regard the shock waves to be surfaces of discontinuity in a good approximation.

The importance of shock waves propagating in a plasma was discovered a few years ago. Since in the interstellar gas and on the surface of the stars flows with supersonic velocities can often be observed, it is evident that the mhd shock waves play an important role in many astrophysical phenomena [1]. The mhd shocks have a special importance in the problems of controlled thermonuclear fusion, since at present, the most powerful method of heating the plasma is the generation of shock waves [2].

II. Conservation laws in magnetohydrodynamics

In the case of an ideal plasma the fundamental equations of magnetohydrodynamics have the form [3]:

$$\operatorname{div} \mathbf{H} = 0 , \tag{1}$$

* Present address: Department of Physics, University of the Heavy Industry, Miskolc, Hungary.

$$\frac{\partial \mathbf{H}}{\partial t} = \operatorname{rot}[\mathbf{v} \times \mathbf{H}], \tag{2}$$

$$\varrho \frac{\partial \mathbf{v}}{\partial t} + \varrho(\mathbf{v}\operatorname{grad})\mathbf{v} = -\operatorname{grad} p - \frac{1}{4\pi}[\mathbf{H} \times \operatorname{rot} \mathbf{H}], \tag{3}$$

$$\frac{\partial \varrho}{\partial t} + \operatorname{div}(\varrho \mathbf{v}) = 0, \tag{4}$$

$$\frac{\partial}{\partial t}\left(\frac{p}{\varrho^{\varkappa}}\right) + \mathbf{v}\operatorname{grad}\left(\frac{p}{\varrho^{\varkappa}}\right) = 0. \tag{5}$$

In these equations $\mathbf{H}$ is the magnetic field strength, $\mathbf{v}$ the flow velocity of the medium, ϱ its density, p its pressure and finally $\varkappa$ is the ratio of the specific heats.

In order to derive the equations of mhd shock waves we write the system of equations (1)—(5) in the form of conservation laws. The eq. (1) is already in this form. It is useful to write eq. (2) in the form

$$\frac{\partial \mathbf{H}}{\partial t} = (\mathbf{H}\operatorname{grad})\mathbf{v} - (\mathbf{v}\operatorname{grad})\mathbf{H} - \mathbf{H}\operatorname{div}\mathbf{v}. \tag{6}$$

Taking eq. (1) into account we can write the components of eq. (6) as

$$\begin{aligned} \frac{\partial H_x}{\partial t} + \operatorname{div}(H_x \mathbf{v}) &= \operatorname{div}(v_x \mathbf{H}), \\ \frac{\partial H_y}{\partial t} + \operatorname{div}(H_y \mathbf{v}) &= \operatorname{div}(v_y \mathbf{H}), \\ \frac{\partial H_z}{\partial t} + \operatorname{div}(H_z \mathbf{v}) &= \operatorname{div}(v_z \mathbf{H}). \end{aligned} \tag{7}$$

The equation of motion (3) can be transformed by means of the identities of vector analysis:

$$\varrho \frac{\partial \mathbf{v}}{\partial t} + \varrho(\mathbf{v}\operatorname{grad})\mathbf{v} = -\operatorname{grad}\left(p + \frac{H^2}{8\pi}\right) + \frac{1}{4\pi}(\mathbf{H}\operatorname{grad})\mathbf{H}, \tag{8}$$

and so the component equations of (8) can be written in the form of conservation laws making use of eqs. (1) and (4):

$$\begin{aligned} \frac{\partial(\varrho v_x)}{\partial t} + \operatorname{div}(\varrho v_x \mathbf{v}) &= -\frac{\partial}{\partial x}\left(p + \frac{H^2}{8\pi}\right) + \frac{1}{4\pi}\operatorname{div}(H_x \mathbf{H}), \\ \frac{\partial(\varrho v_y)}{\partial t} + \operatorname{div}(\varrho v_y \mathbf{v}) &= -\frac{\partial}{\partial y}\left(p + \frac{H^2}{8\pi}\right) + \frac{1}{4\pi}\operatorname{div}(H_y \mathbf{H},), \\ \frac{\partial(\varrho v_z)}{\partial t} + \operatorname{div}(\varrho v_z \mathbf{v}) &= -\frac{\partial}{\partial z}\left(p + \frac{H^2}{8\pi}\right) + \frac{1}{4\pi}\operatorname{div}(H_z \mathbf{H}). \end{aligned} \tag{9}$$

It is clear that the eqs. (9) express the conservation of momentum in a differential form. In a more concise form we have

$$\frac{\partial(\varrho v_i)}{\partial t} = -\frac{\partial T_{ik}}{\partial x_k}, \qquad (i = 1, 2, 3; k = 1, 2, 3,), \tag{10}$$

where

$$T_{ik} = \varrho v_i v_k + \left(p + \frac{H^2}{8\pi}\right)\delta_{ik} - \frac{1}{4\pi} H_i H_k$$

is the complete (three dimensional) stress tensor, $\varrho\, v_i v_k$ is the kinetic tensor and

$$\frac{H^2}{8\pi}\delta_{ik} - \frac{H_i H_k}{4\pi}$$

is the maxwellian stress tensor, δ_{ik} is the Kronecker symbol. Summation convention is understood to refer to indices occurring twice.

Now we derive the differential form of the energy theorem. Taking the scalar product of the velocity and eq. (8), we have

$$\varrho \mathbf{v}\frac{\partial \mathbf{v}}{\partial t} + \varrho \mathbf{v}\{(\mathbf{v}\,\mathrm{grad})\mathbf{v}\} + \mathbf{v}\,\mathrm{grad}\,p + \frac{\mathbf{v}}{8\pi}\,\mathrm{grad}\,H^2 - \frac{\mathbf{v}}{4\pi}\{(\mathbf{H}\,\mathrm{grad})\mathbf{H}\} = 0\,. \tag{11}$$

Making use of the identity

$$\mathrm{div}\{(\mathbf{A}_1\,\mathbf{A}_3)\mathbf{A}_2\} = (\mathbf{A}_1\,\mathbf{A}_3)\mathrm{div}\,\mathbf{A}_2 + \mathbf{A}_1\{(\mathbf{A}_2\,\mathrm{grad})\mathbf{A}_3\} + \mathbf{A}_3\{(\mathbf{A}_2\,\mathrm{grad})\mathbf{A}_1\}$$

and eqs. (1), (2), (4) and (5), eq. (11) can be written in the form

$$\begin{aligned}\frac{\partial}{\partial t}\left(\varrho\frac{v^2}{2} + \frac{p}{\varkappa - 1} + \frac{H^2}{8\pi}\right) + \mathrm{div}\left\{\mathbf{v}\left(\varrho\frac{v^2}{2} + \frac{p}{\varkappa - 1} + \frac{H^2}{8\pi}\right)\right\} = \\ = -\,\mathrm{div}\left\{\mathbf{v}p + \mathbf{v}\frac{H^2}{8\pi} - \frac{(\mathbf{vH})}{4\pi}\mathbf{H}\right\}.\end{aligned} \tag{12}$$

This form of the energy theorem will be used in the following. In order to see clearly its physical meaning, some further transformations will be carried out.

According to our hypotheses, the behaviour of the ideal plasma is governed by the equation of state of the ideal gases

$$\frac{p}{\varrho} = RT\,. \tag{13}$$

For ideal gases we have also

$$c_p - c_v = R\,, \tag{14}$$

where c_p and c_v are the specific heats under constant pressure and constant volume, respectively. From the eqs. (13) and (14) we conclude that

$$\frac{p}{\varkappa - 1} = c_v \varrho T \,. \tag{15}$$

On the left side of the energy theorem the second term inside the brackets is therefore the density of the internal energy of the ideal plasma. We note that the eq. (12) has a similar form even in the case when the plasma does not obey the equation of state of ideal gases. In this case one cannot substitute the concrete form of the energy density and entropy density for $p(\varkappa - 1)^{-1}$ in eq. (12) and $p\varrho^{-\varkappa}$ in eq. (5), respectively.

It is easy to verify that the sum

$$\mathbf{v}\frac{H^2}{8\pi} - \frac{(\mathbf{vH})}{4\pi}\mathbf{H} + \frac{H^2}{8\pi}\mathbf{v}$$

occurring under the divergence operation in the left and right side of eq. (12), is the Poyting vector. Remember that in a medium, moving with velocity $\mathbf{v}$ in a magnetic field $\mathbf{H}$, Ohm's law is given by

$$\mathbf{j} = \varrho_c \mathbf{v} + \varrho\left(\mathbf{E} + \frac{1}{c}[\mathbf{v}\times\mathbf{H}]\right),$$

where $\mathbf{j}$ is the complete current density vector, $\varrho_c \mathbf{v}$ is the convective current density, σ is the electric conductivity of the medium, $\mathbf{E}$ is the electric field strength, and c is the velocity of light in vacuum. Since, according to our hypothesis, the electric conductivity of the ideal plasma is infinitely large, the complete current density remains finite only in the case when

$$\mathbf{E} = -\frac{1}{c}[\mathbf{v}\times\mathbf{H}] \,.$$

Making use of these considerations, the Poynting vector can be written in the form

$$\mathbf{S} = \frac{c}{4\pi}[\mathbf{E}\times\mathbf{H}] = -\frac{1}{4\pi}[[\mathbf{v}\times\mathbf{H}]\times\mathbf{H}] = \frac{H^2}{4\pi}\mathbf{v} - \frac{1}{4\pi}(\mathbf{vH})\,\mathbf{H} \,.$$

After these calculations, the physical meaning of the eq. (12) is very clear. In an arbitrary volume element of the plasma the sum of the kinetic, internal and magnetic field energy can change only on the effect of flow of these types of energy and the work done by the hydrodynamical forces. In the energy

theorem the electric field energy does not occur because in the case of infinite electric conductivity it is negligible as compared to the magnetic energy [4].

Finally, we write down the equation of continuity, which expresses the conservation of mass:

$$\frac{\partial \varrho}{\partial t} + \operatorname{div}(\varrho \mathbf{v}) = 0. \tag{16}$$

III. Fundamental equations of mhd shock waves

A strong discontinuity surface of a shock wave is a surface, on the different sides of which the mhd quantities ($\mathbf{H}$, $\mathbf{v}$, p, ϱ) have different values, that is to say along which the mhd quantities have a finite jump. The equation of the discontinuity surface is

$$\varphi(x, y, z, t) = 0\,. \tag{17}$$

The velocity of the surface characterized by eq. (17) with respect to the coordinate system is given by

$$N = -\frac{\dfrac{\partial \varphi}{\partial t}}{|\operatorname{grad} \varphi|}. \tag{18}$$

The velocity of the discontinuity surface with respect to the plasma is

$$\Theta = N - v_n\,,$$

where v_n is the projection of the plasma velocity to the normal of the surface S. We choose the orientation of the normal so that

$$\Theta = N - v_n < 0 \tag{19}$$

should hold. Along the surface S, v_n and therefore also Θ have a jump. A simple consideration will convince us that Θ has the same sign on both sides of the surfaces. For points on different sides of the surfaces given by (17), the values of $\varphi(x, y, z, t)$ are of different sign. That side of S will be chosen to be positive, where its normal vector will be oriented.

During the flow of the plasma, plasma particles will get across the surface S, say, from the negative to the positive side.

Consider a mhd quantity $\Phi(x, y, z, t)$, its limit taken on the positive side will be denoted by Φ_+, that on the negative side by Φ_-, and the expression

$$[\Phi] = \Phi_+ - \Phi_-$$

will be called the jump of the quantity Φ.

Based on the conservation laws (1), (7), (9), (12) and (16), one can derive conditions that determine jumps of the mhd quantities and the relative velocity Θ. In order to derive these dynamical conditions, the fundamental equations of mhd shock waves, we put the conservation laws in integral form. Therefore we integrate the differential form of the conservation laws over a volume V which moves together with the plasma. We use the formula

$$\frac{d}{dt}\int_V \Phi\, dV = \int_V \left\{\frac{\partial \Phi}{\partial t} + \operatorname{div}(\Phi \mathbf{v})\right\} dV$$

and the Gauss theorem, to obtain the integral forms:

$$\int_F H_n\, df = 0, \tag{20}$$

$$\frac{d}{dt}\int_V \mathbf{H}\, dV = \int_F \mathbf{v}\, H_n\, df, \tag{21}$$

$$\frac{d}{dt}\int_V \varrho\, \mathbf{v}\, dV = -\int_F \left(p + \frac{H^2}{8\pi}\right)\mathbf{n}\, df + \frac{1}{4\pi}\int_F \mathbf{H}\, H_n\, df, \tag{22}$$

$$\frac{d}{dt}\int_V \varrho\, dV = 0, \tag{23}$$

$$\frac{d}{dt}\int_V \left(\frac{1}{2}\varrho v^2 + \frac{p}{\varkappa - 1} + \frac{H^2}{8\pi}\right) dV = -\int_F \left\{\left(p + \frac{H^2}{8\pi}\right) v_n - \frac{(\mathbf{vH})}{4\pi} H_n\right\} df. \tag{24}$$

In these equations V represents a volume moving together with the plasma, F its boundary surface, n the outer normal unit vector of the surface F. The integral conservation laws (20)—(24) have the common form

$$\frac{d}{dt}\int_V \varepsilon\, dV = \int_F \beta_n\, df, \tag{25}$$

where ε and β_n are expressions built up from mhd quantities.

Let us choose the volume V so that the discontinuity surface S is contained in it. From the requirement that the integral conservation laws should hold even in the case when the discontinuity surface goes across the region of integration of eq. (25) we obtain the following expression for the jumps of the quantities ε and β_n on the discontinuity surface [5]:

$$[\varepsilon\Theta] + [\beta_n] = 0. \tag{26}$$

If we substitute ε and β_n by the corresponding quantities, we obtain the fundamental equations of mhd shock from the conservation laws (20)—(24) in the following form:

$$[H_n] = 0\,, \tag{27}$$

$$[\Theta \mathbf{H}] + [\mathbf{v} H_n] = 0\,, \tag{28}$$

$$[\varrho\Theta\,\mathbf{v}] - \left[p + \frac{H^2}{8\pi}\right]\mathbf{n} + \frac{1}{4\pi}[H_n\,\mathbf{H}] = 0\,, \tag{29}$$

$$[\varrho\Theta] = 0, \tag{30}$$

$$\left[\Theta\left(\frac{1}{2}\varrho v^2 + \frac{p}{\varkappa - 1} + \frac{H^2}{8\pi}\right)\right] - \left[\left(p + \frac{H^2}{8\pi}\right)v_n\right] + \left[\frac{(\mathbf{vH})}{9\pi}H_n\right] = 0\,. \tag{31}$$

IV. Zemplén's theorem in magnetohydrodynamics

According to ZEMPLÉN's theorem [6] in hydrodynamics on the positive side of the surface of a strong discontinuity surface the pressure and density are greater than on the negative side. Therefore in hydrodynamics the shock waves can be only compression waves.

We shall show, making use of the fundamental equations (27)—(31) of mhd shock waves, that ZEMPLÉN's theorem holds also in the case of mhd shock waves.

ZEMPLÉN's theorem in magnetohydrodynamics can be formulated as follows.

If $\Theta \neq 0$, i.e. if the surface of strong discontinuity moves with respect to the plasma and the entropy density is greater on the positive side of the discontinuity surface than on its negative side, then

$$\varrho_+ > \varrho_-\,, \quad p_+ > p_- \tag{32}$$

and therefore the mhd shock waves are compression waves.

In order to demonstrate this theorem we start from eq. (31). Since according to (27) and (30), H_n and $\varrho\Theta$ are continuous on the discontinuity surface, these quantities can be put before the jump symbol $[\ldots]$ in every term. Taking this into account we write (31) in the form

$$\begin{aligned}\frac{1}{2}\varrho\Theta\,[v^2] + \frac{\varrho\Theta}{\varkappa - 1}\left[\frac{p}{\varrho}\right] + \frac{\varrho\Theta}{8\pi}\left[\frac{H^2}{\varrho}\right] - [p v_n] - \\ - \frac{1}{8\pi}[H^2\, v_n] + \frac{H_n}{4\pi}[(\mathbf{vH})] = 0\,.\end{aligned} \tag{33}$$

Then if we take the scalar product of eq. (29) with $(\mathbf{v}_+ + \mathbf{v}_-)$ we obtain

$$\varrho\Theta\,[v^2] = [p]\,(v_{n+} + v_{n-}) + \frac{v_{n+} + v_{n-}}{8\pi}\left[H^2\right] - \frac{H_n}{4\pi}\,[\mathbf{H}](\mathbf{v}_+ + \mathbf{v}_-). \quad (34)$$

Substituting this in eq. (33) we get

$$\begin{aligned}&\frac{1}{2}\,[p]\,(v_{n+} + v_{n-}) + \frac{\varrho\Theta}{\varkappa - 1}\left[\frac{p}{\varrho}\right] - [pv_n] + \frac{v_{n+} + v_{n-}}{16\pi}\,[H^2] - \\ &- \frac{H_n}{8\pi}\,[\mathbf{H}]\,(\mathbf{v}_+ + \mathbf{v}_-) + \frac{\varrho\Theta}{8\pi}\left[\frac{H^2}{\varrho}\right] - \frac{1}{8\pi}\,[H^2 v_n] + \frac{H_n}{4\pi}\,[(\mathbf{vH})] = 0.\end{aligned} \quad (35)$$

First of all we transform the terms not containing the magnetic field strength. Writing down the jump symbols in detail according to the definition, after simple calculations we obtain the following identity:

$$\begin{aligned}&\frac{1}{2}\,[p]\,(v_{n+} + v_{n-}) + \frac{\varrho\Theta}{\varkappa - 1}\left[\frac{p}{\varrho}\right] - [pv_n] = \\ &= \frac{\varrho\Theta}{\varkappa - 1}\left[\frac{p}{\varrho}\right] - \frac{1}{2}(p_+ + p_-)\,[v_n].\end{aligned} \quad (36)$$

Since the velocity N of the discontinuity surface is *per definitionem* continuous, we have

$$[v_n] = -\,[\Theta] = -\,\varrho\Theta\left[\frac{1}{\varrho}\right]. \quad (37)$$

Using this result in (36) we get

$$\begin{aligned}&\frac{1}{2}\,[p]\,(v_{n+} + v_{n-}) + \frac{\varrho\Theta}{\varkappa - 1}\left[\frac{p}{\varrho}\right] - [pv_n] = \\ &= \frac{\varrho\Theta}{\varkappa - 1}\left[\frac{p}{\varrho}\right] + \frac{\varrho\Theta}{2}\,(p_+ + p_-)\left[\frac{1}{\varrho}\right].\end{aligned} \quad (38)$$

Now we are going to transform the last five terms in eq. (35). Writing down the terms in detail, after simple calculations we obtain

$$\begin{aligned}&\frac{v_{n+} + v_{n-}}{16\pi}\,[H^2] - \frac{H_n}{8\pi}\,[\mathbf{H}]\,(\mathbf{v}_+ + \mathbf{v}_-) + \frac{\varrho\Theta}{8\pi}\left[\frac{H^2}{\varrho}\right] - \\ &- \frac{1}{8\pi}\,[H^2 v_n] + \frac{H_n}{4\pi}\,[(\mathbf{vH})] = \frac{1}{16\pi}\,(H_-^2\,v_{n-} + H_+^2\,v_{n-} - \\ &- H_+^2\,v_{n+} - H_-^2\,v_{n+}) + \frac{H_n}{8\pi}\,(\mathbf{H}_+ + \mathbf{H}_-)\,(\mathbf{v}_+ - \mathbf{v}_-) + \\ &+ \frac{\varrho\Theta}{8\pi}\left(\frac{H_+^2}{\varrho_+} - \frac{H_-^2}{\varrho_-}\right).\end{aligned} \quad (39)$$

This identity can be transformed by means of eq. (28). The equation (28) itself should be put in a new form, taking (27) and (30) into account:

$$\varrho\Theta\left(\frac{\mathbf{H}_+}{\varrho_+}-\frac{\mathbf{H}_-}{\varrho_-}\right)+H_n(\mathbf{v}_+-\mathbf{v}_-)=0\,. \tag{40}$$

If we take the scalar product of $(\mathbf{H}^+ + \mathbf{H}^-)$ and eq. (40) we obtain:

$$H_n(\mathbf{v}_+-\mathbf{v}_-)(\mathbf{H}_++\mathbf{H}_-)=-\varrho\Theta\left(\frac{H_+^2}{\varrho_+}-\frac{\mathbf{H}_+\mathbf{H}_-}{\varrho_-}+\frac{\mathbf{H}_+\mathbf{H}_-}{\varrho_+}-\frac{H_-^2}{\varrho_-}\right). \tag{41}$$

Making use of (37) and (41), eq. (39) can be put in the form

$$\begin{aligned}&\frac{v_{n+}+v_{n-}}{16\pi}[H^2]-\frac{H_n}{8\pi}[\mathbf{H}](\mathbf{v}_++\mathbf{v}_-)+\frac{\varrho\Theta}{8\pi}\left[\frac{H^2}{\varrho}\right]-\\&-\frac{1}{8\pi}[H^2\,v_n]+\frac{H_n}{4\pi}[(\mathbf{vH})]=\frac{\varrho\Theta}{16\pi}\left[\frac{1}{\varrho}\right][\mathbf{H}]^2\,.\end{aligned} \tag{42}$$

The substitution of (38) and (42) into eq. (35) gives

$$\frac{1}{2}(p_++p_-)\left[\frac{1}{\varrho}\right]+\frac{1}{\varkappa-1}\left[\frac{p}{\varrho}\right]+\frac{1}{16\pi}[\mathbf{H}]^2\left[\frac{1}{\varrho}\right]+0\,. \tag{43}$$

We have chosen the orientation of the normal of the discontinuity surface so that $\Theta<0$ holds. During the motion of the plasma and the discontinuity surface, plasma particles pass the discontinuity surface from the negative to the positive side. According to the second law of thermodynamics we have:

$$\left[\frac{p}{\varrho^\varkappa}\right]=\frac{p_-}{\varrho_+^\varkappa}\left\{\frac{p_+}{p_-}-\left(\frac{\varrho_+}{\varrho_-}\right)^\varkappa\right\}>0\,. \tag{44}$$

If the ratio p^+/p_- is expressed from (43) and substituted into (44), we obtain the following equation:

$$\left[\frac{p}{\varrho^\varkappa}\right]=\frac{p_-}{\varrho_+^\varkappa}\left\{\frac{(\varkappa+1)x-(\varkappa-1)}{\varkappa+1-(\varkappa-1)x}-x^\varkappa+G\frac{x-1}{\varkappa+1-(\varkappa-1)x}\right\}, \tag{45}$$

where

$$x=\frac{\varrho_+}{\varrho_-} \tag{46}$$

and

$$G=\frac{\varkappa-1}{8\pi p_-}[\mathbf{H}]^2\,. \tag{47}$$

An elementary calculation shows that the expression in the brackets of (45) is positive, or in other words eq. (44) holds only in the case, when

$$\frac{\varkappa+1}{\varkappa-1} > \frac{\varrho_+}{\varrho_-} < 1\,,$$

therefore

$$\varrho_+ > \varrho_-\,. \tag{48}$$

Then it follows from eq. (44), that

$$p_+ > p_-\,. \tag{49}$$

It was shown by J. L. SYNGE that in the case of gasdynamic shock waves the temperature is greater on the positive than on the negative side [7]. Based on our results it can easily be shown that in magnetohydrodynamics the jump of the temperature obeys the inequality

$$T_+ > T_-\,. \tag{50}$$

As a matter of fact, we can express the jump of the temperature using the equation of state (13):

$$[T] = \frac{p_-}{R\varrho_+}\left(\frac{p_+}{p_-} - \frac{\varrho_+}{\varrho_-}\right). \tag{51}$$

From the inequalities (44) and (48) we can conclude simply that the right side of (51) is positive, therefore $[T] > 0$.

V. The classification of mhd shock waves

In order to transform the fundamental equations of mhd shocks, we introduce the following notation [8]:

$$\tau = \frac{1}{\varrho}\,; \quad \bar{\tau} = \frac{1}{2}\left(\frac{1}{\varrho_+} + \frac{1}{\varrho_-}\right); \quad \bar{\mathbf{H}} = \frac{1}{2}(\mathbf{H}_+ + \mathbf{H}_-). \tag{52}$$

Then making use of eqs. (27) and (37) the system of eqs. (28)—(30) can be put in the following form:

$$\sigma\bar{\tau}\,[\mathbf{H}] - \bar{\mathbf{H}}\,[v_n] + H_n\,[v] = 0, \tag{53}$$

$$\sigma[\mathbf{v}] - \frac{[p]}{[\tau]}[\tau]\,\mathbf{n} - \frac{1}{4\pi}(\bar{\mathbf{H}}\,[\mathbf{H}])\,\mathbf{n} + \frac{H_n}{4\pi}[\mathbf{H}] = 0, \tag{54}$$

$$\sigma[\tau] - [v_n] = 0\,, \tag{55}$$

where

$$\sigma = \varrho\Theta.$$

It can be observed at once that the system of equations describing strong discontinuities can be transformed into that determining the amplitudes of weak discontinuities [3], if we introduce the following correspondence between the quantities:

$$\sigma \to \varrho(N - v_n); \qquad \bar{\tau} \to \frac{1}{\varrho}; \qquad -\frac{[p]}{[\tau]} \to \varkappa \frac{p}{\varrho}\varrho^2; \qquad \bar{\mathbf{H}} \to \mathbf{H};$$

$$[\mathbf{v}] \to \vec{\lambda}_v; \; ; \; [\tau] \to \varrho^{-2}\lambda_\varrho; \; [\mathbf{H}] \to \vec{\lambda}_H,$$

where $\vec{\lambda}_v$, $\vec{\lambda}$ and $\vec{\lambda}_\varrho$ represent the jumps of the differential quotients of the velocity, magnetic field strength and the density, respectively, on the weak discontinuity surface.

In the eqs. (53)—(55) we consider $[\mathbf{v}]$, $[\mathbf{H}]$ and $[\tau]$ as unknown quantities. The quantity $[p]\,[\tau]^{-1}$, however, which in the limit $[p] \to 0$ and $[\tau] \to 0$ has the limiting value

$$-\varrho^2\frac{dp}{d\varrho} = -\varkappa\frac{p}{\varrho}\varrho^2 = -C_h^2\varrho^2,$$

can be regarded as a given quantity. The system of equations (53)—(55) gives nontrivial solutions for $[\mathbf{v}]$, $[\mathbf{H}]$ and $[\tau]$ only in the case when the determinant of the system of equations vanishes. Expanding the determinant, we obtain the following algebraic equation of the seventh degree for the quantity $\sigma = \varrho\Theta$:

$$\bar{\tau}^2\sigma\left(\bar{\tau}\,\sigma^2 - \frac{H_n^2}{4\pi}\right)\left\{\bar{\tau}\,\sigma^4 + \left(\bar{\tau}\,\frac{[p]}{[\tau]} - \frac{\bar{H}^2}{4\pi}\right)\sigma^2 - \frac{[p]}{[\tau]}\frac{H_n^2}{4\pi}\right\} = 0. \tag{56}$$

It can be verified easily that in the transition

$$\varrho_+ \to \varrho_-; \; \mathbf{v}_+ \to \mathbf{v}_-; \; \mathbf{H}_+ \to \mathbf{H}_-; \; p_+ \to p_-$$

the eq. (56) — disregarding a factor — goes over into the equation for the velocity of the weak discontinuity surface [3].

Different solutions of equations (53)—(55) belong to the different roots of the equation (56). We can investigate different kinds of strong discontinuities on this basis.

a) *Slow and fast shocks*

First of all we consider the case, when the last term on the left side of eq. (56) vanishes. This equation can be put in the form

$$\left(\sigma^2 + \frac{[p]}{[\tau]}\right)\left(\sigma^2 - \frac{H_n^2}{4\pi\,\bar{\tau}}\right) = \sigma^2\left(\frac{\bar{H}^2}{4\pi\bar{\tau}} - \frac{H_n^2}{4\pi\,\bar{\tau}}\right), \tag{57}$$

or

$$\sigma^2\left(\bar{\tau}\,\varrho^2 - \frac{\bar{H}^2}{4\pi}\right) = -\frac{[p]}{[\tau]}\left(\bar{\tau}\,\varrho^2 - \frac{H_n^2}{4\pi}\right). \tag{58}$$

The eq. (57) has two positive roots. σ_f denotes the greater one, and σ_s the smaller one. Remembering that according to ZEMPLÉN's theorem

$$-\frac{[p]}{[\tau]} = \frac{[p]}{[\varrho]}\varrho_+\varrho_- > 0 \tag{59}$$

from eqs. (57) and (58) we obtain easily the following inequalities:

$$\sigma_s^2 \leqq -\frac{[p]}{[\tau]} \leqq \sigma_f^2, \tag{60}$$

$$\sigma_s^2 \leq \frac{H_n^2}{4\pi\bar{\tau}} \leq \sigma_f^2. \tag{61}$$

If σ satisfies eq. (57) then the solution of the eqs. (53)—(55), or in other words, the jumps of the mhd quantities can be expressed in the following way:

$$\begin{aligned}
[\mathbf{H}] &= \alpha\,\bar{\tau}\,\sigma^2(\bar{\mathbf{H}} - H_n\,\mathbf{n}),\\
[\mathbf{v}] &= -\alpha\,\bar{\tau}\,\varrho\left(\frac{H_n}{4\pi}\,\bar{\mathbf{H}} - \bar{\tau}\,\sigma^2\mathbf{n}\right),\\
[\tau] &= -\alpha\bar{\tau}\left(\bar{\tau}\,\sigma^2 - \frac{H_n^2}{4\pi}\right).
\end{aligned} \tag{62}$$

The jump of the pressure can be determined from eq. (29). Taking the scalar product of $\mathbf{n}$ with eq. (29), and making use of eq. (27), we derive the relation

$$[p] = \sigma[v_n] - \frac{1}{4\pi}(\bar{\mathbf{H}}[\mathbf{H}]).$$

If we substitute here the value of $[v_n]$ and $[\mathbf{H}]$ from eq. (62) we obtain the jump of the pressure in the form

$$[p] = \alpha\,\bar{\tau}\,\sigma^2\left(\bar{\tau}\,\sigma^2 - \frac{\bar{\mathbf{H}}^2}{4\pi}\right). \tag{63}$$

We have to determine the jump of the absolute value of the magnetic field strength. From the first equation of (62) we have

$$[H^2] = 2(\mathbf{H}[\mathbf{H}]) = 2\alpha\,\bar{\tau}\,\sigma^2(\bar{H}^2 - H_n^2). \tag{64}$$

Making use of the last equation of (62) we obtain

$$[H^2] = -2\sigma^2 [\tau] \frac{\overline{H}^2 - H_n^2}{\bar{\tau}\sigma^2 - \dfrac{H_n^2}{4\pi}}. \tag{65}$$

Since, according to ZEMPLÉN's theorem, $[\tau] < 0$ it follows from (61) and (65) that in a fast shock wave the magnetic field strength increases, and decreases in a slow shock wave. As the normal component of the magnetic field strength is continuous on the discontinuity surface, we conclude, that the change of the tangential component has the same sign as that of the jump of the absolute value.

A shock wave, which propagates perpendicularly to the direction of the magnetic field strength ($H_n = 0$) is called a perpendicular shock.

It follows from eq. (28), that along a perpendicular shock the tangential component of the magnetic field strength satisfies the equation

$$\frac{H_{t+}}{H_{t-}} = \frac{\Theta_-}{\Theta_+}.$$

A further result concerning perpendicular shocks can be obtained from eq. (28) or directly from (62):

$$[\mathbf{H}] = \alpha\bar{\tau}\,\sigma^2\,\overline{\mathbf{H}}.$$

Finally, if the magnetic field strength is parallel to the normal of the discontinuity surface, then the discontinuity surface is called a parallel shock. In this case, as it can be seen easily using eq. (37), the terms containing magnetic field strengths cancel in eq. (28)—(31), and so we are led to the equations of the ordinary hydrodynamic shock waves.

Because of this circumstance, this kind of mhd shock waves may be called magnetoacoustic shocks.

b) *Transverse shock waves*

The discontinuity surface belonging to the root

$$\sigma_{tr} = \left(\frac{H_n^2}{4\pi\bar{\tau}}\right)^{\frac{1}{2}} \tag{66}$$

of the eq. (56) is called transverse shock. The jumps of the mhd quantities along transverse shocks can be expressed from eqs. (53)—(55) in the form

$$[\mathbf{H}] = \alpha\sigma\,\mathbf{H}\times\mathbf{n}; \quad [\mathbf{v}] = \alpha\frac{H_n}{4\pi}\,\overline{\mathbf{H}}\times\mathbf{n} = \mp\,[\mathbf{H}]\left(\frac{\bar{\tau}}{4\pi}\right)^{\frac{1}{2}};$$
$$[\tau] = 0; \quad [p] = 0. \tag{67}$$

As in this case the density is equal on the two sides of the discontinuity surface, the jumps of the velocity can be given by the formula

$$[\mathbf{v}] = \mp \frac{[\mathbf{H}]}{\sqrt{4\pi\varrho}} \,. \tag{68}$$

This formula corresponds exactly to the relation between $\mathbf{v}$ and $\mathbf{H}$ in the case of Alfvén waves.

The last equation of (67) can be deduced from eq. (29). Indeed, if we take the scalar product of $\mathbf{n}$ and eq. (29), and take (27) and (55) into account, we obtain:

$$[p] + \left[\frac{H^2}{8\pi}\right] = 0. \tag{69}$$

On the other hand, because of eq. (67):

$$[H^2] = 2\,\overline{\mathbf{H}}\,[\mathbf{H}] = 0\,, \tag{70}$$

therefore $[p] = 0$.

It follows from eq. (67) and from the identity

$$\left[\frac{p}{\varrho^\varkappa}\right] = \frac{p_-}{\varrho_+^\varkappa}\left\{\frac{p_+}{p_-} - \left(\frac{\varrho_+}{\varrho_-}\right)^\varkappa\right\} \tag{71}$$

that along a transverse mhd shock the entropy density is continuous. Hydrodynamic shock waves, that would correspond to this kind of mhd shock, do not exist. As a matter of fact, in hydrodynamics, when $\Theta \neq 0$, from the fact that the pressure is continuous it follows that every other hydrodynamical quantity is continuous.

As according to (27) and (70) both the normal component and the absolute value of the magnetic field strength is continuous, therefore the only change of the magnetic field strength during the transition across the discontinuity surface is a rotation around the axis $\mathbf{n}$.

c) *Tangential discontinuity*

Finally, we discuss the case, when the discontinuity surface moves together with the plasma, i.e. when

$$\Theta = N - v_n = 0 \tag{72}$$

If $\mathbf{H}_n \neq 0$, then it follows from the eqs. (27)—(31), that

$$[\mathbf{v}] = 0;\ [\mathbf{H}] = 0;\ [p] = 0;\ [\varrho] \text{ arbitrary.} \tag{73}$$

If $\overline{H}_n = 0$, then only the normal component of the velocity and of the magnetic field strength are continuous, but the tangential components may have a jump on the discontinuity surface. From eq. (29) in the case of $H_n = 0$ and $\Theta = 0$ we obtain

$$\left[p + \frac{H^2}{8\pi}\right] = 0.$$

REFERENCES

1. *С. А. Каплан,* Межзвездная газодинамика, Физматгиз, Москва, 1958.
2. A. C. Kolb, Recent Progress in Shock Wave Research, IV. International Conference on Ionization Phenomena in Gases, Uppsala, 1959.
3. J. Szabó, Magyar Fizikai Folyóirat, **8**, 175, 1960.
4. *С. И. Сыроватский,* УФН, **62**, 247, 1957.
5. *Н. Е. Кочин, И. А. Кибель* и *Н. В. Розе,* Теоретическая гидромеханика, том 2, Гостехиздат, Москва 1955.
6. G. Zemplén, Comptes rendus des Séances de l'Acad. de Paris, **141**, 710, 1905.
7. J. L. Synge, The Relativistic Gas, North-Holland Publ. Co., Amsterdam, 1957.
8. K. O. Friedrichs and H. Kranzer, Nonlinear Wave Motion. Notes on Magnetohydrodynamics, NYO — 6486.

МАГНЕТОГИДРОДИНАМИЧЕСКИЕ УДАРНЫЕ ВОЛНЫ

Я. САБО

Резюме

Исходя из основных уравнений магнетогидродинамики, для идеальной плазмы выводятся теоремы сохранения, а с их помощью выводятся основные уравнения магнетогидродинамических ударных волн для произвольной непрерывной поверхности разрыва. Доказывается, что магнетогидродинамические ударные волны, подобно газодинамическим, являются скачками уплотнения. Наконец исследуются свойства ударных волн различных типов.

BASIC EQUATIONS OF MAGNETOFLUIDO-DYNAMICS IN ANISOTROPIC MEDIA

By

I. ABONYI

INSTITUTE FOR THEORETICAL PHYSICS, ROLAND EÖTVÖS UNIVERSITY, BUDAPEST

(Presented by Z. Gyulai. — Received 20. IX. 1963)

After giving a brief review of the fundamental equations of magnetofluido-dynamics and their limits of validity, the basic equations are derived for the case of media characterized by anisotropic conductivity, electric and magnetic permeability. The material quantities are supposed to be given tensor expressions with a nonvanishing determinant.

§ 1. Introduction

As is well known the basic equations of magnetofluido-dynamics have the form:

$$\frac{\partial \mathfrak{H}}{\partial t} = \operatorname{rot}(\mathfrak{v} \times \mathfrak{H}) + \frac{c^2}{4\pi\sigma} \Delta \mathfrak{H}. \tag{1}$$

$$\operatorname{div} \mathfrak{H} = 0, \tag{2}$$

$$\mathfrak{j} = \frac{c}{4\pi} \operatorname{rot} \mathfrak{H} \tag{3}$$

$$\mathfrak{E} = -\frac{1}{c} \mathfrak{v} \times \mathfrak{H} + \frac{c}{4\pi\sigma} \operatorname{rot} \mathfrak{H}, \tag{4}$$

$$\varrho_e = -\frac{\varepsilon}{4\pi c} \operatorname{div}(\mathfrak{v} \times \mathfrak{H}), \tag{5}$$

$$\mathfrak{j} = \sigma \left(\mathfrak{E} + \frac{1}{c} \mathfrak{v} \times \mathfrak{H} \right), \tag{6}$$

$$\varrho \frac{\partial \mathfrak{v}}{\partial t} + \varrho(\mathfrak{v} \nabla)\, \mathfrak{v} = -\nabla p - \frac{1}{4\pi\varrho} \mathfrak{H} \times \operatorname{rot} \mathfrak{H} + \eta \Delta \mathfrak{v} + \left(\xi + \frac{\eta}{3} \right) \operatorname{grad} \operatorname{div} \mathfrak{v}, \tag{7}$$

$$p = p(\varrho, T), \tag{8}$$

where the usual notation is applied. These equations are obtained from Maxwell's equations, the equation of motion of the fluid (Navier—Stokes equations) and the equation of state, under the following restricting assumptions [1]:

a) the dielectric constant ε of the medium is considered to be a scalar, which is, in many cases, independent of time and position; the magnetic permeability μ is a scalar, too, but its value is everywhere unity.

b) The electric conductivity σ of the fluid is a scalar, its value being very large. In order to assure the scalar character of the electric conductivity one must require the cyclotron frequency $\omega_c = eH(\mathrm{mc})^{-1}$ to be very much smaller than the collision frequency or the frequency of the electromagnetic wave travelling across the fluid [2].

c) The convective current is neglected as compared to the conductive one, and so is the displacement current. This turns out to be a requirement as regards the frequency of the electromagnetic process in the fluid, in the form

$$\varepsilon\omega \ll 4\pi\sigma\,. \tag{9}$$

d) The relativistic effects are excluded in this treatment: $\beta = vc^{-1} \ll 1$.

Though the conditions of these approximations are more or less realizable in quite a wide range of the phenomena and therefore are acceptable for the magnetofluido-dynamic theory, they certainly exclude an important problem, i.e. the inherent anisotropy of a magnetofluido-dynamic medium. Based on simple models one may make definite statements on the electric behaviour of the medium, derive concrete formulae for the electric and magnetic permeability and the electric conductivity. These material quantities turn out to be of a tensorial character (e.g. [2]).

The primary cause of this tensorial character — and that is why we are speaking of inherent anisotropy of a magnetofluido-dynamic medium — is the effect of a magnetic field strong enough to orientate the cyclotron motion of the individual particles. This orientation results in an anisotropy unless it ceases very quickly through the effect of Brownian motion or external noises.

In the following the basic equations of magnetofluido-dynamics are derived for the case of an anisotropic medium, characterized by tensor electric and magnetic permeability and tensor electric conductivity, but the other magnetofluido-dynamic assumptions *c)* and *d)* will be retained.

§ 2. The derivation of the basic equations

We shall use the following notation: dot means differentiation with respect to time, ∂_i differentiation with respect to the coordinate x_i, the summation convention is understood to refer to indices occurring twice and running from 1 to 3. The vector product of two vectors is expressed by means of the Levi-Cività symbol e_{ikl}.

We write down the starting equations. The equations of motion are

$$\varrho \dot{v}_i + \varrho v_r \partial_r v_i = F_i^M + F_i^E, \tag{10}$$

where F_i^M and F_i^E denote the forces of mechanical and electromagnetic origin, respectively. In the expression of F_i^M one may include also the terms of the Navier-Stokes equation, which contain the viscosity. The equation of continuity of the fluid mass is given by

$$\dot{\varrho} + \partial_r (\varrho v_r) = 0, \tag{11}$$

and Maxwell's equations with the approximations, listed in *c)* and *d)* run:

$$e_{ikl} \partial_k H_l = \frac{4\pi}{c} j_i, \tag{12}$$

$$e_{ikl} \partial_k E_l = -\frac{1}{c} B_i, \tag{13}$$

$$\partial_k D_k = 4\pi \varrho^E, \tag{14}$$

$$\partial_k B_k = 0. \tag{15}$$

The material equations are given by the following equations:

$$D_i = \varepsilon_{ik} E_k, \tag{16}$$

$$B_i = \mu_{ik} H_k, \tag{17}$$

$$j_i = \sigma_{ik} \tilde{E}_k = \sigma_{ik} \left\{ E_k + \frac{1}{c} e_{krs} v_r B_s \right\}. \tag{18}$$

Our aim is to derive a set of equations from the set (10)—(18), in which all the variables are expressed in terms of the mechanical quantities (p, ϱ, v_i) and the field variables of the magnetic field (H_i, B_i).

As a first step we derive the equation for B_i.

Let us introduce the reciprocal tensor Σ_{ik} of σ_{ik} by the equation

$$\sigma_{ik} \Sigma_{kl} = \delta_{il}. \tag{19}$$

The criterion of existence of such a tensor Σ_{ik} is

$$\det \{\sigma_{ik}\} \neq 0. \tag{20}$$

Thus we can express the electric field strength E_i from (18):

$$E_p = -\frac{1}{c} e_{prs} v_r B_s + \Sigma_{ps} j_s. \tag{21}$$

Substituting this into (13), we obtain:

$$B_i - e_{ikr}\,\partial_k\,(e_{rps}\,v_p\,B_s) = ce_{ikr}\,\Sigma_{rs}\,\partial_k j_s + ce_{ikr}\,j_s\,\partial_k\,\Sigma_{rs}. \tag{22}$$

Making use of (12) gives

$$j_i = \frac{c}{4\pi}\,e_{ikl}\,\partial_k H_l \tag{23}$$

in (22). We have achieved that this group of Maxwell's equations has the desired property:

$$B_i - e_{ikl}\,\partial_k\,(e_{rps}\,v_p\,B_s) - \frac{c^2}{4\pi}\,e_{ikr}\,e_{spq}\,\partial_k\,(\Sigma_{rs}\,\partial_p H_q) = 0\,. \tag{24}$$

Since in the usual isotropic magnetofluido-dynamics the quantity

$$\frac{c^2}{4\pi\sigma} = \nu_m \tag{25}$$

is introduced as the magnetic viscosity, here it is advantageous to introduce

$$\frac{c^2}{4\pi}\,\Sigma_{ik} = \mathcal{M}_{ik} \tag{26}$$

as the magnetic viscosity tensor. With this quantity (24) can be written in the form

$$B_i - e_{ikr}\,\partial_k\,(e_{rps}\,v_p\,B_s) - e_{ikr}\,e_{spq}\,\partial_k\,(\mathcal{M}_{rs}\,\partial_p H_q) = 0\,. \tag{27}$$

For the sake of homogeneous notation we may express (27) in terms of B_i. Therefore we have to introduce the inverse tensor m_{ik} of the magnetic permeability μ_{ik} with the definition

$$\mu_{ik}\,m_{kl} = \delta_{il}, \tag{28}$$

provided that

$$\det\{\mu_{ik}\} \neq 0\,. \tag{29}$$

So we have

$$H_q = m_{qr}\,B_r$$

and the final form of (27) is:

$$B_i - e_{ikr}\,\partial_k\,(e_{rps}\,v_p\,B_s) - e_{ikr}\,e_{spq}\,\partial_k\,(\mathcal{M}_{rs}\,\partial_p\,m_{qu}\,B_u) = 0\,. \tag{30}$$

We have to transform also the other equations. Substituting (23) into (21) we can express the electric field strength:

$$E_i = -\frac{1}{c}\,e_{irs}\,v_r\,B_s + \frac{1}{c}\,\mathcal{M}_{is}\,e_{skl}\,\partial_k\,(m_{lr}\,B_r), \tag{31}$$

further, using (14) we can give the desired expression for the charge density:

$$\varrho^E = -\frac{1}{4\pi c}\,\partial_r(\varepsilon_{rs}\,e_{spq}\,v_p\,B_q) + \frac{1}{4\pi c}\,\partial_r\big(\varepsilon_{rs}\,e_{skl}\,\partial_k\,(m_{lp}\,B_p)\big). \tag{32}$$

§ 3. Discussion

Finally we write down the set of basic equations of magnetofluido-dynamics in the case of anisotropic media:

$$B_i - e_{ikr}\,\partial_k\,(e_{rps}\,v_p\,B_s) - e_{ikr}\,e_{spq}\,\partial_k\,(\mathscr{M}_{rs}\,\partial_q\,m_{qu}\,B_u) = 0, \tag{33}$$

$$\partial_s\,B_s = 0, \tag{34}$$

$$j_i = \frac{c}{4\pi}\,e_{ikl}\,\partial_k\,H_l, \tag{35}$$

$$E_i = -\frac{1}{c}\,e_{irs}\,v_r\,B_s + \frac{1}{c}\,\mathscr{M}_{is}\,e_{skl}\,\partial_k\,(m_{lr}\,B_r), \tag{36}$$

$$\varrho^E = \frac{1}{4\pi c}\,\partial_r\,(\varepsilon_{rs}\,e_{skl}\,\partial_k\,(m_{lp}\,B_p)) - \frac{1}{4\pi c}\,\partial_r\,(\varepsilon_{rs}\,e_{spq}\,v_p\,B_q). \tag{37}$$

One must add to these material equations the definition of the magnetic viscosity tensor $\mathscr{M}_{ik}$, the equation of state and the equations of motion. It must be stressed, however, that in this case the force of electromagnetic origin has not only one term, the Lorentz force, as in (6), but is given by a generalized expression [3]:

$$F_i^E = -\frac{1}{8\pi}\,(E_s\,E_r\,\partial_i\,\varepsilon_{rs} + H_s\,H_r\,\partial_i\,\mu_{rs}) + \frac{1}{c}\,e_{ikr}\,j_k\,B_r, \tag{38}$$

where the term $\varrho^E E_i$ is neglected because of the approximation condition *d)*.

The specialization of equations (33)—(37) for the case of isotropic material quantities gives again the set of equations (1)—(8).

REFERENCES

1. J. Szabó, Magyar Fizikai Folyóirat, **8**, 175, 1960.
2. J. L. Delcroix, Introduction à la théorie des gas ionisés, Dunod, Paris, 1959, p. 53.
3. K. F. Novobátzky and T. Neugebauer, Elektrodinamika és Optika. Budapest, 1951, p. 143.

ОСНОВНЫЕ УРАВНЕНИЯ МАГНЕТОГИДРОДИНАМИКИ В АНИЗОТРОПНЫХ СРЕДАХ

И. АБОНИ

Р е з ю м е

После обзора основных уравнений магнетогидродинамики и их условий действительности выводятся магнетогидродинамические уравнения среды, обладающей зависящей от места и времени анизотропными электрической проводимостью, диэлектрической и магнитной проницаемостями. Материальные величины представляются в виде заданных тензоров, детерминанты которых не равны нулю.

GENERALIZED LIPPMANN-SCHWINGER EQUATIONS FOR THEORIES WITH MULTIPOLE GHOSTS

By

K. L. NAGY

INSTITUTE FOR THEORETICAL PHYSICS, ROLAND EÖTVÖS UNIVERSITY, BUDAPEST

(Presented by Z. Gyulai. — Received 20. IX. 1963)

Generalized Lippmann-Schwinger equations are derived for theories where it is supposed that indefinite metric and multipole ghost states are present. It is shown that the adiabatic S-matrix gives the transition amplitude also in this case.

1. In electrodynamics the potentials themselves are not observable physical quantities. This was the starting point of Prof. NOVOBÁTZKY's early investigations [1] aiming at the construction of a quantum electrodynamics without the use of the concept of potentials. Investigations of this type were carried out also much more recently (see e.g. [2]).

Quantization by means of potentials, as it is usually done, however, possesses some advantages: first of all it is simple. But this simplicity goes only up to one specific point; the quantization necessarily requires an indefinite metric in the state vector space. This is called the GUPTA-BLEULER formalism [3]. This is one main case, where methods with indefinite metric have been introduced into physics. Since then many other applications have been found.

One of the most interesting points in a theory with an indefinite metric is (in quantum electrodynamics apparently this is not the case) that even in a finite dimensional space it is not sure that the eigenvectors of the Hamiltonian or in general a Hermitian operator form a complete system. If this happens, in order to get a complete system, beside the eigenstates

$$(H - E) \mid E > = 0 ,$$

the dipole ghost states $\mid D >$ satisfying

$$(H - E) \mid D > = E \mid E > ,$$

the tripole ghost states, with

$$(H - E) \mid T > = E \mid D > ,$$

etc. have to be included in the set. How far this series of multipole ghosts goes,

depends on the specific theory. Here some remarks are made referring to the formal theory of scattering supposing the presence of multipole ghost states.

2. The treatment starts with a decomposition of the total Hamiltonian into an unperturbed part H^0 and an interaction term H^1:

$$H = H^0 + H^1.$$

In what follows it is assumed that H^0 and H possess the same continuous spectrum, and H^1 does not cause a level shift. We disregard similarly that a "dressing" operator actually does not exist, this, at the end, in the adiabatic S matrix, leads to renormalization constants.The necessary changes for an ordinary field theory can be found in [4]. Here, however, a much more important new feature emerges. Depending on the actual decomposition of H the state vector system may have the same structure for H^0 and H with a one to one correspondence among unperturbed and perturbed state vectors, or it may not. More accurately, for each H^0 and H the eigenvector system may be complete or incomplete where multipole ghosts emerge. Now according to the actual decomposition, all the four combinations may appear:

a) $H \to$ compl., $\quad H^0 \to$ compl.,

b) $H \to$ compl., $\quad H^0 \to$ incompl.,

c) $H \to$ incompl., $\quad H^0 \to$ compl.,

d) $H \to$ incompl., $\quad H^0 \to$ incompl.

We consider here the fourth case with a one to one correspondence, from which dropping terms corresponding to multipole ghosts, the well-known first case arises as a special case. The mixed cases are to be treated elsewhere.

In this case we have

$$\begin{aligned}
&(H - E_n)|\, E_n > = 0, \qquad (H^0 - E_n)|\, E_n >^0 = 0, \\
&(H - E_n)|\, D_n > = E_n\, |E_n >, \qquad (H^0 - E_n)\,|\, D_n >^0 = E_n\,|\, E_n >^0, \qquad (1) \\
&(H - E_n)\,|T_n > = E_n\,|\, D_n >, \qquad (H^0 - E_n)\,|\, T_n >^0 = E_n\,|\, D_n >^0, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}$$

By applying $(H^0 - E_n)$ on both sides, one can verify that the following — generalized — Lippmann—Schwinger equations are valid ($\varepsilon \to +0$):

$$|E^{\text{in}}_{n\text{out}} > = |\, E_n >^0 + \frac{1}{E_n - H^0 \pm i\varepsilon} H^1\,|\, E_n{}^{\text{in}}_{\text{out}} >,$$

$$|D_n \overset{\text{in}}{\text{out}}> = |D_n>^0 + \frac{1}{E_n - H_0 \pm i\varepsilon} H^1 |D_n \overset{\text{in}}{\text{out}}> -$$

$$- \frac{E_n}{(E_n - H^0 \pm i\varepsilon)^2} H^1 |E_n \overset{\text{in}}{\text{out}}>,$$

$$|T_n \overset{\text{in}}{\text{out}}> = |T_n>^0 + \frac{1}{E_n - H^0 \pm i\varepsilon} H^1 |T_n \overset{\text{in}}{\text{out}}> - \tag{2}$$

$$- \frac{E_n}{(E_n - H^0 \pm i\varepsilon)^2} H^1 |D_n \overset{\text{in}}{\text{out}}> +$$

$$+ \frac{E_n^2}{(E_n - H^0 \pm i\varepsilon)^3} H^1 |E_n \overset{\text{in}}{\text{out}}>,$$

. . . . . .

Such equations were studied in special cases in [5].

The equations possess the following iterative solutions:

$$|E_n \overset{\text{in}}{\text{out}}> = \left(1 + \frac{1}{E_n - H^0 \pm i\varepsilon} H^1 + \frac{1}{E_n - H^0 \pm i\varepsilon} H^1 \frac{1}{E_n - H^0 \pm i\varepsilon} H^1 + \right.$$

$$\left. + \ldots \right) |E_n>^0 \equiv A(H^0, H^1, E_n) |E_n>^0,$$

$$|D_n \overset{\text{in}}{\text{out}}> = A(H^0, H^1, E_n) |D_n>^0 + E_n \frac{\partial A(H^0, H^1, E_n)}{\partial E_n} |E_n>^0,$$

$$|T_n \overset{\text{in}}{\text{out}}> = A(H^0, H^1, E_n) |T_n>^0 + E_n \frac{\partial A(H^0, H^1, E_n)}{\partial E_n} |D_n>^0 +$$

$$+ \frac{E_n^2}{2} \frac{\partial^2 A(H^0, H^1, E_n)}{\partial E_n^2} |E_n>^0, \tag{3}$$

. . . . . .

Next we prove that

$$|E_n, \text{in}> = U_\varepsilon(0, -\infty) |E_n>^0,$$

$$|D_n, \text{in}> = U_\varepsilon(0, -\infty) |D_n>^0, \tag{4}$$

$$|T_n, \text{in}> = U_\varepsilon(0, -\infty) |T_n>^0,$$

. . . . . .

are valid also in this case, where $U_\varepsilon(0, -\infty)$ denotes the usual adiabatic evolution operator, satisfying

$$U_\varepsilon(t, t_0) = 1 - i \int_{t_0}^{t} e^{-\varepsilon|t'|} H^1(t') \, u_\varepsilon(t', t_0) \, dt'$$

with

$$H^1(t) = e^{iH^0 t} H^1 e^{-iH^0 t}. \tag{5}$$

Indeed from (4) for each vector

$$|\text{in}> = |>^0 - i \int_{-\infty}^{0} e^{-\varepsilon t'} e^{iH^0 t'} H^1 e^{-iH^0 t'} |t'> dt'. \tag{6}$$

Since from (1)

$$e^{-iH^0t}\,|E_n>^0 = e^{-iE_nt}\,|E_n>^0,$$

$$e^{-iH^0t}\,|D_n>^0 = \left(|D_n>^0 + |E_n>^0 E_n\frac{\partial}{\partial E_n}\right)e^{-iE_nt},$$

$$e^{-iH^0t}\,|T_n>^0 = \left(|T_n>^0 + |D_n>^0 E_n\frac{\partial}{\partial E_n} + |E_n>^0\frac{E_n^2}{2}\frac{\partial^2}{\partial E_n^2}\right)e^{-iE_nt}, \qquad (7)$$

. . . . . .

the iterative solution of (6) is just (3).

In a similar fashion one can prove

$$|\,\text{out}> = U_\varepsilon\,(0, +\infty)\,|>^0\,.$$

Thus for the S matrix element $<\text{out}\,|\,\text{in}>$ one gets

$$S_{mn} = <m, \text{out}\,|\,n, \text{in}> = {}^0\!<m\,|\,S_\varepsilon\,|\,n>^0,$$

S_ε being the adiabatic S matrix:

$$S_\varepsilon = U_\varepsilon\,(+\infty, -\infty)\,.$$

According to a well-known formula for U, $U(t, 0) = e^{iH^0t}\cdot e^{-iHt}$ we obtain

$$S_{mn} = \lim_{t'\to+\infty}{}^0\!<m|e^{iH^0t'}\,e^{-iHt'}\,|n, \text{in}>.$$

Thus for $|n, in> = |E_n, \text{in}>$, considering (7) and (2)

$$S_{mn} = \lim_{t'\to+\infty} e^{i(E_m-E_n)t'}\;{}^0\!<E_m\,|\,E_n\,\text{in}> =$$

$$= {}^0\!<E_m|E_n>^0 + \lim_{t'\to+\infty}\frac{e^{i(E_m-E_n)t'}}{E_n - E_m + i\varepsilon}\,{}^0\!<E_m\,|H^1\,|E_n, \text{in}> =$$

$$= {}^0\!<E_m\,|\,E_n>^0 - 2\pi i\,\delta\,(E_m - E_n)\,R_{mn},$$

$$R_{mn} = {}^0\!<E_m|H^1\,|E_n, \text{in}> = {}^0\!<E_m\,|\,R\,|E_n>^0,$$

$$S^{(D)}_{mn} = {}^0\!<D_m\,|\,E_n>^0 - 2\pi i\,\delta\,(E_m - E_n)\,R^{(D)}_{mn} -$$
$$- 2\pi i\,E_m\,\delta'\,(E_m - E_n)\,R_{mn},$$

$$R^{(D)}_{mn} = {}^0\!<D_m\,|H^1|\,E_n, \text{in}> = {}^0\!<D_m\,|R|\,E_n>^0,$$

$$S^{(T)}_{mn} = {}^0\!<T_m\,|\,E_n>^0 - 2\pi\,i\,\delta\,(E_m - E_n)\,R^{(T)}_{mn} -$$
$$- 2\pi i\,E_m\,\delta'\,(E_m - E_n)\,R^{(D)}_{mn} -$$
$$- 2\pi i\frac{E_m^2}{2}\,\delta''\,(E_m - E_n)\,R_{mn},$$

$$R^{(T)}_{mn} = {}^0\!<T_m\,|\,H^1\,|\,E_n, \text{in}> = {}^0\!<T_m\,|R|\,E_n>^0,$$

. . . . . .

Note that in the equations higher order poles, derivatives of the δ functions, etc., typical for multipole ghosts [6], appear.

Further consequences in respect of the analytic properties of the amplitudes can be drawn rather easily, and it is intended to deal with them together with the problem of the probabilistic interpretation and the two missing alternatives in a subsequent paper.

REFERENCES

1. K. F. Novobátzky, Zs. f. Phys., **111**, 292, 1938.
2. S. Mandelstam, Ann. Phys., **19**, 1, 1962.
3. S. N. Gupta, Proc. Phys. Soc., *A* **63**, 681, 1950. K. Bleuler, Helv. Phys. Acta, **23**, 567, 1950.
4. B. S. de Witt, Phys. Rev., **100**, 905, 1955.
5. Gy. Solt, Dissertation, unpublished.
6. K. L. Nagy, Acta Phys. Hung., **14**, 15, 1962.

ОБОБЩЕННЫЕ УРАВНЕНИЯ ЛИППМАНА—ШВИНГЕРА В ТЕОРИИ ПОЛЯ С МУЛЬТИПОЛЬНЫМ ПРИЗРАЧНЫМ СОСТОЯНИЕМ

К. Л. НАДЬ

Резюме

В работе выводятся обобщенные уравнения Липпмана—Швингера для теорий, в которых предполагается наличие индефинитной метрики и мультипольных призрачных состояний. Показывается, что и в этом случае переходные амплитуды даются элементами адиабатической S-матрицы.

FERMION SELF-MASSES AND LEHMANN'S SPECTRAL REPRESENTATION

By

G. PÓCSIK

INSTITUTE FOR THEORETICAL PHYSICS, ROLAND EÖTVÖS UNIVERSITY, BUDAPEST

(Presented by Z. Gyulai. — Received 20. IX. 1963)

LEHMANN 's theorem is extended to the case of the spinor self-coupling. In a simplified version of the new theorem the usual self-consistent mass equation holds.

§ 1. Introduction

One of the exact consequences of LEHMANN's spectral representation is the mass theorem. According to this theorem there exists an exact relation between the bare mass and the physical mass containing only the spectral functions. LEHMANN's theorem was originally proved for the PS(PS)-theory [1]. The refinement of the proof was given by MOFFAT [2]. MOFFAT introduced also a $\lambda\Phi^4$ term into the Lagrangian and proved that the $\lambda\Phi^4$ coupling produces most of the observable meson mass. LEHMANN's theorem was investigated for vector particles, too [3]. On the other hand, FORD has deduced these relations from the high-energy limit of the propagator [4].

In this paper fermion self-masses will be investigated, especially for a spinor self-coupling, on the basis of the spectral representation of the propagator, the field equations and the equal-time anticommutators (Lorentz-invariance, spectrality, positive definite norm in Hilbert space, charge conjugation and parity invariance are assumed).

All the meaningless integrals over momentum space are defined by using an invariant cut-off. Under such circumstances it is shown that a large self-mass can arise from the spinor self-coupling.

The theorem proved is in essence the same for general Lagrangians with vanishing bare mass and for γ_5-invariant ones (using only chirality violating propagators). Neglecting the continuous spectrum in the propagator, the usual self-consistent mass equation remain [5, 6].

§ 2. Fermion self-masses

Let us start with a very general field equation of the spin-half particles

$$(i\gamma^\mu \partial_\mu - m)\,\psi(x) = \sum_j O_j\,\psi(x)\,A^j(x), \tag{1}$$

where A^j is a functional of fields interacting with the field $\psi(x)$. O_j means the usual linearly independent matrices, 1, γ_μ, γ_5, $i\gamma_\mu\gamma_5$, $1/2i(\gamma_\mu\gamma_\nu - \gamma_\mu\gamma_\nu)$. The field equation is completed by the canonical commutation rules

$$\begin{aligned} \{\psi(x), \psi(y)\}_{x_0=y_0} &= 0\,, \\ \{\psi(x), \bar{\psi}(y)\}_{x_0=y_0} &= \gamma_0\,\delta^3(x-y), \end{aligned} \tag{2}$$

as well as the equal-time commutators of other fields with ψ vanishing. Proceeding, we note the well-known spectral representations of the two-point functions constructed from $\psi(x)$

$$S^{(\,)\prime}(x) = \int_0^\infty d\mu^2 \left(\varrho_1(\mu^2)\, S^{(\,)}(x;\mu) + \varrho_2(\mu^2)\, \Delta^{(\,)}(x;\mu^2)\right). \tag{3}$$

Here $S^{(\,)}$ means different $S^{(+)}$, $S^{(-)}$, ... functions, etc. In the following we shall study the equations (1, 2, 3).

First of all, the positive definiteness of the norm in Hilbert space leads to the inequalities $\varrho_1(\mu^2) \geqslant 0$, $2\,\mu\varrho_1(\mu^2) \geqslant \varrho_2(\mu^2) \geqslant 0$. The other condition satisfied by $\varrho_1(\mu^2)$ follows from (2) and (3)

$$\int_0^\infty d\mu^2\, \varrho_1(\mu^2) = 1\,. \tag{4}$$

Now, let us turn to the mass theorem. From (1) we have

$$\begin{aligned} &(i\gamma^\mu \partial_\mu - m) < \{\psi(x), \bar{\psi}(y)\} >_0 = \\ &= \sum_j < \{O_j\, \psi(x)\, A^j(x), \bar{\psi}(y)\} >_0 \equiv I(x,y). \end{aligned} \tag{5}$$

On the left hand side, the vacuum expectation value can be expressed by (3). Writing $x_0 = y_0$, we get

$$\begin{aligned} (I(x,y))_{x_0=y_0} = -i \int_0^\infty d\mu^2 \big[\varrho_1(\mu - m)\,(S(x-y;\mu))_{x_0=y_0} + \\ + \varrho_2\,((i\gamma_\nu \partial^\nu - m)\,\Delta(x-y;\mu^2))_{x_0=y_0}\big] = \gamma_0\,\delta^3(x-y) \int_0^\infty d\mu^2 \big[\varrho_1(\mu - m) - \varrho_2\big]. \end{aligned} \tag{6}$$

For a large class of field theories $(I(x,y))_{x_0=y_0}$ vanishes. Namely, if the functional A^j is independent from ψ and does not contain the bilinear form of a field, (6) gives for the self-mass $\delta m = \varkappa - m$

$$\delta m = \int_0^\infty d\mu^2 \big[(\varkappa - \mu)\,\varrho_1 + \varrho_2\big], \tag{7}$$

where $\varkappa$ denotes the observable fermion mass. (7) has been proved by LEHMANN for the PS(PS)-theory [1].

If A^j contains ψ, $A_j = g_j \bar{\psi} O_j \psi$ may be written. In this case we have

$$(I(x,y))_{x_9=y_0} = \sum_j [O_j < \psi(x)\,\bar{\psi}\,(x) >_0 + \\ + Tr\,(< \bar{\psi}\,(x)\,\psi(x) >_0 O_j^T)]\, g_j\, O^j\, \gamma_0\, \delta^3(x-y). \tag{8}$$

We have used the canonical anticommutators (2). The first (second) expectation value is defined as the limit of a T-product taken for $x_0 - y_0 \to +$ $+\ 0 (x_0 - y_0 \to -\ 0)$. In this way one arrives at the equation

$$(I(x,y))_{x_0=y_9} = \frac{1}{2} \sum_j [-\ O_j\, S'_F(0) + Tr\,(O_j\, S'_F(0))] \cdot g_j\, O^j\, \gamma_0\, \delta^3(x-y). \tag{9}$$

(9) can be expressed by the spectral functions. Since $S_F(0;\ \mu) = -\mu \Delta_F(0;\ \mu^2)$, it is clear that $S'_F(0)$ is a number,

$$S'_F(0) = \int_0^\infty d\mu^2\, (\varrho_2 - \mu \varrho_1)\, \Delta_F(0; \mu^2), \tag{10}$$

therefore, (9) can be replaced by

$$(I(x,y))_{x_0=y_0} = -\frac{1}{2}\,(-\ 3g_s + 4g_v + g_p + \\ + 4g_a + 16g_t)\, \gamma_0\, \delta^3(x-y)\, S'_F(0). \tag{11}$$

Here g_s means the coupling constant of the scalar part of the self-action, etc.

Comparing (6) with (11) and making use of (10) it is found that

$$m = \int_0^\infty d\mu^2\, (\mu \varrho_1 - \varrho_2)\, (1 - G \Delta_F(0; \mu^2)), \tag{12}$$

where

$$G = \frac{1}{2}\,(-\ 3g_s + 4g_v + g_p + 4g_a + 16g_t), \tag{13}$$

$$\Delta_F(0;\mu^2) = \frac{2i}{(2\pi)^4} \int (p^2 - \mu^2 + i\varepsilon)^{-1}\, d^4 p = \\ = \frac{1}{8\pi^2} \left[\Lambda^2 - \mu^2 \ln \left(\frac{\Lambda^2}{\mu^2} + 1 \right) \right].$$

Λ is an invariant cut-off momentum. The extension of LEHMANN's theorem for spinor self-coupling is given by (12).

If there exists an asymptotic field of mass $\varkappa$ (there is none for the THIRRING model, [7]) then because of invariance

$$\varrho_1(\mu^2) = Z \delta(\mu^2 - \varkappa^2) + \sigma_1(\mu^2), \tag{14}$$

where σ_1 has no pointwise singularity and in consequence of the inequalities concerning the spectral functions, the constant Z is positive, as well as $2\ \mu\sigma_1 \geqslant \geqslant \varrho_2 \geqslant 0$. Instead of (4) we have

$$Z + \int\limits_{(2\varkappa)^2}^{\infty} d\mu^2\, \sigma_1(\mu^2) = 1, \tag{15}$$

hence $Z \leqslant 1$ ($Z = 1$ is true only if $\sigma_1 \equiv \varrho_2 \equiv 0$). If (14) is valid, the physical mass is determined by (12). The corresponding self-mass is

$$\begin{aligned} \delta m = & \int\limits_{(2\varkappa)^2}^{\infty} d\mu^2 (\sigma_1 \mu - \varrho_2)\big(G\Delta_F(0;\mu^2) - 1\big) + \\ & + \varkappa\Big(\int\limits_{(2\varkappa)^2}^{\infty} d\mu^2\, \sigma_1 + ZG\,\Delta_F(0;\varkappa^2)\Big). \end{aligned} \tag{16}$$

In the first approximation, when the continuous part is neglected one can write

$$\delta m = \varkappa G \Delta_F(0;k^2). \tag{17}$$

This result has been obtained in [6] by using the functional integral method. For a vanishing bare mass (17) has the usual $\varkappa = 0$ and $\varkappa \neq 0$ solutions [5, 6]. If the Lagrangian is γ_5-invariant, these conclusions remain valid, provided that (3) is used. Thus, we see how the treatment based on the LEHMANN representation leads to NAMBU's self-consistent approximation, at the same time we possess the exact formula, (16), too.

REFERENCES

1. H. LEHMANN, Nuovo Cim., **11**, 342, 1954.
2. J. W. MOFFAT, Nucl. Phys., **13**, 150, 1959; **14**, 682, 1959/60.
3. K. JOHNSON, Nucl. Phys., **25**, 435, 1961.
   H. ROLLNIK, On the Self-Mass Problem of Vector Particles, preprint, 1962.
4. K. W. FORD, Nuovo Cim., **14**, 474, 1962.
5. Y. NAMBU, G. JONA—LASINIO, Phys. Rev., **122**, 345, 1961; **124**, 246, 1961;
   H. BANERJEE, Nuovo Cim., **23**, 587, 1962.
6. G. PÓCSIK, Acta Phys. Hung., **13**, 255, 1961; Nuovo Cim., **20**, 201, 1961.
7. K. JOHNSON, Nuovo Cim., **20**, 773, 1961.

СОБСТВЕННЫЕ МАССЫ ФЕРМИ-ЧАСТИЦ И СПЕКТРАЛЬНОЕ ПРЕДСТАВЛЕНИЕ ЛЕМАННА

ДЬ. ПОЧИК

Р е з ю м е

Теорема Леманна распространяется на случай спинорной самосвязи. В упрощённом варианте новой теоремы обычное самосогласованное уравнение масс удовлетворяется.

ON THE BEHAVIOUR OF GREEN FUNCTIONS AT SMALL DISTANCES

By

G. DOMOKOS and P. SURÁNYI

JOINT INSTITUTE FOR NUCLEAR RESEARCH, LABORATORY OF THEORETICAL PHYSICS, DUBNA, USSR

(Presented by Z. Gyulai. — Received 8. XI. 1963)

The asymptotic behaviour of the four-point Green function at small distances is investigated. The behaviour depends critically on the renormalizability of the field theory. For nonrenormalizable interactions the Bethe-Salpeter equation gives a well-defined solution, however, with an essential singularity at the origin. The physical implications of the results are discussed.

1. Introduction

The behaviour of certain quantities at small distances as resulting from quantum field theories has been the subject of investigation by many authors, both because of the principal interest of the problem and for more practical reasons, e.g. so as to be able to predict the high-energy behaviour of observable quantities, like transition probabilities etc. No definite and final answer has been obtained as yet and it is commonly believed that by answering the question as to the high-energy (or small-distance) behaviour of Green functions (GF) or S-matrix elements one would settle the problem of self-consistency of a quantum field theory as well.

It has been known for a long time that consistency requirements impose rather severe restrictions on the two- and three-point GF [1] but no such restrictions have been found so far for higher-order GF; in particular, the four-point GF is of special interest, because of its connection with the S-matrix element of elastic scattering, and so its close relation to experiments.

In the present article we study the four point field theoretical GF

$$G(x_1 \ldots x_4) = <0 \,|\, T(\Phi(x_1) \ldots \Phi(x_4)) \,|\, 0>,$$

where $\Phi(x)$ is a scalar field, with the help of the Bethe-Salpeter (BS) equation.

Taking for the BS kernel the lowest nonvanishing approximation, we find a close connection between the behaviour of the GF at small distances and the type of the interaction: We find different behaviours according to whether the dimension of the coupling constant involved is a negative power

of length ("superrenormalizable" interactions), a positive power of length (nonrenormalizable interaction) or dimensionless (renormalizable interactions).

At the end of the paper we discuss some possible physical implications of our results.

2. Transformation of the Bethe-Salpeter equation

The four-point GF is translation-invariant, therefore its Fourier transform depends on three independent momenta. Denoting the one-particle momenta by $p_1, \ldots, p_4$ we write:

$$p_1 = p + \frac{1}{2}E,$$

$$p_2 = -p + \frac{1}{2}E,$$

$$p_3 = q - \frac{1}{2}E,$$

$$p_4 = -q - \frac{1}{2}E,$$

so that the Fourier transform of $G(x_1, \ldots, x_4)$ is considered to depend on p, q, E. Energy-momentum is conserved: $\sum_{i=0}^{4} p_i = 0$, E being the total four-momenta in CMS. The function $G(p, q \mid E)$ satisfies the BS equation:

$$\left[\left(p + \frac{1}{2}E\right)^2 - m^2\right]\left[\left(p - \frac{1}{2}E\right)^2 - m^2\right] G(p, q \mid E) = \\ = \delta(p - q) + \frac{g^2}{(2\pi)^4 i}\int K(p, p' \mid E)\, G(p' q\, E)\, dp' . \tag{2.1}$$

We are interested in the behaviour of $G(p, q \mid E)$ for $|p| \to \infty$. This means that in (2,1) we can neglect E as compared to P. (We do not neglect the mass terms for the moment, in order to avoid "infrared" difficulties.) Dropping everywhere the variable E, we get

$$(p^2 - m^2)^2\, G(p, q) = \delta(p - q) + \\ + \frac{g^2}{(2\pi)^4 i} \cdot \int dp'\, K(p, p')\, G(p', q) . \tag{2.2}$$

We now want to go over to an Euclidean metric in (2.2), suggested by WICK [2].

This is possible if

$$K(p, p' \mid E = 0) \equiv K(p, p')$$

considered as a function of $\mathbf{p}$, $\mathbf{p}'$, p^0, $p^{0\prime}$, is analytic in p^0, $p^{0\prime}$ in the first and third quadrants of the complex p^0 and $p^{0\prime}$ planes, for real values of $\mathbf{p}$, $\mathbf{p}'$. We cannot prove this property in general for a BS kernel; for those types of kernels we are going to use, the condition is satisfied.

Hence, we obtain in Euclidean metric:

$$(p^2 + m^2)^2 G(p, q) = \delta(p - q) + \\ + \frac{g^2}{(2\pi)^4} \int dp' K(p, p') G(p', q), \tag{2.3}$$

where p, q already mean Euclidean vectors:

$$p^2 = p^{0^2} + \mathbf{p}^2 \text{ etc.}$$

We now observe that our eq. (2.3) is invariant under the four-dimensional rotation group, R_4, so it is natural to expand G and K in terms of the four-dimensional spherical harmonics:

$$G(p, q) = \sum_{nlm} Z_{nl}^m(\hat{p})\, G_{nl}(p^2, q^2)\, Z_{nl}^{m*}(\hat{q}), \tag{2.4}$$

where $Z_{nl}^m(\hat{e})$ is a normalized four-dimensional spherical harmonic, as defined e.g. in ref. [2], depending on the unit vector $\hat{e}$.

Inserting expansion (2.4) and a corresponding one for the kernel into eq. (2.3) we get for the radial functions:

$$(p^2 + m^2)^2 G_{nl}(p, q) = \frac{\delta(p - q)}{\sqrt{p^3 q^3}} + \\ + \frac{g^2}{(2\pi)^4} \int_0^\infty p^{3\prime}\, dp'\, K_{nl}(p, p')\, G_{nl}(p', q), \tag{2.5}$$

where the letters p, p', q in eq. (2.5) mean the moduli of the corresponding four-vectors. In what follows, we specialize the kernels as mentioned in the introduction and analyze the properties of the solutions of (2.5).

3. Lowest-order kernels

Our BS kernels are defined as the sum of the contributions of diagrams, regular at the point $E^2 = 4m^2$ in the complex E^2 plane. Assuming that the BS kernel

$$\frac{g^2}{(2\pi)^4 i} K(p, p' \mid E, g^2)$$

can be expanded in powers of the coupling constant g^2, the first few terms of the expansion are given by the contribution of the diagrams of Fig. 1. The diagrams of Fig. 1 are to be understood in a rather symbolic sense: some of the terms may be absent, internal lines may not represent the same particle as the external ones etc., depending on the type of interaction, we assume. Nevertheless, there is one important feature of the kernels in lowest order,

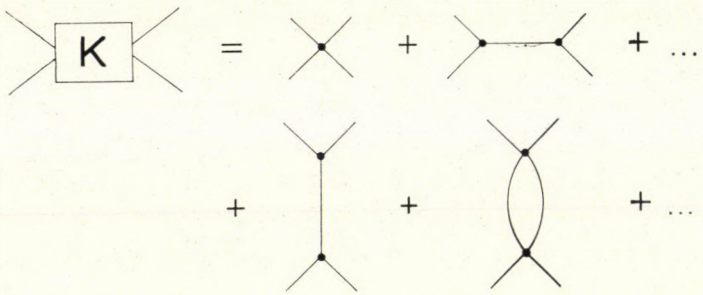


Fig. 1

we can immediately read off from Fig. 1. This feature is that the lowest-order approximations to a BS kernel is either of the form

$$K(p, p' \mid E) = f(E^2)\, P(E, p, p') \tag{3.1}$$

or

$$K(p, p'/E) = g((p - p')^2) Q(E, p, p'), \tag{3.2}$$

where the functions f and g possess a spectral representation of the form

$$\left.\begin{matrix} f(z) \\ g(z) \end{matrix}\right\} = \frac{1}{\pi} \int_{\mu_0}^{\infty} \frac{dx}{x - z} \left\{\begin{matrix} Im\, f(x + i\,0), \\ Im\, g(x + i\,0), \end{matrix}\right. \tag{3.3}$$

while P and Q are — invariant — polynomials in all of their variables. [In (3.3) we ignored subtractions for the time being.]

It has already been shown in ref. [3] that contributions of the type (3.1) differ from zero at certain discrete (n, l) values only and can be treated separately (they are essentially subtraction terms in the GF), so we confine our attention to kernels of the type (3.2). If $E = 0$, then Q is a polynomia in the invariants p^2, p'^2 and p, p'. Now, a function of the type (3.3) possesses a very simple expansion in terms of our Z-functions.

In fact, after a simple calculation we obtain [3]

$$g((p - p')^2) = \sum_{nlm} Z_{nl}^m(\hat{p})\, g_n\, Z_{nl}^{m*}(\hat{p}'). \tag{3.4}$$

where

$$g_n = \int_0^\infty r^3\, dr\, \frac{J_n(pr)}{pr}\, V(r)\, \frac{J_n(p'r)}{p'r}\,, \tag{3.5}$$

$$V(r) = \frac{1}{\pi} \int_{\mu 0}^{\infty} dx\, x^3\, \gamma(x)\, \frac{K_1(xr)}{xr}, \tag{3.6}$$

with the notation

$$\gamma(x) = Img(x + i0)\,.$$

With the help of eqs. (3.4)—(3.6) and the identity:

$$-\left[\frac{a^2}{dr^2} + \frac{3}{r}\,\frac{d}{dr} + \frac{1-n^2}{r^2}\right] \frac{J_n(pr)}{pr} = p^2\, \frac{J_n(pr)}{pr}\,, \tag{3.7}$$

we see that a kernel of the type (3.2) at $E = 0$ can be expressed as the Hankel transform of $V(r)$ and its derivatives; besides, we may notice that K_{nl} does not depend on n and l explicity but on n only, consequently so does G_{nl} as well.

4. The behaviour of the Green function at small distances

Representing the GF in the form a Hankel transform, like (3.5), results in:

$$G_n(p, p') = \int_0^\infty r^3\, dr \int_0^\infty r'^3\, dr' \times$$
$$\times\, \frac{J_n(pr)}{pr}\, G_n(r, r')\, \frac{J_n(p'r')}{p'r'}\,. \tag{4.1}$$

We see from (3.7) that $G_n(r, r')$ satisfies a differential equation:

$$D_n^2\, G_n \equiv \left[\frac{d^2}{dr^2} + \frac{3}{r}\,\frac{d}{dr} - \left(m^2 + \frac{n^2-1}{r^2}\right)\right]^2 G_n(r, r') =$$
$$= \frac{\delta(r - r')}{\sqrt{r^3 r'^3}} + \frac{q^2}{(2\pi)^4}\, W(r)\, G(r, r') \tag{4.2}$$

If r' is kept fixed at finite values, then in order to investigate the asymptotic behaviour for $r \to 0$ we can drop the δ-function in (4.2). $W(r)$ is the "effective potential" and is of the form:

$$W(r) = \tilde{Q}(D_n)\, V(r)\,, \tag{4.3}$$

where $\tilde{Q}(D_n)$ is some polynomial in the differential operator defined by (4.2).

Now, as it is well known, the asymptotic behaviour of $G(r, r')$ for $r \to 0$ depends on the behaviour of $W(r)$ for small r. The latter can be found, however, by simple dimensional considerations. In fact, as one can see immediately from (4.2) the quantity $g^2 W(r)$ must be of the dimension L^{-4} (in units, where $\hbar = c = 1$). On the other hand, from (3.6) and (4.3) one finds that $W(r)$ may asymptotically for $r \to 0$ be described by a function of the following form:

$$W(r) = O(r^{-4-\alpha}). \tag{4.4}$$

Now, if the coupling constant in question has the dimension L^{η} we immediately obtain:

$$\alpha = 2\eta. \tag{4.5}$$

On the other hand, as eq. (4.2) shows, $G(r, r')$ for $r \to 0$ behaves like a power function if $\alpha \leqslant 0$. For $\alpha > 0$ $G(r, r')$ has an essential singularity at the origin; as shown in ref. [3] it behaves as follows:

$$G(r, r') \sim \exp \sigma r^{-\alpha/4}, \qquad (r \to 0)$$

with

$$\sigma = \frac{4g^{1/2}}{\alpha} \exp iN\frac{\pi}{2},$$

$$(N - \text{integer}).$$

5. Discussion

The three classes of solutions $\alpha < 0$, $\alpha = 0$, $\alpha > 0$ correspond physically by (4.5) to superrenormalizable, renormalizable and nonrenormalizable interactions. The difference between superrenormalizable and renormalizable interactions is that for the latter the interaction at small distances is of the same order of magnitude as the kinetic term: while for the former the interaction is weaker and only the kinetic terms survive as $r \to 0$.

(Mathematically the difference appears in the analytic properties of G_n as a function of n: for $\alpha < 0$, G_n is meromorphic in n, for $\alpha = 0$ it has a series of branch cuts [3]).

As an example of the third class ($\alpha > 0$), one can consider the following model. Two scalar particles interact through the exchange of a fermion pair. Assuming the interaction to be of the form:

$$\mathscr{L}_{\text{int}} = f(\Phi \partial_\mu \Phi)(\bar{\psi} \gamma\mu \psi),$$

one obtains $\alpha = 4$ in (4.4) and

$$G(r, r') \sim e^{\frac{\sqrt{f}}{r}\varrho}$$

$$(\varrho = \pm 1, \pm i),$$

depending on the solution we choose. At any rate, one can show that there is one and only one solution which satisfies the integral equation (2.5).

A solution of the type indicated above has, however, rather strange and unusual properties. It is obviously singular in the coupling constant, but this is probably not a property one can observe experimentally. However, when transformed back to momentum space, one observes that the corresponding scattering amplitude has an essential singularity in the momentum transfer

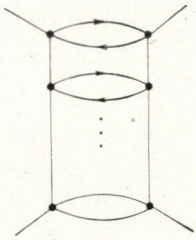


Fig. 2

$t = (p - p')^2$ (equivalently, by a crossing transformation, one can regard this function as a multiperipheral graph, contributing to a high energy process — Fig. 2 — where p plays the role of the CM momentum). The strange behaviour of this scattering amplitude

$$T(t) = T(t, s)\,|_{s=0} \sim e^{i\sqrt{f|t|}}$$

for $t \to \infty$ implies that this amplitude does not satisfy a dispersion relation with a finite number of subtractions. The situation reminds one of the case of a hard core potential, with radius $R = \sqrt{f}$; from quantum mechanics one knows that it is not the scattering amplitude $T(t)$ itself, but rather $e^{-ikR}T$ which satisfies a dispersion relation. At any rate, one sees that the coordinate-space method gives a unique solution; while if one tried to iterate the one-loop diagram, one would need new subtraction constants at every step of iteration.

It is fairly obvious that our procedure of determining the asymptotic behaviour of the Green function is far from being rigorous or complete. (In particular, our approximation to the BS kernel is at least at first sight- very crude, and one cannot even estimate the effects of a "better" kernel on the solution. Nonetheless, we believe our results are quite suggestive as to the essential difference between interactions of different types. It is not excluded that in consequence of similar investigations one has to review the common opinion about the negligible role of weak interactions in our "usual" world.

The authors would like to express their sincere thanks to Profs. D. I. Blokhintsev and N. N. Bogoliubov for their interest in the present work.

They are also grateful for a series of very interesting discussions with Prof. M. CINI during his visit to Dubna and have learnt with pleasure that a group of Italian physicists has come independently to much the same conclusions [4] as the authors in [3] and in the present work.

REFERENCES

1. H. LEHMANN, K. SYMANZIK and F. ZIMMERMANN, Nuovo Cim., **2**, 425, 1955.
2. G. C. SICK, Phys. Rev., **96**, 1124, 1954.
3. G. DOMOKOS and P. SURÁNYI, Nucl. Phys. (to be published); and preprint JINR E—1400, 1963.
4. A. BASTAI, L. BERTOCCHI, S. FUBINI and G. FURLAN, Preprint CERN 7229/TH 374, 1963; A. BASTAI, L. BERTOCCHI, G. FURLAN and M. TONIN, preprint CERN 7081/TH 370, 1963.

О ПОВЕДЕНИИ ФУНКЦИИ ГРИНА НА МАЛЫХ РАССТОЯНИЯХ

Г. ДОМОКОШ и П. ШУРАНИ

Р е з ю м е

Исследуется асимптотическое поведение четырёхполюсной функции Грина на малых расстояниях. Это поведение критически зависит от возможности ренормализации теории поля. Для неренормализуемых взаимодействий уравнение Бете—Сольпитера дает полностью определённое решение, однако последнее имеет значительную сингулярность при начале координат. Результаты дискутируются с физической точки зрения.

ON A SLIGHT MODIFICATION OF HORI'S STRONG-COUPLING METHOD

By

G. HEBER

JOINT INSTITUTE FOR NUCLEAR RESEARCH, DUBNA, USSR*

(Received 20. IX. 1963)

In order to be able to perform in HORI's method the necessary number of functional differentiations we propose to expand the vacuum expectation value of the S-matrix (with external sources) and the generating functional of the S-matrix in Volterra series. The coefficients of the Volterra series for the generating functional Ω can be calculated from functional integrals by making extensive use of the distribution-theoretical definition of the products of singular functions. The coefficients of the Volterra series for the S-matrix, which have direct physical meaning, appear as infinite series of certain integrals over the coefficients of the Ω-series. The main difficulty is summing up this series. Presently the author cannot solve this difficulty, see, however [6].

1. Introduction

Performing some calculations by means of the HORI method [1] the author found the following difficulty [2, 3]: In order to deduce a physically acceptable Feynman amplitude one has to operate with a big number of functional differentiations (with respect to the external sources) of the generating functional Ω of the vacuum expectation value of the S-matrix. However, Ω depends in interesting cases in a very complicated way on the external sources. Therefore one practically can not calculate higher orders of Feynman amplitudes with HORI's method. Here the author proposes to modify the HORI method slightly: Do not calculate the generating functional Ω in a closed form but write it as a Volterra series (with respect to the external sources). The coefficients of this series can be written down as functional integrals and calculated (in our example) exactly by means of distribution analysis. If the vacuum-expectation-value of the S-matrix is also written as a Volterra series, the coefficient functions of this series can be represented as infinite series of certain integrals over the coefficient functions of the Ω-functional. The main difficulty of the present method is to sum up this infinite series. This difficulty we cannot solve here.

* On leave of absence from University of Leipzig.

2. Volterra series for Ω

In HORI's work [1] the generating functional Ω is introduced; in some simple examples, HORI could give closed expressions for this functional by means of functional integration. — HORI's definition of Ω reads as follows:

$$\Omega[\varrho, \varrho^+, j] = \iiint \exp\left[-i\int (g\psi^+\psi\Phi - \varrho^+\psi - \varrho\psi^+ - j\Phi)\, d^4x\right] D(\psi, \psi^+, \Phi). \quad (1)$$

Here and throughout this work we use the example of one real (Φ) and one complex (ψ, ψ^+) scalar field, interacting by the term $g\psi^+\psi\Phi$ in the Lagrangian. The terms $\varrho^+\psi + \varrho\psi^+ + j\Phi$ represent the interaction of the fields with the external sources ϱ, ϱ^+, j. The free Lagrangian would read $L_0 = \psi^+(\Box - m^2)\psi + \Phi(\Box - \mu^2)\Phi$. The symbol $D(\psi^+, \psi, \Phi)$ in (1) stands for the integration element in the space of the functions ψ, ψ^+, Φ. As already stated, we propose to represent Ω as a Volterra series (see e.g. RZEWUSKI [4]):

$$\Omega[\varrho, \varrho^+, j] = \sum_{l,m,n=0}^{\infty} \frac{(i)^{l+m+n}}{l!\, m!\, n!} \int\ldots\int d^4x_1 \ldots d^4x_n \cdot d^4y_1 \ldots d^4y_l \cdot d^4z_1 \ldots d^4z_m \times$$
$$\times\, \varphi_{n,l,m}(x_1 \ldots x_n; y_1 \ldots y_l; z_1 \ldots z_m) \cdot e_{n,l,m}(x_1 \ldots z_m), \quad (2)$$

with

$$e_{n,l,m}(x_1 \ldots z_m) = \varrho(x_1) \ldots \varrho(x_n)\, \varrho^+(y_1) \ldots \varrho^+(y_l)\, j(z_1) \ldots j(z_m).$$

The functions $\varphi_{n,l,m}(x_1 \ldots z_m)$, of course, determine the functional Ω. They are symmetric with respect to the variables $x_1 \ldots x_n$; $y_1 \ldots y_l$, $z_1 \ldots z_m$. (Here and in the following we assume, that our Volterra series converges uniformly at the "point" $\varrho = \varrho^+ = j = 0$; whether this is true, is, however, not certain.) By comparing eqs. (1) and (2) one gets the expressions:

$$(i)^{m+l+m} \cdot \varphi_{n,l,m}(x_1 \ldots z_m) =$$
$$= \lim_{\substack{\varrho\to 0 \\ \varrho^+\to 0 \\ j\to 0}} \frac{\delta^{(n+l+m)}\,\Omega}{\delta\varrho(x_1) \ldots \delta\varrho(x_n) \cdot \delta\varrho^+(y_1) \ldots \delta\varrho^+(y_l) \cdot \delta j(z_1) \ldots \delta j(z_m)} =$$
$$= (i)^{n+l+m} \iiint \psi^+(x_1) \ldots \psi^+(x_n)\, \psi(y_1) \ldots \psi(y_l)\, \Phi(z_1) \ldots$$
$$\ldots \Phi(z_m) \exp\left(-ig\int \psi^+\psi\Phi\, d^4x\right) D(\psi^+, \psi, \Phi). \quad (3)$$

The functional integrals (3) we shall calculate directly. The representation (2) for Ω enables one to perform any number of functional integrations of Ω very easily. It is also very useful, because at the end of the calculation of Feynman amplitudes (which always is our aim, of course) always the limiting process $\varrho \to 0$, $\varrho^+ \to 0$, $j \to 0$ has to be performed. Let us calculate at first φ_{000} defined by

$$\varphi_{000} = \iiint \exp\left(-ig\int \psi^+\psi\Phi\, d^4x\right) D(\psi^+, \psi, \Phi). \quad (4)$$

Using the definition of the δ-functional

$$\delta[f(x)] = \int \exp\left(-2\pi i \int f(x)\,\Phi(x)\,d^4x\right) D(\Phi), \tag{5}$$

we can at once perform the functional integration with respect to Φ. This gives:

$$\varphi_{000} = \int\int \delta\left[\frac{g}{2\pi}\,\psi^+\,\psi\right] D(\psi^+, \psi).$$

Writing $\psi = re^{ia}$ and introducing the lattice space x_k instead of the continuous x-space, one has

$$\varphi_{000} = \lim_k \Pi \left(\int_0^\infty \int_0^{2\pi} \delta\left(\frac{2\pi}{g}\,r_k^2\right) r_k\,dr_k\,da_k \right).$$

(The symbol $\lim \Pi_k$ means: Take the product over all points of the lattice space, then go to the limit of the x-continuum.) That is, we have: $\varphi_{000} = \lim_k \Pi(\omega_k)$;

$$\omega_k = 2\pi \int_0^\infty \delta\left(\frac{g}{2\pi}\,r_k^2\right) r_k\,dr_k = \frac{2\pi^2}{g} \int_0^\infty \delta(y)\,dy.$$

Now we write

$$\int_0^\infty \delta(y)\,dy = \int_{-\infty}^\infty \vartheta(y)\,\delta(y)\,dy, \quad \text{where } \vartheta(y) = \begin{cases} 1 & \text{for } y > 0, \\ 0 & \text{for } y < 0. \end{cases}$$

The product of the improper functions $\delta(y)$ and $\vartheta(y)$ in the integrand must be treated by means of distribution analysis. In the work [5] of GÜTTINGER we find the rule

$$\vartheta(y) \cdot \delta(y) = -\,c_0\,\delta(y), \tag{6}$$

where c_0 is a finite, but quite arbitrary constant. Therefore we get

$$\omega_k = -\,\frac{2\pi^2}{g}\,c_0$$

and

$$\varphi_{000} = \lim_k \Pi \left(-\,\frac{2\pi^2}{g}\,c^2 \right). \tag{7}$$

From (7) one would like to conclude, that φ_{000} is a highly singular quantity. But using the rules for products of the type $\lim_k \Pi f(x_k)$ determined in [2], we arrive at the expression:

$$\varphi_{000} = \exp\left(c_1^2 \ln\left(-\,\frac{2\pi^2}{g}\,c_0\right)\right), \tag{8}$$

where c_1 is another finite, arbitrary constant, which stems from GÜTTINGER's rule (see [5])

$$\delta^2(y) = c_1\,\delta(y). \tag{9}$$

We especially point out that φ_{000} for $g \to 0$ and arbitrary c_1, c_0 is singular (see also [3]).

Next let us discuss the function

$$\varphi_{110}(x_1, y_1) = \iiint \psi^+(x_1)\,\psi(y_1) \exp\left(-ig\int \psi^+\psi\Phi\, d^4x\right) D(\psi^+, \psi, \Phi). \tag{10}$$

Note that all functions of type $\varphi_{n,0,0}$ $(n \neq 0)$ and $\varphi_{0,l,0}$ $(l \neq 0)$ and $\varphi_{n,l,0}$ with $n \neq l$ are zero because they contain at least one integral of the type $\int_0^{2\pi} e^{ia} da$. The function is zero for all $x_1 = y_1$, as in that case also the integral $\int_0^{2\pi} e^{ia}\, da$ appears when $\psi = re^{ia}$ is introduced. For $x_1 = y_1$, however, one gets with $\psi(x_1) = r_1\, e^{ia_1}$

$$\varphi_{110}(x_1, x_1) = \lim \prod_k{}' \left(\int_0^\infty \int_0^{2\pi} \delta\left(\frac{g}{2\pi} r_k^2\right) r_k\, dr_k\, da_k \right) \cdot \int_0^\infty \int_0^{2\pi} r_1^2\, \delta\left(\frac{g}{2\pi} r_1^2\right) r_1\, dr_1\, da_1 .$$

(The symbol $\prod_k{}'$ means: Take the product over all points of the lattice space with the exception of the point x_1.) That means, we can write:

$$\varphi_{110}(x_1, x_1) = \varphi_{000} \cdot \frac{\int_0^\infty r^2\, \delta\left(\frac{g}{2\pi}\, r^2\right) d(r^2)}{-\,c_0 \dfrac{2\pi}{g}} .$$

Now we have

$$\int_0^\infty r^2\, \delta\left(\frac{g}{2\pi}\, r^2\right) d(r^2) = \left(\frac{2\pi}{g}\right)^2 \int_{-\infty}^{+\infty} y\, \vartheta(y)\, \delta(y)\, dy.$$

Once more we use formula (6) and get for our integral

$$-\,c_0 \left(\frac{2\pi}{g}\right)^2 \int_{-\infty}^{+\infty} \delta(y) \cdot y \cdot dy = 0,$$

Thus, we arrive at the result:

$$\varphi_{110}(x_1, y_1) \equiv 0. \tag{11}$$

By a quite similar calculation we can show that

$$\varphi_{n,n,0}(x_1 \ldots y_n) \equiv 0 \quad \text{for } n \neq 0. \tag{12}$$

Next we are interested in the function

$$\varphi_{002}(z_1, z_2) = \iiint \Phi(z_1)\,\Phi(z_2) \exp\left(-ig\int \psi^+\,\psi\Phi\, d^4x\right) D(\psi, \psi^+, \Phi). \qquad (13)$$

Let us first suppose $z_1 \neq z_2$. Then we get

$$\varphi_{002} = \lim_k \Pi'' \left(-\frac{2\pi^2}{g}\, c_0\right) \left(\frac{-i\pi}{g} \int_0^\infty \delta^{(1)}(y)\, dy\right)^2,$$

where

$$\delta^{(1)}(y) = \frac{d}{dy}\,\delta(y), \quad \text{i.e.}$$

$$\int \Phi_k \exp(-2\pi i y\, \Phi_k)\, d\Phi_k = \frac{i}{2\pi}\,\frac{d}{dy}\,\delta(y).$$

Now we have

$$\int_0^\infty \delta^{(1)}(y)\, dy = \int_{-\infty}^{+\infty} \vartheta(y)\,\delta^{(1)}(y)\, dy = -c_0 \int_{-\infty}^{+\infty} \delta^{(1)}(y)\, dy - c_1 \int_{-\infty}^{+\infty} \delta(y)\, dy = -c_1$$

(see [5]). This leads to

$$\varphi_{002}(z_1, z_2) = \varphi_{000}\left(-\frac{c_1^2}{4\pi^2 c^2}\right) \quad \text{for } z_1 \neq z_2. \qquad (14a)$$

For $z_1 = z_2$, however, we get

$$\varphi_{002}(z_1, z_1) = \lim_k \Pi' \left(-\frac{2\pi^2}{g}\, c_0\right) \left(-\frac{1}{2g} \int_0^\infty \delta^{(2)}(y)\, dy\right).$$

With [5] we have

$$\delta^{(2)}(y)\,\vartheta(y) = -c_0\,\delta^{(2)}(y) - c_1\,\delta^{(1)}(y) - c_2\delta(y)$$

(c_0, c_1, c_2 finite, arbitrary constants), so that $\int_0^\infty \delta^{(2)}(y)dy = -c_2$. Finally we have:

$$\varphi_{002}(z_1, z_1) = -\varphi_{000} \cdot \frac{c_2}{4\pi^2 c_0}. \qquad (14b)$$

As the last example for the calculation of coefficients from the Volterra series (2) we take

$$\varphi_{112}(x_1, y_1, z_1, z_2) = \iiint \psi^+(x_1)\,\psi(y_1)\,\Phi(z_1)\,\Phi(z_2) \exp\left(-ig\int \psi^+\,\psi\,\Phi d^4x\right) D(\psi^+, \psi, \Phi). \qquad (15)$$

The functional integration over Φ can be performed in the same manner as in the function φ_{002}. Because of integration over the phases of the functions

* Π''_k means: leave out the points $x_k = z_1$ and $x_k = z_2$ from the product!

ψ^+ and ψ we get the result

$$\varphi_{112} = 0 \text{ for } x_1 \neq y_1.$$

Calculations similar to those performed earlier show, that φ_{112} is different from zero only, if $x_1 = y_1 = z_1$ or $x_1 = y_1 = z_2$ or even $x_1 = y_1 = z_1 = z_2$. More exactly we get:

$$\varphi_{112}(x_1, y_1, z_1, z_2) = \begin{cases} \varphi_{000}\left(-\frac{c_1}{2\pi g c_0}\right) \text{for} \begin{cases} x_1 = y_1 = z_1, & \text{or} \\ x_1 = y_1 = z_2, & \text{or} \\ x_1 = y_1 = z_1 = z_2 \end{cases} \\ 0 \qquad\qquad \text{otherwise} \end{cases} \tag{16}$$

Higher φ-functions can be calculated in a very similar manner; they all can be calculated exactly.

3. Connection between the φ-functions and Feynman amplitudes

It turns out to be useful to expand also the vacuum-expectation-value of the S-matrix in a Volterra series:

$$S_{\text{vac}}[\varrho, \varrho^+, j] = \sum_{n,l,m} \frac{(i)^{n+l+m}}{n!\, l!\, m!} \int \dots \int d^4 x_1 \dots d^4 z_m \times \\ \times T_{n,l,m}(x_1, \dots z_m)\, e_{n,l,m}(x_1, \dots, z_m) \tag{17}$$

The quantities $e_{n,l,m}$ are the same as in equ. (2). The functions $T_{n,l,m}$, of course, determine the functional $S_{\text{vac}}[\varrho, \varrho^+, j]$. They are (up to a constant) already the Feynman amplitudes, because by definition e.g.

$$\chi(x_1, y_1) = \lim_{\substack{\varrho \to 0 \\ \varrho^+ \to 0 \\ j \to 0}} S_{\text{vae}}^{-1} \frac{\delta^2 S_{\text{vac}}}{\delta \varrho^+(x_1)\, \delta \varrho(y_1)} = \\ = -\frac{T_{110}(x_1, y_1)}{T_{000}}. \tag{18}$$

(Here one can see once more, that the expansion (17) is useful.)

To establish the relation between the coefficients $T_{n,l,m}$ and $\varphi_{n,l,m}$, we start from the equation

$$S_{\text{vac}} = \exp(-i\omega)\, \Omega, \tag{19}$$

with

$$\omega = \iint \left(\frac{\delta}{\delta \varrho(x)} G_k(x, y) \frac{\delta}{\delta \varrho^+(y)} + \frac{\delta}{\delta j(x)} G_\mu(x, y) \frac{\delta}{\delta j(y)} \right) d^4 x\, d^4 y$$

and

$$G_\lambda(x,y) = \delta(x-y)\,(\Box_y - \lambda^2)$$

(see [1]).

Consider now the functional

$$\Phi_{n,l,m}[\varrho,\varrho^+,j] = \int\ldots\int d^4x_1\ldots d^4z_m\,\varphi_{n,l,m}(x_1,\ldots,z_m)\,e_{n,l,m}(x_1,\ldots,z_m).$$

It can be proved easily that

$$\frac{\delta\Phi_{n,l,m}}{\delta\varrho(x)} = n\int\ldots\int d^4x_2\ldots d^4z_m\cdot\varphi_{n,l,m}(x,x_2,\ldots,z_m)\,e_{n-1,l,m}(x_2,\ldots,z_m)$$

and

$$\iint\frac{\delta}{\delta\varrho(x)}\,G(x,y)\,\frac{\delta}{\delta\varrho^+(y)}\,d^4x\,d^4y\,\Phi_{n,l,m} =$$

$$= n\cdot l\int\ldots\int d^4x_2\ldots d^4x_n\,d^4y_2\ldots d^4y_l\,d^4z_1\ldots d^4z_m\times$$

$$\times G(x,y)\,\varphi_{n,l,m}(x,x_2,\ldots,x_n;\;y_1,y_2,\ldots,y_l;\;z_1,\ldots,z_m)\times$$

$$\times\varrho(x_2)\ldots\varrho(x_n)\,\varrho^+(y_2)\ldots\varrho^+(y_l)\,j(z_1)\ldots j(z_m).$$

Using this formula one can write up easily for each $T_{n,l,m}$ a series of the following type:

$$T_{000} = \varphi_{000} + i\iint G_k(x_1,y_1)\,\varphi_{110}(x_1;y_1)\,d^4x_1\,d^4y_1 +$$

$$+ i\iint G_\mu(z_1,z_2)\,\varphi_{002}(z_1,z_2)\,d^4z_1\,d^4z_2 -$$

$$- \iiiint G_k(x_1,y_1)\,G_\mu(z_1,z_2)\,\varphi_{112}(x_1,y_1;\;z_1,z_2)\,d^4x_1\,d^4y_1\,d^4z_1\,d^4z_2 + \ldots \qquad (20)$$

$$T^{1,1,0}(x_1,y_1) = \varphi_{110}(x_1,y_1) + i\iint G_k(x_2,y_2)\,\varphi_{220}(x_1,x_2;y_1,y_2)\,d^4x_2\,d^4y_2 +$$

$$+ i\iint G_\mu(z_1,z_2)\,\varphi_{112}(x_1;y_1;z_1,z_2)\,d^4z_1\,d^4z_2 -$$

$$- \frac{1}{2}\iiiint G_k(x_2,y_2)\,G_k(x_3,y_3)\,\varphi_{330}(x_1,x_2,x_3;y_1,y_2,y_3)\,d^4x_2\,d^4x_3\,d^4y_2\,d^4y_3 -$$

$$- \iiiint G_k(x_2,y_2)\,G_\mu(z_1,z_2)\,\varphi_{222}(x_1,x_2;y_1,y_2;z_1,z_2)\,d^4x_2\,d^4y_2\,d^4z_1\,d^4z_2 -$$

$$- \frac{1}{2}\iiiint G_\mu(z_1,z_2)\,G_\mu(z_3,z_4)\,\varphi_{114}(x_1;y_1;z_1,z_2,z_3,z_4)\,d^4z_1\,d^4z_2\,d^4z_3\,d^4z_4 + \ldots \qquad (21)$$

The rule for the construction of these infinite series can easily be concluded. It is of no great value to write up the general series for $T_{n,l,m}$, but of course, this could be done.

4. Discussion

In order to examine the value of the formalism proposed above, we introduce into (21) the φ-functions from section 2 of the present paper and remember, that T_{110} has (up to a constant, see (18)) the meaning of the propagation function for the ψ-particles.

In the series (21) the zeroth-order term vanishes ($\varphi_{110} \equiv 0$). From the two first-order terms only one does not vanish (φ_{112}), but it is different from zero only for $x_1 = y_1$. That means, we do not have any propagation. This is well understandable: Propagation is represented by the term L_0 of the Lagrangian. We have in Ω, however, only the interaction part of L. This part does not contain derivatives.

Therefore propagation should appear only in higher orders. In the next order the term $\varphi_{330} \equiv 0$. φ_{114} is only different from zero, if $x_1 = y_1$. However, φ_{222} is different from zero also for $x_1 \neq y_1$, namely for $x_1 = y_2$, $x_2 = y_1$. But performing the integrations over x_2 and y_2 the factor $\delta(x_2 - y_2)$ appears in $G_k(x_2, y_2)$. That means, only the points $x_2 = y_2$ survive and T_{110} is also in this order different from zero only for $x_1 = y_1$.

One can prove the same property of T_{110} also in higher orders. It would mean, that one can get a real propagation function only after summing up the whole infinite series (21). How to extract physical information in spite of this difficulty, is shown in [6].

5. Acknowledgement

This work was completed at the Joint Institute for Nuclear Research, Dubna. It is a pleasure to the author, to thank the Direction of the Institute for the kind invitation, the hospitality and the interest in this work.

REFERENCES

1. S. Hori, Nucl. Phys., **30**, 644, 1962.
2. G. Heber, preprint TUL 4, Leipzig 1963.
3. G. Heber, preprint TUL 5, Leipzig 1963.
4. J. Rzewuski, CERN-preprint 3212/Th. 251.
5. W. Güttinger, Progr. Theor. Phys., **13**, 612, 1955.
6. G. Heber and H. J. Kaiser, Dubna-preprint E 1500 (1964).

О НЕБОЛЬШОЙ МОДИФИКАЦИИ МЕТОДА СИЛЬНОЙ СВЯЗИ ХОРИ

Г. ХЕБЕР

Р е з ю м е

Для проведения необходимых операций функционального дифференцирования в методе Хори автором предлагается разложение вакуумного среднего значения S-матрицы (с внешними источниками) и производящего функционала S-матрицы в ряд Вольтерры. Коэффициенты ряда Вольтерры для производящего функционала Ω могут быть вычислены из функциональных интегралов путём широкого применения определения произведений особенных функций в теории обобщенных функций. Коэффициенты ряда Вольтерры для S-матрицы, имеющие прямой физический смысл, появляются в виде бесконечного ряда некоторых интегралов по коэффициентам ряда Ω. Основная трудность заключается в суммировании данного ряда. Автор был не в силах сразу же решить эту проблему.

MODEL WITH SUPERCONDUCTING SOLUTION IN QUANTUM FIELD THEORY II

By

G. KUTI and G. MARX

INSTITUTE FOR THEORETICAL PHYSICS, ROLAND EÖTVÖS UNIVERSITY, BUDAPEST

(Presented by Z. Gyulai. — Received 20. IX. 1963)

The energy spectrum of a self-coupled scalar field with two separate ground states is calculated using the variational method of RITZ. Two types of interacting particles are obtained. Their energies turn out to be finite after a suitable mass renormalization. Finally the possible physical consequences of the model are discussed.

§ 1. Introduction

Science, since its birth, has been searching for the ultimate unique "element" to which the whole variety of matter in the universe may be reduced. After centuries spent in speculation, the theory of general relativity gave the first realistic promise to explain besides the gravitational field also the electromagnetic and other matter fields. Already one or two decades earlier, besides A. EINSTEIN, H. WEYL, E. SCHRÖDINGER and other eminent scientists, also K. F. NOVOBÁTZKY was engaged in the investigation of the mathematical possibilites of such a unified field theory. Nowadays many scientists try to realize this idea, which is thousands of years old, in the frame of the quantum field theory. Also the present paper was born under the influence of this tendency: a simple example is investigated, which seemed to be apt to explain two different particles as two forms of one and the same field.

Let us consider a self-coupled scalar field [1] which is described by the field equation (1):

$$\Box \varphi + \frac{1}{2} \mu_0^2 \varphi - \lambda_0^2 \varphi^3 = 0 . \tag{1}$$

In the classical field theory this has several homogeneous — force-free — ground states, namely the solutions $\varphi \equiv 0$ and $\varphi \equiv \pm \mu_0/\lambda_0 \sqrt{2}$. The latter are stable. The excitations of the field with small energy give rise to oscillations of small amplitude around the one or the other stable ground state. It may be expected that the quantum field theoretical treatment will lead to the conclusion that the scalar field described by (1) will account for two types of particles (quanta), which interact with each other. Investigating a model which in principle is similar D. J. BLOKHINTSEV was the first to propose the

use of the field with two ground states for the description of two different elementary particles [2]. BLOKHINTSEV has made use of the lattice space approximation where in the field energy of the type

$$H = \int \left[\frac{1}{2} \pi^2 + \frac{1}{2} (\nabla \varphi)^2 + V(\varphi) \right] d^3 \mathbf{r} \tag{2}$$

the term containing $(\bar{\nabla}\varphi)^2$ — which produces the coupling between the anharmonic oscillators, oscillating in different points of the space — is considered to be a perturbation.

In a previous paper [3] we considered as a perturbation the terms φ^3 and φ^4 in the expression of $V(\varphi)$ which couple oscillators of different wave number. In this treatment the interaction induced by the zero point fluctuations of the field was omitted. It was not possible to see what kind of inference might be drawn as regards orthogonality of one kind of particle states to the other, as information was given only on the orthogonality of the unperturbed eigenfunctions. Indeed, R. HAAG has shown [4] that the "distance" between the exact and the unperturbed eigenstates in the space of the states may be large, and it may happen, too, that the two states do not even belong to the same separable Hilbert space.

The results obtained by mathematical methods, which were not completely satisfying, called the attention to some very exciting physical possibilities. The results of the calculations hinted at the fact that interactions between particles of the same type are *strong* while they are *weak* between two different particles. One might thus hope to find that besides the different families of elementary particles also the different types of couplings are inherent properties of one and the same field. The variety of the forms of states and interactions have then to be reduced to the multiplicity of the ground state of the field.

With this in mind it seems desirable to investigate the energy spectrum of the field with several ground states, making use of more exact and more accurate mathematical tools than the perturbation theory. In the present paper the variational method is applied to the case of the self-coupled scalar field with degenerate ground state that has been successfully used by L. I. SCHIFF [5] in the case of the scalar field with nondegenerate ground state. The mathematical structure of the energy eigenfunctions is restricted by several assumptions. From among the restricted ensemble of functions the optimal eigenfunctions of the Hamiltonian are determined. It will be shown in the course of the calculations that the expansion in a series in powers of the coupling constant λ_0 can be avoided. This expansion is, as a matter of fact, forbidden, as we have pointed out earlier [3]. We shall see that the zero point fluctuations give rise to an important interaction among the field oscillators of different wave number.

§ 2. The Hamiltonian

The Lagrangian density of the field equation (1) is given by

$$L = -\frac{1}{2}\partial_\mu \varphi \partial_\mu \varphi + \frac{1}{4}\mu_0^2 \varphi^2 - \frac{1}{4}\lambda_0^2 \varphi^4 , \tag{3}$$

where μ_0 and λ_0 are real and positive constants. It is worth while to note that the sign of the term containing φ^2 is different from the usual one. The operator conjugated canonically to φ (in the $\hbar = c = 1$ system of units) is

$$\pi = \frac{\partial L}{\partial \dot{\varphi}} = \dot{\varphi} .$$

The commutation relations for the same time are taken in the usual form:

$$\begin{aligned} [\varphi(\mathbf{r}, t), \varphi(\mathbf{r}', t)] = [\pi(\mathbf{r}, t), \pi(\mathbf{r}', t)] = 0 , \\ [\varphi(\mathbf{r}, t), \pi(\mathbf{r}', t)] = i\,\delta(\mathbf{r} - \mathbf{r}') . \end{aligned} \tag{4}$$

The energy operator and the field momentum are

$$H = \int (\pi\dot{\varphi} - L)\, d^3\mathbf{r} = \int \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\varphi)^2 - \frac{1}{4}\mu_0^2\varphi^2 + \frac{1}{4}\lambda_0^2\varphi^4\right] d^3\mathbf{r} , \tag{5}$$

$$\mathbf{P} = -\frac{1}{2} \cdot \int (\pi \cdot \nabla\varphi + \nabla\varphi \cdot \pi)\, d^3\mathbf{r} . \tag{6}$$

The Lagrangian and the Hamiltonian are invariant with respect to the symmetry transformation

$$U^{-1}\varphi U = -\varphi , \qquad U^{-1}\pi U = -\pi . \tag{7}$$

Naturally, U is unitary and $U^2 = 1$ therefore Hermitian. The operators H, $\mathbf{P}$ and U commute. Our task is to determine their eigenvalues and common eigenfunctions.

The terms of the Hamiltonian (5) which do not contain differentiation, can be interpreted as a "potential energy density" curve represented by:

$$V(\varphi) = -\frac{1}{4}\mu_0^2\varphi^2 + \frac{1}{4}\lambda_0^2\varphi^4 . \tag{8}$$

The shape of this expression is shown in Fig. 1. One can see that in the neighbourhood of the equilibrium solution $\varphi = 0$ oscillations cannot exist. Let us carry out the substitution:

$$\Phi(x) = \varphi(x) + \frac{1}{\sqrt{2}}\frac{\mu_0}{\lambda_0} , \quad \Pi(x) = \frac{\partial L}{\partial \dot{\Phi}} = \pi(x) . \tag{9}$$

Then the "potential energy density" has the form

$$V(\Phi) = -\frac{1}{16}\frac{\mu_0^4}{\lambda_0^2} + \frac{1}{2}\mu_0^2\Phi^2 - \frac{\mu_0\lambda_0}{\sqrt{2}}\Phi^3 + \frac{1}{4}\lambda_0^2\Phi^4. \tag{10}$$

The equilibrium solution $\Phi = 0$ is already stable and under the effect of excitations of small amplitude the field will oscillate around $\Phi = 0$.

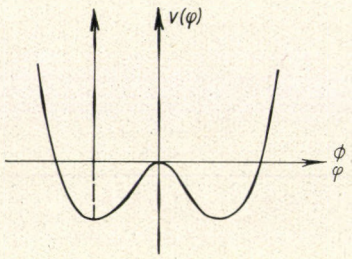


Fig. 1

Let us enclose the field in a cube of volume Ω and require periodicity on its boundary. In this case the field energy has the form

$$H = -\frac{1}{16}\frac{\mu_0^4}{\lambda_0^2}\Omega + \int_\Omega \left[\frac{1}{2}\Pi^2 + \frac{1}{2}(\nabla\Phi)^2 + \frac{1}{2}\mu_0^2\Phi^2 - \frac{\mu_0\lambda_0}{\sqrt{2}}\Phi^3 + \frac{1}{4}\lambda_0^2\Phi^4\right] d^3\mathbf{r}. \tag{11}$$

The field operators are expanded into series:

$$\Phi(\mathbf{r},t) = \frac{1}{\sqrt{\Omega}}\sum_{\mathbf{k}} q_{\mathbf{k}}(t)\cdot e^{i\mathbf{k}\cdot\mathbf{r}},$$

$$\Pi(\mathbf{r},t) = \frac{1}{\sqrt{\Omega}}\sum_{\mathbf{k}} p_{\mathbf{k}}(t)\cdot e^{-i\mathbf{k}\cdot\mathbf{r}}.$$

Here the notation

$$k_x = \frac{2\pi}{\Omega^{1/3}}\cdot S_1,\quad k_y = \frac{2\pi}{\Omega^{1/3}}\cdot S_2,\quad k_z = \frac{2\pi}{\Omega^{1/3}}\cdot S_3\quad (S_i = 0, \pm 1, \pm 2, \ldots) \tag{12}$$

is used. The operators $q_{\mathbf{k}}$ and $p_{\mathbf{k}}$ are canonical variables which obey the restrictions

$$q_{\mathbf{k}}^+ = q_{-\mathbf{k}},\quad p_{\mathbf{k}}^+ = p_{-\mathbf{k}}$$

because of the hermiticity of Φ and Π. In order to introduce Hermitian variables let us decompose the $\mathbf{k}$ space in the following way:

$$\mathbf{k}\in K \quad \text{if} \begin{cases} 1.\ k_z > 0,\ k_x \text{ and } k_y \text{ are arbitrary or} \\ 2.\ k_z = 0,\ k_x > 0,\ k_y \text{ is arbitrary or} \\ 3.\ k_z = k_x = 0,\ k_y > 0\,. \end{cases}$$

Let us now introduce the quantities

$$x_{\mathbf{k}} = \frac{1}{\sqrt{2}}(q_{\mathbf{k}} + q_{-\mathbf{k}}), \quad y_{\mathbf{k}} = \frac{i}{\sqrt{2}}(q_{-\mathbf{k}} - q_{\mathbf{k}}),$$

$$u_{\mathbf{k}} = \frac{1}{\sqrt{2}}(p_{\mathbf{k}} + p_{-\mathbf{k}}), \quad v_{\mathbf{k}} = \frac{i}{\sqrt{2}}(p_{\mathbf{k}} - p_{-\mathbf{k}}),$$

where $\mathbf{k} \in K$, and

$$x_0 = q_0, \qquad u_0 = p_0.$$

The variables $x_{\mathbf{k}}, y_{\mathbf{k}}, u_{\mathbf{k}}, v_{\mathbf{k}}, x_0$ and y_0 are Hermitian operators. On the basis of the commutation relations (4) we have

$$[x_{\mathbf{k}}, u_{\mathbf{k}'}] = [y_{\mathbf{k}}, v_{\mathbf{k}'}] = i\,\delta_{\mathbf{k}\mathbf{k}'},$$

all the other commutators vanish. The field oscillators the wave number of which belongs to K have two, and the only wave with $\mathbf{k} = 0$ has one polarization state.

In the following let us use the $x - y$ representation, in which $x_{\mathbf{k}}$ and $y_{\mathbf{k}}$ are diagonalized, $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are represented by differential operators. If

$$< x_0 \ldots x_{\mathbf{k}} \ldots y_{\mathbf{k}} \ldots | > = \psi(x_0 \ldots x_{\mathbf{k}} \ldots y_{\mathbf{k}} \ldots),$$

then

$$< x_0 \ldots x_{\mathbf{k}} \ldots y_{\mathbf{k}} \ldots | x_{\mathbf{k}'} | > = x_{\mathbf{k}'}\,\psi(x_0 \ldots x_{\mathbf{k}} \ldots y_{\mathbf{k}} \ldots),$$

$$< x_0 \ldots x_{\mathbf{k}} \ldots y_{\mathbf{k}} \ldots | y_{\mathbf{k}'} | > = y_{\mathbf{k}'}\,\psi(x_0 \ldots x_{\mathbf{k}} \ldots y_{\mathbf{k}} \ldots),$$

$$< x_0 \ldots x_{\mathbf{k}} \ldots y_{\mathbf{k}} \ldots | u_{\mathbf{k}'} | > = -i\frac{\partial}{\partial x_{\mathbf{k}'}} \cdot \psi(x_0 \ldots x_{\mathbf{k}} \ldots y_{\mathbf{k}} \ldots),$$

$$< x_0 \ldots x_{\mathbf{k}} \ldots y_{\mathbf{k}} \ldots | v_{\mathbf{k}'} | > = -i\frac{\partial}{\partial y_{\mathbf{k}'}}\psi(x_0 \ldots x_{\mathbf{k}} \ldots y_{\mathbf{k}} \ldots).$$

Sometimes it is advantageous to use polar coordinates instead of $x_{\mathbf{k}}$ and $y_{\mathbf{k}}$

$$x_{\mathbf{k}} = z_{\mathbf{k}} \cos \Theta_{\mathbf{k}}, \quad y_{\mathbf{k}} = z_{\mathbf{k}} \sin \Theta_{\mathbf{k}}, \quad x_0 = x_0. \tag{13}$$

We should like to express the operators H, P and U in terms of these coordinates. To begin with, let us consider the case of U. Taking (7) and (9) into account:

$$U^{-1}\Phi U = -\Phi + \sqrt{2}\,\frac{\mu_0}{\lambda_0}, \quad U^{-1}\Pi U = -\Pi$$

or

$$U^{-1}x_0 U = -x_0 + \sqrt{2}\cdot\frac{\mu_0}{\lambda_0}\sqrt{\Omega}\,, \quad U^{-1}x_{\mathbf{k}}U = -x_{\mathbf{k}}, \quad U^{-1}y_{\mathbf{k}}U = -y_{\mathbf{k}}\ (\mathbf{k} \in K).$$

Making U act on the state function results in

$$U\psi(x_0 \ldots x_{\mathbf{k}} \ldots y_{\mathbf{k}} \ldots) = \psi\left(-x_0 + \sqrt{2}\,\frac{\mu_0}{\lambda_0}\sqrt{\Omega} \ldots - x_{\mathbf{k}} \ldots - y_{\mathbf{k}} \ldots\right) \quad (14)$$

that means that U is of the following form:

$$U = \exp\left(\sqrt{2} \cdot \frac{\mu_0}{\lambda_0}\sqrt{\Omega} \cdot \frac{\partial}{\partial x_0}\right) \cdot P, \quad (15)$$

where P is the "parity operator" in the $x-y$ space, which reflects the coordinates with respect to the origin.

Now let us turn to the case of the momentum operator. Taking (6) into account, we have

$$\mathbf{P} = -\frac{1}{2}\int (\Pi \cdot \nabla\Phi + \nabla\Phi \cdot \Pi)\, d^3\mathbf{r} = -i\sum_{\mathbf{k}} \mathbf{k} p_{\mathbf{k}} q_{\mathbf{k}} = -i\sum \mathbf{k}(p_{\mathbf{k}} q_{\mathbf{k}} - p_{-\mathbf{k}} q_{-\mathbf{k}}).$$

(If now and in the following we do not give the range of the summation, this always refers to K). Finally, introducing the polar coordinates (13), after some calculations we obtain:

$$\mathbf{P} = \sum i\mathbf{k} \cdot \frac{\partial}{\partial \Theta_{\mathbf{k}}}. \quad (16)$$

Finally let us consider the operator H. Introducing polar coordinates also in (11) we obtain after an elementary calculation

$$H = -\frac{1}{16}\frac{\mu_0^4}{\lambda_0^2} \cdot \Omega - \frac{1}{2}\frac{\partial^2}{\partial x_0^2} + \frac{1}{2}\mu_0^2 x_0^2 - \frac{1}{2}\sum\left[\frac{1}{z_{\mathbf{k}}} \cdot \frac{\partial}{\partial z_{\mathbf{k}}}\left(z_{\mathbf{k}} \cdot \frac{\partial}{\partial z_{\mathbf{k}}}\right) + \right.$$

$$\left. + \frac{1}{z_{\mathbf{k}}^2} \cdot \frac{\partial^2}{\partial \Theta_{\mathbf{k}}^2} - (\mu_0^2 + \mathbf{k}^2) z_{\mathbf{k}}^2\right] + \frac{\lambda_0^2}{\Omega^2}\int_{\Omega}\left[\frac{1}{\sqrt{2}} x_0 + \sum z_k \cos(\mathbf{k} \cdot \mathbf{r} + \Theta_{\mathbf{k}})\right]^4 d^3\mathbf{r} -$$

$$- \left(\frac{2}{\Omega}\right)^{3/2} \cdot \frac{\mu_0 \lambda_0}{\sqrt{2}} \cdot \int_{\Omega}\left[\frac{1}{\sqrt{2}} x_0 + \sum z_k \cos(\mathbf{k} \cdot \mathbf{r} + \Theta_{\mathbf{k}})\right]^3 d^3\mathbf{r}. \quad (17)$$

The integrations can be performed if we take (17) into account. The expression for the Hamiltonian obtained in this way, however, contains too many terms, therefore its exact form is given in the Appendix.

Now we are left with the task of determining the common eigenfunctions and the corresponding eigenvalues of the operators (15), (16) and (17).

§ 3. The separation of the field oscillators

In order to determine the eigenfunctions of H, $\mathbf{P}$ and U we apply a version of the method of RITZ, adapted by L. I. SCHIFF [5]. We wish to choose the optimal eigenfunction from the functions of the form

$$\psi = \psi_0(x_0) \cdot \prod_{\mathbf{k} \in K} \psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}}). \tag{18}$$

We take the correlation of the field oscillators with the wave number $\mathbf{k}$ and $-\mathbf{k}$ into account but suppose that in first approximation the partial waves with $\mathbf{k}$ and $\mathbf{k}'$ differing in magnitude or in direction are independent of each other. This is a reasonable assumption if one wishes the investigated problem to be practically solvable. The normalization of (18) is secured by the validity of

$$<\psi_{\mathbf{k}} | \psi_{\mathbf{k}}> = \int_0^{2\pi} \int_0^{\infty} |\psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}})|^2 \cdot z_{\mathbf{k}} \, dz_{\mathbf{k}} \, d\Theta_{\mathbf{k}} = 1 \qquad (\mathbf{k} = 0 \text{ and } \mathbf{k} \in K). \tag{19}$$

Since (18) has to be an eigenstate of $\mathbf{P}$ the eigenvalue equation

$$\mathbf{P}\psi = \sum i\mathbf{k} \cdot \frac{\partial}{\partial \Theta_{\mathbf{k}}} \cdot \psi = \psi \cdot \sum i\mathbf{k} \cdot \psi_{\mathbf{k}}^{-1} \cdot \frac{\partial \psi_{\mathbf{k}}}{\partial \Theta_{\mathbf{k}}} = \mathbf{K}\psi$$

must be fulfilled. This is the case if $\psi_{\mathbf{k}}$ has the form

$$\psi_{\mathbf{k}} = f_{\mathbf{k}}(z_{\mathbf{k}}) \cdot e^{-in_{\mathbf{k}} \cdot \Theta_{\mathbf{k}}}, \tag{20}$$

where $n_{\mathbf{k}}$ is a constant. The expression (18) constructed from functions of the form (20), is a momentum eigenfunction and

$$\mathbf{K} = \Sigma n_{\mathbf{k}} \cdot \mathbf{k} \tag{21}$$

is the corresponding momentum eigenvalue. Because of

$$\psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}} + 2\pi) = \psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}}) \tag{22}$$

the numbers $n_{\mathbf{k}}$ are integers ($n_{\mathbf{k}} = 0, \pm 1, \pm 2, \ldots$). Since at the same time (18) has to be an U-eigenfunction, too, there is

$$\psi_0\left(-x_0 + \sqrt{2} \cdot \frac{\mu_0}{\lambda_0} \sqrt{\Omega}\right) = \varepsilon_0 \cdot \psi_0(x_0), \qquad (\varepsilon_0 = \pm 1), \tag{23}$$

$$\psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}} + \pi) = \varepsilon_{\mathbf{k}} \psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}}), \qquad (\varepsilon_{\mathbf{k}} = \pm 1). \tag{24}$$

If we take (20) into account, (24) is automatically fulfilled, $\varepsilon_{\mathbf{k}} = (-1)^{n_{\mathbf{k}}}$, therefore from the validity of (20) and (23) it follows that the function (18)

is an eigenfunction of U and

$$\varepsilon = \varepsilon_0 \cdot (-1)^{\Sigma n_{\mathbf{k}}} \tag{25}$$

is the corresponding eigenvalue.

The function ψ obtained in this way will not be the eigenfunction of the Hamiltonian, since H produces coupling between the field oscillators. It will be supposed, however, that considering one oscillator only, (i.e. particle states described by the momentum pair $\mathbf{k}$ and $-\mathbf{k}$) the function (18) may give a good approximation. The functions $\psi_0(x_0)$, $f_{\mathbf{k}}(z_{\mathbf{k}})$ are to be calculated so as to optimize this approximation. In order to achieve this let us calculate the average energy. The validity of (24) makes the evaluation of the integral $< \psi \mid H \mid \psi >$ very easy, since then

$$\int_0^{2\pi} \cos \Theta_{\mathbf{k}} \mid \psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}}) \mid^2 d\Theta_{\mathbf{k}} = 0 ,$$

$$\int_0^{2\pi} \sin \Theta_{\mathbf{k}} \mid \psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}}) \mid^2 d\Theta_{\mathbf{k}} = 0 .$$

If one takes this into account, after a longer calculation based on the Appendix one obtains

$$\begin{aligned}
&< \psi \mid H \mid \psi > = < \psi_0 \mid -\frac{1}{16}\frac{\mu_0^4}{\lambda_0^2}\Omega - \frac{1}{2}\frac{\partial^2}{\partial x_0^2} + \frac{1}{2}\mu_0^2 x_0^2 - \\
&- \frac{1}{\sqrt{2}}\frac{\mu_0 \lambda_0}{\sqrt{\Omega}} x_0^3 + \frac{\lambda_0^2}{4\Omega} x_0^4 \mid \psi_0 > + \Sigma < \psi_{\mathbf{k}} \mid -\frac{1}{2}\left[\frac{1}{z_{\mathbf{k}}} \cdot \frac{\partial}{\partial z_{\mathbf{k}}}\left(z_{\mathbf{k}} \cdot \frac{\partial}{\partial z_{\mathbf{k}}}\right) + \right. \\
&\left. + \frac{1}{z_{\mathbf{k}}^2} \cdot \frac{\partial^2}{\partial \Theta_{\mathbf{k}}^2}\right] + \frac{1}{2}(\mu_0^2 + \mathbf{k}^2) z_{\mathbf{k}}^2 + \frac{3}{8} \cdot \frac{\lambda_0^2}{\Omega} z_{\mathbf{k}}^4 \mid \psi_{\mathbf{k}} > + \\
&+ \frac{3\lambda_0^2}{4\Omega} \sum_{\mathbf{k}} \sum_{\mathbf{k}'}{}' < \psi_{\mathbf{k}} \mid z_{\mathbf{k}}^2 \mid \psi_{\mathbf{k}} > < \psi_{\mathbf{k}'} \mid z_{\mathbf{k}'}^2 \mid \psi_{\mathbf{k}'} > + \\
&+ \frac{3}{2} \cdot \frac{\lambda_0^2}{\Omega} < \psi_0 \mid x_0^2 \mid \psi_0 > \cdot \sum < \psi_{\mathbf{k}} \mid z_{\mathbf{k}}^2 \mid \psi_{\mathbf{k}} > - \\
&- \frac{3}{\sqrt{2}}\frac{\mu_0 \lambda_0}{\sqrt{\Omega}} \cdot < \psi_0 \mid x_0 \mid \psi_0 > \cdot \sum < \psi_{\mathbf{k}} \mid z_{\mathbf{k}}^2 \mid \psi_{\mathbf{k}} > .
\end{aligned} \tag{26}$$

The optimal form of ψ_0 and $\psi_{\mathbf{k}}$ will be determined by minimizing (26) under the subsidiary condition (19). Denoting the Lagrange multiplicators by E_0' and $E_{\mathbf{k}}$, the condition of the minimum is

$$< \psi \mid H \mid \psi > - E_0' [< \psi_0 \mid \psi_0 > - 1] - \sum E_{\mathbf{k}} [< \psi_{\mathbf{k}} \mid \psi_{\mathbf{k}} > - 1] = \min .$$

or

$$\frac{\delta < \psi \,|\, H \,|\, \psi >}{\delta \psi_{\mathbf{k}}^*} = E_{\mathbf{k}} \cdot \psi_{\mathbf{k}} \,. \tag{27}$$

Let us take (26) into account for $\mathbf{k} = 0$:

$$\left| -\frac{1}{2}\frac{d^2}{dx_0^2} + \frac{1}{4}\frac{\lambda_0^2}{\Omega}x_0^4 - \frac{\mu_0\lambda_0}{\sqrt{2\Omega}}x_0^3 + \frac{1}{2}\left(\mu_0^2 + 3\frac{\lambda_0^2}{\Omega}\sum < \psi_{\mathbf{k}} \,|\, z_{\mathbf{k}}^2 \,|\, \psi_{\mathbf{k}} >\right)x_0^2 - \right.$$
$$\left. - \frac{3\mu_0\lambda_0}{\sqrt{2\Omega}} \cdot \left(\sum < \psi_{\mathbf{k}} \,|\, z_{\mathbf{k}}^2 \,|\, \psi_{\mathbf{k}} >\right) x_0 - \frac{1}{16}\frac{\mu_0^4}{\lambda_0^2}\Omega\right] \psi_0(x_0) = E_0' \psi_0(x_0), \tag{28}$$

and for $\mathbf{k} \in K$:

$$\left[-\frac{1}{2}\left\{\frac{1}{z_{\mathbf{k}}} \cdot \frac{\partial}{\partial z_{\mathbf{k}}} \cdot \left(z_{\mathbf{k}} \cdot \frac{\partial}{\partial z_{\mathbf{k}}}\right) + \frac{1}{z_{\mathbf{k}}^2} \cdot \frac{\partial^2}{\partial \Theta_{\mathbf{k}}^2}\right\} + \left\{\frac{1}{2}\mu_0^2 + \frac{1}{2}\mathbf{k}^2 + \right.\right.$$
$$+ \frac{3\lambda_0^2}{2\Omega} \cdot \sum < \psi_{\mathbf{k}'} \,|\, z_{\mathbf{k}'}^2 \,|\, \psi_{\mathbf{k}'} > + \frac{3\lambda_0^2}{2\Omega} < \psi_0 \,|\, x_0^2 \,|\, \psi_0 > -$$
$$\left.\left. - \frac{3\mu_0\lambda_0}{\sqrt{2\Omega}} < \psi_0 \,|\, x_0 \,|\, \psi_0 > \right\} z_{\mathbf{k}}^2 - \right. \tag{29}$$
$$\left. - \frac{3\lambda_0^2}{2\Omega} < \psi_{\mathbf{k}} \,|\, z_{\mathbf{k}}^2 \,|\, \psi_{\mathbf{k}} > z_{\mathbf{k}}^2 + \frac{3\lambda_0^2}{8\Omega} z_{\mathbf{k}}^4 \right] \psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}}) = E_{\mathbf{k}} \cdot \psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}}) \,.$$

The eqs. (28) and (29) form a coupled system of equations determining the functions $\psi_0(x_0)$ and $\psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}})$ — or essentially $f_{\mathbf{k}}(z_{\mathbf{k}})$. Since on the left side the expressions in the brackets [...] are Hermitian operators the numbers $E_{\mathbf{k}}$ are real.

The solution of the system of equations is a difficult task in spite of the approximation (18). The solution cannot be obtained without approximation since (28) and (29) describe anharmonic oscillators.

Let us begin with the eq. (28). Formally this is the Schrödinger equation of a one-dimensional anharmonic oscillator

$$\left[-\frac{1}{2}\frac{d^2}{dx_0^2} + V(x_0)\right] \psi_0(x_0) = E_0' \psi_0(x_0) \,, \tag{30}$$

where the potential energy

$$V(x_0) = -\frac{1}{16}\frac{\mu_0^4}{\lambda_0^2}\Omega + \frac{1}{4}\frac{\lambda_0^2}{\Omega}x_0^4 - \frac{\mu_0\lambda_0}{\sqrt{2\Omega}} \cdot x_0^3 + \frac{1}{2}\mu_0^2 x_0^2 +$$
$$+ \bar{B}\left[\frac{3\lambda_0^2}{2\Omega}\left(x_0 - \frac{\mu_0}{\lambda_0}\sqrt{\frac{\Omega}{2}}\right)^2 - \frac{3}{4}\mu_0^2\right] \tag{31}$$

is considered to be the superposition of a parabola of the fourth order and of a second one of second order. The latter is characterized by the quantity

$$\bar{B} = \sum < \psi_{\mathbf{k}} | z_{\mathbf{k}}^2 | \psi_{\mathbf{k}} > , \tag{32}$$

or in other words, it arises from the zero point fluctuation of the partial waves with $\mathbf{k} \neq 0$. Let us perform a coordinate transformation which corresponds to the reverse of (9), i.e. to the displacement of the origin into unstable equilibrium:

$$y_0 = x_0 - \frac{\mu_0}{\lambda_0} \sqrt{\frac{\Omega}{2}} . \tag{33}$$

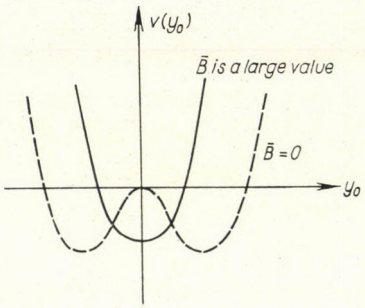


Fig. 2

In the new coordinate system the "potential energy" of the field oscillator with $\mathbf{k} = 0$ is given by

$$V(x_0) = V(y_0) = -\frac{1}{4}\left(\mu_0^2 - 6\frac{\lambda_0^2}{\Omega} \cdot \bar{B}\right) y_0^2 + \frac{\lambda_0^2}{4\Omega} \cdot y_0^4 - \frac{3}{4}\mu_0^2 \bar{B} . \tag{34}$$

We see that in the case of $\bar{B} = 0$ (neglecting the other oscillators) $V(y_0)$ has a maximum at $y_0 = 0$ and two minima at $y_0 = \pm \mu_0/\lambda_0 \sqrt{\Omega/2}$. The ground state of the oscillator is degenerate. This situation is modified by the occurrence of $\bar{B}$. The influence of the other oscillators if $\mu_0^2 < 6\lambda_0^2 \bar{B}/\Omega$ shows itself in the potential energy having a single minimum at $y_0 = 0$, with a decreased value $V_{\min} = -3\mu_0^2 \bar{B}/4$ (Fig. 2). But if $\mu_0^2 > 6\lambda_0^2 \bar{B}/\Omega$ i.e. if the influence of the other oscillators is small, the two minima remain, solely they wander towards the location of the local maximum $y_0 = 0$ and their depths are somewhat decreased (Fig. 3). In the following let us suppose that the parameter μ_0 occurring in the field equation (1) is so large that the latter situation is realized, i.e. the duplication of the ground state survives. Let us introduce the following notations:

$$\bar{\mu}^2 = \mu_0^2 - 6\frac{\lambda_0^2}{\Omega} \cdot \bar{B} > 0 , \tag{35}$$

and

$$\bar{\gamma} = \frac{\lambda_0}{\bar{\mu}} \cdot \sqrt{\frac{2}{\Omega}} \,. \tag{36}$$

Making use of the new parameters one gets

$$V(y_0) = -\frac{1}{4}\bar{\mu}^2\left[y_0^2 - \frac{1}{2}\bar{\gamma}^2 \cdot y_0^4 + 3\bar{B}\,(1 + 3\bar{\gamma}^2\,\bar{B})\right].$$

The locations of minima of $V(y_0)$ are $y_0 = \pm\bar{\gamma}^{-1}$ and here

$$V(\bar{\gamma}^{-1}) = V(-\bar{\gamma}^{-1}) = -\frac{1}{8}\bar{\mu}^2 \cdot \bar{\gamma}^{-2} - \frac{3}{4}\bar{\mu}^2 \cdot \bar{B}\,(1 + 3\bar{\gamma}^2\,\bar{B}) \,.$$

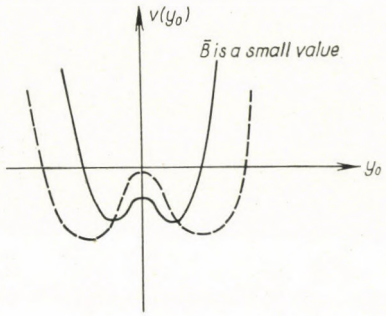


Fig. 3

If we write $\bar{B} = 0$ (i.e. if we neglect the influence of the zero point fluctuations of the oscillators with $\mathbf{k} \neq 0$ on the oscillator with $\mathbf{k} = 0$) (38) goes over into an expression that may be obtained by substituting $\Phi = \Pi = 0$ in (11); this is the ground state energy of the classical scalar field:

$$\lim_{b=0} V(\bar{\gamma}^{-1}) = -\frac{1}{16} \cdot \frac{\mu_0^4}{\lambda_0^2} \cdot \Omega = H\,[\Phi = 0, \Pi = 0] \,. \tag{37}$$

Introducing the new variable

$$\xi = y_0 \pm \bar{\gamma}^{-1} \,,$$

one obtains

$$V(y_0) = V(\xi) = \frac{1}{2}\bar{\mu}^2\left[\xi^2 \mp \bar{\gamma} \cdot \xi^3 + \frac{1}{4}\bar{\gamma}^2 \cdot \xi^4\right] + V(\bar{\gamma}^{-1}) \,. \tag{38}$$

We see that the anharmonicity in the oscillator, described by eq. (30), originates in $\bar{\gamma}$, i.e. in the appearence of the coupling constant λ_0. By means of the renormalisation (35) we took into account the influence of the zero point

fluctuations on the other oscillators. The values of $\bar{\mu}$ and $\bar{\gamma}$ depend on the form of $\psi_{\mathbf{k}}$, as resulting from $\bar{B}$, therefore they are different in different states of the field.

Let us denote the excitation energy of the oscillator by writing

$$E_{\mathbf{0}} = E'_{\mathbf{0}} - V(\bar{\gamma}^{-1}) .$$

Using this notation the eigenvalue equation for the oscillator with $\mathbf{k} = 0$ has the form:

$$\left\{ -\frac{1}{2}\frac{d^2}{dy_0^2} + \frac{1}{8}\bar{\mu}^2 \cdot \bar{\gamma}^2 (y_0^2 - \bar{\gamma}^{-2})^2 \right\} \psi_{\mathbf{0}}(y_0) = E_{\mathbf{0}}\, \psi_{\mathbf{0}}(y_0) . \tag{39}$$

Also in the differential equation of the oscillator with nonvanishing wave number the coefficients contain $\psi_{\mathbf{0}}$ and $\psi_{\mathbf{k}'}$, therefore also the $\psi_{\mathbf{k}}$ is influenced by the state of the other waves. On the other hand, this differential equation is formally relatively simple:

$$\begin{aligned} &\left\{ -\frac{1}{2}\left[\frac{1}{z_{\mathbf{k}}} \cdot \frac{\partial}{\partial z_{\mathbf{k}}} \cdot \left(z_{\mathbf{k}} \cdot \frac{\partial}{\partial z_{\mathbf{k}}} \right) + \frac{1}{z_{\mathbf{k}}^2} \cdot \frac{\partial^2}{\partial \Theta_{\mathbf{k}}^2} \right] + \frac{1}{2}\bar{\omega}_{\mathbf{k}}^2 \cdot z_{\mathbf{k}}^2 + \right. \\ &\left. + \frac{3}{4}\bar{\mu}^2 \cdot \bar{\gamma}^2 \left[\frac{1}{4} z_{\mathbf{k}}^4 - < \psi_{\mathbf{k}} | z_{\mathbf{k}}^2 | \psi_{\mathbf{k}} > z_{\mathbf{k}}^2 \right] \right\} \psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}}) = E_{\mathbf{k}} \cdot \psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}}) , \end{aligned} \tag{40}$$

where

$$\bar{\omega}_{\mathbf{k}}^2 = \mathbf{k}^2 + \mu_0^2 + \frac{3\lambda_0^2}{\Omega}[< \psi_{\mathbf{0}} | x_0^2 | \psi_{\mathbf{0}} > + \bar{B}] - \frac{3\sqrt{2} \cdot \mu_0 \lambda_0}{\sqrt{\Omega}} < \psi_{\mathbf{0}} | x_0 | \psi_{\mathbf{0}} > .$$

Let us take into account that $\psi_{\mathbf{0}}$ is the eigenfunction of U, too, so:

$$U\, \psi_{\mathbf{0}}(y_0) \equiv \psi_{\mathbf{0}}(-y_0) = \varepsilon_{\mathbf{0}}\, \psi_{\mathbf{0}}(y_0) \qquad \varepsilon_0 = \pm 1 ,$$

therefore

$$< \psi_{\mathbf{0}} | x_0 | \psi_{\mathbf{0}} > = < \psi_{\mathbf{0}}(y_0) | y_0 + \frac{\mu_0}{\lambda_0}\sqrt{\frac{\Omega}{2}} | \psi_{\mathbf{0}}(y_0) > = \frac{\mu_0}{\lambda_0}\sqrt{\frac{\Omega}{2}} ,$$

$$< \psi_{\mathbf{0}} | x_0^2 | \psi_{\mathbf{0}} > = < \psi_{\mathbf{0}}(y_0) | y_0^2 + 2y_0 \cdot \frac{\mu_0}{\lambda_0}\sqrt{\frac{\Omega}{2}} + \frac{\mu_0^2}{\lambda_0^2}\frac{\Omega}{2} | \psi_{\mathbf{0}}(y_0) > =$$

$$= \frac{\mu_0^2}{\lambda_0^2} \cdot \frac{\Omega}{2} + < \psi_{\mathbf{0}} | y_0^2 | \psi_{\mathbf{0}} > ,$$

and finally

$$\bar{\omega}_{\mathbf{k}}^2 = \mathbf{k}^2 + \bar{\mu}^2 \left[1 - \frac{3}{2}(1 - \bar{\gamma}^2 < \psi_{\mathbf{0}} | y_0^2 | \psi_{\mathbf{0}} >) \right] . \tag{41}$$

Based on the former differential equations one can determine ψ_0 and $\psi_{\mathbf{k}}$ and construct the eigenfunction (18) from them and one can evaluate the

approximate value of the energy. On the basis of (26)

$$W = <\psi | H | \psi> = -\frac{1}{8}\bar{\mu}^2 \cdot \bar{\gamma}^{-2} + E_0 + \sum E_k + \\ + \frac{3}{8}\bar{\mu}^2 \cdot \bar{\gamma}^2 \left\{ \sum (<\psi_k \; z_k^2 \; \psi_k>)^2 - (\sum <\psi_k | z_k^2 | \psi_k>)^2 - \right. \\ \left. - 2 <\psi_0 | y_0^2 | \psi_0> \cdot \sum <\psi_k | z_k^2 | \psi_k> \right\}. \tag{42}$$

The physical meaning of each term of the formula is evident. The first term describes the classical ground state energy (37) of the scalar field in a renormalized form. The E_0 and E_k stand for the energy of the anharmonic oscillators. Finally the terms standing in the $\{\ldots\}$ bracket and proportional to λ_0^2 describe the interaction of the vibrations of the different oscillators. Into this term the products of the squared amplitudes $<\psi_k | z_k^2 | \psi_k>$ of the *different* ($\mathbf{k} = 0$ and $\mathbf{k} \in K$) field oscillators enter.

The formula (42) makes possible to try to explore in practice the energy spectrum of the scalar field.

§ 4. The ground state

Our first task is the evaluation of the energy of the ground state. The function, describing the ground state is:

$$\psi^0 = \psi_0^0(y_0) \cdot \prod_{k \in K} \psi_k^0(z_k, \Theta_k). \tag{43}$$

The values of the state-dependent parameters $\bar{\mu}$, $\bar{B}$ and $\bar{\gamma}$ are denoted in the ground state by μ, B and γ, i.e.

$$B = \sum <\psi_k^0 | z_k^2 | \psi_k^0>, \\ \mu^2 = \mu_0^2 - 6\frac{\lambda_0^2}{\Omega} \cdot B, \tag{44}$$

$$\gamma = \frac{\lambda_0}{\mu}\sqrt{\frac{2}{\Omega}}. \tag{45}$$

Further let us introduce the notation

$$g = \frac{\lambda_0}{\sqrt{\mu^3 \Omega}} = \frac{\gamma}{\sqrt{2\mu}}, \tag{46}$$

$$Z = 6\mu B. \tag{47}$$

One can now write

$$\mu^2 = \frac{\mu_0^2}{1 + g^2 Z}. \tag{48}$$

The potential energy of the field oscillator (38) with $\mathbf{k} = 0$ describes a harmonic oscillator in the case of $\gamma = 0$, the eigenfunction of which is of the form

$$\psi_0^0 = \left(\frac{\mu}{\pi}\right)^{\frac{1}{4}} \cdot e^{-\frac{1}{2}\mu y_0^2},$$

and $E_0 = \mu/2$ is the corresponding eigenvalue. The finite terms in (39) proportional to γ^2 spoil the harmonicity. We should like to determine the energy up to the order g^2, therefore we make use of the first approximation of the perturbation theory. To start with we use as eigenfunction such a combination which is an eigenfunction also of U:

$$\psi_0^0 = N\left[e^{-\frac{1}{2}\mu(y_0+\gamma^{-1})^2} + \varepsilon_0 \cdot e^{-\frac{1}{2}\mu(y_0-\gamma^{-1})^2}\right] \qquad (\varepsilon_0 = \pm 1). \tag{49}$$

This gives the value of E_0 up to the order g^2 correctly:

$$E_0 = \frac{<\psi_0| -\frac{1}{2}\frac{d^2}{dy_0^2} + V(y_0) - V(\gamma^{-1})\,|\psi_0>}{<\psi_0|\psi_0>} = \frac{1}{2}\mu\left(1 + \frac{3}{8}g^2\right) + 0(g^4) + \varepsilon_0 \Delta_0, \tag{50}$$

where

$$\Delta_0 = 0(f), \quad f = e^{-g^{-2}}.$$

Instead of (50) we may obtain a more accurate value if instead of (49) the form of ψ_0^0 is taken in the following way:

$$\psi_0^0 = N\left[e^{-\frac{1}{2}a_0\mu(y_0+\gamma^{-1})^2} + \varepsilon_0 \cdot e^{-\frac{1}{2}a_0\mu(y_0-\gamma^{-1})^2}\right],$$

a_0 is given by the requirement

$$\frac{<\psi^0|H|\psi^0>}{<\psi^0|\psi^0>} = \min.$$

This would give an expression of the type

$$a_0 = 1 + ag^2 + 0(g^4).$$

One can show, however, that a influences the value of W only in the order g^4. So in the given approximation we may choose $a = 1$.

Knowing that ψ_0^0 has the form (49) one finds

$$< \psi_0^0 \,|\, y_0^2 \,|\, \psi_0^0 > = \gamma^{-2} + \frac{1}{2}\,\mu^{-1} + \varepsilon_0 \cdot 0\,(f)\,. \tag{51}$$

Taking this into account we can write down the differential equation which determines $\psi_{\mathbf{k}}$:

$$\left\{ -\frac{1}{2}\left[\frac{1}{z_{\mathbf{k}}} \cdot \frac{\partial}{\partial z_{\mathbf{k}}} \cdot \left(z_{\mathbf{k}} \cdot \frac{\partial}{\partial z_{\mathbf{k}}}\right) + \frac{1}{z_{\mathbf{k}}^2} \cdot \frac{\partial^2}{\partial \Theta_{\mathbf{k}}^2}\right] + \frac{1}{2}\,\omega_{\mathbf{k}}^2\, z_{\mathbf{k}}^2 + \right.$$
$$\left. + \frac{3}{2}\,\mu^3 g^2 \left[\frac{1}{4}\, z_{\mathbf{k}}^4 - < \psi_{\mathbf{k}}^0 \,|\, z_{\mathbf{k}}^2 \,|\, \psi_{\mathbf{k}}^0 > z_{\mathbf{k}}^2\right]\right\} \psi_{\mathbf{k}}^0 = E_{\mathbf{k}}^0\, \psi_{\mathbf{k}}^0\,, \tag{52}$$

where

$$\omega_{\mathbf{k}}^2 = \mathbf{k}^2 + \mu^2 \left[1 + \frac{3}{2}\, g^2\right] + \varepsilon_0 \cdot 0(f)\,. \tag{53}$$

If we write $g = 0$ before the finite expression, we gain the Schrödinger equation of a two-dimensional harmonic oscillator, the solution of which is

$$\psi_{\mathbf{k}}^0 = \left(\frac{\omega_{\mathbf{k}}}{\pi}\right)^{\frac{1}{2}} \cdot e^{-\frac{1}{2}\omega_{\mathbf{k}} \cdot z_{\mathbf{k}}^2}\,, \tag{54}$$

with

$$E_{\mathbf{k}}^0 = \omega_{\mathbf{k}}\,. \tag{55}$$

Making use of (54) one can obtain the correct value of $E_{\mathbf{k}}^0$ up to the order g^2 by evaluating the mean value:

$$E_{\mathbf{k}}^0 = \omega_{\mathbf{k}} \left[1 - \frac{3}{4}\, g^2 \left(\frac{\mu}{\omega_{\mathbf{k}}}\right)^3 + 0(g^4)\right].$$

Further, also the mean value of $z_{\mathbf{k}}^2$ may be calculated which we shall use in the following:

$$< \psi_{\mathbf{k}}^0 \,|\, z_{\mathbf{k}}^2 \,|\, \psi_{\mathbf{k}}^0 > = \frac{1}{\omega_{\mathbf{k}}}\,. \tag{56}$$

Here, as a matter of fact, also the $\omega_{\mathbf{k}}$ contain the coupling constant. According to (53):

$$\omega_{\mathbf{k}} = \varepsilon_{\mathbf{k}} \left[1 + \frac{3}{4}\, g^2 \left(\frac{\mu}{\varepsilon_{\mathbf{k}}}\right)^2\right] + \varepsilon_0 \cdot 0(f)\,, \tag{57}$$

where ε_k (with given μ) is already independent of the coupling constant:

$$\varepsilon_{\mathbf{k}} = \sqrt{\mu^2 + \mathbf{k}^2}\,. \tag{58}$$

Finally, on the basis of (42) the ground state energy correct up to the order g^2 can be calculated:

$$W_0 = <\psi^0 | H | \psi^0> = -\frac{\mu}{16\, g^2} + \frac{1}{2}\mu + \sum \varepsilon_{\mathbf{k}} \left[1 - \frac{3}{4}\left(\frac{\mu}{\varepsilon_{\mathbf{k}}}\right)^2\right] + \\ + \frac{3}{4} g^2 \cdot \mu \left\{\frac{1}{4} - \left[\sum \frac{\mu}{\varepsilon_{\mathbf{k}}}\right]^2 + \frac{3}{4} \sum \left(\frac{\mu}{\varepsilon_{\mathbf{k}}}\right)^3\right\} + 0(g^4) + \varepsilon_0 \cdot 0(f)\,. \tag{59}$$

This gives the value of the lowest energy level with the parity $\varepsilon = +1$. The energy level of parity $\varepsilon = -1$, which is also a discrete one, is located above the former, the difference between their energies being

$$\Delta = \mu \cdot 0(f). \tag{60}$$

The energy of the ground state is given — as in the case of all the fields — by a divergent expression.

§ 5. The excited states

The differential equations (39) and (40) which determine the eigenfunctions, are formally separated but the parameters $\bar{\mu}$ and $\bar{\omega}_{\mathbf{k}}$ depend on the wave function, so in reality they describe a very strong coupling of the oscillators. This causes serious complications especially in the determination of the excited states: it is very difficult to ensure the orthogonality of the eigenfunctions of higher quantum numbers to those of smaller quantum numbers, since because of the excitation the separated energy operator of the oscillator, and therefore also the form of the separated eigenvalue equation itself changes via the coupling among the oscillators hidden in the parameters $\bar{\mu}$ and $\bar{\omega}_{\mathbf{k}}$.

In order to elucidate the circumstances, let us introduce the quantities μ and g instead of $\bar{\mu}$ and $\bar{\gamma}$ into the eq. (39) and (40) which are valid also for arbitrary excited states. Naturally we have to take into account that the meaning of these parameters is given by means of the functions $\psi_{\mathbf{0}}^0$ and $\psi_{\mathbf{k}}^0$, which are characteristic of the ground state. We obtain

$$\left\{-\frac{1}{2}\frac{d^2}{dy_0^2} + \frac{g^2\mu}{16}(\delta Z - g^{-2} + 2\mu y_0^2)^2\right\} \psi_{\mathbf{0}}(y_0) = E_{\mathbf{0}}\, \psi_{\mathbf{0}}(y_0)\,, \tag{61}$$

$$\left\{-\frac{1}{2}\left[\frac{1}{z_{\mathbf{k}}} \cdot \frac{\partial}{\partial z_{\mathbf{k}}} \cdot \left(z_{\mathbf{k}} \cdot \frac{\partial}{\partial z_{\mathbf{k}}}\right) + \frac{1}{z_{\mathbf{k}}^2} \cdot \frac{\partial^2}{\partial \Theta_{\mathbf{k}}^2}\right] + \frac{1}{2}\bar{\omega}_{\mathbf{k}}^2 \cdot z_{\mathbf{k}}^2 + \right. \\ \left. + \frac{3}{2} g^2 \mu^3 \left[\frac{1}{4} z_{\mathbf{k}}^4 - <\psi_k | z_{\mathbf{k}}^2 | \psi_k> z_{\mathbf{k}}^2\right]\right\} \psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}}) = E_{\mathbf{k}}\, \psi_{\mathbf{k}}(z_{\mathbf{k}}, \Theta_{\mathbf{k}}), \tag{62}$$

where

$$\bar{\omega}_{\mathbf{k}}^2 = \omega_{\mathbf{k}}^2 + \frac{1}{2} \cdot \mu^2 g^2 (\delta X + \delta Z) . \tag{63}$$

Using these results, the expression of the field energy is — according to (42) — of the following form:

$$W = <\psi | H | \psi> = -\frac{\mu}{16 g^2}(1 - g^2 \delta Z)^2 + E_0 + \sum E_{\mathbf{k}} +$$

$$+ \frac{3}{4} g^2 \cdot \mu^3 \left\{ \sum (<\psi_{\mathbf{k}} | z_{\mathbf{k}}^2 | \psi_{\mathbf{k}}>)^2 - \left(\sum <\psi_{\mathbf{k}} | z_{\mathbf{k}}^2 | \psi_{\mathbf{k}}>\right)^2 - \right. \tag{64}$$

$$\left. - 2 <\psi_0 | y_0^2 | \psi_0> \sum <\psi_{\mathbf{k}} | z_{\mathbf{k}}^2 | \psi_{\mathbf{k}}> \right\} ,$$

and the excitation energy

$$W_e = (E_0 - E_0^0) + \sum (E_{\mathbf{k}} - E_{\mathbf{k}}^0) - \frac{g^2 \cdot \mu}{24} Z (\delta X + \delta Z) +$$

$$+ \frac{g^2 \mu}{48} [\delta Y - 2 \delta X \cdot \delta Z - 4 \delta Z^2] + \frac{\mu}{8} \delta Z - \frac{1}{4} g^2 \cdot \mu^2 \delta Z <\psi_0^0 | y_0^2 | \psi_0^0> , \tag{65}$$

therefore in first approximation

$$W_e = (E_0 - E_0^0) + \sum (E_{\mathbf{k}} - E_{\mathbf{k}}^0) - \frac{g^2 \mu}{24} Z (\delta X + \delta Z)$$

$$+ \frac{g^2 \mu}{48} [\delta Y - 2 \delta X \cdot \delta Z - 4 \delta Z^2 - 6 \delta Z] + g^2 \delta Z \cdot 0(f) . \tag{66}$$

Here

$$\delta X = 6 \mu [<\psi_0 | y_0^2 | \psi_0> - <\psi_0^0 | y_0^2 | \psi_0^0>] , \tag{67}$$

$$\delta Y = 36 \mu^2 \sum [(<\psi_{\mathbf{k}} | z_{\mathbf{k}}^2 | \psi_{\mathbf{k}}>)^2 - (<\psi_{\mathbf{k}}^0 | z_{\mathbf{k}}^2 | \psi_{\mathbf{k}}^0>)^2] , \tag{68}$$

$$\delta Z = 6 \mu \sum [<\psi_{\mathbf{k}} | z_{\mathbf{k}}^2 | \psi_{\mathbf{k}}> - <\psi_{\mathbf{k}}^0 | z_{\mathbf{k}}^2 | \psi_{\mathbf{k}}^0>] . \tag{69}$$

Because of the strong coupling between the field oscillators initiated by these parameters we do not undertake here to explore the accurate excitation spectrum. We are satisfied to give upper limits for the excitation energies. This will be achieved if we calculate $\langle \psi \mid H \mid \psi \rangle$ by means of conveniently orthogonalized functions.

Let us consider first the excitation of the oscillation with $\mathbf{k} = 0$. The two functions, orthogonal to ψ^0 will be constructed in the following form:

$$\psi^{(1,0)} = \psi_0^1 (y_0) \cdot \prod_{\mathbf{k} \in K} \psi_{\mathbf{k}}^0 (z_{\mathbf{k}}, \Theta_{\mathbf{k}}) , \tag{70}$$

where

$$\psi_0^1(y_0) = N\left[(y_0 + \gamma^{-1}) \cdot e^{-\frac{1}{2}\mu(y_0+\gamma^{-1})^2} + \varepsilon \cdot (y_0 - \gamma^{-1}) \cdot e^{-\frac{1}{2}\mu(y_0-\gamma^{-1})^2}\right], \quad (71)$$

and N is a constant of normalization.

The functions $\psi_{\mathbf{k}}$ are not modified as compared to the ground state, is spite of the fact, that the excitation of the oscillator with $\mathbf{k} = 0$ modifies also the eq. (62) through the parameters $\bar{\omega}_{\mathbf{k}}$. Thus (70) is convenient exclusively to supply an upper limit for the excitation energy. Since the $\psi_{\mathbf{k}}$ remain unchanged, $\delta Y = \delta Z = 0$. On the other hand, because of (71) one has

$$\delta X = 6 + \varepsilon_0 \cdot 0(f).$$

Calculating the mean value of the bracket $\{\ldots\}$ in (61) with the use of (71) one obtains

$$E_0 = \frac{3}{2}\mu\left[1 + \frac{5}{8}g^2 + 0(g^4) + \varepsilon \cdot 0(f)\right] = E_0^0 + \mu + \frac{3}{4}g^2 \cdot \mu + 0(g^4) + \varepsilon \cdot 0(f).$$

The displacement of $E_{\mathbf{k}}$ in this approximation is due to the change in $\omega_{\mathbf{k}}$,

$$E_{\mathbf{k}} - E_{\mathbf{k}}^0 = \frac{1}{2}(\bar{\omega}_{\mathbf{k}}^2 - \omega_{\mathbf{k}}^2) < \psi_{\mathbf{k}}^0 \,|\, z_{\mathbf{k}}^2 \,|\, \psi_{\mathbf{k}}^0 > = \frac{g^2 \cdot \mu^2}{4\omega_{\mathbf{k}}}\delta X.$$

Making use of these results, we calculate the excitation energy on the basis of (66). The expressions $\Sigma(E_{\mathbf{k}} - E_{\mathbf{k}}^0)$ and $Z\,\delta\,X$ are divergent, but

$$\sum (E_{\mathbf{k}} - E_{\mathbf{k}}^0) - \frac{g^2 \cdot \mu}{24} \cdot Z\,\delta X = 0.$$

So the upper limit of W_e is

$$W_e = E_0 - E_0^0 = \mu\left[1 + \frac{3}{4}g^2 + 0(g^4) + \varepsilon \cdot 0(f)\right]. \quad (72)$$

The excited state of an oscillator with $\mathbf{k} \neq 0$ will be described by the function

$$\psi^{(1,\mathbf{k})} = \psi_0^0(y_0) \cdot \frac{\psi_{\mathbf{k}}^1(z_{\mathbf{k}}, \Theta_{\mathbf{k}})}{\psi_{\mathbf{k}}^0(z_{\mathbf{k}}, \Theta_{\mathbf{k}})} \cdot \prod_{\mathbf{k}' \in K} \psi_{\mathbf{k}'}^0(z_{\mathbf{k}'}, \Theta_{\mathbf{k}'}), \quad (73)$$

where

$$\psi_{\mathbf{k}}^1 = \frac{\omega_{\mathbf{k}}}{\sqrt{\pi}} \cdot z_{\mathbf{k}} \cdot e^{-\frac{1}{2}\omega_{\mathbf{k}} \cdot z_{\mathbf{k}}^2 - i\Theta_{\mathbf{k}}}. \quad (74)$$

This is an exact eigenfunction of $\mathbf{P}$ and U with $\mathbf{k}$ and $-\varepsilon_0$ as respective eigenvalues. The function (73) is evidently orthogonal to (43), (70) and also to

$\psi^{(1k')}$. It supplies, however, only an upper limit for the energy eigenvalue, since it does not take into account the change of the state of the other oscillators due to the coupling with the excited oscillator.

Now

$$\delta X = 0\,,$$

$$< \psi_{\mathbf{k}} \,|\, z_{\mathbf{k}}^2 \,|\, \psi_{\mathbf{k}} > = \frac{2}{\omega_{\mathbf{k}}}\,,$$

therefore in our approximation (neglecting the terms of the order of g^2)

$$\delta Y = 12 \left(\frac{3\mu}{\omega_{\mathbf{k}}}\right)^2, \qquad \delta Z = 6 \left(\frac{\mu}{\omega_{\mathbf{k}}}\right).$$

The strongly modified value of $E_{\mathbf{k}}$ obtained as a mean value is

$$E_{\mathbf{k}} = E_{\mathbf{k}}^0 + \omega_{\mathbf{k}}\,.$$

Also the E_0 and $E_{\mathbf{k}'}$ change, but only due to the modification of $\bar{\mu}$ and $\bar{\omega}_{\mathbf{k}'}$ in the order of g^2:

$$E_{\mathbf{0}} = E_{\mathbf{0}}^0 + \frac{3}{4} g^2 \cdot \frac{\mu^2}{\omega_{\mathbf{k}}} \cdot \left(1 + 3\,\frac{\mu}{\omega_{\mathbf{k}}}\right),$$

$$E_{\mathbf{k}'} = E_{\mathbf{k}'}^0 + \frac{3}{2} g^2 \cdot \frac{\mu^3}{\omega_{\mathbf{k}'}} \cdot \frac{1}{\omega_{\mathbf{k}}} \qquad (\mathbf{k} \neq \mathbf{k}')\,.$$

Now the divergent terms in the formula (66) cancel again since

$$\sum (E_{\mathbf{k}'} - E_{\mathbf{k}'}^0) - \frac{g^2 \cdot \mu}{24} Z \cdot \delta Z = \omega_{\mathbf{k}} - \frac{3}{2} g^2 \cdot \frac{\mu^3}{\omega_{\mathbf{k}}^2}\,.$$

Thus the excitation energy up to the order of g^2 is

$$W_e = \varepsilon_{\mathbf{k}} \cdot \left[1 + \frac{3}{4} g^2 \left(\frac{\mu}{\varepsilon_{\mathbf{k}}}\right)^2\right].$$

We see therefore that in the neighbourhood of $W_e = \mu$ a continuous spectrum begins. So it is correct to consider (70) and (73) as the one-quantum excitations of the scalar field while the two discrete ψ^0 states describe a degenerate vacuum state. Based on our estimation one can state that in the case of bounded μ and g also the one-particle excitations are bounded, therefore the substitution (48) has solved the mass renormalization. μ is the renormalized mass, a finite value, which in a realistic theory must be determined, on the basis

of experience. The Z is, however, a divergent constant. Comparing (47), (44) and (55) we have

$$Z = 6\mu B = 6 \sum \frac{\mu}{\omega_{\mathbf{k}}},$$

or transforming the summation into integration

$$Z = 6\mu B = 3 \sum_{\text{all } \mathbf{k}} \frac{\mu}{\omega_{\mathbf{k}}} - 3 = \frac{3\Omega\mu}{(2\pi)^3} \cdot \int \frac{d^3\mathbf{k}}{\sqrt{\mathbf{k}^2 + \mu'^2}} - 3,$$

which contains a quadratically divergent integral besides Ω. This means that both parameters μ_0 and λ_0, which occur in the field equation (1), cannot be chosen finite at the same time since (45) cannot be satisfied for finite values of μ, μ_0 and λ_0.

If we excite the oscillator with $\mathbf{k} = 0$ to the quantum number n_0,the oscillators with $\mathbf{k} \neq 0$ to the quantum number $n_{\mathbf{k}}$, it can be shown that

$$E_0 = \left(\frac{1}{2} + n_0\right)\mu + 0(g^2), \qquad < \psi_0 \,|\, y_0^2 \,|\, \psi_0 > = \frac{n_0 + 1}{\mu} + \gamma^{-2} + 0(f),$$

$$E_{\mathbf{k}} = (1 + n_{\mathbf{k}})\,\omega_{\mathbf{k}} + 0(g^2), \quad < \psi_{\mathbf{k}} \,|\, z_{\mathbf{k}}^2 \,|\, \psi_{\mathbf{k}} > = \frac{n_{\mathbf{k}+1}}{\varepsilon_{\mathbf{k}}} + 0(g^2).$$

In the case of physically realistic excitations

$$n_0 + \Sigma n_{\mathbf{k}} < \infty.$$

Then the terms of δX, δY, δZ and $\Sigma(E_{\mathbf{k}} - E_{\mathbf{k}}^0)$ which do not contain g, are finite. Because of the change in $\bar{\omega}_{\mathbf{k}}$, all the $E_{\mathbf{k}}$-s are displaced by an extra amount

$$\Delta E_{\mathbf{k}} = \frac{1}{2}(\bar{\omega}_{\mathbf{k}}^2 - \omega_{\mathbf{k}}^2) < \psi_{\mathbf{k}} \,|\, z_{\mathbf{k}}^2 \,|\, \psi_{\mathbf{k}} > =$$

$$= \frac{g^2}{4}\frac{\mu^2}{\omega_{\mathbf{k}}} \cdot (\delta X + \delta Z) + \frac{g^2 \cdot \mu^2}{4}(\delta X + \delta Z) \cdot \frac{n_{\mathbf{k}}}{\omega_{\mathbf{k}}},$$

the sum of which just cancels the term proportional to Z occurring in eq. (66):

$$\sum \Delta E_{\mathbf{k}} = \frac{g^2}{24}\mu^2 Z\,(\delta X + \delta Y) + \frac{g^2\mu^2}{4}(\delta X + \delta Z) \cdot \sum \frac{n_{\mathbf{k}}}{\omega_{\mathbf{k}}}.$$

We see that the mass renormalization makes also the higher states (with several quanta) finite.

It is a very important fact that the functions of such a simple form as (18) are not suitable to take into account the interactions of physical particles

[6]. The states — obtained from the solutions of the eqs. (61), (62) — containing more than one quanta correspond to "dressed" but noninteracting particles. It turns out from the fact that these states are eigenfunctions of the momentum $\mathbf{P}$ with eigenvalue $\sum_{\mathbf{k}} \sum_{s=1}^{n_{\mathbf{k}}} \varepsilon^s_{n_{\mathbf{k}}} \cdot \mathbf{k}$ ($\varepsilon^s_{\mathbf{k}}$ can be ± 1) and with average energy $n_0 \omega_0 + \Sigma n_{\mathbf{k}} \omega_{\mathbf{k}}$ above the vacuum value apart from terms of order $0(g^2)$ but they are not eigenstates of the Hamiltonian.

The excitation energy of the state containing two noninteracting particles differs in the order $0(g^2)$ from the sum of the excitation energies of the one-

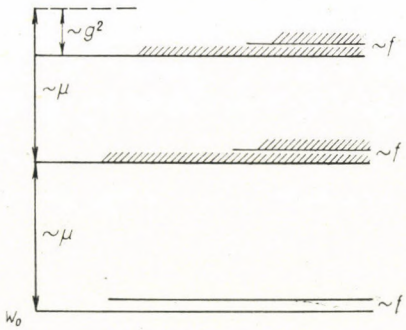


Fig. 4

particle states. Further, every state splits up into two energy levels of opposite parity and differing in the order $0(f)$. The energy spectrum of the noninteracting particles is shown in Fig. 4.

To investigate the two-particle scattering it is necessary to use a wave-packet of amplitude $a(\mathbf{k}, \mathbf{k}')$ which is built up from functions corresponding to noninteracting physical particle-pairs with momenta $\mathbf{k}$ and $\mathbf{k}'$. One has to find such a scattering amplitude $a(\mathbf{k}, \mathbf{k}')$ that minimalizes the mean value of the Hamiltonian. One can interpret the two-particle scattering with finite λ_0, therefore $g \to 0$ if $\Omega \to \infty$. The authors would like to return to the problem of the two-particle scattering and solutions of the coupled eq. (61), (62) in a subsequent paper [6].

§ 6. The weak interaction

The approximate degeneracy of the ground state (and the n-particle excitations) is a consequence of the "potential barrier" of the oscillator with $\mathbf{k} = 0$ shown in Fig. 3. Because of this $\varphi(x)$ may oscillate in a stable manner around the amplitude values

$$< | y_0 | > = \pm \gamma^{-1}, \qquad < | x_{\mathbf{k}} | > = < | y_{\mathbf{k}} | > = 0,$$

i.e. around the values

$$<|\varphi(x)|> = \frac{1}{\sqrt{2}} \frac{\mu}{\lambda_0},$$

$$<|\varphi(x)|> = -\frac{1}{\sqrt{2}} \frac{\mu}{\lambda_0}.$$

For instance let us choose the wave function so that in the state characterized through

$$<|\varphi(x)|> = \frac{1}{\sqrt{2}} \frac{\mu}{\lambda_0}$$

there be a unique particle (the α-particle) of momentum $\mathbf{k}$:

$$\psi_\alpha = N \cdot e^{-\frac{1}{2}\mu(y_0+\gamma^{-1})^2} \cdot z_{\mathbf{k}} \cdot e^{-i\Theta_{\mathbf{k}}} \cdot \prod_{\mathbf{k}' \in K} \cdot e^{-\frac{1}{2}\varepsilon_{\mathbf{k}'} \cdot z_{\mathbf{k}'}^{\circ}}.$$

If we write $-\gamma^{-1}$ instead of γ^{-1}, we obtain the particle (the β-particle) ψ_β excited in the state characterized through

$$<|\varphi(x)>| = -\frac{1}{\sqrt{2}} \frac{\mu}{\lambda_0}.$$

The ψ_α and ψ_β are not eigenstates of the parity operator U. One can construct the eigenfunctions, corresponding to the parity values $\varepsilon = \pm 1$, starting from ψ_α and ψ_β:

$$\psi_\varepsilon = \frac{1}{\sqrt{2}}(\psi_\alpha - \varepsilon \psi_\beta).$$

According to the results obtained in § 5 and § 6:

$$H \psi_\varepsilon = (W_0 + \omega_{\mathbf{k}} - \varepsilon \Delta_{\mathbf{k}}) \psi_\varepsilon.$$

Taking this into account, the solution of the wave equation

$$H \psi = i \frac{\partial \psi}{\partial t},$$

which obeys the initial condition

$$\psi(0) = \psi_\alpha$$

has the form

$$\psi(t) = e^{-\frac{i}{\hbar}(W_0+\omega_{\mathbf{k}})t} [\psi_\alpha \cdot \cos \Delta_1 t + i \psi_\beta \sin \Delta_1 t].$$

Therefore the α-particle slowly transforms into a β-particle, with the period of the order

$$\tau = \frac{2\pi}{\Delta_1} = \mu \cdot 0(f^{-1}) .$$

This transition, which is slow in the case of $g < 1$ because of $f \ll 1$, is called *weak interaction* in our model. The period $\tau \sim \mu \cdot f^{-1}$ can be called the lifetime of an α-particle (or β-particle, respectively). One can see, that the same non-linearity, the same bare coupling constant produces two interactions in our model, with quite different orders of magnitude and character. Substituting by means of the eq. (36)—(46), one obtains

$$f = e^{-\frac{1}{g^2}} = \exp\left(-\frac{4\pi^3 \cdot \mu^3 \cdot \delta^3(0)}{\lambda_0^2}\right) \sim < \psi_\alpha | \psi_\beta > .$$

This is the renormalized coupling constant of the *weak interaction*. The question as to the coupling constant in terms of the λ_0 of the *strong interaction* of the α particles (or of the β particles, respectively) can be answered only by the more detailed study of the higher excitations.

§ 7. The asymmetric model

In the paper [3] we have made a more detailed study of the case of the asymmetric potential valley when the non-linear field equation

$$\Box \varphi + \mu_0^2 \varphi + K_0^2 \varphi^2 - \lambda_0^2 \varphi^3 = 0$$

has two different stable and constant solutions in the classical theory. In this case the symmetry (7) does not occur (because of the quadratic term). Then even if the momenta $\mathbf{k}$ are identical, the energy of the particles of the type a will be essentially different from that of the particles of the type b, therefore the degeneracy of the ground state leads to two types of particles essentially different from each other with different rest masses. The one is excited from the ground state of the type a, the other from the ground state of the type b by substituting $H_n(z_\mathbf{k}) \cdot \exp(-cz_\mathbf{k}^2)$ instead of $\exp(-cz_\mathbf{k}^2)$. The mathematical treatment of this case is a considerably more difficult problem: the expression $<\psi |H| \psi>$ calculated by means of the supposition (18) turns out to be much more intricate. In the present work we do not deal with the mathematical elaboration of such an asymmetric model, but we make some remarks as to the properties to be expected of such a model.

One may expect in an asymmetric model that from the ground state of the type a the excitation spectrum of the type

$$E = W_a + \Sigma n_{\mathbf{k}} \sqrt{\mu^2 + \mathbf{k}^2},$$

and from the ground state of the type b the excitation spectrum of the type

$$E = W_b + \Sigma n_{\mathbf{k}} \sqrt{M^2 + \mathbf{k}^2}$$

may be obtained, if the interaction terms proportional to g^2 are neglected. The interaction of the μ-particles with each other, and the interaction of the M-particles with each other is a strong one, characterized by the coupling constant λ_0, while the M-particles can transmute into the μ-particles via a weak interaction characterized by the coupling constant $f = \exp(-g^{-2})$.

The essential property of the model is, that *either* μ-particles *or* M-particles occur in the world, the two types cannot exist simultaneously. There is no energy eigenvalue of the type $W_0 + \sqrt{\mu^2 + \mathbf{k}^2} + \sqrt{M^2 + \mathbf{k}^2}$. If we start with a state of the field characterized by $<\varphi> = 0$, then all the excitations of small energy correspond to particles of the type $M(E < E_{\text{crit}})$. We can arrange that the potential of the field is stressed up to the value $<\varphi> = a$. This means an energy surplus $\Delta = W_a - W_b$ over the ground state. This surplus energy, however, cannot be localized in the space (φ is constant also now), is not connected with momentum and is not of a quantized nature ($E_b + 2\Delta$, $E_b + 3\Delta$, ... are not eigenvalues). Therefore the ground state b cannot be considered to be a state containing a real particle (it is a "spurion"-state). The existence of the Δ spurion, however, is important, because the rest energy of the quanta of the field change from M to μ (because of the "particle-spurion interaction"). If the spurion is present only the particles with mass μ occur, if it does not appear, they occur only with mass M. Since the spurion cannot be localized, this is valid for the whole space, universally. This is the obstacle in the way of the attempts at an explanation of the two types of elementary particles, which really do occur simultaneously in our universe, by means of such a model.

In his model in the zeroth approximation BLOKHINTSEV removes the coupling of the anharmonic oscillators, vibrating at different points of the space [2]. It is thus possible to start from the fact that in a domain of the space, where $<\varphi> = 0$ (ground state of the type b, there is no spurion) the particles M occur, while in another region of the space, where $<\varphi> = a$ (the spurion exists) the particles μ occur. In the region of transition, however, where $<\varphi>$ increases from zero to a, the term $(\nabla\varphi)^2$ produces an energy density, greatly different from zero in a zone which in its turn isolates the two types of particles by means of a zone of amorphous matter (i.e. which has no particle

structure). Such a phenomenon could occur only in the jets, but not in the case of metastable particles.

It seems to be problematic, how one can solve the analogous problem in other spurion theories (e.g. [7]). If the spurion has no momentum, it cannot be localized in the space and in this case it is difficult to imagine how it can interact differently with different particles (e.g. some of the nucleons are excited by it to become hyperons while others are practically uninfluenced) when it is "equally far away" from each type of particle.

§ 8. Conclusion

A peculiarity of the non-linear scalar field described by the field equation (1) is the degeneracy of the ground state. In the classical field theory both $\varphi = 1/\sqrt{2}\ \mu_0/\lambda_0$ and $\varphi = -1/\sqrt{2}\ \mu_0/\lambda_0$ are solutions of the field equation, they represent the stable ground states. Also in quantum field theory there is an approximate degeneracy of the ground state (and of all the excited states). In the present paper this problem has been treated on the basis of an assumption proposed elsewhere by L. I. SCHIFF, that the eigenfunctions can be approximated by the products of functions that describe the different oscillators of the field. This means that the correlation of the particles with momentum $\mathbf{k}$ and $-\mathbf{k}$ is taken into account, the particles of different wave numbers $\mathbf{k}$ interact with each other only by means of their zero point fluctuation. This interaction is divergent, but one can take it into account after a mass renormalization. (In this respect we pass over the results achieved in paper I.)

The model makes it possible to describe two kinds of particles as the excitations of one and the same field. After the charge renormalization the non-linearity λ_0 causes a strong interaction between the particles of the type a and also between the particles of the type b, but the transmutation of the particles a into particles b (or vice versa) is a weak interaction. The paper does not go into greater details as regards the mathematical analysis of the interactions.

The physical application of the model is restricted by the fact that the field can have excitations either of the type b, or the type a, but not both of them simultaneously. This may be a difficulty of all the theories, based on vacuum degeneracy (spurion theories).

The authors are indebted to Prof. L. I. SCHIFF (Stanford University), Dr. F. KÁROLYHÁZY and Dr. K. L. NAGY (Budapest, L. Eötvös University) for valuable discussions.

Appendix

The Hamiltonian is of the form

$$H = -\frac{1}{16}\frac{\mu_0^4}{\lambda_0^2}\Omega - \frac{1}{2}\frac{\partial^2}{\partial x_0^2} + \frac{1}{2}\mu_0^2 x_0^2 - \frac{1}{2}\cdot\sum\left[\frac{1}{z_\mathbf{k}}\cdot\frac{\partial}{\partial z_\mathbf{k}}\left(z_\mathbf{k}\cdot\frac{\partial}{\partial Z_\mathbf{k}}\right) + \right.$$

$$\left. + \frac{1}{z_\mathbf{k}^2}\frac{\partial^2}{\partial\Theta_\mathbf{k}^2} - (\mu_0^2 + \mathbf{k}^2)\, z_\mathbf{k}^2\right] + \mathrm{H_{int}}\,,$$

where

$$\mathrm{H_{int}} = \frac{\lambda_0^2}{\Omega^2}\cdot\int_\Omega\left|\frac{1}{\sqrt{2}}x_0 + \sum z_\mathbf{k}\cos(\mathbf{k}\cdot\mathbf{r} + \Theta_\mathbf{k})\right|^4 d^3\mathbf{r} - \left(\frac{2}{\Omega}\right)^{3/2}\cdot$$

$$\cdot\frac{\mu_0\lambda_0}{\sqrt{2}}\cdot\int_\Omega\left|\frac{1}{\sqrt{2}}x_0 + \sum z_\mathbf{k}\cos(\mathbf{k}\cdot\mathbf{r} + \Theta_\mathbf{k})\right|^3 d^3\mathbf{r}\,.$$

We introduce the notation $z_0 = 1/\sqrt{2}\, x_0$. The symbol $\sum_\mathbf{k}$ denotes a summation which will be extended in the following also to $\mathbf{k} = 0$, besides K. Performing the calculation of the corresponding powers in the integrals

$$\mathrm{Hint} = \frac{\lambda_0^2}{\Omega^2}\left[\sum_\mathbf{k} z_\mathbf{k}^4\cdot\int_\Omega\cos^4(\mathbf{k}\cdot\mathbf{r} + \Theta_\mathbf{k})\, d^3\mathbf{r} + 3\sum_\mathbf{k}\sum_\mathbf{l} z_\mathbf{k}^2\cdot z_\mathbf{l}^2\cdot\right.$$

$$\cdot\int_\Omega\cos^2(\mathbf{k}\cdot\mathbf{r} + \Theta_\mathbf{k})\cos^2(\mathbf{l}\cdot\mathbf{r} + \Theta_\mathbf{l})\, d^3\mathbf{r} +$$

$$+ 6\sum_\mathbf{k}\sum_\mathbf{l}\sum_\mathbf{m} z_\mathbf{k}^2\cdot z_\mathbf{l}\cdot z_\mathbf{m}\cdot\int_\Omega\cos^2(\mathbf{k}\cdot\mathbf{r} + \Theta_\mathbf{k})\cos(\mathbf{l}\cdot\mathbf{r} + \Theta_\mathbf{l})\cos(\mathbf{m}\cdot\mathbf{r} + \Theta_\mathbf{m})\, d^3\mathbf{r} +$$

$$+ 4\sum_\mathbf{k}\sum_\mathbf{l} z_\mathbf{k}^3\cdot z_\mathbf{l}\cdot\int_\Omega\cos^3(\mathbf{k}\cdot\mathbf{r} + \Theta_\mathbf{k})\cdot\cos(\mathbf{l}\cdot\mathbf{r} + \Theta_\mathbf{l})\, d^3\mathbf{r} +$$

$$+ \sum_\mathbf{k}\sum_\mathbf{l}\sum_\mathbf{m}\sum_\mathbf{n} z_\mathbf{k}\cdot z_\mathbf{l}\cdot z_\mathbf{m}\cdot z_\mathbf{n}\cdot\int_\Omega\cos(\mathbf{k}\cdot\mathbf{r} + \Theta_\mathbf{k})\cos(\mathbf{lr} + \Theta_\mathbf{l})\cos(\mathbf{mr} +$$

$$\left. + \Theta_\mathbf{m})\cos(\mathbf{nr} + \Theta_\mathbf{n})\, d^3\mathbf{r}\right] -$$

$$- \left(\frac{2}{\Omega}\right)^{3/2}\cdot\frac{\mu_0\lambda_0}{\sqrt{2}}\left[\sum_\mathbf{k} z_\mathbf{k}^3\cdot\int_\Omega\cos^3(\mathbf{k}\cdot\mathbf{r} + \Theta_\mathbf{k})\, d^3\mathbf{r} + 3\sum_\mathbf{k}\sum_\mathbf{l} z_\mathbf{k}^2\cdot z_\mathbf{l}\cdot\right.$$

$$\cdot\int_\Omega\cos^2(\mathbf{k}\cdot\mathbf{r} + \Theta_\mathbf{k})\cos(\mathbf{l}\cdot\mathbf{r} + \Theta_\mathbf{l})\, d^3\mathbf{r}$$

$$\left. + \sum_\mathbf{k}\sum_\mathbf{l}\sum_\mathbf{m} z_\mathbf{k}\cdot z_\mathbf{l}\cdot z_\mathbf{m}\cdot\int_\Omega\cos(\mathbf{k}\cdot\mathbf{r} + \Theta_\mathbf{k})\cos(\mathbf{l}\cdot\mathbf{r} + \Theta_\mathbf{l})\cos(\mathbf{m}\cdot\mathbf{r} + \Theta_\mathbf{m})\, d^3\mathbf{r}\right].$$

In the sums two of the summation indices may not coincide. The evaluation of the trigonometric integrals is simple but of considerable length. The calculations give:

$$\int \cos^4(\mathbf{k}\cdot\mathbf{r}+\Theta_{\mathbf{k}})\,d^3\mathbf{r}=\begin{cases}\frac{3}{8}\Omega & \mathbf{k}\neq 0\,,\\ \Omega & \mathbf{k}=0\,.\end{cases}$$

If $\mathbf{k}\neq\mathbf{l}$:

$$\int \cos^2(\mathbf{k}\cdot\mathbf{r}+\Theta_{\mathbf{k}})\cos^2(\mathbf{l}\cdot\mathbf{r}+\Theta_{\mathbf{l}})\,d^3\mathbf{r}=\begin{cases}\frac{1}{4}\Omega & \mathbf{l}\neq 0 \text{ and } \mathbf{k}\neq 0\,,\\ \frac{1}{2}\Omega & \mathbf{k}=0 \text{ or } \mathbf{l}=0\,.\end{cases}$$

If $\mathbf{k},\mathbf{l},\mathbf{m}$ are different:

$$\int \cos^2(\mathbf{k}\cdot\mathbf{r}+\Theta_{\mathbf{k}})\cos(\mathbf{l}\cdot\mathbf{r}+\Theta_{\mathbf{l}})\cdot\cos(\mathbf{m}\cdot\mathbf{r}+\Theta_{\mathbf{m}})\,d^3\mathbf{r}=$$

$$=\frac{1}{8}\Omega\{\cos 2\Theta_{\mathbf{k}}\cdot\cos\Theta_{\mathbf{l}}\cdot\cos\Theta_{\mathbf{m}}[\delta(2\mathbf{k}-\mathbf{l}-\mathbf{m})+\delta(2\mathbf{k}-\mathbf{l}+\mathbf{m})]+$$

$$+\delta(2\mathbf{k}+\mathbf{l}-\mathbf{m})]+$$

$$+\sin 2\Theta_{\mathbf{k}}\cdot\sin\Theta_{\mathbf{l}}\cdot\cos\Theta_{\mathbf{m}}\cdot[\delta(2\mathbf{k}-\mathbf{l}-\mathbf{m})+\delta(2\mathbf{k}-\mathbf{l}+\mathbf{m})-\delta(2\mathbf{k}+\mathbf{l}-\mathbf{m})]+$$

$$+\sin 2\Theta_{\mathbf{k}}\cos\Theta_{\mathbf{l}}\cdot\sin\Theta_{\mathbf{m}}[\delta(2\mathbf{k}-\mathbf{l}-\mathbf{m})+\delta(2\mathbf{k}-\mathbf{m}+\mathbf{l})-\delta(2\mathbf{k}+\mathbf{m}-l)]+$$

$$+\cos 2\Theta_{\mathbf{k}}\cdot\sin\Theta_{\mathbf{l}}\cdot\sin\Theta_{\mathbf{m}}[\delta(2\mathbf{k}-\mathbf{l}+\mathbf{m})+\delta(2\mathbf{k}+\mathbf{l}-\mathbf{m})-\delta(2\mathbf{k}-\mathbf{l}-\mathbf{m})]\},$$

where

$$\delta(\mathbf{k}+\mathbf{l}+\mathbf{m})=\begin{cases}1 & \mathbf{k}+\mathbf{l}+\mathbf{m}=0\,,\\ 0 & \mathbf{k}+\mathbf{l}+\mathbf{m}\neq 0\,.\end{cases}$$

if $\mathbf{k}\neq\mathbf{l}$:

$$\int \cos^3(\mathbf{k}\cdot\mathbf{r}+\Theta_{\mathbf{k}})\cdot\cos(\mathbf{l}\cdot\mathbf{r}+\Theta_{\mathbf{l}})\,d^3\mathbf{r}=$$

$$=\frac{1}{8}\Omega[\cos\Theta_{\mathbf{k}}\cdot\cos 2\Theta_{\mathbf{k}}\cdot\cos\Theta_{\mathbf{l}}\cdot\delta(3\mathbf{k}-\mathbf{l})-\sin\Theta_{\mathbf{k}}\sin 2\Theta_{\mathbf{k}}\cos\Theta_{\mathbf{l}}\delta(3\mathbf{k}-\mathbf{l})+$$

$$+\cos\Theta_{\mathbf{k}}\cdot\sin 2\Theta_{\mathbf{k}}\sin\Theta_{\mathbf{l}}\cdot\delta(3\mathbf{k}-\mathbf{l})+\sin\Theta_{\mathbf{k}}\cdot\cos 2\Theta_{\mathbf{k}}\cdot\sin\Theta_{\mathbf{l}}\delta(3\mathbf{k}-\mathbf{l})]\,.$$

When $\mathbf{k},\mathbf{l},\mathbf{m},\mathbf{n}$ are different:

$$\int \cos(\mathbf{k}\cdot\mathbf{r}+\Theta_{\mathbf{k}})\cos(\mathbf{l}\cdot\mathbf{r}+\Theta_{\mathbf{l}})\cdot\cos(\mathbf{m}\cdot\mathbf{r}+\Theta_{\mathbf{m}})\cdot\cos(\mathbf{n}\cdot\mathbf{r}+\Theta_{\mathbf{n}})\,d^3\mathbf{r}=$$

$$=\frac{1}{8}\Omega\{\cos\Theta_{\mathbf{k}}\cos\Theta_{\mathbf{l}}\cos\Theta_{\mathbf{m}}\cos\Theta_{\mathbf{n}}[\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n})+\delta(\mathbf{k}-\mathbf{l}-\mathbf{m}-\mathbf{n})+$$

$$+\delta(\mathbf{k}-\mathbf{l}+\mathbf{m}+\mathbf{n})+\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n})+\delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n})+$$

$$+\delta(\mathbf{k}-\mathbf{l}-\mathbf{m}+\mathbf{n})+\delta(\mathbf{k}-\mathbf{l}+\mathbf{m}-\mathbf{n})]$$

$$+ \sin \Theta_{\mathbf{k}} \cdot \sin \Theta_{\mathbf{l}} \cdot \cos \Theta_{\mathbf{m}} \cdot \cos \Theta_{\mathbf{n}} \cdot [\delta(\mathbf{k}-\mathbf{l}-\mathbf{m}-\mathbf{n}) + \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}+\mathbf{n}) -$$
$$- \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n}) + \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}-\mathbf{n}) -$$
$$- \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n}) - \delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n})]$$

$$+ \sin \Theta_{\mathbf{k}} \cdot \cos \Theta_{\mathbf{l}} \cdot \sin \Theta_{\mathbf{m}} \cdot \cos \Theta_{\mathbf{n}} \cdot [\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n}) + \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}-\mathbf{n}) -$$
$$- \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n}) - \delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n}) +$$
$$+ \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}+\mathbf{n}) - \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}-\mathbf{n})]$$

$$+ \cos \Theta_{\mathbf{k}} \cdot \sin \Theta_{\mathbf{l}} \cdot \sin \Theta_{\mathbf{m}} \cdot \cos \Theta_{\mathbf{n}} \cdot [\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n}) + \delta(\mathbf{l}-\mathbf{k}-\mathbf{m}-\mathbf{n}) -$$
$$- \delta(\mathbf{l}-\mathbf{k}+\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n}) - \delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n}) +$$
$$+ \delta(\mathbf{l}-\mathbf{k}-\mathbf{m}+\mathbf{n}) - \delta(\mathbf{l}-\mathbf{k}+\mathbf{m}-\mathbf{n})] +$$

$$+ \sin \Theta_{\mathbf{k}} \cdot \cos \Theta_{\mathbf{l}} \cdot \cos \Theta_{\mathbf{m}} \cdot \sin \Theta_{\mathbf{n}} \cdot [\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n}) + \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}-\mathbf{n}) -$$
$$- \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}+\mathbf{l}-\mathbf{n}+\mathbf{m}) - \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n}) +$$
$$+ \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}-\mathbf{n}) - \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}+\mathbf{n})]$$

$$+ \cos \Theta_{\mathbf{k}} \cdot \sin \Theta_{\mathbf{l}} \cdot \cos \Theta_{\mathbf{m}} \cdot \sin \Theta_{\mathbf{n}} \cdot [\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n}) + \delta(\mathbf{l}-\mathbf{k}-\mathbf{m}-\mathbf{n}) -$$
$$- \delta(\mathbf{l}-\mathbf{k}+\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n}) - \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n}) +$$
$$+ \delta(\mathbf{l}-\mathbf{k}+\mathbf{m}-\mathbf{n}) - \delta(\mathbf{l}-\mathbf{k}-\mathbf{m}+\mathbf{n})] +$$

$$+ \cos \Theta_{\mathbf{k}} \cdot \cos \Theta_{\mathbf{l}} \cdot \sin \Theta_{\mathbf{m}} \cdot \sin \Theta_{\mathbf{n}} \cdot [\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n}) +$$
$$+ \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}-\mathbf{n}) - \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n}) -$$
$$- \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}-\mathbf{n}) - \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}+\mathbf{n})]$$

$$+ \sin \Theta_{\mathbf{k}} \cdot \sin \Theta_{\mathbf{l}} \cdot \sin \Theta_{\mathbf{m}} \cdot \sin \Theta_{\mathbf{n}} \cdot [\delta(\mathbf{k}-\mathbf{l}-\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}-\mathbf{n}) -$$
$$- \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n}) - \delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n}) - \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}-\mathbf{n}) -$$
$$- \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n})]\},$$

$$\int \cos^3(\mathbf{k}\cdot\mathbf{r} + \Theta_{\mathbf{k}})\, d^3\mathbf{r} = \begin{cases} \Omega & \mathbf{k} = 0, \\ 0 & \mathbf{k} \neq 0. \end{cases}$$

If $\mathbf{k} \neq \mathbf{l}$:

$$\int \cos^2(\mathbf{k}\cdot\mathbf{r} + \Theta_{\mathbf{k}}) \cdot \cos(\mathbf{l}\cdot\mathbf{r} + \Theta_{\mathbf{l}})\, d^3\mathbf{r} = \frac{1}{4}\Omega\,[2\delta(\mathbf{l}) + \cos 2\Theta_{\mathbf{k}} \cdot \cos \Theta_{\mathbf{i}} \cdot$$
$$\cdot \delta(2\mathbf{k} - \mathbf{l}) + \sin 2\Theta_{\mathbf{k}} \cdot \sin \Theta_{\mathbf{l}} \cdot \delta(2\mathbf{k} - \mathbf{l})].$$

When $\mathbf{k}, \mathbf{l}, \mathbf{m}$ are different:

$$\int \cos(\mathbf{k}\cdot\mathbf{r}+\Theta_{\mathbf{k}})\cos(\mathbf{l}\cdot\mathbf{r}+\Theta_{\mathbf{l}})\cos(\mathbf{m}\cdot\mathbf{r}+\Theta_{\mathbf{m}})\,d^3\mathbf{r} =$$

$$= \frac{1}{4}\Omega\{\cos\Theta_{\mathbf{k}}\cdot\cos\Theta_{\mathbf{l}}\cdot\cos\Theta_{\mathbf{m}}\cdot[\delta(\mathbf{k}+\mathbf{l}-\mathbf{m})+\delta(\mathbf{k}-\mathbf{l}-\mathbf{m})+\delta(\mathbf{k}-\mathbf{l}+\mathbf{m})]+$$

$$+\sin\Theta_{\mathbf{k}}\cdot\sin\Theta_{\mathbf{l}}\cdot\cos\Theta_{\mathbf{m}}\cdot[\delta(\mathbf{k}-\mathbf{l}-\mathbf{m})+\delta(\mathbf{k}-\mathbf{l}+\mathbf{m})-\delta(\mathbf{k}+\mathbf{l}-\mathbf{m})]+$$

$$+\sin\Theta_{\mathbf{k}}\cdot\cos\Theta_{\mathbf{l}}\cdot\sin\Theta_{\mathbf{m}}\cdot[\delta(\mathbf{k}+\mathbf{l}-\mathbf{m})+\delta(\mathbf{k}-\mathbf{l}-\mathbf{m})-\delta(\mathbf{k}-\mathbf{l}+\mathbf{m})]+$$

$$+\cos\Theta_{\mathbf{k}}\cdot\sin\Theta_{\mathbf{l}}\cdot\sin\Theta_{\mathbf{m}}\cdot[\delta(\mathbf{k}+\mathbf{l}-\mathbf{m})+\delta(\mathbf{l}-\mathbf{k}-\mathbf{m})-\delta(\mathbf{l}-\mathbf{k}+\mathbf{m})]\}.$$

Making use of these integrals we obtain:

$$\mathrm{H}_{\mathrm{int}} = \frac{3}{8}\cdot\frac{\lambda_0^2}{\Omega}\sum z_{\mathbf{k}}^4+\frac{\lambda_0^2}{\Omega}z_0^4+\frac{3\lambda_0^2}{4\Omega}\sum_{\mathbf{k}\neq\mathbf{l}}\sum z_{\mathbf{k}}^2\cdot z_{\mathbf{l}}^2+\frac{3\lambda_0^2}{\Omega}z_0^2\sum z_{\mathbf{k}}^2+$$

$$+\frac{3\lambda_0^2}{4\Omega}\sum_{\mathbf{k}}\sum_{\mathbf{l}}\sum_{\mathbf{m}} z_{\mathbf{k}}^2\cdot z_{\mathbf{l}}\cdot z_{\mathbf{m}}\{\cos 2\Theta_{\mathbf{k}}\cdot\cos\Theta_{\mathbf{l}}\cdot\cos\Theta_{\mathbf{m}}\cdot[\delta(2\mathbf{k}-\mathbf{l}-\mathbf{m})+$$

$$+\delta(2\mathbf{k}-\mathbf{l}+\mathbf{m})+\delta(2\mathbf{k}+\mathbf{l}-\mathbf{m})]+$$

$$+\sin 2\Theta_{\mathbf{k}}\cdot\sin\Theta_{\mathbf{l}}\cos\Theta_{\mathbf{m}}[\delta(2\mathbf{k}-\mathbf{l}-\mathbf{m})+\delta(2\mathbf{k}-\mathbf{l}+\mathbf{m})-\delta(2\mathbf{k}+\mathbf{l}-\mathbf{m})]+$$

$$+\sin 2\Theta_{\mathbf{k}}\cdot\cos\Theta_{\mathbf{l}}\cdot\sin\Theta_{\mathbf{m}}[\delta(2\mathbf{k}-\mathbf{l}-\mathbf{m})+\delta(2\mathbf{k}+\mathbf{l}-\mathbf{m})-\delta(2\mathbf{k}-\mathbf{l}+\mathbf{m})]+$$

$$+\cos 2\Theta_{\mathbf{k}}\cdot\sin\Theta_{\mathbf{l}}\cdot\sin\Theta_{\mathbf{m}}[\delta(2\mathbf{k}-\mathbf{l}+\mathbf{m})+\delta(2\mathbf{k}+\mathbf{l}-\mathbf{m})-\delta(2\mathbf{k}-\mathbf{l}-\mathbf{m})]\}+$$

$$+\frac{\lambda_0^2}{2\Omega}\cdot\sum_{\mathbf{k}}\sum_{\mathbf{l}} z_{\mathbf{k}}^3\cdot z_{\mathbf{l}}\{\cos\Theta_{\mathbf{k}}\cos 2\Theta_{\mathbf{k}}\cdot\cos\Theta_{\mathbf{l}}\cdot\delta(3\mathbf{k}-\mathbf{l})-\sin\Theta_{\mathbf{k}}\cdot\sin 2\Theta_{\mathbf{k}}\cdot$$

$$\cdot\cos\Theta_{\mathbf{l}}\,\delta(3\mathbf{k}-\mathbf{l})+\cos\Theta_{\mathbf{k}}\sin 2\Theta_{\mathbf{k}}\sin\Theta_{\mathbf{l}}\,\delta(3\mathbf{k}-\mathbf{l})+$$

$$+\sin\Theta_{\mathbf{k}}\cdot\cos 2\Theta_{\mathbf{k}}\sin\Theta_{\mathbf{l}}\,\delta(3\mathbf{k}-\mathbf{l})\}+$$

$$+\frac{\lambda_0^2}{8\Omega}\sum_{\mathbf{k}}\sum_{\mathbf{l}}\sum_{\mathbf{m}}\sum_{\mathbf{n}} z_{\mathbf{k}}z_{\mathbf{l}}z_{\mathbf{m}}z_{\mathbf{n}}\{\cos\Theta_{\mathbf{k}}\cos\Theta_{\mathbf{l}}\cos\Theta_{\mathbf{m}}\cos\Theta_{\mathbf{n}}\cdot[\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n})+$$

$$+\delta(\mathbf{k}-\mathbf{l}-\mathbf{m}-\mathbf{n})+\delta(\mathbf{k}-\mathbf{l}+\mathbf{m}+\mathbf{n})+\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n})+$$

$$+\delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n})+\delta(\mathbf{k}-\mathbf{l}-\mathbf{m}+\mathbf{n})+\delta(\mathbf{k}-\mathbf{l}+\mathbf{m}-\mathbf{n})]+$$

$$+\sin\Theta_{\mathbf{k}}\sin\Theta_{\mathbf{l}}\cos\Theta_{\mathbf{m}}\cos\Theta_{\mathbf{n}}[\delta(\mathbf{k}-\mathbf{l}-\mathbf{m}-\mathbf{n})+\delta(\mathbf{k}-\mathbf{l}+\mathbf{m}+\mathbf{n})-$$

$$-\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n})+\delta(\mathbf{k}-\mathbf{l}-\mathbf{m}+\mathbf{n})+\delta(\mathbf{k}-\mathbf{l}+\mathbf{m}-\mathbf{n})-$$

$$-\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n})-\delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n})]+$$

$$
\begin{aligned}
&+ \sin\Theta_{\mathbf{k}} \cos\Theta_{\mathbf{l}} \sin\Theta_{\mathbf{m}} \cos\Theta_{\mathbf{n}} [\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n}) + \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}-\mathbf{n}) - \\
&\quad - \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n}) - \delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n}) + \\
&\quad + \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}+\mathbf{n}) - \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}-\mathbf{n})] + \\
&+ \cos\Theta_{\mathbf{k}} \sin\Theta_{\mathbf{l}} \sin\Theta_{\mathbf{m}} \cos\Theta_{\mathbf{n}} [\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n}) + \delta(\mathbf{l}-\mathbf{k}-\mathbf{m}-\mathbf{n}) - \\
&\quad - \delta(\mathbf{l}-\mathbf{k}+\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n}) - \delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n}) + \\
&\quad + \delta(\mathbf{l}-\mathbf{k}-\mathbf{m}+\mathbf{n}) - \delta(\mathbf{l}-\mathbf{k}+\mathbf{m}-\mathbf{n})] + \\
&+ \sin\Theta_{\mathbf{k}} \cos\Theta_{\mathbf{l}} \cos\Theta_{\mathbf{m}} \sin\Theta_{\mathbf{n}} [\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n}) + \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}-\mathbf{n}) - \\
&\quad - \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n}) - \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n}) + \\
&\quad + \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}-\mathbf{n}) - \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}+\mathbf{n})] + \\
&+ \cos\Theta_{\mathbf{k}} \sin\Theta_{\mathbf{l}} \cos\Theta_{\mathbf{m}} \sin\Theta_{\mathbf{n}} [\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n}) + \delta(\mathbf{l}-\mathbf{k}-\mathbf{m}-\mathbf{n}) - \\
&\quad - \delta(\mathbf{l}-\mathbf{k}+\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n}) - \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n}) + \\
&\quad + \delta(\mathbf{l}-\mathbf{k}+\mathbf{m}-\mathbf{n}) - \delta(\mathbf{l}-\mathbf{k}-\mathbf{m}+\mathbf{n})] + \\
&+ \cos\Theta_{\mathbf{k}} \cos\Theta_{\mathbf{l}} \sin\Theta_{\mathbf{m}} \sin\Theta_{\mathbf{n}} [\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n}) + \\
&\quad + \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}-\mathbf{n}) - \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n}) - \\
&\quad - \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}-\mathbf{n}) - \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}+\mathbf{n})] + \\
&+ \sin\Theta_{\mathbf{k}} \sin\Theta_{\mathbf{l}} \sin\Theta_{\mathbf{m}} \sin\Theta_{\mathbf{n}} [\delta(\mathbf{k}-\mathbf{l}-\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}-\mathbf{n}) - \\
&\quad - \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}+\mathbf{n}) - \delta(\mathbf{k}+\mathbf{l}+\mathbf{m}-\mathbf{n}) - \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}-\mathbf{n}) - \\
&\quad - \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}+\mathbf{n}) + \delta(\mathbf{k}+\mathbf{l}-\mathbf{m}-\mathbf{n})]\} - \\
&- 2 \cdot \frac{\mu_0 \lambda_0}{\sqrt{\Omega}} z_0^3 - 3 \frac{\mu_0 \lambda_0}{\sqrt{\Omega}} z_0 \cdot \sum z_{\mathbf{k}}^2 - \frac{3\mu_0 \lambda_0}{2\sqrt{\Omega}} \sum_{\mathbf{k}} \sum z_{\mathbf{k}}^2 z_{\mathbf{l}} \cdot \\
&\quad \cdot [\cos 2\Theta_{\mathbf{k}} \cos\Theta_{\mathbf{l}} \delta(2\mathbf{k}-\mathbf{l}) + \sin 2\Theta_{\mathbf{k}} \sin\Theta_{\mathbf{l}} \cdot \delta(2\mathbf{k}-\mathbf{l})] - \\
&\quad - \frac{\mu_0 \lambda_0}{2\sqrt{\Omega}} \sum_{\mathbf{k}} \sum_{\mathbf{l}} \sum_{\mathbf{m}} \cdot z_{\mathbf{k}} z_{\mathbf{l}} z_{\mathbf{m}} \cdot \{\cos\Theta_{\mathbf{k}} \cos\Theta_{\mathbf{l}} \cos\Theta_{\mathbf{m}} \cdot \\
&\quad \cdot [\delta(\mathbf{k}+\mathbf{l}+\mathbf{m}) + \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}) + \delta(\mathbf{k}-\mathbf{l}+\mathbf{m})] + \\
&+ \sin\Theta_{\mathbf{k}} \sin\Theta_{\mathbf{l}} \cdot \cos\Theta_{\mathbf{m}} \cdot [\delta(\mathbf{k}-\mathbf{l}-\mathbf{m}) + \delta(\mathbf{k}-\mathbf{l}+\mathbf{m}) - \delta(\mathbf{k}+\mathbf{l}-\mathbf{m})] + \\
&+ \sin\Theta_{\mathbf{k}} \cos\Theta_{\mathbf{l}} \sin\Theta_{\mathbf{m}} \cdot [\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}) + \delta(\mathbf{k}-\mathbf{l}-\mathbf{m}) - \delta(\mathbf{k}-\mathbf{l}+\mathbf{m})] + \\
&+ \cos\Theta_{\mathbf{k}} \sin\Theta_{\mathbf{l}} \sin\Theta_{\mathbf{m}} \cdot [\delta(\mathbf{k}+\mathbf{l}-\mathbf{m}) + \delta(\mathbf{l}-\mathbf{k}-\mathbf{m}) - \delta(\mathbf{l}-\mathbf{k}+\mathbf{m})]\} .
\end{aligned}
$$

REFERENCES

1. I. Goldstone, Nuovo Cimento, **19**, 154, 1961.
2. D. I. Blokhintsev, preprint, Dubna, 1962.
3. G. Marx, Acta Phys. Hung., **14**, 27, 1962.
4. R. Haag, Danske Vidensk. Sels. Mat. Phys. Medd., **29**, No. 12, 1955.
5. L. I. Schiff, preprint, Stanford University, 1962.
6. G. Kuti, Acta Phys. Hung., in preparation.
7. H. P. Dürr and W. Heisenberg, Zeitschrift für Naturforschung, **16a**, 726, 1961.

МОДЕЛЬ С РЕШЕНИЕМ СВЕРХПРОВОДИМОСТИ В КВАНТОВОЙ ТЕОРИИ ПОЛЯ. II.

Г. КУТИ и Г. МАРКС

Р е з ю м е

Рассматривается нелинейное скалярное поле, обладающее двумя различными основными состояниями. Основное и первое возбуждённое состояния скалярного поля определяются методом Ритца. Применяются функции состояния в виде произведения функций состояния отдельных парциальных волн. После ренормировки масс энергия покоя квантов скалярного поля оказывается конечной.

REMARKS ON PION DECAY

By

K. NAGY

INSTITUTE FOR THEORETICAL PHYSICS, ROLAND EÖTVÖS UNIVERSITY, BUDAPEST

(Presented by Z. Gyulai. — Received 20. IX. 1963)

The branching ratio of the decay processess $\pi^- \to e^- + \tilde{\nu}$, $\pi^- \to \mu^- + \tilde{\nu}$, is determined, supposing that $\nu \neq \nu'$ and the muon-neutrino mass $m_{\nu'}$ is different from zero. Numerical values are given for the mass ratios $m_{\nu'}/m_e = 5, 10, 20, 30$ and 40. $m_{\nu'} \sim (5-10)m_e$ seems to be compatible with the recent experimental results.

The $\pi^- \to e^- + \tilde{\nu}$ decay process was discovered by FAZZINI et al. [1] in 1958. This experiment has given an excellent proof for the existence of the universal Fermi interaction. Namely, according to this universal $V - A$ interaction this decay process also exists and is 10^{-4} times as rare as the usual $\pi^- \to \mu^- + \tilde{\nu}'$ decay. The process goes through a virtual nucleon antinucleon pair, schematically in the following way

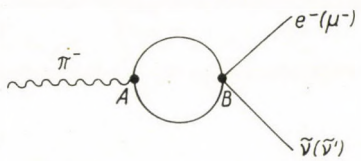


Fig. 1

where there stands in A a strong, in B a weak vertex. Supposing that the weak coupling constants $(e\ \nu)$ and $(\mu\nu')$ are the same, the branching ratio is independent of the strong pion-nucleon interaction. This ratio has been calculated first by FINKELSTEIN and RUDERMAN [2]. They found

$$R\left(\frac{\pi \to e\nu}{\pi \to \mu\nu'}\right) = \frac{m_e^2}{m_\mu^2}\left(\frac{m_\pi^2 - m_e^2}{m_\pi^2 - m_\mu^2}\right)^2. \tag{1}$$

Substituting the mass values $m_\mu = 206{,}771\ m_e$; $m_\pi = 273{,}183\ m_e$ respectively, one obtains $R = 1{,}282 \cdot 10^{-4}$. BERMAN [3] and KINOSHITA [4] later have taken into account also higher-order radiation corrections, from which $R = 1{,}233 \cdot 10^{-4}$. The experimental results

$$R_{\text{exp}} = (1{,}28 \pm 0{,}30) \cdot 10^{-4} \quad \text{ASHKIN et al. [5], 1959,}$$
$$R_{\text{exp}} = (1{,}21 \pm 0{,}07) \cdot 10^{-4} \quad \text{ANDERSON et al. [6], 1960,}$$

confirm the theoretical value extremely well.

In the theoretical papers quoted above it is supposed that the neutral particles created together with the electrons and muons are the same. According to the measurements of DANBY et al. [7] the electron-neutrino is different from the muon-neutrino. In an earlier paper [8] we have supposed that the muon and the muon-neutrino possess also another type of interaction producing a muon-neutrino mass different from zero. Whether such an interaction does or does not exist, has to be decided by further investigations. At any rate it seems very tempting to suppose that $m_{\nu'} \neq 0$. We wish to find here a limit for $m_{\nu'}$ from an appropriate experimental result. The branching ratio $(\pi \to e\ \nu)/(\pi \to \mu\nu')$ too seems to give an upper limit for $m_{\nu'}$. In the following the ratio $R\,(\pi \to e\nu/\pi \to \mu\nu')$ will be determined supposing that $\nu \neq \nu'$ and $m_{\nu'}$ is different from zero.

Let us study first the process $\pi^- \to e^- + \tilde{\nu}$. The transition probability for this process is

$$w(\pi \to e\nu) \sim \Sigma \mid M \mid^2 \varrho(E)\,, \tag{2}$$

where Σ denotes summation over the final lepton spin states. M is the transition matrix element of the weak Hamiltonian between the initial and final physical states determined by the strong interaction:

$$M = < e^-, \tilde{\nu} \mid H_{\text{weak}} \mid \pi > = A_\alpha\, \overline{u}_e\, \gamma_\alpha\, (1 + \gamma_5)\, v_\nu\,, \tag{3}$$

where u_e and v_ν are the lepton spinor amplitudes, A_α is the matrix element of the axial vector current between the states $\mid \pi>$ and $\mid 0>$. A_α is proportional to the four-momenta P_α of the pion [9]. Furthermore in the rest system of the pion

$$M = A_4\, \overline{u}_e\, \gamma_4\, (1 + \gamma_5)\, v_\nu\,. \tag{4}$$

The summation over the spin states can be carried out by means of the energy projection operator as usual, which gives:

$$\sum_{e,\tilde{\nu}} \mid M \mid^2 = \mid A_4 \mid^2 \frac{1}{2 p_4 q_4}\, (\mathrm{Sp}\hat{p}\, \gamma_4\, \hat{q}\, \gamma_4 + \mathrm{Sp}\hat{p}\, \gamma_4\, \gamma_5\, \hat{q}\, \gamma_4)\,, \tag{5}$$

where p_α and q_β denote the momenta of the electron and antineutrino, respectively, $\hat{p} = p_\alpha\, \gamma_\alpha$, $\hat{q} = q_\beta\, \gamma_\beta$.

Using the well-known rules

$$\begin{aligned}
&\gamma_\alpha\, \gamma_\beta + \gamma_\beta\, \gamma_\alpha = 2\delta_{\alpha\beta}\,,\\
&\mathrm{Sp}\, \gamma_\alpha = 0 \quad (\alpha = 1, 2, 3, 4)\,,\\
&\mathrm{Sp}\, \gamma_5 = 0\,,\\
&\mathrm{Sp}\, (\gamma_\alpha\, \gamma_\beta) = 4\delta_{\alpha\beta}\,,\\
&\mathrm{Sp}\, (\gamma \text{ matrices of an odd number}) = 0\,,
\end{aligned} \tag{6}$$

the traces can be easily calculated, and thus

$$\sum_{e,\tilde{\nu}} |M|^2 = |A_4|^2 \, 2\left(1 - \frac{pq\,c^2}{E_e E_{\tilde{\nu}}}\right). \tag{7}$$

Here p and q are the absolute values of the corresponding momenta, E_e the electron energy and $E_{\tilde{\nu}}$ the antineutrino energy. In the rest system of the pion $p = q$. Since the mass of the electron-neutrino is zero, $E_{\tilde{\nu}} = cq$, thus

$$\sum_{e,\tilde{\nu}} |M|^2 = |A_4|^2 \, 2\left(1 - \frac{pc}{E_e}\right). \tag{8}$$

Using

$$m_\pi c^2 = E_e + E_{\tilde{\nu}} \tag{9}$$

one obtains

$$\sum_{e,\tilde{\nu}} |M|^2 = |A_4|^2 \, 4 \frac{m_e^2}{m_\pi^2 + m_e^2}. \tag{10}$$

Let us determine now the density function $\varrho(E)$.

$$\varrho(E) = \frac{4\pi V p^2}{h^3} \frac{1}{\left(\frac{dE}{dp}\right)_{E=E_e+E_{\tilde{\nu}}}}, \tag{11}$$

$\left(\frac{dE}{dp}\right)_{E=E_e+E_{\tilde{\nu}}}$ can be written as

$$\left(\frac{dE}{dp}\right)_{E=E_e+E_{\tilde{\nu}}} = \frac{dE_e}{dp} + \frac{dE_\nu}{dq}\frac{dq}{dp} = c\left(1 + \frac{cp}{E_e}\right). \tag{12}$$

Using (9), (12), (11) gives

$$\varrho(E) \sim \frac{c(m_\pi^2 - m_e^2)(m_\pi^2 + m_e^2)}{8m_\pi^4}. \tag{13}$$

The transition probability is the product of equ. (10) and (12):

$$w(\pi \to e\nu) \sim \frac{c}{2} |A_4|^2 \frac{m_e^2}{m_\pi^4} (m_\pi^2 - m_e^2)^2. \tag{14}$$

The transition probability for the process $\pi^- \to \mu^- + \tilde{\nu}'$ with $m_{\nu'} \neq 0$ can be calculated analogously. The result is:

$$w(\pi \to \mu\nu') \sim \frac{c}{2} |A_4|^2 \frac{1}{m_\pi^4} [(m_\pi^2 - m_\mu^2)^2 + m_{\nu'}^2 (m_{\nu'}^2 - 2m_\pi^2 - 2m_\mu^2)] \times \tag{15}$$

$$\times [m_\mu^2 (m_\pi^2 - m_\mu^2) + m_{\nu'}^2 (2m_\mu^2 + m_\pi^2 - m_{\nu'}^2)].$$

Here $|A_4|$ is, of course, the same as in (14).

The branching ratio is given by the quotient of the two transition probabilities, therefore

$$R\left(\frac{\pi \to e\nu}{\pi \to \mu\nu'}\right) =$$

$$= \frac{m_e^2 (m_\pi^2 - m_e^2)^2}{[(m_\pi^2 - m_\mu^2)^2 + m_{\nu'}^2 (m_{\nu'}^2 - 2m_\pi^2 - 2m_\mu^2)]^{1/2} [m_\mu^2 (m_\pi^2 - m_\mu^2) + m_{\nu'}^2 (2m_\mu^2 + m_\pi^2 - m_{\nu'}^2)]}. \quad (16)$$

In order to get an easily comparable formula let us write appropriately

$$R\left(\frac{\pi \to e\nu}{\pi \to \mu\nu'}\right) = R_0 \, r \,, \quad (17)$$

where R_0 is the ratio for $m_{\nu'} = 0$:

$$R_0 = \frac{m_e^2}{m_\mu^2} \left(\frac{m_\pi^2 - m_e^2}{m_\pi^2 - m_\mu^2}\right)^2, \quad (18)$$

and r expresses the correction from $m_{\nu'} \neq 0$:

$$\frac{1}{r} = \left[1 + m_{\nu'}^2 \frac{m_{\nu'}^2 - 2m_\pi^2 - 2m_\mu^2}{(m_\pi^2 - m_\mu^2)^2}\right]^{\frac{1}{2}} \left[1 + m_{\nu'}^2 \frac{2m_\mu^2 + m_\pi^2 - m_{\nu'}^2}{m_\mu^2 (m_\pi^2 - m_\mu^2)}\right]. \quad (19)$$

The numerical values of r for different $m_{\nu'}$ are tabulated below.

$m_{\nu'}/m_e =$	0	5	10	20	30	40
$r =$	1	0,99996	1,00001	1,00259	1,01621	1,05981

As one can observe, for $m_{\nu'}/m_e = 5{,}10$; $r \sim 1$, $R \sim R_0$. For $m_{\nu'}/m_e = 20$, $R = 1{,}32 \cdot 10^{-4}$ which is already outside the experimental range of accuracy, thus this value is definitely excluded by the experiments. This is valid a fortiori for higher values of $m_{\nu'}$. DUDZIAK et al. [10] measuring the maximal energy of the electrons emerging from a muon decay and using the formula

$$E_{e\max} = \frac{m_\mu^2 + m_e^2 - m_{\nu'}^2}{2m_\mu} c^2 \,,$$

have concluded, that $m_{\nu'} \sim 8\, m_e$. BAHCALL and CURTIS [11] concluded from the lifetime of the muon that the value $m_{\nu'} \sim 5\, m_e$ is the most probable.

The above results do not permit us to determine the possible value of $m_{\nu'}$ unambiguously, but the comparison with the experimental results allows

the range $m_{\nu'} \sim (5 - 10)\, m_e$ in accordance with [10] and [11] and also with another result of ours [12] dealing with the energy spectrum of the electrons in muon decay.

Here the higher-order radiation corrections have not been taken into account. It is conceivable that these corrections restrict the value of $m_{\nu'}$ more accurately. Investigations of this type are in progress and will be published in a subsequent paper.

REFERENCES

1. T. Fazzini, G. Fidecaro, A. W. Merrison, H. Paul and A. V. Tollestrup, Phys. Rev. Letters, **1**, 247, 1958.
2. M. Ruderman and R. Finkelstein, Phys. Rev., **76**, 1458, 1949.
3. S. M. Berman, Phys. Rev. Letters, **1**, 468, 1959.
4. T. Kinoshita, Phys. Rev. Letters, **2**, 477, 1959.
5. J. Ashkin, T. Fazzini, G. Fidecaro, N. H. Lipman, A. W. Merrison and H. Paul, Nuovo Cimento, **14**, 1266, 1959.
6. H. L. Anderson, T. Fujii, R. H. Miller and L. Tau, Phys. Rev., **119**, 2050, 1960.
7. G. Danby, J. M. Gaillard, K. Goulianos, L. M. Lederman, N. Mistry, M. Schwartz and J. Steinberg, Phys. Rev. Letters, **9**, 36, 1962.
8. K. Nagy, Conference on Theor. Phys. Balatonföldvár, 1961. Wiss. Zeitschr. d. Karl Marx Univ. Leipzig, **11**, 389, 1962.
9. N. N. Khuri, Proc. of the 1960 Ann. Intern. Conf. on High Energy Physics at Rochester, 513, 1960.
10. W. Dudziak, R. Sagane and J. Veder, Phys. Rev., **114**, 336, 1959.
11. J. Bahcall and R. B. Curtis, Nuovo Cimento, **21**, 422, 1961.
12. K. Nagy, Acta Phys. Hung., **17**, 163, 1964.

ЗАМЕЧАНИЯ О РАСПАДЕ ПИОНОВ

К. НАДЬ

Резюме

Определяется отношение переходных вероятностей распадов $\pi^- \to e^- + \tilde{\nu}$, $\pi^- \to \mu^- + \tilde{\nu}'$ с предположением, что $\nu \neq \nu'$ и масса мюон-нейтрино $m_{\nu'}$ отлична от нуля. Численное значение отношения распада определялось для значения масс $m_{\nu'}/m_e = 5$, 10, 20, 30 и 40. по сравнению с экспериментальными данными вероятно, что $m_{\nu'} \sim (5—10)\, m_e$.

MUON DECAY WITH TWO KINDS OF NEUTRINOS

By

K. NAGY

INSTITUTE FOR THEORETICAL PHYSICS, ROLAND EÖTVÖS UNIVERSITY, BUDAPEST

(Presented by Z. Gyulai — Received: 20. IX. 1963)

The energy spectrum of the electrons originated in the muon decay is calculated by assuming a muon-neutrino of non-vanishing rest mass. The change of the spectrum caused by this assumption is also discussed.

In recent years many papers have dealt with the problem of the muon-neutrino. It was first concluded from theoretical considerations that the neutrinos originating from the decay $\pi \to \mu + \nu'$ and the β-decay are different. Last year the assumption of the existence of the two kinds of neutrinos has been experimentally justified by DANBY et al. [1]. In the latest theoretical articles several authors attempt to construct the theory of weak interactions including these different neutrinos [2]. There is hope that on the basis of the new theory one can explain the smaller coupling constant of the strangeness non-conserving weak interactions and the e — μ mass difference.

The question arises as to the fundamental property in which the two neutrinos differ. In recent papers the possibility was discussed that the muon doublet (thus also the muon-neutrino) participates in an interaction hitherto unknown and consequently the rest mass $m_{\nu'}$ of the muon-neutrino is finite [3]. Perhaps this will lead to the understanding of the e — μ mass difference, too. According to the investigations [4], the present experimentaldata are consistent with the assumption of $m_{\nu'} \neq 0$ and give $m_{\nu'} \sim (5 - 10)m_e$.

In the following we shall deal with the muon decay and examine the consequence of the assumption $m_{\nu'} \neq 0$ for the energy spectrum of the electrons originating from the decay. It will be shown that the parts of high energy of the spectrum depend very sensitively on the muon-neutrino mass.

Thus we shall determine the transition probability of the process $\mu^- \to e^- + \tilde{\nu} + \nu'$ as the function of the energy of the outgoing electron. The interaction is described by the Hamiltonian

$$H = g \int \left(\overline{\psi}_e \gamma_\alpha \frac{1 + \gamma_5}{2} \psi_\nu \right) \left(\overline{\psi}_{\nu'} \gamma_\alpha \frac{1 + \gamma_5}{2} \psi_\mu \right) dV + \text{herm. conj.}, \qquad (1)$$

where g is the coupling constant, ψ_i ($i = e, \mu, \nu, \nu'$) are the field operators

of the leptons participating in the process, $\bar{\psi} = \psi^+ \gamma_4$, γ_α ($\alpha = 1, 2, 3, 4$) are the well-known Dirac matrices and $\gamma_5 = \gamma_1 \gamma_2 \gamma_3 \gamma_4$. In the representation used here all the γ_α are Hermitian, $\gamma_\alpha^+ = \gamma_\alpha$. As usual, we expand the lepton fields

$$\psi_i = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}_i} \left(a_{\mathbf{k}_i} u_i(k) e^{i\mathbf{k}_i \mathbf{r}} + b^+_{\mathbf{k}_i} v_i(k) e^{-i\mathbf{k}_i \mathbf{r}} \right). \tag{2}$$

Here V denotes the normalization volume, a (b) annihilates a particle (antiparticle) and a^+, b^+ are the corresponding creation operators.

In the lowest order of the perturbation theory the transition probability is given by

$$w = \frac{2\pi}{\hbar} \overline{|H_{fi}|^2} \varrho(E) . \tag{3}$$

H_{fi} is the matrix element of the transition $i \to f$ and the dash means an average over initial and a summation over final spin states. $\varrho(E)$ denotes the density of the final states which can be easily determined by phase space considerations. First of all, let us compute the transition matrix element H_{fi}. From (1) and (2) we have

$$\begin{aligned} H_{fi} &= < e^-, \tilde{\nu}, \nu' \,|\, H \,|\, \mu^- > = \\ &= \frac{g}{V^2} \left(\bar{u}_e \gamma_\alpha \frac{1+\gamma_5}{2} v_\nu \right) \left(\bar{u}_{\nu'} \gamma_\alpha \frac{1+\gamma_5}{2} u_\mu \right) \int e^{i(\mathbf{k}_\mu - \mathbf{k}_e - \mathbf{k}_\nu - \mathbf{k}_{\tilde{\nu}})\mathbf{r}} \, dV = \\ &= \frac{g}{V} \left(\bar{u}_e \gamma_\alpha \frac{1+\gamma_5}{2} v_\nu \right) \left(\bar{u}_{\nu'} \gamma_\alpha \frac{1+\gamma_5}{2} u_\mu \right) . \end{aligned} \tag{4}$$

Here we have used that if

$$\mathbf{k}_\mu = \mathbf{k}_e + \mathbf{k}_\nu + \mathbf{k}_{\nu'} \tag{5}$$

then the integral is equal to V, otherwise it is zero. (5) expresses just the momentum conservation. Now, we introduce the following abbreviations

$$\hbar \mathbf{k}_e = \mathbf{p} ; \quad \hbar \mathbf{k}_{\tilde{\nu}} = \mathbf{q}' ; \quad \hbar \mathbf{k}_{\nu'} = \mathbf{q} ; \quad \hbar \mathbf{k}_\mu = \mathbf{p}_\mu \tag{6}$$

and choose the rest system of the muon as the frame of reference. Then $\mathbf{p}_\mu = 0$. In this frame of reference instead of (5) one can write

$$\mathbf{p} + \mathbf{q} + \mathbf{q}' = 0 . \tag{7}$$

Let us calculate $|H_{fi}|^2$ average over initial and sum over final spin states, we obtain

$$\begin{aligned} |H_{fi}|^2 = \frac{g^2}{V^2} &\left(v_\nu^+ \frac{1+\gamma_5}{2} \gamma_\alpha \gamma_4 u_e \right) \left(\bar{u}_e \gamma_\beta \frac{1+\gamma_5}{2} v_\nu \right) \times \\ &\times \left(u_\mu^+ \frac{1+\gamma_5}{2} \gamma_\alpha \gamma_4 u_{\nu'} \right) \left(\bar{u}_{\nu'} \gamma_\beta \frac{1+\gamma_5}{2} u_\mu \right) . \end{aligned} \tag{8}$$

$$\overline{|H_{fi}|^2} = \frac{1}{2} \sum_{e,\tilde{\nu},\nu',\mu}^{\varepsilon=1,2} |H_{fi}|^2 =$$

$$= -\frac{g^2 ic}{128 V^2 E_{\tilde{\nu}} E E_{\nu'}} \mathrm{Sp}\left[(m_e c^2 - ic\hat{p})\gamma_\beta(1+\gamma_5)\hat{q}'\gamma_4\gamma_\alpha\gamma_4\right] \times \tag{9}$$

$$\times \mathrm{Sp}\left[(m_{\nu'} c^2 - ic\hat{q})\gamma_\beta(1+\gamma_5)\gamma_\alpha\gamma_4\right].$$

E, $E_{\tilde{\nu}}$ and $E_{\nu'}$ are the energies of the electron, antineutrino and muon-neutrino, respectively. The traces can be easily evaluated by using the rules

$$\begin{aligned}
&\gamma_\alpha\gamma_\beta + \gamma_\beta\gamma_\alpha = 2\delta_{\alpha\beta}\,,\\
&\mathrm{Sp}\,(\gamma_\alpha) = 0 \quad \mathrm{Sp}\,(\gamma_5) = 0\,,\\
&\mathrm{Sp}\,(\gamma_\alpha\gamma_\beta) = 4\delta_{\alpha\beta} \quad \mathrm{Sp}(1) = 4\,,\\
&\mathrm{Sp}\,(\text{matrices of an odd number}) = 0\,,\\
&\mathrm{Sp}\,(\gamma_5 \cdot \gamma_\alpha \text{ matrices of number} < 4) = 0\,.
\end{aligned} \tag{10}$$

Finally, it follows that

$$\overline{|H_{fi}|^2} = \frac{g^2 c^2 p}{2V^2 E}\left(\frac{E}{pc^2} - \frac{q}{E_{\nu'}}\cos(e,\nu')\right), \tag{11}$$

where $p = |\mathbf{p}|$, $q = |\mathbf{q}|$ and (e, ν') denotes the angle between the momenta of the muon-neutrino and the electron (Fig. 1). It seems to be expedient to express (11) by $\cos(e, \bar{\nu})$. Using (7), in an elementary way one gets the relation

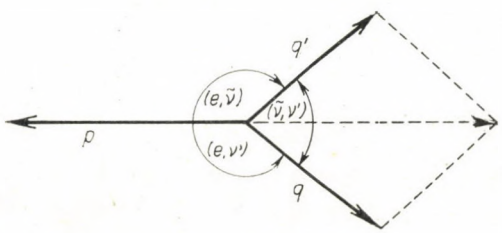


Fig. 1

$$-q\cos(e,\nu') = p + q'\cos\vartheta\,, \tag{12}$$

where $\vartheta = (e, \tilde{\nu})$. Substituting (12) into (11) we can write

$$\overline{|H_{fi}|^2} = \frac{g^2}{2V^2 E}\left(E + \frac{pc^2(p + q'\cos\vartheta)}{E_{\nu'}}\right). \tag{13}$$

In the following the density $\varrho(E_0)$ of the final states will be determined. Here E_0 means the total energy of the leptons originating from the decay

which, because of the energy conservation, is equal to the rest energy of the muon

$$E_0 = m_\mu c^2 = E + E_{\tilde{\nu}} + E_{\nu'} . \tag{14}$$

The number of the states in unit energy interval can be easily computed by means of the phase space: the corresponding phase volume is divided by the volume of the phase cell representing the states. In this manner we have

$$\varrho(E_0) = \frac{8\pi^2 V^2}{h^6} p^2 \, dp q'^2 \left(\frac{dq'}{dE_0} \right)_p \sin\vartheta \, d\vartheta . \tag{15}$$

By using the energy equation q' can be expressed as a function of E_0, E and p and the derivation can be easily accomplished. From

$$E_0 - E - E_{\tilde{\nu}} = E_{\nu'} = c \sqrt{(\mathbf{p} + \mathbf{q}')^2 + m_{\nu'}^2 c^2}$$

it follows that

$$q' = \frac{(E_0 - E)^2 - c^2 p^2 - m_{\nu'}^2 c^4}{2c(E_0 - E + cp\cos\vartheta)} . \tag{16}$$

Let us differentiate (16) with respect to E_0 (p is constant), we get

$$\left(\frac{dq'}{dE_0} \right)_p = \frac{(E_0 - E)^2 + c^2 p^2 + m_{\nu'}^2 c^4 + 2cp\,(E_0 - E)\cos\vartheta}{2c\,(E_0 - E + cp\cos\vartheta)^2} . \tag{17}$$

At this stage of the calculation it is expedient to introduce the following abbreviations

$$2\,ca = (E_0 - E)^2 + c^2 p^2 + m_{\nu'}^2 c^4 , \tag{18,a}$$

$$2\,cb = (E_0 - E)^2 - c^2 p^2 - m_{\nu'}^2 c^4 , \tag{18,b}$$

$$A = cp[E(E_0 - E) + c^2 p^2 + bc] , \tag{18,c}$$

$$B = E(E_0 - E)^2 - Ebc + (E_0 - E) p^2 c^2 , \tag{18,d}$$

$$C = (E_0 - E)cp , \tag{18,e}$$

$$D = (E_0 - E)^2 - bc , \tag{18,f}$$

$$x = \cos\vartheta . \tag{18,g}$$

Now we substitute (16), (17) and the abbreviations (18,a)–(18,g) into (15). Hence

$$\varrho(E_0) = \frac{8\pi^2 V^2}{h^6} p^2 b^2 \, dp \, \frac{[a + (E_0 - E)\,px]}{(E_0 - E + cpx)^4} \, d(-x) . \tag{19}$$

Furthermore, from (18,a)—(18,g) it is clear that

$$a + px(E_0 - E) = \frac{1}{c}(Cx + D),$$

therefore (19) can be rewritten in the form

$$\varrho(E_0) = \frac{8\pi^2 V^2}{h^6 c} p^2 b^2 \frac{(Cx + D)}{(E_0 - E + cpx)^4} d(-x). \tag{20}$$

By using the abbreviations (18,a)---(18,g), we rewrite also (13)

$$\overline{|H_{fi}|^2} = \frac{g^2}{2V^2 E} \frac{Ax + B}{Cx + D}. \tag{21}$$

The transition probability is given by (3), substituting (20) and (21) into (3), it is found that

$$w = \frac{16\pi^4 g^2 b^2 p^2 dp}{h^7 cE} \left\{ \frac{K}{2} \int_{-1}^{1} \frac{dx}{(cpx + E_0 - E)^3} - E_0 bc \int_{-1}^{1} \frac{dx}{(cpx + E_0 - E)^4} \right\}, \tag{22}$$

where

$$K = E_0^2 - m_e^2 c^4 - m_{\nu'}^2 c^4. \tag{23}$$

After having calculated the integrals, (22) has the form

$$w(E)\, dE = \frac{4\pi^4 g^2}{h^7 c^6} \sqrt{E^2 - m_e^2 c^4}\, F^2 \left\{ K(E_0 - E) - \frac{FE_0}{3} [3(E_0 - E)^2 + E^2 - m_e^2 c^4] \right\}, \tag{24}$$

here

$$F = \frac{E_0^2 - 2E_0 E + m_e^2 c^4 - m_{\nu'}^2 c^4}{E_0^2 - 2E_0 E + m_e^2 c^4}. \tag{25}$$

In case of $m_{\nu'} \neq 0$ equation (24) describes the energy spectrum of the electrons originating from the muon decay.

In the following we shall examine the domain $E \gg m_e c^2$. For this purpose we introduce two abbreviations:

$$\varepsilon = \frac{E}{E_0}; \qquad \mu = \frac{m_{\nu'}}{m_\mu}. \tag{26}$$

The quantities K, F defined by (23) and (25) can be expressed by ε and μ. In case of $E \gg m_e c^2$ we have

$$K = E_0^2 (1 - \mu^2), \tag{27}$$

$$F = 1 - \frac{\mu^2}{1 - 2\varepsilon}. \tag{28}$$

Substituting (26), (27) and (28) into (24), and neglecting the rest energy of the electron ($m_e c^2$ is small compared with E), we are led to the following energy spectrum of the electron:

$$w(\varepsilon)\, d\varepsilon = \frac{4\pi^4 g^2}{h^7 c^6} E_0^5 I(\varepsilon)\, d\varepsilon\,, \tag{29}$$

and

$$I(\varepsilon) = \varepsilon \left(1 - \frac{\mu^2}{1-2\varepsilon}\right)^2 \left[(1-\mu^2)(1-\varepsilon) - \left(1 - \frac{\mu^2}{1-2\varepsilon}\right)\left((1-\varepsilon)^2 + \frac{\varepsilon^2}{3}\right)\right]. \tag{30}$$

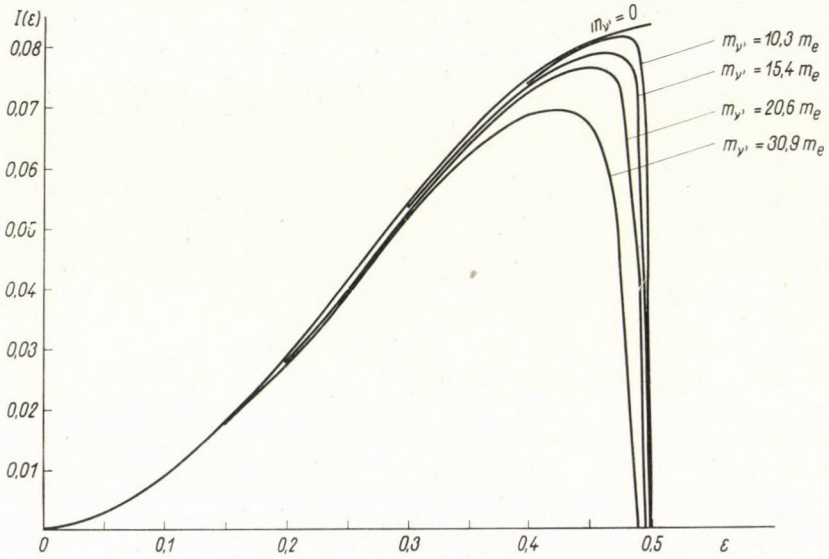


Fig. 2

At present we are interested mainly in the influence of the parameter μ on the energy spectrum. In order to make easier the survey, we expand $I(\varepsilon)$ in powers of μ^2. A very simple calculation yields the result

$$I(\varepsilon) = \varepsilon^2 \left[1 - \frac{4}{3}\varepsilon - \mu^2 - \frac{\mu^4}{(1-2\varepsilon)^2} + \frac{\mu^6}{(1-2\varepsilon)^3}\left(1 - \frac{2}{3}\varepsilon\right)\right]. \tag{31}$$

Equation (31) shows very clearly the dependence of the energy spectrum on the parameter μ. The functions $I(\varepsilon)$ belonging to the values $\mu = 0$; 0,05; 0,075; 0,10 and 0,15 are represented in Fig. 2.

The curves $\mu \neq 0$ fall off at the edge of the spectrum and cut the energy axis before the value $\varepsilon = \frac{1}{2}$ is reached. In the general case treated here the maximum energy of the electron is as follows

$$E_{\max} = \frac{(m_\mu^2 + m_e^2 - m_\nu^2)\, c^2}{2 m_\mu}\,. \tag{32}$$

From (32) it is obvious that if the parameter μ increases, the energy $E_{\max}$ decreases, in other words, the curves belonging to larger values μ are cut off before those of smaller μ. The shapes of the curves are in qualitative agreement with those of the empiric curves [5]; indeed, the energy spectrum falls off at high energies in contrast to the theoretical curve corresponding to the case $\mu = 0$. At any rate, this shows that the hypothesis $m_{\nu'} \neq 0$ does not contradict to the experiments. A rough comparison with the experimental data leads to the conclusion that the curves belonging to the smaller values μ indicated above are nearer to reality. The muon-neutrino masses corresponding to the five values of μ mentioned above are

$$\frac{m_{\nu'}}{m_e} = 0\,; \quad 10{,}3\,; \quad 15{,}4\,; \quad 20{,}6\,; \quad 30{,}9\,.$$

In the calculation of the transition probability radiation corrections have not been taken into account. Naturally, the theoretical estimation of them is needed and one may hope to obtain the correct mass $m_{\nu'}$ only if their contributions are known. These problems will be treated in a subsequent paper.

We mention merely that the branching ratio $(\pi \to e + \nu)/(\pi \to \mu + \nu')$ [4] leads to the conclusion $m_{\nu'} \leqslant 10\, m_e$.

REFERENCES

1. G. Danby, J. M. Gaillard, K. Goulianos, L. M. Lederman, N. Mistry, M. Schwartz and J. Steinberg, Phys. Rev. Letters, **9**, 36, 1962.
2. Z. Maki, M. Nakagawa and S. Sakata, Progr. Theor. Phys., **28**, 870, 1962.
3. G. Marx and K. L. Nagy, Nucl. Phys., **12**, 125, 1959; K. Nagy, Conference on Theoretical Physics, Balatonföldvár, 1961; Wiss. Ztschr. d. Karl Marx Univ. Leipzig, **11**, 389, 1962.
4. W. Dudziak, R. Sagane and J. Vedder, Phys. Rev., **114**, 336, 1959; J. Bahcall and R. B. Curtis, Nuovo Cimento, **21**, 422, 1961; K. Nagy, Acta Phys. Hung., **17**, 157, 1964.
5. M. M. Block, E. Fiorini, T. Kikuchi, G. Giacomelli and S. Ratti, Proc. Annual Intern. Conf. on High Energy Physics at Rochester, 533, 1960.

МЮОН-РАСПАД С ДВУМЯ РАЗЛИЧНЫМИ НЕЙТРИНО

К. НАДЬ

Р е з ю м е

Предполагая массу покоя мюон-нейтрино отличной от нуля, рассматривается, что данное предположение к каким изменениям приводит в энергетическом спектре электронов, появляющихся в мюон-распаде.

INTERSTELLAR NEUTRINO DENSITY AND COSMOGONY

By

L. Fodor,* Zs. Kövesy** and G. Marx

INSTITUTE FOR THEORETICAL PHYSICS, ROLAND EÖTVÖS UNIVERSITY, BUDAPEST

(Presented by Z. Gyulai. — Received 20. IX. 1963)

The cosmic neutrino density hitherto produced by the stars is estimated. Experimental determination of this quantity would be of great importance and any discrepancy between estimated and experimental value would be of great consequence for cosmogony. A brief review of the indirect (gravitational) and direct (atomphysical) possibilities of observation is given.

§ 1

Recently it has become evident that the neutrinos play an important role in certain periods of the evolution of the stars, and that they have to be considered as one of the most common constituents of the matter of the Universe.

In the stars heavy elements are being built up from hydrogen, this process being accompanied by the reaction

$$p^+ \to n + e^+ + \nu\,. \tag{1}$$

One can estimate easily, how many neutrinos have been produced by the stars up to now. According to data obtained by astronomical observations about $^2/_3$ of the known matter of the Universe is hydrogen, the rest is represented by the heavy elements. This means, that if formerly the baryonic charge of the Universe was carried exclusively by protons, up to now 10—20% of them have transmuted into neutrons. According to (1) the birth of every n is accompanied by the birth of a ν. The energy of the β neutrinos is greater than or equal to 1 MeV. Consequently, the mass concentration of the neutrinos of thermonuclear origin (as compared to the atomic mass density ϱ^*) is about 0.1‰ today. Since, according to the astronomical estimation, $\varrho^* \sim 2 \cdot 10^{-29}$ g cm^{-3}, the mass density of β neutrinos turns out to be $\varrho_\beta \sim 10^{-32}$ g cm^{-3}, and the intensity $I_\beta(\nu)$ of the overall neutrino radiation $I_\beta(\nu) \sim 10^6\ \nu$ cm^{-2}s^{-1} [1].

* Permanent address: Observatory of the Hungarian Academy of Sciences, Budapest-Szabadsághegy.

** Permanent address: Central Research Institute of Physics, Budapest.

The actual density is probably overestimated by having taken the mean energy of the neutrinos to be 1 MeV, because the energy of the neutrinos, the origin of which is far from here, may be considerably decreased by the DOPPLER shift. But it is not probable that this would cause a change in the order of magnitude.

Naturally, the β^+-decays accompanying the thermonuclear reactions are not the uniquely possible sources of the neutrinos. In the case of a direct $(ev)(ev)$ interaction a considerable part of the thermal radiation inside a celestial body of a temperature above 10^8 °K is carried by ν-$\bar{\nu}$ pairs, independently of the chemical composition [2, 3, 4, 5]. Let us estimate the neutrino contribution to this thermal radiation and the density of $\bar{\nu}$ of thermal origin. The calculation can be carried out on the basis of a special star model. Let us choose the model proposed by C. HAYASHI and R. C. CAMERON and worked out by them in detail for a giant star of mass $15.6\ M_\odot = 3.1 \cdot 10^{34}$ g [6]. On

Table

Period	Million years	$E(\gamma)$	$E_{\beta}(\nu)$	$E_{\beta}(\bar{\nu})$	$E_{\mathrm{th}}(\nu\bar{\nu})$	ΣE
		in 10^{50} ergs units				
Contraction	1	—	—	—	—	—
H → He	156	410	28	—	—	440
He → C	12	69	3	—	10^{-3}	72
C → Ne	0.05	60	10	—	10	80
Ne → Fe	10^{-3}	0.3	—	—	24	24
Explosion	10^{-6}	0.1	0.1	10^{-3}	—	0.2

hand of this model, they describe the development in the course of time of the stellar material, from the initial state, containing 90% hydrogen, to the supernova explosion. The neutrino luminosity of the star is determined for β neutrinos by the thermonuclear reactions occurring in the stars and for the thermal ν-$\bar{\nu}$ pair radiation by the distribution of the density and temperature in the center of the star. In the supernova state also such elements are produced (and ejected) which show β^--decay, so they are $\bar{\nu}$-active (e.g. the natural radioactive elements on the Earth are of this kind).

The energies radiated during certain periods of the evolution of a giant star, calculated according to the HAYASHI—CAMERON model are summarized in the Table. It can be seen, that between the formation of the star and the supernova explosion altogether about 2‰ of the rest energy of the stellar matter has been liberated as radiation. The radiated energy is distributed in the following way: in the form of thermal optical radiation $540 \cdot 10^{50}$ erg (88%), in the form of β neutrinos $42 \cdot 10^{50}$ erg (7%), in the form of β anti-neutrinos $0.01 \cdot 10^{50}$ erg ($\ll 1\%$) and in the form of thermal neutrino pairs

$34 \cdot 10^{50}$ erg (5%). Our results mean that if all the stars obeyed the HAYASHI—CAMERON model exactly and if all the stars reached the final stage of their evolution, the mass density $\varrho(\nu)$ of the neutrino radiation being present in the Universe would be $^4/_3$ times as great as $\varrho_\beta(\nu)$, while the antineutrino density $\varrho(\bar{\nu})$ would be $^1/_3$ $\varrho_\beta(\nu)$.

Most stars have not yet reached the final stage of their evolution, and the most common stars are dwarfs, the evolution of which is not yet understood in enough detail to allow a comprehensive calculation. It is probable, however, that in the case of a dwarf star the temperature is lower and the degeneracy is stronger than in the case of a giant star. Therefore the ratio $E_{\text{th}}(\nu\bar{\nu})/E_\beta(\nu)$ is smaller in the case of the dwarf stars than it would be according to the HAYASHI—CAMERON model. (The processes of high ν-$\bar{\nu}$ productivity are slowed down and other processes of small productivity become dominating [7, 8, 9].) It seems to be certain that

$$\varrho(\nu) \lesssim 10^{-32}\,\text{g cm}^{-3}, \quad \varrho(\bar{\nu}) \leq \frac{1}{3}\varrho(\nu) \tag{2}$$

is a good estimation of the density of the neutrinos occurring in the Universe. The average energy of a neutrino is not greater than 1 MeV, only a small fraction of them has energies of 2—3 MeV. (We should like to give a more accurate estimation in another paper [10].)

§ 2

One may ask whether there is another possible source which may make an essential contribution to the present neutrino density? It is not probable that another reaction does exist, which has escaped notice and which could produce a thermal ν-$\bar{\nu}$ pair radiation of an intensity greater than 1 erg $\text{g}^{-1}\,\text{s}^{-1}$ in the circumstances that reign in the interior of the dwarf stars (degenerated gas of 10^7—10^8 °K temperature). One can rather imagine that besides the short-lived pre-supernovae, there is a considerable amount of matter in the Universe which is in the ideal gas state and is of about 10^9 °K temperature. Here one may think of certain types of variable stars in a certain interval of the period of their variation; or rather of the central parts of our Galaxy or other galaxies. Since the centers of the galaxies make up a considerable part of the matter of the Universe, their contribution to $\varrho(\nu)$ may be important.

According to some hypotheses, the state of the matter in the central regions of the galaxies is yet more singular than that of supernovae. In the centers of the galaxies the density may be higher than that of nuclear matter and the Fermi energy of the nucleons may be greater than the mass of the pion or the mass difference between the hyperon and the nucleon. (State of degener-

ate hyperon gas [11].) According to AMBARTSUMIAN [12] the evolution of the galactic centers leads to the decay of this state, while in HOYLE's opinion [13] it leads to the catastrophic formation of such a state. The occurrence of such catastrophic transitions seems to be verified by astronomical observations (ejection of matter out of the central regions of galaxies, the strange properties of the radio source 3C273B, etc.). If temperatures occur which are characterized by the relation $kT \gg m_\pi \cdot c^2$ a considerable part of the matter can go over into mesons, the decay of which produces high energy neutrinos.

According to some other theories all the matter of the Universe was in such a hot state about 10^{10} years ago and the decay of this state could produce a high neutrino intensity [14, 2].

To procure direct observational evidence for the occurrence of such a hot state of the cosmic matter seems to be hopeless at first sight. But the cosmic neutrino background may offer a check for the existence of this state: if $\varrho(\nu)$ turns out to be larger than the value given by [2], this fact could scarcely be explained without supposing the existence of "hot matter". A more straightforward proof of its existence would be the observation of cosmic muon-neutrinos, as these are produced only by processes, where the energy involved is greater than 100 MeV. These neutrinos mostly arise from the pion decay:

$$\begin{aligned} &\pi^+ \to \mu^+ + \nu_\mu, \quad \pi^- \to \mu^- + \bar{\nu}_\mu, \\ &\mu^+ \to e^+ + \bar{\nu} + \bar{\nu}_\mu, \quad \mu^- \to e^- + \bar{\nu} + \nu_\mu. \end{aligned} \tag{3}$$

The hot matter therefore produces β neutrinos and μ antineutrinos in numbers of equal order of magnitude. One cannot imagine an exception from this rule. In matter of high atomic number Z and in which the electron gas is degenerated, the capture processes

$$\mu^- + Z \to (Z - 1) + \nu$$

may become predominant and may diminish the $\bar{\nu}$ radiation. The ν radiation, however, survives always besides the ν_μ-$\bar{\nu}_\mu$ radiation. It could be diminished only by the reactions

$$\mu^- + \mu^+ \to 2\gamma$$

but this would need an unimaginably large muon density ($n > 10^{24}$ $\mu^\pm$ cm^{-3}). (The γ radiation would be converted into ν-$\bar{\nu}$ pairs even in this case [8].)

§ 3

If the density of neutrinos is comparable to that of the stellar matter, the gravitational effect of the neutrinos may play an important role in the formation of the cosmologic structure of the Universe. On the other hand,

once this structure is established, it offers a new possibility for the estimation of the value of the astronomically unobservable mass density of the neutrinos. Accepting the cosmological model of general relativity, and taking into account the HUBBLE constant and the lowest limit of the age of star systems, we are led to the value [2]

$$\bar{\varrho}(\nu) < 10^{-28}\ \mathrm{gcm}^{-3}\,.$$

Independent of the uncertain validity of the homogeneous model the neutrino density may appear in connection with the stability of cluster of galaxies. A cluster of galaxies of mass M, by means of its gravitational field, may give rise to an inhomogeneity in the otherwise homogeneous neutrino sea. In a region of the space with gravitational potential $\Phi(r)$ a neutrino with momentum p has the energy

$$\varepsilon = cp\left[1 + \frac{\Phi(r) - \Phi(\infty)}{c^2}\right].$$

Therefore the density of completely degenerate neutrino gas (this assumption is made for the sake of simplicity) is modified by the gravitational potential to

$$\varrho_\nu(r) = \varrho_\nu(\infty)\left[1 + \frac{\Phi(r) - \Phi(\infty)}{c^2}\right]^{-3}.$$

Since the gravitational potential of the cluster with mass M at a large distance varies according to

$$\Phi_*(r) - \Phi_*(\infty) \sim -\frac{kM}{r} \ll c^2\,,$$

(where k is the gravitational constant) we obtain

$$\varrho_\nu(r) - \varrho_\nu(\infty) \sim \frac{3kM}{c^2}\,\frac{\varrho_\nu(\infty)}{r}\,. \tag{4}$$

Therefore the inhomogeneity contributes to the gravitational potential at the center of the cluster by

$$\delta\Phi(0) \sim -k\int_0^{D/2} \frac{\varrho_\nu(r) - \varrho_\nu(\infty)}{r}\,dV \sim \frac{6\pi k^2}{c^2}\,MD\,\varrho_\nu(\infty)\,.$$

If therefore there is only one cluster with mass M in the space and the diameter D of the region occupied by the neutrinos tends to infinity, $\delta\Phi(0)$ diverges. We may use, however, the mean distance of the clusters of galaxies instead of D. On doing this we conclude that the contribution of the neutrino inhomo-

geneity to the potential becomes comparable to the potential of the cluster itself, if

$$\varrho_\nu(\infty) \sim \frac{c^2}{4\pi k DR} \sim 10^{-24}\,\mathrm{g\,cm^{-3}}.$$

According to stellar statistical observations some clusters are stable in spite of the fact that an instability would follow from their dynamical data (density and velocity values) [15]. Supposing $\delta\Phi(0) \sim 10\,\Phi_*(0)$, the induced inhomogeneity of the neutrino sea can explain the stabilization. One has to stress, however, that this requires a neutrino density which is a million times as high as the observed density of astronomical matter. According to the opinion of some researchers such a high density cannot be excluded in principle. The accelerating role of the inhomogeneity in the neutrino sea at a gravitational collapse would be also worth investigating.

§ 4

From what has been said in the foregoing paragraphs, it is evident that the empirical determination of the value and the character of the astronomical neutrino density would be very important.

PONTECORVO and SMORODINSKY [2] were the first to call the attention to the fact that the background values of the experiments of REINES and COWAN, and of DAVIS give an upper limit for the density of ν and $\bar{\nu}$ in the region between 1 and 10 MeV. They state that this is smaller than $10^{-24}\,\mathrm{g\,cm^{-3}}$. One cannot, however, expect the energy of the neutrinos of astronomical origin to be so high. The mentioned data do not give any information as to the soft component of the neutrino radiation. (E.g. provided the energy spectrum is a thermal one, the measurements referred to above would allow a neutrino temperature of $20 \cdot 10^6$ °K, which leads to an integral neutrino density $\varrho(\nu) \sim 91\,\mathrm{g\,cm^{-3}}$.) For the observation of the soft component WEINBERG [17] proposed the investigation of the β-spectrum of H^3; the occupied $\bar{\nu}$ states (because of the Pauli principle) slow down, whereas the occupied ν states (via the forced $\nu + H^3 \to He^3 + e^-$ decay) precipitate the β-decay and deform its spectrum:

$$f(E) = CF_Z(E)\,E\sqrt{E^2-1}\,(E_0-E)^2\,[1 - N_{\bar{\nu}}(E_0-E)], \text{ if } E < E_0,$$

$$f(E) = CF_Z(E)\,E\sqrt{E^2-1}\,(E_0-E)^2\,N_\nu(E-E), \text{ if } E > E_0.$$

(Here E is the energy of the electron, E_0 its maximal energy in m_0c^2 units, N_ν and $N_{\bar{\nu}}$ represent the occupation probability of the ν and $\bar{\nu}$ states, and

F_Z is the Coulomb factor.) Comparison with the observed β-spectrum of tritium leads to a neutrino temperature of $T < 10^6$ °K and a density of the soft component (in the KeV region) of $\varrho(\nu) + \varrho(\bar{\nu}) < 10^{-9}$ g cm^{-3}. Though this value lies very far from the astronomically interesting region, the measurement of the β-spectrum offers the best possibility for the detection of the interstellar neutrinos and antineutrinos in the region under 1 MeV (Fig. 1).

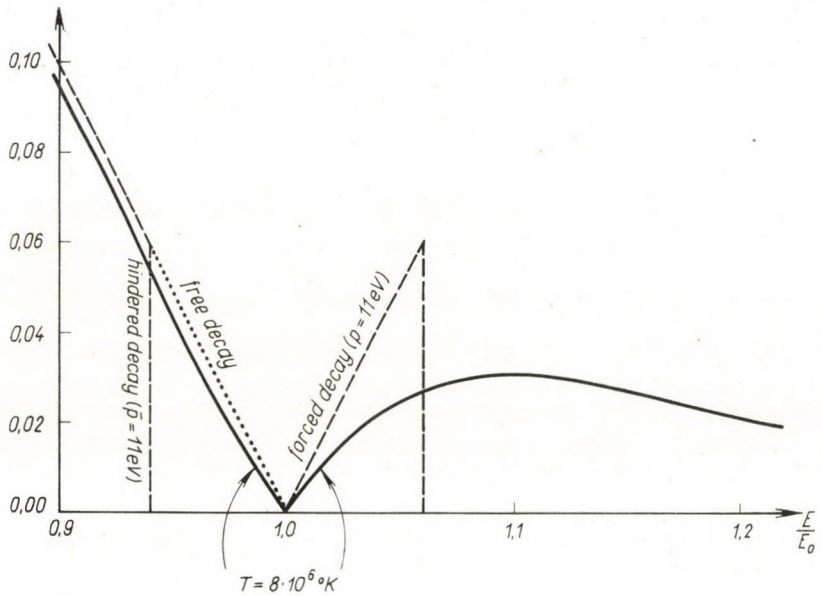


Fig. 1

Measurement of the β-spectrum is not suitable for the observation of the μ neutrinos. Here the experimental possibilities are much more restricted. From the reaction

$$\bar{\nu}_\mu + p^+ \rightarrow n + e^+$$

one obtains for the density of the $\bar{\nu}_\mu$-s of GeV energy the value $\varrho(\bar{\nu}_\mu) < < 10^{-28}$ gcm^{-3} [18]. Such high energies, however, cannot occur with astronomical abundance even in "hot" matter. In the case, when $\varepsilon(\nu_\mu) < m_\mu c^2$, the only possibility for the direct detection is given by the decay spectrum of the muon [16]. The occupation of the ν_μ states accelerates the μ^+ decay (via the reaction $\nu_\mu + \mu^+ \rightarrow e^+ + \nu$), whereas the occupation of the $\bar{\nu}_\mu$ states slows it down (because of the Pauli principle). The energy spectrum of the μ decay (Fig. 2 and enlarged parts in Fig. 3) is represented by

$$f(E) = 192 \int_{1/2-E}^{1/2} \left(-\frac{1}{2} + E + F\right)(1 - E - F)\left(1 - N_{\bar{\nu}_\mu}(F)\right) dF, \text{ if } E < \frac{1}{2},$$

$$f(E) = 192 \int_{E+1/2}^{\infty} \left(\frac{1}{2} - E + F\right)(1 - E + F)\, N_{\nu_\mu}(F)\, dF\,, \quad \text{if} \quad E > \frac{1}{2}\,. \qquad (5)$$

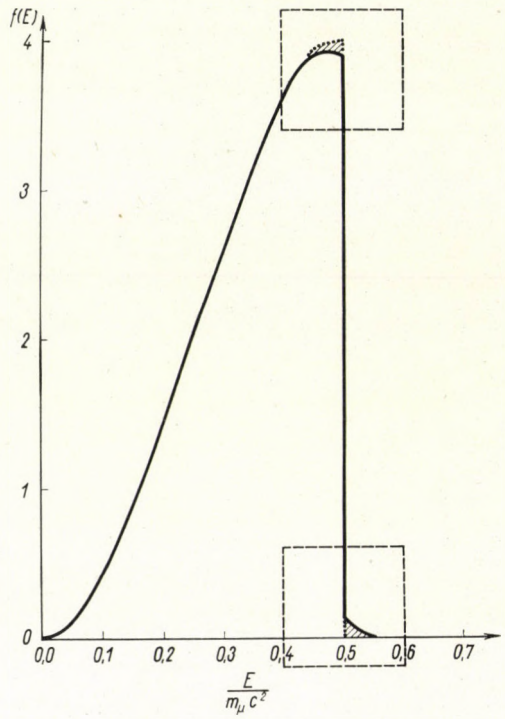


Fig. 2

Here E is the energy of the emitted positron, F is the energy of the muon-neutrino in $m_\mu c^2$ units. In principle, eq. (5) makes possible to determine the ν_μ and $\bar{\nu}_\mu$ background, but this method is very inaccurate, the upper limit for the density of the low energy muon-neutrinos obtained from the present experimental datas being $\varrho(\nu_\mu) + \varrho(\bar{\nu}_\mu) < 10^6$ g cm^{-3}. One does not obtain a more accurate value by comparing the measured and calculated values of the lifetime of μ [16] either.

Taking into account that the detection methods are far more inaccurate in the case of the μ neutrinos than in the case of β neutrinos, it is evident that research should be concentrated on the latter. The "everyday" ν radiation of the stars can well be estimated (§ 1), any observed departure from this estimated value would point to the existence of "hot matter". As we have seen in § 2 one cannot imagine such without ν emission. As a matter of fact, in the case of ν measurements the Sun, in the case of $\bar{\nu}$ measurements the Earth,

causes unpleasant background effects [1]. (The ν_μ measurement is not disturbed by the atmospheric background.) In this respect the hardness ($\varepsilon(\nu) \gg 1$ MeV) of the ν radiation emitted by the "hot matter" could help.

In order to estimate optimistically the practical possibilities, let us consider e.g. the center of our Galaxy. At our Earth an intensity $I(\nu) \sim 10^{10}$ ν cm^{-2} s^{-1} can be said to be observable with the help of the present techniques for neutrinos of several MeV energy. (For neutrinos of an energy about GeV

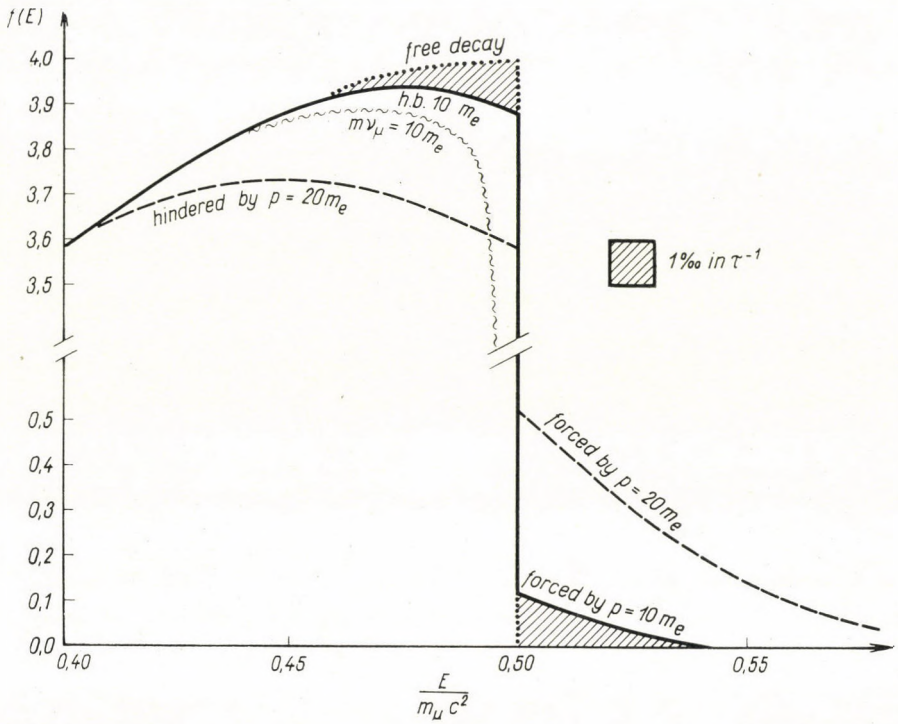


Fig. 3

the absorption cross section is a million times as high as for neutrinos of energy about MeV.) This means that the neutrino radiation of a hyperstar at the center of our Galaxy could be observed, if its neutrino luminosity would be about $10^{48} - 10^{50}$ erg s^{-1}, which could be supplied already by a specific energy liberation $\sim 10^{10}$ erg g^{-1} s^{-1}. Such values certainly occur in certain periods of the evolution of a hyperstar. (This needs a mean temperature $1 - 2 \cdot 10^9$ °K inside the hyperstar.)

The neutrino pulse emitted by a hypernova explosion could be detected at a distance of a 100 million light years, if the neutrino luminosity of the hypernova reached the value 10^{55} erg s^{-1}. This is not unimaginable at all, since we know that the total energy set free in such an explosion may exceed

the value of 10^{60} erg [20]. The question is only, whether the mean free path of the β neutrinos is not very small compared with the dimensions of the hyperstar. As CHIU has suggested recently [21], under special conditions the mean energy loss of a dense hyperstar can be caused by the muon-neutrino luminosity. In this case the possibilities of observation would be much more unfavourable, as we have shown just above. In spite of all these problems, it can not be excluded, that the observation of neutrino radiation will give a direct proof of the hypothesis of hypernova explosion and gravitational collapse.

Appendix
Neutrino production in the Hayashi—Cameron model

The giant star considered in the HAYASHI—CAMERON model remains in the main sequence of the HERTZSPRUNG—RUSSEL diagram during its first period of evolution, for $1{,}56 \cdot 10^8$ years. The temperature in its central region reaches only $5{,}85 \cdot 10^7$ °K, therefore in this period exclusively the H $\rightarrow$ He transmutation produces neutrinos. Meanwhile the hydrogen content of the convective core of the star decreases from the initial 90% to 62%. This means that $7{,}3 \cdot 10^{33}$ g hydrogen is fusioned into helium, i.e. about $N = 2{,}2 \cdot 10^{57}$ neutrinos are produced. A neutrino obtains an energy of $0{,}4 \cdot 10^{-6}$ erg in the $p - p$ fusion and $1{,}3 \cdot 10^{-6}$ erg in the C—N cycle. Since in the case of a giant star this latter process plays an overwhelming role, from the produced energy of $4{,}4 \cdot 10^{52}$ erg an amount of $E(\nu) = 2{,}85 \cdot 10^{51}$ erg is carried away by the neutrino radiation.

When the star has exhausted its central H content the gravitational contraction begins. When the central temperature reaches $1{,}5 \cdot 10^8$ °K, the reaction 3 He $\rightarrow$ C starts, which does not produce any neutrinos. Therefore one needs to take into account only the neutrino production of the H shell which surrounds the He shell. During $3{,}37 \cdot 10^5$ years about $0{,}027\ M_\odot = 8{,}4 \cdot 10^{32}$ g hydrogen has been transformed into helium exclusively in the C—N cycle (since the temperature is higher than $1{,}6 \cdot 10^7$ °K in the whole zone), therefore $E(\nu) = 3{,}25 \cdot 10^{50}$ erg. If the temperature at the centre is $3{,}16 \cdot 10^8$ °K also the thermal neutrino radiations starts, by means of the Compton neutrino production:

$$\gamma + e^- \rightarrow e^- + \nu + \bar{\nu}\,. \tag{6}$$

Taking into account the temperature distribution given by different models, and the cross section of the process (6) given in [6], we obtain the estimate

$$6{,}4 \cdot 10^{46}\ \text{erg} < E(\nu, \bar{\nu}) < 3 \cdot 10^{48}\ \text{erg}.$$

This is scarcely 5% of the energy supplied by the thermonuclear β neutrinos.

In the central zone the exhaustion of He again causes gravitational contraction. With the increasing temperature a great variety of nuclear reactions starts, which makes the estimation of the neutrino emission rather complicated. In this period, which lasts about $5 \cdot 10^4$ years, the luminosity of the star — supposing a central zone containing only pure C^{12} or Ne^{20} — can be characterized by

$$L_{\text{opt}} = 3.5 \cdot 10^{28} \text{ erg s}^{-1}, \quad \text{or} \quad 4.6 \cdot 10^{38} \text{ erg s}^{-1},$$

$$L_{\nu} = 1.02 \cdot 10^{38} \text{ erg s}^{-1}, \quad \text{or} \quad 2.5 \cdot 10^{39} \text{ erg s}^{-1}.$$

Thus the neutrino emission reaches the order of magnitude of the optical radiation, and is about equally composed of the processes (1) and (6). The emitted neutrino energy during the whole period is

$$E_s(\nu) \sim E_{\text{th}}(\nu, \bar{\nu}) \sim 10^{51} \text{ erg},$$

which is nearly $^1/_3$ of the optical energy.

Between $8 \cdot 10^8$ °K and 10^9 °K there is no considerable liberation of nuclear energy, the thermal radiation of the star must be supplied by the gravitational contraction. The evolution is extraordinarily quick, it scarcely lasts 200 years. During this time the thermonuclear β-decays do not produce neutrinos at a rate which could be compared to that produced during the former periods, $E_{\beta}(\nu) \approx 0$. On the other hand, because of (6)

$$E_{\text{th}}(\nu, \bar{\nu}) = 3.5 \cdot 10^{46} \text{ erg}.$$

(If there would be a nuclear energy source, which could produce the luminosity given above of the star for a longer time, the period of the evolution considered could be taken to have lasted longer and then $E_{\text{th}}(\nu, \bar{\nu})$ would have a larger value.

At central temperature if 10^9 °K the Ne →Fe transition series is initiated. Calculating with a neon core of mass $M = 0.114\ M_{\odot} = 3.42 \cdot 10^{33}$ g, then Ne → Fe reaction chain represents the liberation of $2.4 \cdot 10^{51}$ erg energy. The optical luminosity of the star core is smaller by an order of magnitude than its neutrino luminosity, therefore the process (6) and the reactions

$$\gamma + Z \rightarrow Z + e^- + e^+, \qquad e^- + e^+ \rightarrow \nu + \bar{\nu}$$

are the main forms of the energy emission. Thus in this period

$$E_{\text{th}}(\nu, \bar{\nu}) = 2.4 \cdot 10^{51} \text{ erg}.$$

During the formation of the Fe-core of mass 0.11 $M_\odot$ — 0.15 $M_\odot$ which is the final period of the stellar evolution the neutron ratio increases to more than 50%, therefore the thermonuclear reactions are accompained again by β-decays. The number of the neutrinos produced in these decays can be neglected as compared to that of the other processes. The core is heated from 10^9 °K to $5 \cdot 10^9$ °K, the central density increases from 10^5 g cm^{-3} to $125 \cdot 10^5$ g cm^{-3} and the internal energy increases from $6.1 \cdot 10^{16}$ erg g^{-1} to $30.5 \cdot 10^{16}$ erg g^{-1}. This corresponds to a total energy increase of $1.2 \cdot 10^{51}$ erg. Because of the virial theorem this is just equal to the emitted energy of which only a few percent is optical radiation. Therefore

$$E_{\text{th}}(\nu, \bar{\nu}) = 1.2 \cdot 10^{51} \text{ erg}.$$

At this stage the supernova explosion may occur, during which also heavy elements (and the neutrons) are produced. These emit $\bar{\nu}$-radiation. But if we suppose that the concentration of the heavy elements inside a star, the mass of which is several times that of the sun, is similar to the terrestrial one, the rate of production of the elements heavier than Fe is so slow that $E_\beta(\nu)$, due to their production, and $E_\beta(\bar{\nu})$, due to their subsequent radioactive decay (e.g. natural terrestrial radioactivity), are very small:

$$E_\beta(\nu) < 10^{49} \text{ erg}, \qquad E_\beta(\bar{\nu}) < 10^{48} \text{ erg}.$$

The values obtained here are summarized in the Table at the end of § 4.

Note added in proof: The most accurate upper limit of the cosmic neutrino flux in the MeV region is given by the recent measurement of RAYMOND DAVIS (BNL-Preprint No. 7660): $\varrho(\nu) < 10^{-28}$ g cm^{-3}.

REFERENCES

1. G. MARX and N. MENYHÁRD, Mitteilungen der Sternwarte Budapest—Szabadsághegy, No. 48 (1960); Sciences, **131,** 299, 1960.
2. B. PONTECORVO and YA. SMORODINSKI, JETP, **41,** 239, 1961.
3. G. GAMOW and M. SCHÖNBERG, Phys. Rev., **59,** 539, 1941.
4. B. PONTECORVO, JETP, **36,** 1615, 1959.
5. H. Y. CHIU, Annales of Physics, **15,** 1, 1961; Phys. Rev., **122,** 1317, 1961.
6. CH. HAYASHY and R. C. CAMERON, Astrophys. J., **136,** 166, 1961.
7. G. MARX and J. NÉMETH, Proc. XI. Int. Conf. on High Energy Physics Geneva, 1962.
8. G. MARX and J. NÉMETH, Mitteilungen der Sternwarte Budapest—Szabadsághegy, No. 52. 1962.
9. G. MARX and T. NAGY, Acta Phys. Hung., **16,** 141, 1963.
10. L. FODOR, Mitteilungen der Sternwarte Budapest—Szabadsághegy, to be printed.
11. V. A. AMBARTSUMIAN and G. S. SAAKYAN, Astron. J. USSR, **37,** 193, 1960.
12. V. A. AMBARTSUMIAN, e.g. Proc. Solvei Conf., Bruxelles 1958, p. 241.
13. F. HOYLE and W. FOWLER, Nature, **197,** 533, 1963,
14. R. A. ALPHER and R. C. HERMAN, Rev. Mod. Phys., **22,** 153, 1950.

15. G. DE VANCOULEURS, Astrophys. J., **66,** 629, 1961.
16. G. MARX, Nuovo Cim., **30**, 1555, 1963.
17. S. WEINBERG, preprint, 1962.
18. B. PONTECORVO and A. E. CHUDAKOV, Proc. XI. Conf. on High Energy Physics 1962, p. 415;
M. SCHWARTZ et al., ibid. p. 809;
S. MIYAKE et al, Kyoto, Proc., 318, 1962.
19. F. HOYLE and W. A. FOWLER, Monthly Notices RAS., **125,** 169, 1963.
20. M. SCHMIDT, Nature, **197,** 1040, 1963.
21. H. Y. CHIU, preprint, 1963.

КОСМИЧЕСКАЯ ЧАСТОТА НЕЙТРИНО И КОСМОГОНИЯ

Л. ФОДОР, Ж. КОВЕШИ и Г. МАРКС

Р е з ю м е

Оценивается плотность космических нейтрино как продуктов термоядерных реакций звёзд и супернов. Наблюдение большей плотности нейтрино указывало бы на сигнулярное (дозвездное) состояние материи. Подробно истолкуются гравитационное влияние нейтринного фона и непосредственные возможности наблюдения мягкой компоненты нейтринного излучения двух видов.

DETERMINATION OF NUCLEAR MATRIX ELEMENTS FROM THE MEASUREMENT OF ε/β^+

By

I. MONTVAY

INSTITUTE FOR THEORETICAL PHYSICS, ROLAND EÖTVÖS UNIVERSITY, BUDAPEST

(Presented by Z. Gyulai. — Received 20. IX. 1963)

The calculation of transition probabilities of β-decay in the case of forbidden transitions — mostly of higher forbidden ones — will be simpler, if we use matrix elements, containing only the radial part of nuclear wave functions, instead of the traditional reduced matrix elements. In this way also the ratio ε/β^+ can be calculated in all the cases; respectively the measurement of the ratio ε/β^+ can be used for the determination of nuclear matrix elements.

1. Introduction

The ratio of electron capture to positron emission can contain valuable information concerning the nuclei which take part in the decay. It is thus important to know the prediction of the theory for the value of the ratio ε/β^+. In the case of allowed transitions this ratio is independent of the nuclear matrix elements and therefore the calculation does not encounter difficulties. The situation is completely different in the case of forbidden transitions. It was pointed out for example by P. DEPOMMIER and R. BOUCHEZ [8] that the ratio is independent of matrix elements also for unique-forbidden transitions. According to the point of view of H. BRYSK and M. E. ROSE [7] it is not possible to give an exact value for ε/β^+ in the case of higher, non-unique forbidden transitions, because of the great number of unknown matrix elements. These authors have found that in general the ratio ε/β^+ must increase for forbidden transitions in comparison with the allowed value. Here too, and in other papers (see for example [9]) generally, there are exact calculations of the ratio ε/β^+ for allowed and unique-forbidden cases, reliable estimations for first non-unique forbidden transitions, and only approximate statements for higher forbidden ones. In this case the great number of reduced matrix elements gives rise to the difficulties.

In the nucleus the nucleons are in energy, parity and angular momentum eigenstate. Consequently, the only unknowns in the wave functions are the radial parts of small and large components, respectively, since we do not know the exact shape of the potential valley, representing the nucleus. As only the integrals of wave functions appear in the transition probabilities, the mean

life of a nucleus depends on the following four integrals:

$$
\begin{aligned}
R_1 &= R(vk) = \int_0^R dr\, r^2\, M_n^{j_v l_v *}(r)\, M_p^{j_k l_k}(r)\,, \\
R_2 &= R(vk') = \int_0^R dr\, r^2\, M_n^{j_v l_v *}(r)\, N_p^{j_k l_k}(r)\,, \\
R_3 &= R(v'k) = \int_0^R dr\, r^2\, N_n^{j_v l_v *}(r)\, M_p^{j_k l_k}(r)\,, \\
R_4 &= R(v'k') = \int_0^R dr\, r^2\, N_n^{j_v l_v *}(r)\, N_p^{j_k l_k}(r)\,.
\end{aligned}
\tag{1}
$$

Here $M_p^{j_k l_k}(r)$ and $N_p^{j_k l_k}(r)$ are the radial parts of the wave function of the proton in the initial state. Similarly, the radial functions $M_n^{j_v l_v}(r)$ and $N_n^{j_v l_v}(r)$ belong to the neutron in the final state. (We shall deal in more detail with the wave functions of particles, taking part in the decay, in the following section.) The quantities R_k are real, if we choose the phase factors of eigenvectors in the usual way. The upper limit of the integration R is the radius of the nucleus. We note that the transition probability does not depend on integrals like

$$
\int_0^R dr\, r^2\, M_n^{j_v l_v *}(r)\, M_p^{j_k l_k}(r)\, f(r)\,.
$$

If we calculate the transition probability of β-decay — in allowed and also in forbidden cases — we must compute the coefficients of the quantities $R(vk)$, $R(vk')$, $R(v'k)$, $R(v'k')$. In the following sections we shall deal first with β transitions in general, later with the help of the results obtained we calculate the value of the ratio ε/β^+ in a case of arbitrary forbiddenness. Finally, we examine the possibility of using the measured value of the ratio ε/β^+ for the determination of the radial matrix elements R_k.

2. Transitions between angular momentum and parity eigenstates

Let us deal first with processes, when besides the nucleons also the leptons are in angular momentum, parity (and naturally) energy eigenstates. For example the wave function of an electron, bound in the Coulomb field and having an energy ε[1] and parity $(-1)^{l_2}$, an absolute value of angular momentum j_2, with the z-component M_2 [1] has the form

$$
f_{n_2}^{j_2 l_2 M_2}(\mathbf{r}) = \begin{bmatrix} i g_{n_2}^{\varkappa_2}(r)\, \Omega_{j_2 l_2 M_2} \\ -f_{n_2}^{\varkappa_2}(r)\, \Omega_{j_2 l_2' M_2} \end{bmatrix}, \tag{2}
$$

[1] We use units, where $\hbar = c = m_e = 1$ (m_e = electron mass).

where

$$\Omega_{jlM}(\hat{r}) = \begin{bmatrix} (lm\,\frac{1}{2} - \frac{1}{2} \mid jM)\, Y_{l,M-\frac{1}{2}}(\hat{r}) \\ (lm\,\frac{1}{2} - \frac{1}{2} \mid jM)\, Y_{l,M+\frac{1}{2}}(\hat{r}) \end{bmatrix}; \tag{3}$$

and henceforth always $l' = 2j - l$ as well as

$$g^{\varkappa_2}_{n_2}(r) = -\frac{\sqrt{\Gamma(2\gamma_2 + n' + 1)}}{\Gamma(2\gamma_2 + 1)\sqrt{n'!}} \sqrt{\frac{1+\varepsilon}{4N(N-\varkappa_2)}} \left(\frac{2Z}{Na_0}\right)^{3/2} e^{-\frac{Zr}{Na_0}} \left(\frac{2Zr}{Na_0}\right)^{\gamma_2 - 1} \times$$

$$\times \left[-n' F\left(-n'+1, 2\gamma_2+1; \frac{2Zr}{Na_0}\right) + (N-\varkappa_2) F\left(-n', 2\gamma_2+1; \frac{2Zr}{Na_0}\right), \right. \tag{4}$$

$$f^{\varkappa_2}_{n_2}(r) = -\frac{\sqrt{\Gamma(2\gamma_2 + n' + 1)}}{\Gamma(2\gamma_2 + 1)\sqrt{n'!}} \sqrt{\frac{1-\varepsilon}{4N(N-\varkappa_2)}} \left(\frac{2Z}{Na_0}\right)^{3/2} e^{\frac{-Zr}{Na_0}} \left(\frac{2Zr}{Na_0}\right)^{\gamma_2 - 1} \times$$

$$\times \left[n' F\left(-n'+1, 2\gamma_2+1; \frac{2Zr}{Na_0}\right) + (N-\varkappa_2) F\left(-n', 2\gamma_2+1; \frac{2Zr}{Na_0}\right) \right].$$

The possible values of the energy

$$\varepsilon = \frac{1}{\sqrt{1 + \frac{(Z\alpha)^2}{(n_2 - |\varkappa_2| + \sqrt{\varkappa_2^2 - (Z\alpha)^2})^2}}},$$

and

$$\varkappa_2 = \begin{cases} -l_2 - 1 & \text{if} \quad j_2 = l_2 + \frac{1}{2} \\ l_2 & \text{if} \quad j_2 = l_2 - \frac{1}{2} \end{cases} \tag{5a}$$

$$n_2 = 1, 2, 3, \ldots \qquad l_2 = 0, 1, \ldots, n_2 - 1$$

$$j_2 = \frac{1}{2}, \frac{3}{2}, \ldots, n_2 - \frac{1}{2} \qquad M_2 = -j_2, -j_2+1, \ldots, j_2.$$

The wave functions of states with continuous energy are:

$$f^{j_2 l_2}_{\varepsilon}(\mathbf{r}) = \begin{bmatrix} i g^{j_2 l_2}_{\varepsilon}(r)\, \Omega_{j_2 l_2 M_2} \\ -f^{j_2 l_2}_{\varepsilon}(r)\, \Omega_{j_2 l_2' M_2} \end{bmatrix}, \tag{6}$$

where

$$g_{\varepsilon}^{j_2 l_2}(r) = -\frac{ie^{\nu\frac{\pi}{2}}|\Gamma(\gamma_2+i\nu)|}{r\,\Gamma(2\gamma_2+1)}\sqrt{\frac{1}{p\pi}}\sqrt{1-\varepsilon}\,(2pr)^{\gamma_2}\times$$

$$\times 2\,Re\left[e^{-ipr+i\xi}(\gamma_2+i\nu)\,F(\gamma_2+1+i\nu, 2\gamma_2+1; 2ipr\right], \tag{7}$$

$$f_{\varepsilon}^{j_2 l_2}(r) = \frac{ie^{\nu\frac{\pi}{2}}|\Gamma(\gamma_2+i\nu)|}{r\,\Gamma(2\gamma_2+1)}\sqrt{\frac{1}{p\pi}}\sqrt{1+\varepsilon}\,(2pr)^{\gamma_2}\times$$

$$\times 2i\,Im\left[e^{-ipr+i\xi}(\gamma_2+i\nu)\,F(\gamma_2+1+i\nu, 2\gamma_2+1; 2ipr)\right].$$

We used the notations

$$\gamma_2 = \sqrt{\varkappa_2^2-(\alpha Z)^2}\,; \qquad n' = n_2 - |\varkappa_2|\,;$$

$$N = \sqrt{n_2^2 - 2n'\left(|\varkappa_2| - \sqrt{\varkappa_2^2-(\alpha Z)^2}\right)}; \qquad \mu = \frac{Z\alpha}{\sqrt{\varepsilon^2-1}}\,; \tag{7a}$$

$$\nu = -Z\alpha\frac{\varepsilon}{\sqrt{\varepsilon^2-1}}\,; \qquad e^{2i\xi} = \frac{\varkappa_2 - i\mu}{\gamma_2+i\nu}\,.$$

It is very advantageous for our calculations, that (2) and (6) are of entirely the same form, moreover the wave functions of the neutrino and the nucleons can be brought to this form too:

$$f_q^{j_1 l_1 M_1}(\mathbf{r}) = (2\pi)^{-3/2}\,q\,\frac{1}{\sqrt{2}}\begin{bmatrix} g_{l_1}(qr)\,\Omega_{j_1 l_1 M_1} \\ -g_{l_1'}(qr)\,\Omega_{j_1 l_1' M_1}\end{bmatrix}, \tag{8}$$

where

$$g_l(x) = (2\pi)^{3/2}\,i^l\,J_{l+\frac{1}{2}}(x)\cdot x^{-\frac{1}{2}}, \tag{9}$$

and $E_\nu = q$ is the energy of the neutrino. The wave functions of the nucleons are the following:

$$F_n^{j_v l_v M_v}(\mathbf{r}) = \begin{bmatrix} M_n^{j_v l_v}(r)\,\Omega_{j_v l_v M_v} \\ N_n^{j_v l_v}(r)\,\Omega_{j_v l_v' M_v}\end{bmatrix};$$

$$F_p^{j_k l_k M_k}(\mathbf{r}) = \begin{bmatrix} M_p^{j_k l_k}(r)\,\Omega_{j_k l_k M_k} \\ N_p^{j_k l_k}(r)\,\Omega_{j_k l_k' M_k}\end{bmatrix}. \tag{10}$$

Here j_1, M_1, $(-1)^{l_1}$; j_ν, M_ν, $(-1)^{l_\nu}$; j_k, M_k, $(-1)^{l_k}$ are the absolute value and z-component of the angular momentum and the parity of the neutrino, neutron and proton, respectively.

The transition probability per unit time is:

$$w_{ba} = 2\pi\,\delta(E)\,|M_{ba}|^2, \tag{11}$$

and if we use (6), (8) and (10) we get

$$M_{ba} = \frac{G}{\sqrt{2}}\int (F_n^{j_\nu l_\nu M_\nu +}(\mathbf{r})\,\gamma_4\,\gamma_\alpha\,(1+\lambda\gamma_5)\,F_p^{j_k l_k M_k}(\mathbf{r}))\,(f_q^{j_1 l_1 M_1 +}(\mathbf{r})\,\gamma_4\,\gamma_\alpha\,\times$$

$$\times\,(1+\gamma_5) f_\varepsilon^{j_2 l_2 M_2}(\mathbf{r}))\,dv\,. \tag{12}$$

Let us introduce the following real quantities:

$$\begin{aligned} i^{-l_1} g^*_{l_1}(R) &= G(1); \qquad i^{-l_1-1} g^*_{l'_1}(R) = F(1');\\ g^{\varkappa_2}_{\varepsilon}(R) &= g(2); \qquad f^{\varkappa_2}_{\varepsilon}(R) = f(2'). \end{aligned} \tag{13}$$

Using the fact that r^2 times the radial part of the lepton wave functions remains inside the comparatively small nucleus unaltered, and further taking into account the relation

$$\gamma_4\,\gamma_\alpha\,(1+\lambda\gamma_5) = \sigma_\alpha \begin{bmatrix} \lambda-1 & \\ -1 & \lambda \end{bmatrix} + \delta_{\alpha 4}\,\sigma_\alpha\,(1-\lambda)\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \tag{14a}$$

where

$$\sigma = (i\sigma_1,\, i\sigma_2,\, i\sigma_3,\, I);$$

we get from (12) and (1) the form of M_{ba}.

Introducing the simple notations[2]

$$\begin{Bmatrix} f_1(\mathbf{x}) \\ f_2(\mathrm{y}) = f_1(x) f_2(y) f(xy) \\ f(\,) \end{Bmatrix}$$

and

$$\begin{Bmatrix} a_1(\mathbf{x}) \\ a_2(\mathrm{y}) \\ a\ (\,) \end{Bmatrix} \times \begin{Bmatrix} b_1(\mathbf{x}) \\ b_2(\mathrm{y}) \\ b\ (\,) \end{Bmatrix} = \begin{Bmatrix} a_1(\mathbf{x})\,|\,b_1(\mathbf{x}) \\ a_2(\mathrm{y})\,|\,b_2(\mathrm{y}) \\ a\ (\,)\,|\,b\ (\,) \end{Bmatrix} = \begin{Bmatrix} b_1(\mathbf{x})\,|\,a_1(\mathbf{x}) \\ b_2(\mathrm{y})\,|\,a_2(\mathrm{y}) \\ b\ (\,)\,|\,a\ (\,) \end{Bmatrix}$$

[2] The properties of these "quantities" (the calculation rules) can be easily seen from the definition.

we can write

$$M_{ba} = \frac{G_q}{2(2\pi)^{3/2}} i^{l_1+1} \Bigg[\begin{cases} R(\mathbf{vk}) + R(\mathbf{v'k'}) \\ (G(\mathbf{1}) + iF(\mathbf{1'}))\,(g(\mathbf{2}) - if(\mathbf{2'})) + \\ \lambda <> + (1-\lambda) <>_4 \end{cases} \tag{14b}$$

$$+ \begin{cases} - R(\mathbf{vk'}) - R(\mathbf{v'k}) \\ (G(\mathbf{1}) + iF(\mathbf{1'}))\,(g(\mathbf{2}) - if(\mathbf{2'})) \\ - i <> + i(1-\lambda) <>_4 \end{cases} \Bigg];$$

thus

$$|M_{ba}|^2 = \frac{G^2 q^2}{4(2\pi)^3} \Bigg[\begin{cases} R(\mathbf{vk}) + R(\mathbf{v'k'}) \,|\, R(\mathbf{vk}) + R(\mathbf{v'k'}) \\ [G(\mathbf{1})\,g(\mathbf{2}) + F(\mathbf{1'})f(\mathbf{2'}) \,|\, G(\mathbf{1})\,g(\mathbf{2}) + F(\mathbf{1'})f(\mathbf{2'})] + \\ \lambda <> + (1-\lambda) <>_4 \,|\, \lambda <> + (1-\lambda) <>_4 \end{cases}$$

$$+ [F(\mathbf{1'})\,g(\mathbf{2}) - G(\mathbf{1})f(\mathbf{2'}) \,|\, F(\mathbf{1'})\,g(\mathbf{2}) - G(\mathbf{1})f(\mathbf{2'})] +$$

$$+ \begin{cases} R(\mathbf{vk'}) + R(\mathbf{v'k}) \,|\, R(\mathbf{vk'}) + R(\mathbf{v'k}) \\ [G(\mathbf{1})\,g(\mathbf{2}) + F(\mathbf{1'})f(\mathbf{2'}) \,|\, G(\mathbf{1})\,g(\mathbf{2}) + F(\mathbf{1'})f(\mathbf{2'})] + \\ + [F(\mathbf{1'})\,g(\mathbf{2}) - G(\mathbf{1})f(\mathbf{2'}) \,|\, F(\mathbf{1'})\,g(\mathbf{2}) - G(\mathbf{1})f(\mathbf{2})] \\ - <> + (1-\lambda) <>_4 \,|\, - <> + (1-\lambda) <>_4 \end{cases} - \tag{14}$$

$$- \begin{cases} R(\mathbf{vk}) + R(\mathbf{v'k'}) \,|\, R(\mathbf{vk'}) + R(\mathbf{v'k}) \\ 2\{[F(\mathbf{1'})\,g(\mathbf{2}) - G(\mathbf{1})f(\mathbf{2'}) \,|\, G(\mathbf{1})\,g(\mathbf{2}) + F(\mathbf{1'})f(\mathbf{2'})] - \\ - [G(\mathbf{1})\,g(\mathbf{2}) + F(\mathbf{1'})f(\mathbf{2'}) \,|\, F(\mathbf{1'})\,g(\mathbf{2}) - G(\mathbf{1})f(\mathbf{2'})]\} \\ \lambda <> + (1-\lambda) <>_4 \,|\, - <> + (1-\lambda) <>_4 \end{cases} \Bigg].$$

Here we have introduced for the angular integrals the notations:[3]

$$<vk\,12> = \int d\hat{r}\,(\Omega_v^+\,\sigma_\alpha\,\Omega_k)\,(\Omega_1^+\,\sigma_\alpha\,\Omega_2) =$$

$$= (2\pi)^{-1} \sum_{a' j_3 \varkappa} (j_3)^{-1} T_{j_3}(vk\,12|a')\,(j_v\,M_v\,j_k - M_k\,|\,j_3 - \varkappa) \times$$

$$\times (j_1\,M_1\,j_2 - M_2 | j_3 \varkappa)(-)^{M_1+M_k}, \tag{15}$$

where

$$\Omega_v = \Omega_{j_v l_v M_v} \quad \text{and so on.}$$

In (15) we have used also the following notations:

$$(j) = (2j+1)^{1/2}$$

as well as

$$T_{j_3}(vk\,12|a') = \sum_a (j_3)\,(l_v\,0\,l_k\,0\,|\,a'\,0)\,(l_1\,0\,l_2\,0\,|a'\,0)\,W\left(j_v\,l_v\,j_1\,l_1;\frac{1}{2}\,a\right) \times$$

$$\times W(l_v\,l_k\,l_1\,l_2;\ a'\,a)\,W\left(j_k\,l_k\,j_2\,l_2;\ \frac{1}{2}\,a\right) \times \tag{16}$$

$$\times W(j_v j_k j_1 j_2; j_3\,a)\,(j_v)\,(j_k)\,(j_1)\,(j_2)\,(l_v)\,(l_k)\,(l_1)\,(l_2);$$

[3] The computation of these integrals is briefly outlined in the Appendix.

Similarly

$$<vk\,12>_4 = \int d\hat{r}\,(\Omega_v^+\,\sigma_4\,\Omega_k)\,(\Omega_1^+\,\sigma_4\,\Omega_2) =$$

$$= (2\pi)^{-1} \sum_{j_3 \varkappa} (j_3)^{-1}\, S_{j_3}(vk\,12)\,(-)^{M_1+Mk}\,(j_v M_v j_k - M_k | j_3 - \varkappa)\,(j_1\, M_1 j_2 - M_2 | j_3\, \varkappa),$$

and

$$S_{j_3}(vk\,12) = (-)^{1+j_1+j_2+j_v+j_k} \frac{1}{2(j_3)}\,(l_v\,0\,l_k\,0\,|\,j_3\,0)\,(l_1\,0\,l_2\,0\,|\,j_3\,0) \times \qquad (18)$$

$$\times\, W\left(j_1 j_2 l_1 l_2\,;\;\; j_3 \frac{1}{2}\right) W\left(j_v j_k l_v l_k;\;\; j_3 \frac{1}{2}\right) \cdot (j_v)\,(j_k)\,(j_1)\,(j_2)\,(l_v)\,(l_k)\,(l_1)\,(l_2).$$

In (14) which gives with (11) the transition probability between angular momentum and parity eigenstates, besides the integrals (15) and (17) also others occur:

$$\begin{aligned} <vk\,12'> &= \int d\hat{r}\,(\Omega_v^+\,\sigma_\alpha\,\Omega_k)\,(\Omega_1^+\,\sigma_\alpha\,\Omega_{2'}); \\ <v'\,k'\,12> &= \int d\hat{r}\,(\Omega_{v'}^+\,\sigma_\alpha\,\Omega_{k'})\,(\Omega_1^+\,\sigma_\alpha\,\Omega_2); \\ <v'\,k'\,1'\,2>_4 &= \int d\hat{r}\,(\Omega_{v'}^+\,\sigma_4\,\Omega_{k'})\,(\Omega_{1'}^+\,\sigma_4\,\Omega_2); \\ <v'\,k'\,1'\,2'>_4 &= \int d\hat{r}\,(\Omega_{v'}^+\,\sigma_4\,\Omega_{k'})(\Omega_{1'}^+\,\sigma_4\,\Omega_{2'}) \end{aligned} \qquad (18a)$$

and so on. Here for example $\Omega_{v'} = \Omega_{j_v l_v' M_v}$. For these integrals we can immediately make use of (15) and (17) taking care to use the appropriate l.

3. Selection rules

Let us examine the case when only one nucleon takes part in the decay[4] (this nucleon being outside a closed shell, which applies both to neutron and proton) and investigate which values of angular momentum and parity are possible for leptons if the angular momentum and parity of the nucleon is given in the initial and the final state. In other words for fixed j_v, l_v, j_k, l_k we look for j_1, l_1, j_2, l_2 values, for which the transition probability (1) can be different from zero.

These selection rules, expressing the conservation law of the angular momentum and parity, can be read from the form (16) and (18) of $T_{j_3}(vk12/a')$ and $S_{j_3}(vk12)$ respectively. We deal with the following cases:

$$G_r: \begin{cases} \Delta I = r \text{ thus } |j_v - j_k| = r \text{ and } |l_v - l_k| = |l_v' - l_k'| = r \\ 1 + \Delta\Pi = (-1)^r \text{ and } |l_v' - l_k| = r \pm 1 \text{ thus } |l_v - l_k'| = r \mp 1 \end{cases} \qquad (19)$$

[4] The problem of more complicated nuclei can be reduced to this case.

(the upper symbol is the G_r^+ the lower the G_r^- case).

$$M_r: \begin{cases} \Delta I = r \text{ thus } |j_v - j_k| = r \text{ and } |l'_v - l_k| = |l_v - l'_k| = r \\ 1 + \Delta\Pi = (-1)^{r+1} \text{ and } |l_v - l_k| = r \pm 1 \text{ thus } |l'_v - l'_k| = r \mp 1 \end{cases} \tag{20}$$

(the upper symbol is the M_r^+, the lower the M_r^- case).

The connection between our notation and the denomination used generally in the literature is the following:

Table 1

Notation	ΔI	$\Delta\Pi$	Denomination
G_0	0	no	allowed
M_0	0	yes	first, non-unique forbidden
G_1	1	yes	first, non-unique forbidden
M_1	1	no	allowed
G_2	2	no	second, non-unique forbidden
M_2	2	yes	first, unique forbidden
G_3	3	yes	third, non-unique forbidden
M_3	3	no	second, unique forbidden
and so on			

On the basis of the properties of Clebsch—Gordan and Racah coefficients and from (18) $S_{j_3}(vk12)$ can be no zero, if

$$l_v - l_k = j_3 + \text{an even number; } l_1 - l_2 = j_3 + \text{an even number} \tag{21}$$

and the following triangles exist:

$$\Delta(l_v l_k j_3),\ \Delta(j_v j_k j_3),\ \Delta(l_1 l_2 j_3),\ \Delta(j_1 j_2 j_3).$$

In the same way from (16): $T_{j_3}(vk12/a') \neq 0$ is possible, if

$$\begin{aligned} l_v - l_k &= a' + \text{an even number;} \\ l_1 - l_2 &= a' + \text{an even number} \end{aligned} \tag{22}$$

and

$$\Delta(l_v l_k a'),\ \Delta(j_v j_k j_3),\ \Delta(l_1 l_2 a'),\ \Delta(j_1 j_2 j_3)$$

as well as

$$\Delta(j_v j_1 a),\ \Delta(l_v l_1 a),\ \Delta(l_k l_2 a) \text{ and } \Delta(j_k j_2 a) \tag{22a}$$

exist.

Table 2

		S									
	$(j_1 j_2 L_1 L_2)$	...	$\left(\frac{5}{2}-\frac{5}{2}\,2-2\right)$ $\left(\frac{3}{2}-\frac{3}{2}\,2-2\right)$ $\left(\frac{5}{2}-\frac{3}{2}\,3-1\right)$ $\left(\frac{3}{2}-\frac{3}{2}\,2-2\right)$	$\left(\frac{3}{2}-\frac{3}{2}\,2-1\right)$ $\left(\frac{3}{2}-\frac{3}{2}\,1-2\right)$ $\left(\frac{5}{2}-\frac{3}{2}\,2-1\right)$ $\left(\frac{3}{2}-\frac{1}{2}\,2-1\right)$	$\left(\frac{3}{2}-\frac{3}{2}\,1-1\right)$ $\left(\frac{1}{2}-\frac{1}{2}\,1-1\right)$ $\left(\frac{3}{2}-\frac{1}{2}\,20\right)$ $\left(\frac{3}{2}-\frac{1}{2}\,1-1\right)$	$\left(\frac{1}{2}-\frac{1}{2}\,10\right)$ $\left(\frac{1}{2}-\frac{1}{2}\,0-1\right)$ $\left(\frac{3}{2}-\frac{1}{2}\,10\right)$ $\left(\frac{1}{2}-\frac{1}{2}\,10\right)$	$\left(\frac{1}{2}-\frac{1}{2}\,00\right)$ $\left(\frac{1}{2}-\frac{1}{2}\,1-1\right)$ $\left(\frac{1}{2}-\frac{1}{2}\,11\right)$ $\left(\frac{1}{2}-\frac{1}{2}\,00\right)$	$\left(\frac{1}{2}-\frac{1}{2}\,10\right)$ $\left(\frac{3}{2}-\frac{1}{2}\,10\right)$ $\left(\frac{1}{2}\quad\frac{1}{2}\,01\right)$ $\left(\frac{1}{2}\quad\frac{1}{2}\,10\right)$	$\left(\frac{3}{2}\,\frac{1}{2}\,11\right)$ $\left(\frac{3}{2}\,\frac{1}{2}\,20\right)$ $\left(\frac{1}{2}\,\frac{1}{2}\,11\right)$ $\left(\frac{3}{2}\,\frac{1}{2}\,11\right)$	$\left(\frac{3}{2}\,\frac{1}{2}\,21\right)$ $\left(\frac{5}{3}\,\frac{3}{2}\,21\right)$ $\left(\frac{3}{2}\,\frac{3}{2}\,12\right)$ $\left(\frac{3}{2}\,\frac{3}{2}\,21\right)$	$\left(\frac{5}{2}\,\frac{3}{2}\,22\right)$ $\left(\frac{5}{2}\,\frac{3}{2}\,31\right)$ $\left(\frac{3}{2}\,\frac{3}{2}\,22\right)$ $\left(\frac{5}{2}\,\frac{5}{2}\,22\right)$	...
G_r^+ and G_r^- vk12 and v′k′12	$(j_3\ a')$	...	(4, 0) (6, 2) (8, 4) . . .	(3, 0) (5, 2) (7, 4) . . .	(2, 0) (4, 2) (6, 4) . . .	(1, 0) (3, 2) (5, 4) . . .	(0, 0) (2, 2) (4, 4) . . .	(1, 2) (3, 4) (5, 6) . . .	(0, 2) (2, 4) (4, 6) . . .	(1, 4) (3, 6) (5, 8) . . .	(0, 4) (2, 6) (4, 8) . . .	...
M_r^+ and M_r^- vk′12 and v′k12												
M_r^+ vk 12	$(j_3\ a')$	...	(5, 1) (7, 3) (9, 5) . . .	(4, 1) (6, 3) (8, 5) . . .	(3, 1) (5, 3) (7, 5) . . .	(2, 1) (4, 3) (6, 5) . . .	(1, 1) (3, 3) (5, 5) . . .	(0, 1) (2, 3) (4, 5) . . .	(1, 3) (3, 5) (5, 7) . . .	(0, 3) (2, 5) (4, 7) . . .	(1, 5) (3, 7) (5, 9) . . .	...
M_r^- v′k′ 12												
G_r^+ vk′ 12												
G_r^- v′k 12												
M_r^- vk 12	$(j_3\ a')$	...	(3, —1) (5, 1) (7, 3) . . .	(2, —1) (4, 1) (6, 3) . . .	(1, —1) (3, 1) (5, 3) . . .	(0, —1) (2, 1) (4, 3) . . .	(1, 1) (3, 3) (5, 5) . . .	(0, 1) (2, 3) (4, 5) . . .	(1, 3) (3, 5) (5, 7) . . .	(0, 3) (2, 5) (4, 7) . . .	(1, 5) (3, 7) (5, 9) . . .	...
M_r^+ v′k′ 12												
G_r^+ v′k 12												
G_r^- vk′ 12												

The case $j_3 = a'$ refers to S_{j_3} $(vk\ 12)$.

Notations: $(A, B) \to (j_3 = r + A,\ a' = r + B)$;

$(abcd) \to (j_1 = a,\ j_2 = j_3 + b,\ L_1 = c,\ L_2 = j_3 + d)$, $(a + 1, j_3 + b - 1,\ c + 1, j_3 + d - 1)$, $(a + 2, j_3 + b - 2, c + 2, j_3 + d - 2), \ldots$

It can be seen that if (22) is fulfilled, there is always an a for which (22a) is valid, so the selection rules for $T_{j_3}(vk12|a')$ can be read from (22) alone.

It is natural, that if we are interested in selection rules for $S_{j_3}(v'k12)$ for instance, then we must write l_v' in (21) instead of l_v, but this would be the only change. Thus it is enough to examine (21) and (22). The results are summarized in Table 2.

4. The determination of ε/β^+

We use our results obtained in the second section for the calculation of the transition probability of β^+-decay and K-capture. If w_{ba} is the transition probability into a state with fixed value of the angular momentum and parity, the total probability of the β^+-decay, for unpolarized nuclei is given by

$$\Gamma_{\beta+} = \int dE_v \sum_{M_v M_k} (2j_k+1)^{-1} \sum_{\substack{j_1 l_1 M_1 \\ j_2 l_2 M_2}} \int d\varepsilon \, w_{ba}. \tag{23}$$

Since M_1, M_2, M_v, M_k appear only in $<vk12>$ and $<vk12>_4$ we can sum over them with the help of (15) and (17):

$$\begin{aligned}
&\sum_{M_1 M_2 M_v M_k} <vk\,1\,2> <vk\,12> = (2\pi)^{-2} \sum_{a' a'' j_3} T_{j_3}(vk\,12|a')\, T_{j_3}(vk\,12|a''); \\
&\sum_{M_1 M_2 M_v M_k} <vk\,12> <vk\,12>_4 = (2\pi)^{-2} \sum_{a' j_3} T_{j_3}(vk\,12|\,a')\, S_{j_3}(vk\,12); \\
&\sum_{M_1 M_2 M_v M_k} <vk\,12>_4 <vk\,12>_4 = (2\pi)^{-2} \sum_{j_3} S_{j_3}(vk\,12)\, S_{j_3}(vk\,12).
\end{aligned} \tag{24}$$

The same holds for $v'k'1'2'$ too (only we have to write v', k', $1'$, $2'$ instead of v, k, 1, 2). We can integrate with the help of the δ-function, and get

$$E_v = q = \varepsilon_0 - \varepsilon, \qquad p = \sqrt{\varepsilon^2 - 1}, \tag{25}$$

where ε_0 means the energy difference between the initial and the final state of the nucleus.

In (23) we must sum over j_1 and j_2 from 0 to ∞. Remembering that if a term of the sum is not zero, it is by at least two orders of magnitude larger than the following, it is enough to retain the first term which is not equal to zero, the following ones are negligible.

We now estimate the order of magnitude of the terms. For this purpose we calculate G^2, F^2, g^2 and f^2. Using the fact that for the energies occurring

$pR \ll 1$, $qR \ll 1$ and that

$$J_{l+\frac{1}{2}}(x) \cdot x^{-\frac{1}{2}} = \frac{1}{\sqrt{2}} \left(\frac{x}{2}\right)^l \sum_0^\infty \frac{(-1)^k \left(\frac{x}{2}\right)^{2k}}{k!\, \Pi\left(l + k + \frac{1}{2}\right)}, \tag{26a}$$

$$F(a, \alpha; x) = 1 + \frac{a}{\alpha} x + \frac{a(a+1)}{\alpha(\alpha+1)} \frac{x^2}{2!} + \ldots \tag{26b}$$

we get with the help of (7), (9) and (13)

$$G^2 = \frac{16\pi^2}{[(2l_1+1)!!]^2} (qR)^{2l_1} + 0\left[(qR)^{2l_1+2}\right];$$

$$F^2 = \frac{16\pi^2}{[(2l_1'+1)!!]^2} (qR)^{2l_1'} + 0\left[(qR)^{2l_1'+2}\right];$$

$$\begin{aligned} g^2 = {} & p(\varepsilon-1)(pR)^{2\gamma_2-2} \Gamma^{-2}(2\gamma_2+1)\, 2^{2\gamma_2} e^{\pi\nu} |\Gamma(\gamma_2+i\nu)|^2 \times \\ & \times \frac{1}{\pi} \Big\{ (\gamma_2+\varkappa_2)^2 + (\nu+\mu)^2 - \frac{2pR}{2\gamma_2+1} \left[(\gamma_2+\varkappa_2)(\nu-\mu) + \right. \\ & + 2\nu(\gamma_2+\varkappa_2)^2 + 2\nu(\nu+\mu)^2 - (\gamma_2-\varkappa_2)(\nu+\mu)\Big] + \\ & + (pR)^2 \Big\{ (2\gamma_2+1)^{-2} \left[(\gamma_2-\varkappa_2)^2 + (\mu-\nu)^2 + 4\nu^2(\gamma_2+\varkappa_2)^2 + 4\nu^2(\mu-\nu)^2 + \right. \\ & + 4\nu(\gamma_2+\varkappa_2)(\nu-\mu) - 4\nu(\gamma_2-\varkappa_2)(\nu+\mu)\Big] - \\ & - \frac{2\gamma^2+5}{2\gamma_2+1} \left[(\gamma_2+\varkappa_2)^2 + (\nu+\mu)^2\right] + \frac{4}{(2\gamma_2+1)(\gamma_2+1)} \left[\nu^2(\gamma_2+\varkappa_2)^2 + \right. \\ & + \nu^2(\nu+\mu)^2 + \nu(\gamma_2+2)(\gamma_2+\varkappa_2)(\nu-\mu) - \\ & \left.\left. - \nu(\gamma_2+2)(\gamma_2-\varkappa_2)(\nu+\mu)\right]\Big\}\Big\} + 0\left[(pR)^{2\gamma_2+1}\right]; \end{aligned} \tag{26}$$

$$\begin{aligned} f^2 = {} & p(\varepsilon+1)(pR)^{2\gamma_2-2} \Gamma^{-2}(2\gamma_2+1) 2^{2\gamma_2} e^{\pi\nu} |\Gamma(\gamma_2+i\nu)|^2 \frac{1}{\pi} \Big\{ (\gamma_2-\varkappa_2)^2 + \\ & + (\nu-\mu)^2 - \frac{2pR}{2\gamma_2+1} \left[(\gamma_2-\varkappa_2)(\nu+\mu) + 2\nu(\gamma_2-\varkappa_2)^2 + \right. \\ & + 2\nu(\nu-\mu)^2 - (\gamma_2+\varkappa_2)(\nu-\mu)\Big] + (pR)^2 \Big\{ (2\gamma^2+1)^{-2} \left[(\gamma_2+\varkappa_2)^2 + \right. \\ & + (\mu+\nu)^2 + 4\nu^2(\gamma_2-\varkappa_2)^2 + 4\nu^2(\mu+\nu)^2 + 4\nu(\gamma_2-\varkappa_2)(\nu+\mu) - \\ & - 4\nu(\gamma_2+\varkappa_2)(\nu-\mu)\Big] - \frac{2\gamma_2+5}{2\gamma_2+1} \left[(\gamma_2-\varkappa_2)^2 + (\nu-\mu)^2\right] + \\ & + \frac{4}{(2\gamma_2+1)(\gamma_2+1)} \left[\nu^2(\gamma_2-\varkappa_2)^2 + \nu^2(\nu-\mu)^2 + \right. \\ & + \nu(\gamma_2+2)(\gamma_2-\varkappa_2)(\nu+\mu) - \nu(\gamma_2+2)(\gamma_2+ \\ & \left.\left. + \varkappa_2)(\nu-\mu)\right]\Big\}\Big\} + 0\left[(pR)^{2\gamma_2+1}\right]. \end{aligned}$$

We can see from (25) that, as a first approximation, it is enough to retain only the lowest power of (pR) and (qR), respectively, because $\varkappa R < 10^{-2}$ is always fulfilled (for small Z in (26) it is better to take into account the first three powers of (pR), since if Z tends to zero, the coefficients of the first two terms do so too). If we consider the lowest powers, the order of magnitude for example of $G^2(g^2+f^2)$ is determined by the number $2(l_1+\gamma_2-1)$, since $(\varkappa R)$ is to this power. So the order of magnitude is determined approximately by

$$k=\begin{cases} l_1+l_2 & \text{if} \quad l_2=j_2+\dfrac{1}{2}\,, \\ \\ l_1+l_2+1 & \text{if} \quad l_2=j_2-\dfrac{1}{2}\,. \end{cases} \tag{27}$$

Now we can turn to the calculation of the probability of β^+ decay. We introduce the notation:

$$A_{\beta+}=\frac{G^2}{2\,(2\pi)^4\,(j_K)^2}\,, \tag{28}$$

so we get from (23), (24) and (1)

$$\Gamma_{\beta+}=A_{\beta+}\,(r_{11}\,R_1^2+2r_{14}\,R_1\,R_4+r_{44}\,R_4^2+r_{22}\,R_2^2+2r_{23}\,R_1\,R_3+r_{33}\,R_3^2)=$$

$$=A_{\beta+}\left\{\begin{array}{l} R(\mathbf{vk})+R(\mathbf{v'k'})\,|\,R(\mathbf{vk})+R(\mathbf{v'k'}) \\ \sum\limits_{j_1L_1j_2L_2}\int d\varepsilon\,[G(\mathbf{1})\,g(\mathbf{2})\,|\,G(\mathbf{1})\,g(\mathbf{2})+G(\mathbf{1})f(\mathbf{2'})\,|\,G(\mathbf{1})f(\mathbf{2'})] \\ \sum\limits_{j_3}\big[\lambda\sum\limits_{a'}T_{j_3}(\ |\,a')+(1-\lambda)\,S_{j_3}(\ )\,\big|\,\lambda\sum\limits_{a''}T_{j_3}(\,|\,a'')+(1-\lambda)\,S_{j_3}(\ )\big] \end{array}\right. +$$

(29)

$$+\left\{\begin{array}{l} R(\mathbf{vk'})+R(\mathbf{v'k})\,|\,R(\mathbf{vk'})+R(\mathbf{v'k}) \\ \sum\limits_{j_1L_1j_2L_2}\int d\varepsilon\,[G(\mathbf{1})\,g(\mathbf{2})\,|\,G(\mathbf{1})\,(\mathbf{2})+G(\mathbf{1})f(\mathbf{2'})\,|\,G(\mathbf{1})f(\mathbf{2'})] \\ \sum\limits_{j_3}\big[-\sum\limits_{a'}T_{j_3}(\ |\,a')+(1-\lambda)\,S_{j_3}(\ )\,\big|-\sum\limits_{a''}T_{j_3}(\ |\,a'')+(1-\lambda)\,S_{j_3}(\ )\big] \end{array}\right. -$$

$$-2\left\{\begin{array}{l} R(\mathbf{vk})+R(\mathbf{v'k'})\,|\,R(\mathbf{vk'})+R(\mathbf{v'k}) \\ \sum\limits_{j_1L_1j_2L_2}\int d\varepsilon\,[G(\mathbf{1})\,g(\mathbf{2})\,|\,G(\mathbf{1})f(\mathbf{2'})-G(\mathbf{1})\,f(\mathbf{2'})\,|\,G(1)\,g(\mathbf{2})] \\ \sum\limits_{j_3}\big[\lambda\sum\limits_{a'}T_{j_3}(\ |\,a')+(1-\lambda)\,S_{j_3}(\ )\,\big|-\sum\limits_{a''}T_{j_3}(\ |\,a'')+(1-\lambda)\,S_{j_3}(\ )\big] \end{array}\right\}.$$

Thus for example

$$r_{11} = r\,(vk\,vk) =$$

$$= \begin{Bmatrix} 1\,(\mathbf{vk})\,|\,1(\mathbf{vk}) \\ \sum\limits_{j_1 L_1 j_2 L_2} \int d\varepsilon\,[G\,(\mathbf{1})\,g\,(\mathbf{2})\,|\,G\,(\mathbf{1})\,g\,(\mathbf{2}) + G\,(\mathbf{1})f\,(\mathbf{2}')\,|\,G\,(\mathbf{1})f\,(\mathbf{2}')] \\ \sum\limits_{j_3}\left[\lambda \sum\limits_{a'} T_{j_3}(\ |\,a') + (1-\lambda)\,S_{j_3}(\ )\,|\,\lambda \sum\limits_{a''} T_{j_3}(\ |\,a'') + (1-\lambda)\,S_{j_3}(\ )\right] \end{Bmatrix} \tag{30}$$

where

$$1(vk) \equiv 1.$$

Here the summation runs over values of $j_1, j_2, L_1, L_2, a, a', j_3$ which satisfy the selection rules included in Table 2. Naturally on the basis of (27) we must take into account only the first nonvanishing term, for this k is the least.

The probability of the K-capture can be calculated in entirely the same way for the wave functions (2) and (6) are of completely the same form:

$$\Gamma_K = \int dE_\nu \sum_{M_\nu M_k} (2j_k + 1)^{-1} \sum_{j_1 l_1 M_1 M_2} (2j_2 + 1)^{-1} w_{ba}\,. \tag{31}$$

With the help of (11), (14) and (24) and integrating over the neutrino energy E_ν we get $q = E_\nu = (\varepsilon_0 + 1)\,(1 - b/1 + \varepsilon_0)$, (where b is the binding energy of the electron on the K shell). The expression Γ_K for the transition probability way, of K-capture is similar to Γ_{β^+}:

$$\Gamma_K = A_k\,(\varrho_{11}\,R_1^2 + 2\varrho_{14}\,R_1\,R_4 + \varrho_{44}\,R_4^2 + \varrho_{22}\,R_2^2 + 2\varrho_{23}\,R_2\,R_3 + \varrho_{33}\,R_3^2) =$$

$$= A_k \cdot \left[\begin{Bmatrix} R\,(\mathbf{vk}) + R\,(\mathbf{v'k'})\,|\,R\,(\mathbf{vk}) + R\,(\mathbf{v'k'}) \\ \sum\limits_{j_1 L_1}^{1} [G\,(\mathbf{1})\,1\,(\mathbf{K})\,|\,G(\mathbf{1})1(\mathbf{K})] + \dfrac{1-\varepsilon_1}{1+\varepsilon_1} \sum\limits_{j_1 L_1}^{2} [G\,(\mathbf{1})\,1\,(\mathbf{K'})\,|\,G\,(\mathbf{1})\,1\,(\mathbf{K'})] \\ \sum\limits_{j_3}\left[\lambda \sum\limits_{a'} T_{j_3}(\ |\,a') + (1-\lambda)\,S_{j_3}(\ )\,|\,\lambda \sum\limits_{a''} T_{j_3}(\ |\,a'') + (1-\lambda)\,S_{j_3}(\ )\right] \end{Bmatrix} + \right.$$

$$+ \begin{Bmatrix} R\,(\mathbf{vk'}) + R\,(\mathbf{v'k})\,|\,R\,(\mathbf{vk'}) + R\,(\mathbf{v'k}) \\ \sum\limits_{j_1 L_1}^{1} [G\,(\mathbf{1})1\,(\mathbf{K})\,|\,G\,(\mathbf{1})1\,(\mathbf{K})] + \dfrac{1-\varepsilon_1}{1+\varepsilon_1} \sum\limits_{j_1 L_1}^{2} [G\,(\mathbf{1})\,1\,(\mathbf{K'})\,|\,G\,(\mathbf{1})\,1\,(\mathbf{K'})] \\ \sum\limits_{j_3}\left[-\sum\limits_{a'} T_{j_3}(\ |\,a') + (1-\lambda)\,S_{j_3}(\ )\,|\,-\sum\limits_{a''} T_{j_3}(\ |\,a'') + (1-\lambda)\,S_{j_3}(\ )\right] \end{Bmatrix} + \tag{32}$$

$$\left. + 2 \begin{Bmatrix} R\,(\mathbf{vk}) + R\,(\mathbf{v'k'})\,|\,R\,(\mathbf{vk'}) + R\,(\mathbf{v'k}) \\ \left\{ \sum\limits_{j_1 L_1}^{1|2} [G\,(\mathbf{1})\,1\,(\mathbf{K})\,|\,G\,(\mathbf{1})\,1\,(\mathbf{K'})] - \sum\limits_{j_1 L_1}^{2|1} [G\,(\mathbf{1})\,1\,(\mathbf{K'})\,|\,G\,(\mathbf{1})\,1\,(\mathbf{K})] \right\} \sqrt{\dfrac{1-\varepsilon_1}{1+\varepsilon_1}} \\ \sum\limits_{j_3}\left[\lambda \sum\limits_{a'} T_{j_3}(\ |\,a') + (1-\lambda)\,S_{j_3}(\ )\,|\,-\sum\limits_{a''} T_{j_3}(\ |\,a'') + (1-\lambda)\,S_{j_3}(\ )\right] \end{Bmatrix} \right]$$

Here

$$1(K) = 1\left(j_2 = \frac{1}{2}, \quad l_2 = 0\right) \equiv 1$$

and

$$1(K') = 1\left(j_2 = \frac{1}{2}, \quad l_2' = 1\right) \equiv 1;$$

as well as

$$A_k = \frac{G^2 g_1^{-1}(R)^2}{2(j_k)^2 (2\pi)^4} (\varepsilon_0 + 1)^2 \left(1 - \frac{b}{1+\varepsilon_0}\right)^2; \tag{33}$$

and using (4)

$$g_1^{-1}(r) = \left(\frac{2Z}{a_0}\right)^{3/2} \sqrt{\frac{1+\varepsilon_1}{2\Gamma(\varepsilon+1)}}\, e^{-\varrho_1/2}\, \varrho_1^{\varepsilon_1 - 1}; \tag{34}$$

$$f_1^{-1}(r) = -\sqrt{\frac{1-\varepsilon_1}{1+\varepsilon_1}}\, g_1^{-1}(r)$$

here

$$\varepsilon_1 = \sqrt{1-(\alpha Z)^2} \qquad \text{and} \qquad \varrho_1 = \frac{2Zr}{a_0}.$$

In (32) we must sum according to the selection rules given in Table 2. Since now $j_2 = 1/2$, $l_2 = 0$ and $l_2' = 1$ are fixed, we must sum only over j_1 and L_1, in particular for values where $j_2 = 1/2$ and $L_2 = 0$ in the case of $\sum^1$ (because these terms are connected with g^2), and with the condition $j_2 = 1/2$ and $L_2 = 1$ in the case of $\sum^2$ (these terms are connected with f^2).

We can see from the expression (32) and (29) that the ratio of the K-capture to the positron emission is not independent of the nuclear matrix elements in the general case. Namely

$$\Gamma = \frac{\Gamma_K}{\Gamma_{\beta+}} = A_t \, \frac{\varrho_{11} R_1^2 + 2\varrho_{14} R_1 R_4 + \varrho_{44} R_4^2 + \varrho_{22} R_2^2 + 2\varrho_{23} R_2 R_3 + \varrho_{33} R_3^2}{r_{11} R_1^2 + 2r_{14} R_1 R_4 + r_{44} R_4^2 + r_{22} R_2^2 + 2r_{23} R_2 R_3 + r_{33} R_3^2},$$

where

$$A_t = \frac{A_k}{A_{\beta+}} = \frac{1}{2} g_1^{-1}(R)^2 (1+\varepsilon_0)^2 \left(1 - \frac{b}{1+\varepsilon_0}\right)^2. \tag{35a}$$

Thus if we want to get the ratio ε/β^+ for a concrete nucleus, we must compute the coefficients ϱ_{ik} and r_{ik} and then form the following quantities:

$$\varrho_{11}/r_{11},\ \varrho_{14}/r_{14},\ \varrho_{44}/r_{44},\ \varrho_{22}/r_{22},\ \varrho_{23}/r_{23},\ \varrho_{33}/r_{33}. \tag{36a}$$

It is easy to prove that

$$A_t \cdot \min \frac{\varrho_{ik}}{r_{ik}} \leq \Gamma \leq A_t \cdot \max \frac{\varrho_{ik}}{r_{ik}} \tag{36}$$

constitutes (in special cases possibly narrow) limits for Γ. The theory does not predict any more than this inequality for the ratio ε/β^+ when the nuclear matrix elements are not known. At most we can add that Γ is probably nearer to $A_t \cdot \varrho_{11}/r_{11}$ than, for example, to $A_t \cdot \varrho_{44}/r_{44}$, since the first value belongs to integrals containing the large components of wave functions, while the second belongs to integrals containing only small components. (Naturally it is possible to give for ε/β^+ a more exact value, if one of the coefficients ϱ_{ik} and the corresponding r_{ik} are larger than the others by orders of magnitude).

The numerical results are given for the case of ${}_{33}As^{71} \rightarrow {}_{32}Ge^{71*}$ transition. (This is a $5/2^- \rightarrow 5/2^-$ transition: [10]—[11]. So $j_k = 5/2$, $l_k = 3$ and $j_v = 5/2$, $l_v = 3$ consequently this is a G_0 transition, Table 1.)

$$A_t = 1{,}17 \cdot 10^{-17},$$

$$\begin{aligned} r_{11} &= 2{,}06 \cdot 10^{-17}, & r_{14} &= -2{,}41 \cdot 10^{-17}, & r_{44} &= 4{,}71 \cdot 10^{-17}, \\ r_{22} &= 7{,}42 \cdot 10^{-16}, & r_{23} &= 1{,}85 \cdot 10^{-16}, & r_{33} &= 3{,}28 \cdot 10^{-17}, \\ \varrho_{11} &= 6{,}02, & \varrho_{14} &= -7{,}05, & \varrho_{44} &= 13{,}8, \\ \varrho_{22} &= 1{,}93 \cdot 10^{-1}, & \varrho_{23} &= 4{,}82 \cdot 10^{-2}, & \varrho_{33} &= 8{,}52 \cdot 10^{-1}. \end{aligned} \tag{37a}$$

Thus

$$A_t \frac{\varrho_{11}}{r_{11}} = A_t \frac{\varrho_{14}}{r_{14}} = A_t \frac{\varrho_{44}}{r_{44}} = 3{,}41$$

and (37b)

$$A_t \cdot \frac{\varrho_{22}}{r_{22}} = A_t \cdot \frac{\varrho_{23}}{r_{23}} = A_t \cdot \frac{\varrho_{33}}{r_{33}} = 3{,}02 \cdot 10^{-3}\,.$$

The measured value of Γ is 1.76. This is much nearer to the first value than to the second one. This fact is easy to understand, because the value 3.41 is obtained from the coefficients of integrals, containing only the large components of the wave functions.

Using the measured value of Γ we get the following relation between the nuclear matrix elements:

$$R_2^2 + 2R_2 R_3 \cdot 0{,}25 + R_3^2 \cdot 4{,}4 = R_1^2 \cdot 0{,}026 + 2R_1 R_4\, 0{,}03 + R_4^2 \cdot 0{,}1\,. \tag{37}$$

Lately a whole series of articles have dealt with the M_0 transitions [3]—[6]. Here there occur also 50% differences between the measured and calculated values. The simplest M_0 transition is:

$$j_v = \frac{1}{2}, \quad l_v = 0; \quad j_k = \frac{1}{2}, \quad l_k = 1\,.$$

Such a transition has as yet not been observed, thus the values $Z = 81$ and $A = 200$ are taken from the data of the decay of Tl^{200} showing a $2^- \to 2^+$ transition). In this case we get for the coefficients the values

$$A_t \cdot \frac{\varrho_{11}}{r_{11}} = 0{,}305; \quad A_t \cdot \frac{\varrho_{14}}{r_{14}} = 0{,}068; \quad A_t \frac{\varrho_{44}}{r_{44}} = 0{,}321;$$
$$A_t \frac{\varrho_{22}}{r_{22}} = 3{,}57; \quad A_t \frac{\varrho_{23}}{r_{23}} = 3{,}64; \quad A_t \frac{\varrho_{33}}{r_{33}} = 4{,}36\,. \tag{38}$$

In the cases, when Γ depends on the matrix elements, we can use the measurement of the ratio ε/β^+ to determine the nuclear matrix elements (see (37)). There are only four unknown matrix elements, thus it is enough to make four independent measurements to determine their values. (For example: $\lg f_t$, ε/β^+, the energy spectrum, angular correlation.) At the same time these four quantities represent the maximal information concerning the nuclear structure and available from β-decay, since in the transition probabilities there are only these four quantities characterizing the nucleus.

I should like to thank Dr. K. NAGY and Dr. G. MARX for many helpful discussions and advice, which have made it possible for me to write my diploma work and this article.

Appendix

$$<vk\,12> = \int d\hat{r}\,(\Omega_v^+\,\sigma_\alpha\,\Omega_k)\,(\Omega_1^+\,\sigma_\alpha\,\Omega_2) =$$
$$= \int d\hat{r} \sum_{\mu_1\mu_2\mu_v\mu_k} Y^*_{l_v m_v}(\hat{r})\,Y^*_{l_1 m_1}(\hat{r})\,Y_{l_2 m_2}(\hat{r})\,Y_{l_k m_k}(\hat{r}) \times$$
$$\times \left(l_v\, m_v\, \frac{1}{2}\, \mu_v \,|\, j_v\, M_v\right)\left(l_k\, m_k\, \frac{1}{2}\, \mu_k |\, j_k M_k\right)\left(l_1\, m_1\, \frac{1}{2}\, \mu_1 \,|\, j_1\, M_1\right) \times$$
$$\times \left(l_2\, m_2\, \frac{1}{2}\, \mu_2 \,|\, j_2\, M_2\right) \sigma_\alpha\,(\mu_v\,\mu_k)\,\sigma_\alpha\,(\mu_1\,\mu_2),$$

because

$$\sigma_\alpha\,(\mu_v\,\mu_k) = \text{ the } \mu_v\,\mu_k - \text{ th element of } \sigma_\alpha;$$
$$\sigma_\alpha\,(\mu_v\,\mu_k)\,\sigma_\alpha\,(\mu_1\,\mu_2) = 2(-)^{1-\mu_1-\mu_2}\,\delta_{-\mu_v,\,\mu_1}\,\delta_{-\mu_k,\,\mu_2}$$

and

$$Y_{l_1 m_1}(\hat{r})\,Y_{l_2 m_2}(\hat{r}) = \sum_{LM} \frac{(l_1)\,(l_2)}{\sqrt{4\pi}\,(L)}\,(l_1\,0\,l_2\,0\,|\,L\,0)\,(l_1\,m_1\,l_2\,m_2\,|LM)\,Y_{LM}(\hat{r}),$$

and

$$Y^*_{lm} = (-)^m\,Y_{l-m} \text{ as well as } \int d\hat{r}\,Y^*_{lm}\,Y_{l'm'} = \delta_{ll'}\,\delta_{mm'}.$$

We can write

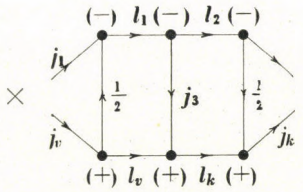
$$<vk\,12> = (2\pi)^{-1}(l_v)(l_k)(l_1)(l_2)\sum_{j_3}(l_v\,0\,l_k\,0\,|\,j_3\,0)(l_1\,0\,l_2\,0\,|\,j_3\,0)(j_3)^{-2}\times$$

$$\times\sum_{\substack{\mu_1\mu_2\varkappa\\ m_1m_2m_vm_k}}(-)^{1-\mu_1-\mu_2-m_1-m_k}(l_v m_v\frac{1}{2}\mu_1\,|\,j_v\,M_v)(l_k m_k\frac{1}{2}\mu_2\,|\,j_k\,M_k)(l_1\,m_1\frac{1}{2}-\mu_1\,|\,j_1\,M_1)$$

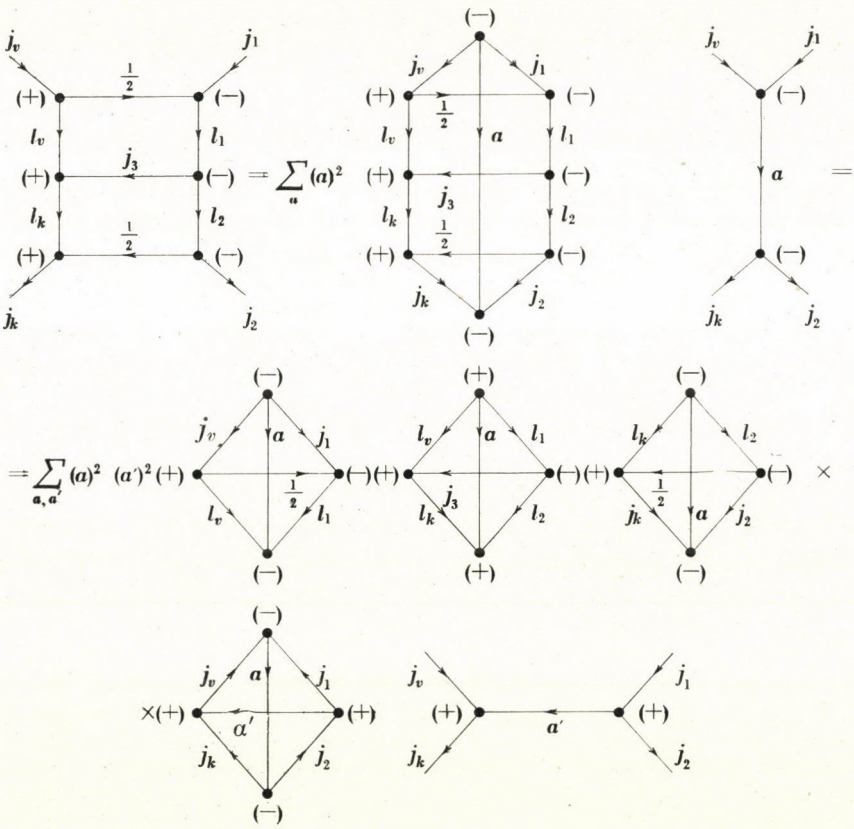
$$\times(l_2\,m_2\,\frac{1}{2}-\mu_2\,|\,j_2\,M_2)(l_v-m_v\,l_k\,m_k\,|\,j_3\,\varkappa)(l_1-m_1\,l_2\,m_2\,|\,j_3-\varkappa).$$

According to [2]

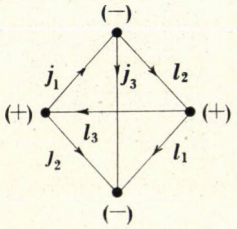
$$\sum_{\substack{\mu_1\mu_2\varkappa\\ m_1m_2m_vm_k}}(\ldots) = (j_v)(j_k)(j_1)(j_2)(j_3)^2(-)^{l_1+l_k+l_2+l_v+j_3+M_v+M_1}\times$$



The graph can be drawn in the following way:

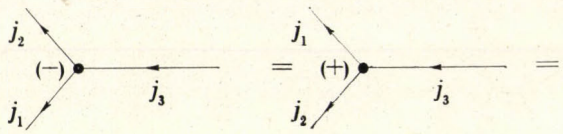


Using this and taking into account



$$= \begin{Bmatrix} j_1 j_2 j_3 \\ l_1 l_2 l_3 \end{Bmatrix} = (-)^{j_1+j_2+l_1+l_2} W(j_1 j_2 l_2 l_1;\ j_3 l_3),$$

and



$$= \begin{pmatrix} j_1 j_2 j_3 \\ m_1 m_2 - m_3 \end{pmatrix} = (j_3)^{-1} (-)^{j_2-j_1+m_3} (j_1 m_1 j_2 m_2 | j_3 m_3)$$

we get (15). $<vk\ 12>_4$ can be calculated in a similar way.

REFERENCES

1. A. I. Ahiezer and V. B. Bereszteckij, Kvantumelektrodinamika (in Hungarian) Akadémiai Kiadó, Budapest, 1961.
2. G. V. Jefimov, preprint.
3. M. L. Perlman, Phys. Rev., **110,** 381, 1958.
4. J. Konijn, Nucl. Phys., **16,** 683, 1960.
5. J. Konijn, Nucl. Phys., **31,** 406, 1962.
6. D. Berényi, Dissertation of candidature.
7. H. Brysk and M. E. Rose, Rev. Mod. Phys., **30,** 1169, 1958.
8. P. Depommier and R. Bouchez, Journ. de Phys. Radium, **21,** 456, 1960.
9. D. S. Harmer and M. L. Perlman, Phys. Rev., **122,** 218, 1961.
10. S. Thulin and J. Moreau, Ark. Fysik, **8,** 219, 1954.
11. W. E. Graves and A. C. G. Mitchell, Phys. Rev., **97,** 1038, 1955.

ОПРЕДЕЛЕНИЕ МАТРИЧНЫХ ЭЛЕМЕНТОВ ЯДРА ПО ИЗМЕРЕНИЮ ε/β^+

И. МОНТВАЙ

Р е з ю м е

Подсчёт переходных вероятностей β-переходов, запрещённых в произвольном порядке — главным образом в высшем — принципиально упрощается в случае использования вместо приведённых матричных элементов таких, в которые входят в различных комбинациях просто радиальные части волновых функций нуклона. В таких случаях всегда имеется возможность для вычисления отношения ε/β^+, то-есть определение матричных элементов ядра сводится к измерению отношения ε/β^+.

ZU DEM PROBLEM DES ABSOLUTEN DIAMAGNETISMUS UND DER SUPRALEITUNG

Von

TH. NEUGEBAUER

INSTITUT FÜR THEORETISCHE PHYSIK DER ROLAND EÖTVÖS UNIVERSITÄT, BUDAPEST

(Vorgelegt von Z. Gyulai. — Eingegangen 20. IX. 1963)

Alle bis jetzt veröffentlichten Theorien der Supraleitung trachteten nur die unendliche Leitfähigkeit zu erklären, der absolute Diamagnetismus wurde dann entweder durch weitere Annahmen in das Modell eingeführt, oder es wurde durch verwickelte analytische Methoden gezeigt, dass die Theorie auch diesen erklärt. Die vorliegende Arbeit geht dagegen direkt von dem MEISSNER—OCHSENFELD-Effekt aus und untersucht zuerst die Frage, ob ein Übergang zwischen den gewöhnlichen diamagnetischen Erscheinungen und dem absoluten Diamagnetismus möglich ist. Das Verhalten der kondensierten aromatischen Verbindungen und des Graphits wird diesbezüglich betrachtet. Danach wird auf eine neue Regel bezüglich der Verteilung der supraleitenden Eigenschaften im periodischem System aufmerksam gemacht, mit deren Hilfe und der bezüglich des Diamagnetismus im ersten Teil erhaltenen Ergebnisse eine neue Theorie des absoluten Diamagnetismus und der Supraleitung folgt.

Einleitung

Zur naturwissenschaftlichen Erklärung der Supraleitungserscheinungen wurden schon sehr viele Theorien aufgestellt, und es wäre deshalb ganz unmöglich, diese hier alle auch nur ganz kurz zu besprechen. Wir verweisen deshalb auf die über diesen Gegenstand erschienenen zahlreichen zusammenfassenden Darstellungen und wollen jetzt nur einige wesentliche Punkte hervorheben:

Man kann sagen, dass alle bis jetzt veröffentlichten Theorien der unendlich grossen Leitfähigkeit von zwei verschiedenen Grundannahmen ausgegangen sind. Die erste solche Hypothese war, dass im supraleitenden Zustand Bohrsche Quantenbahnen irgendwie makroskopische Dimensionen annehmen. Da sich ein Elektron im Grundzustande auf solch einer Bahn selbstverständlich eine unendlich lange Zeit hindurch bewegen kann, ohne dabei Energie zu verlieren, so ist damit die unendlich grosse Leitfähigkeit erklärt. Die erste solche Theorie, die noch vor der Entdeckung der Quantenmechanik veröffentlicht wurde, ist die von BENEDICKS [1]. Selbstverständlich ist diese schon längst veraltet, aber die meisten modernen Theorien gehen ebenfalls von diesem Grundgedanken aus. Eine zweite solche Auffassung war, dass die Leitungselektronen unterhalb des Sprungpunktes eine Art festes Kristallgitter aufbauen. Dieses Gitter kann man dann innerhalb des Gitters der Metallionen widerstandslos verschieben. Aus dieser Auffassung würde also wieder eine

unendlich grosse Leitfähigkeit folgen. Der Sprungpunkt der Supraleiter wäre dann der »Schmelzpunkt« dieses Elektronengitters, ober den dann infolge der auftretenden Unordnung der Ohmsche Widerstand erscheint. Diese Theorie rührt von R. DE L. KRONIG [2] her. In seinen späteren Veröffentlichungen über diesen Gegenstand nahm dann der erwähnte Verfasser an, dass nicht ein ganzes Elektronengitter, sondern nur lineare Elektronenketten im Supraleiter vorhanden sind. Alles das war noch vor der Entdeckung der Fermi—Diracschen Statistik und der Nullpunktsenergie. Die HEISENBERGsche Theorie [3] ist eigentlich eine Modernisierung dieser Hypothese, nach der angenommen wird, dass nicht alle Leitungselektronen, sondern nur diejenigen, welche die grössten Energien besitzen und deshalb von den Einschränkungen der Fermistatistik schon ziemlich frei sind, ein Übergitter im Supraleiter aufbauen. Der Sprungpunkt ist dann wieder der Schmelzpunkt dieses Übergitters.

Die LONDONsche Theorie [4] und ihre Ergänzungen durch LAUE [5] und andere Autoren liefert dagegen keine naturwissenschaftliche Erklärung für die Supraleitung, und das war ja auch nicht ihr Ziel, sondern trachtet nur die Maxwellschen Differentialgleichungen innerhalb des Supraleiters richtig zu formulieren und dabei auch den MEISSNER—OCHSENFELD-Effekt in dieses Gleichungssystem einzuarbeiten.

Jetzt haben die Physiker das grösste Vertrauen zu der Theorie von BARDEEN, COOPER und SCHRIEFFER [6], deren Grundgedanke es ist, dass ein sich in der Nähe der Fermioberfläche befindendes Elektron ein virtuelles Phonon emittiert und dieses dann von einem zweiten Elektron absorbiert wird, wobei letzteres entsprechend dem Impulserhaltungssatz gestreut wird. Das entspricht einer Zweiteilchenwechselwirkung der erwähnten Elektronen und diese ist anziehend (wie das von den Verfassern gezeigt wird), wenn die Energiedifferenz zwischen den zwei bei diesem Prozess eine Rolle spielenden Elektronenzuständen kleiner als die Energie des virtuellen Phonons ist. In diesem Falle kann also die erwähnte Wechselwirkung die Coulombabstossung überkompensieren und so die Supraleitungserscheinungen erklären. Tatsächlich erklärt diese Theorie von BARDEEN, COOPER und SCHRIEFFER viele sich auf die Supraleiter beziehenden Erfahrungen ganz einwandfrei. Wir wollen hier nur zwei solche Erscheinungen erwähnen: Einerseits den von MAXWELL und unabhängig von ihm von REYNOLDS und seinen Mitarbeitern [7] entdeckten Isotopeneffekt, nach dem T_c prop. $M^{-1/2}$ ist, wo T_c die Sprungpunkte und M die Massen der Atome der zu demselben chemischen Element gehörenden Isotopen bedeutet. Eine zweite solche wichtige Erfahrung ist die, dass bei der Anregung von Elektronen aus dem supraleitenden in den Normalzustand ein endliches Energieintervall auftritt. Am Sprungpunkt verschwindet dann diese Energiedifferenz. Eigentlich kann man auch diese zuletzt besprochene Theorie in die erwähnten Gedankengänge einordnen, welche die Supraleitung durch das Makroskopischwerden von Bohrschen Bahnen zu erklären trachten. Nach

dieser Auffassung werden also Quantenerscheinungen im supraleitenden Zustande in unserer makroskopischen Welt bemerkbar. Vollständigkeitshalber sei nur noch erwähnt, dass man auch oft versucht hat, die Superfluiditätserscheinungen des flüssigen Heliums irgendwie mit den Supraleitungserscheinungen in Zusammenhang zu bringen.

Alle die erwähnten Theorien trachteten ursprünglich nur danach die unendlich grosse Leitfähigkeit zu erklären. Den MEISSNER—OCHSENFELD-Effekt (nach dem der einfach zusammenhängende Supraleiter im supraleitenden Zustand alle magnetischen Kraftlinien aus seinem Inneren verdrängt), enthielten sie nicht. Zur Erklärung des MEISSNER-Effektes musste man entweder weitere zusätzliche Hypothesen einführen oder aber es gelang nur durch verwickelte analytische Betrachtungen nachträglich zu zeigen oder wenigstens wahrscheinlich zu machen, dass das eingeführte Modell des Supraleiters auch den MEISSNER—OCHSENFELD-Effekt erklärt. Es ist jedoch ganz sicher, dass der MEISSNER—OCHSENFELD-Effekt eine ebenso fundamentale Erfahrung bezüglich des supraleitenden Zustandes ist wie das Verschwinden des Ohmschen Widerstandes. (Die Erfahrung, dass in die supraleitenden Legierungen einige magnetische Kraftlinien hineinfrieren, hat nur sekundäre Ursachen und rührt daher, dass Legierungen doch nicht ganz homogen sind.) Dass man zuerst die unendlich grosse Leitfähigkeit der Supraleiter und erst viel später den MEISSNER-Effekt entdeckt hat, ist nur eine Folge davon, dass elektrische Messungen am leichtesten durchführbar sind. Wir wollen hier jedoch den umgekehrten Weg einschlagen und den ganzen Erscheinungskomplex der Supraleitung ausgehend von dem MEISSNER—OCHSENFELD-Effekt, also von dem absoluten Diamagnetismus der Supraleiter, betrachten.

§ 1

Die erste Frage, die wir beantworten müssen, ist die, ob der absolute Diamagnetismus der Supraleiter und die gewöhnlichen diamagnetischen Erscheinungen, die ja eine allgemeine Eigenschaft jeder Materie sind, zwei ihrem physikalischen Inhalte nach ganz verschiedene Erscheinungen sind, die nur wegen dem übereinstimmenden Vorzeichen der Suszeptibilität zueinander mathematisch analog sind, oder ob es sich bei beiden um wesentlich die selbe physikalische Erscheinung handelt und deshalb Übergänge zwischen ihnen möglich sein würden.

Da ein Supraleiter alle magnetischen Kraftlinien aus seinem Inneren herauswirft, so muss bekannterweise in ihm die Permeabilität $\mu = 0$ sein. Also haben wir

$$\mu = 1 + 4\pi\chi = 0 \tag{1}$$

und daraus folgt für die Suszeptibilität des absolut diamagnetischen Körpers das bekannte Resultat

$$\chi = -1/4\pi. \tag{2}$$

Die gewöhnliche diamagnetische Suszeptibilität ist demgegenüber immer um viele Grössenordnungen kleiner. In der Literatur wird meistens die diamagnetische Suszeptibilität von einem Grammatom oder einem Grammolekül Materie angegeben, und diese ist dann von der Grössenordnung 10^{-5}—-10^{-6}. Für Argon beträgt sie z. B. $-21{,}5 \cdot 10^{-6}$, woraus für ein Argonatom $-3{,}55 \cdot 10^{-29}$ (in cgs-Einheiten) folgt. Berechnen wir dagegen die Suszeptibilität von einem cm^3 fester Materie [um sie mit (2) vergleichen zu können], so erhalten wir das Resultat, dass diese höchstens von der Grössenordnung $10^{-6} - 10^{-7}$ sein kann. Daraus folgt also, dass zwischen den erwähnten zwei Erscheinungen ein wenigstens fünf Grössenordnungen betragender Unterschied besteht.

Bevor wir noch die Frage besprechen, dass unter welchen Umständen der gewöhnliche Diamagnetismus wesentlich grösser werden könnte, wollen wir noch das Problem betrachten, ob man (2) auch aus den für den gewöhnlichen Diamagnetismus hergeleiteten Formeln erhalten könnte. Bezeichnen wir mit $\varkappa$ die diamagnetische Suszeptibilität eines Teilchens (eines Atoms, eines Moleküls usw.) und mit N die Zahl dieser Teilchen in der Volumeneinheit, so folgt bekannterweise für die Suszeptibilität von einem cm^3

$$\chi = \frac{\varkappa N}{1 - \frac{4\pi}{3}\varkappa N}. \tag{3}$$

Wird jetzt $\varkappa$ eine sehr grosse negative Zahl, so erhält man

$$\chi = -3/4\pi, \tag{4}$$

also ein Resultat, das zwar grössenordnungsmässig mit (2) übereinstimmt, jedoch einen anderen numerischen Faktor enthält. Doch ist dieser Widerspruch nur scheinbar, wie wir das gleich zeigen wollen. Der Faktor $4\pi/3$ im Nenner von (3) rührt vom Lorentzschen inneren Felde her. Bekannterweise haben wir nämlich, wenn wir annehmen, dass ein Atom in einem kugelförmigen Hohlraum enthalten ist, für die auf dieses Atom einwirkende magnetische Feldintensität

$$H_i = H + \frac{4\pi}{3} M, \tag{5}$$

wo H das äussere magnetische Feld und M das in der Volumeneinheit induzierte magnetische Moment ist. Das zweite Glied auf der rechten Seite von

(5) ist das Lorentzsche innere Feld. Bekanntermeise kann man aus (5) die Formel (3) leicht herleiten.

Abgesehen davon, dass der Wert der Konstante des Lorentzschen inneren Feldes auch im einfachen Fall eines kugelförmigen Hohlraums fraglich ist, wie das neuere besonders im Zusammenhang mit der Theorie der ferroelektrischen Erscheinungen unternommene Untersuchungen zeigen, wird die Annahme eines kugelförmigen Hohlraumes in unserem Falle, in dem wir annehmen, dass die Suszeptibilität sehr gross ist, recht schlecht sein. Eine grosse diamagnetische Suszeptibilität verlangt nämlich eine sehr ausgedehnte Elektronenbahn, also ein flaches Gebilde. Es wird also jedenfalls eine den tatsächlich auftretenden Verhältnissen viel näher stehende Hypothese sein, wenn wir annehmen, dass unser fragliches Teilchen in einem in der Richtung der Kraftlinien sehr kurzen und in der darauf senkrechten Richtung sehr breiten Hohlraum liegt. In dessen Inneren haben wir bekanntermeise jedoch statt (5) für die Feldintensität

$$H_i = H + 4\,\pi M\,. \tag{6}$$

Beachten wir jetzt weiter, dass $M = \chi H$ und ausserdem, dass $M = \varkappa H_i N$ ist, so erhalten wir, wenn wir diese zwei Formeln einander gleichsetzen und dann für H_i (6) einführen,

$$\chi = \frac{\varkappa N}{1 - 4\pi\varkappa N}\,. \tag{7}$$

Wenn wir jetzt wieder den Grenzübergang durchführen, dass $\varkappa$ eine sehr grosse negative Zahl wird, so folgt statt (4)

$$\chi = -1/4\pi, \tag{8}$$

also tatsächlich ein mit (2) übereinstimmendes Resultat. Damit haben wir also bewiesen, dass man die Formel für die Suszeptibilität des absolut diamagnetischen Körpers auch aus der Theorie des gewöhnlichen Diamagnetismus herleiten kann.

Aus (7) würde folgen, dass man daraus (8) auch dann erhält, wenn $\varkappa$ eine sehr grosse jedoch positive Zahl ist und daraus würde weiter das absurde Resultat folgen, dass auch eine sehr grosse paramagnetische Suszeptibilität einen absoluten Diamagnetismus verursachen kann. Dieses Paradoxon ist jedoch nur scheinbar. Der Paramagnetismus ist nämlich ein Orientierungseffekt, bei dem Sättigungserscheinungen auftreten müssen, also sind in diesem Falle unsere Formeln nicht mehr anwendbar.

§ 2

Wir wollen jetzt im Zusammenhang mit unserem Problem die Fälle betrachten, in denen die diamagnetische Suszeptibilität (numerisch) verhältnismässig gross ist. Bei aliphatischen Verbindungen bewährt sich in guter Näherung die Regel, dass sich ihre diamagnetische Suszeptibilität additiv aus denen ihrer Atome oder Ionen aufbaut. Bei Kettenmolekülen (besonders bei solchen, in welchen keine oder nur wenig Doppelbindungen enthalten sind) ist die Suszeptibilität entlang der Achse (numerisch) ein ganz wenig grösser, senkrecht darauf dagegen ein wenig kleiner. Wir sehen also, dass aliphatische Verbindungen bezüglich unseres Problems ganz uninteressant sind.

Ganz anders sind jedoch die Verhältnisse bei aromatischen Verbindungen. In unserer Tabelle [8] bezeichnen wir mit χ_1 die Suszeptibilitäten eines Mols senkrecht zur Ebene der Moleküle (also in senkrechter Richtung auf die Benzolringe), mit χ_2 diejenigen entlang ihrer Längsachse und mit χ_3 deren Werte senkrecht zu dieser Längsachse in der Ebene des Moleküls.

	$\chi_1 \cdot 10^{+6}$	$\chi_2 \cdot 10^{+6}$	$\chi_3 \cdot 10^{+6}$
Benzol	— 91,2	—37,3	—37,3
Naphtalin	—187,2	—39,4	—43,0
Anthracen	—275,5	—45,9	—52,7

Wir sehen also, dass die diamagnetische Suszeptibilität dieser kondensierten aromatischen Verbindungen senkrecht zur Molekülebene abnormal gross ist. Nach dem Kekuléschen Modell enthält der Benzolring abwechselnd einfache und Doppelbindungen. Jetzt wissen wir es jedoch bereits aus quantenmechanischen Überlegungen, dass diese Doppelbindungen tatsächlich nicht vorhanden sind. Sie sind viel mehr entlang des ganzen Ringes verschmiert. Eben die diamagnetische Anisotropie ist ein interessanter Beweis für diese Auffassung. Die überzähligen Bindungen verursachenden Elektronen (die sogenannten π-Elektronen) bewegen sich frei entlang des Ringes, beschreiben also eine viel grössere Bahn als die ein Atom umkreisenden Elektronen. Aus der bekannten Formel von Pascal für die diamagnetische Suszeptibilität eines Mols

$$\chi_M = -\frac{e^2 L}{6mc^2} \sum_i \overline{R_i^2}, \tag{9}$$

in der alle Symbole die gewohnte Bedeutung haben, sehen wir gleich, dass, wenn die Elektronen eine grössere Bahn beschreiben, der absolute Wert der diamagnetischen Suszeptibilität dann ebenfalls viel grösser wird. In unserem Falle ist beim Anthracen diese Bahn am grössten, weil dieser aus drei zusammengewachsenen Benzolringen besteht.

Eigentlich ist unsere Formel (9) über alle Orientierungen der Elektronenbahnen gemittelt und bezieht sich deshalb auf Gase. In dem hier betrachteten Fall sollte man eigentlich statt (9) die noch nicht gemittelte Formel

$$\chi_M = -\frac{e^2 L}{4mc^2} \sum_i \overline{R_i^2} \tag{10}$$

benützen. Das ändert jedoch nichts am Wesen der Sache. Es wäre interessant, wenn uns auch die Suszeptibilitäten von aus noch mehr zusammenwachsenen Benzolringen aufgebauten Verbindungen, also z. B. von Pyren und Chrysen (4 Ringe), von Benzpyren und Perylen (5 Ringe) oder sogar von Coronen und Ovalen [9] bekannt wären. Doch scheinen solche Messungen abgesehen von der erwähnten einen Ausnahme nicht vorzuliegen. Nach unserer Auffassung repräsentieren die entlang der Ringe sich bewegenden Elektronen eigentlich einen »Suprastrom« von molekularer Dimension. Unmittelbar können wir denselben selbstverständlich nicht messen, weil wir ja an ein Molekül keine Elektroden anlegen können. Es ist noch interessant zu bemerken, dass dieser »Suprastrom« auch noch bei Zimmertemperatur fliesst. Das rührt daher, dass diese Benzolringe durch Valenzkräfte zusammengehalten werden, und die sind so stark, dass sie bei normaler Temperatur von der thermischen Bewegung noch nicht zerrissen werden.

Ausser den erwähnten Beispielen bezüglich der anomal grossen diamagnetischen Suszeptibilität gibt es noch zwei feste Körper, welche abnormal stark diamagnetisch sind. Der eine ist das Wismut und der andere der Graphit. Beim Wismut wird diese Erscheinung von der Brillouinzonenstruktur verursacht und ist deshalb bezüglich unseres Problems uninteressant. Der Graphit besitzt dagegen ein Schichtengitter und die ihn aufbauenden einzelnen senkrecht zur hexagonalen Achse orientierten Schichten bestehen eigentlich aus lauter zusammengewachsenen Benzolringen. Diese Schichten repräsentieren also eigentlich ein Extremum einer kondensierten aromatischen Verbindung. Ein Unterschied besteht jedoch in einem gewissen Sinne darin, dass die erwähnten überschüssigen Elektronen hier für die die einzelnen Schichten zusammenhaltenden Kräfte verantwortlich sind. Tatsächlich beträgt die diamagnetische Suszeptibilität des Graphits entlang der hexagonalen Achse (also senkrecht zur Ebene der Benzolringe) nach den Messungen von KRISHNAN und GANGULI [10] $-22 \cdot 10^{-6}$ und senkrecht zu dieser nur $-0{,}5 \cdot 10^{-6}$. Der Graphit ist also diamagnetisch extrem anisotrop, wie wir das nach unseren theoretischen Betrachtungen auch erwartet haben. Es sei nur noch erwähnt, dass die entlang der hexagonalen Achse gemessene Suszeptibilität etwas ungewiss ist, und von dem mehr oder weniger perfekten Aufbau des Kristalls abhängt, ganz im Einklange mit unserer Auffassung. Einige Autoren haben tatsächlich auch schon eine sechzigfache Anisotropie gemessen.

Die Leitfähigkeit des Graphits ist noch extremer anisotrop. Ebenfalls nach den Messungen von KRISHNAN und GANGULI [11] und auch nach anderen Autoren besitzt der Graphit senkrecht zur hexagonalen Achse (also in der Ebene der Benzolringe) eine zehntausendmal so grosse Leitfähigkeit als parallel zu dieser Achse. Deshalb haben schon mehrere Forscher das elektrische Verhalten des Graphits als das eines zweidimensionalen Metalls, als eines Halbleiters mit der Anregungsenergie gleich Null usw. zu deuten versucht.

Nach unserer Auffassung wäre dagegen der Graphit, wenn die sich in den von den Benzolringen zusammengesetzten Schichten bewegenden Elektronen nicht auch diese Schichten zusammenhalten würden, ein zweidimensionaler Supraleiter (und zwar auch bei normaler Temperatur). Infolge der erwähnten Ursache kann jedoch weder die diamagnetische Suszeptibilität noch die elektrische Leitfähigkeit ihren extremen Wert erreichen, jedoch sind aber beide anomal gross. Der Graphit ist nach unserer Auffassung eigentlich die extrem kondensierte aromatische »Verbindung«, anf dessen »Molekülen« wir schon tatsächlich Elektroden anlegen und so ihre Leitfähigkeit messen können.

§ 3

Jetzt wollen wir die Frage besprechen, ob wir die im vorangehenden erhaltenen Ergebnisse doch irgendwie zur Erklärung der bei den tatsächlichen Supraleitern auftretenden Verhältnissen heranziehen könnten. Die erste Frage, die wir besprechen müssen, ist die, wie sich die Supraleitungserscheinungen im periodischen System verteilen. Eine Zusammenfassung unserer diesbezüglichen Kenntnisse wurde von MATTHIAS [12] veröffentlicht. Wir wollen dagegen hier auf eine noch kaum beachtete Tatsache bezüglich der supraleitenden Elemente aufmerksam machen: Bekanntermeise kann man in einer gewissen Näherung doch Radien für die in einem Metall in die Elektronenflüssigkeit eingebetteten Metallionen angeben. Wenn wir jetzt von dem Atomvolumen des Metalls diese Volumina der Ionen subtrahieren, so erhalten wir den für die Leitungselektronen zur Verfügung stehenden Raum. Wenn wir den Letzteren noch durch die Zahl der Leitungselektronen pro Metallion dividieren, so erhalten wir den Begriff des Elektronenvolumens V_E im Metall. Wir entnehmen die in der folgenden Tabelle enthaltenen diesbezüglichen Daten aus dem Werke von BILTZ [13], die wir mit noch einer Spalte ergänzt haben, welche die Sprungpunkte T_c der Supraleiter aus dem Werke von SCHOENBERG [14] enthält.

Es wurde schon von MEISSNER und SCHUBERT [15] darauf aufmerksam gemacht, dass das Atomvolumen der Supraleiter zwischen engen Grenzen schwankt. Aus unserer Tabelle kann man dagegen die noch viel auffallendere Regel ablesen, dass die Elektronenvolumina der Supraleiter alle recht klein sind. Das bedeutet also, dass in ihnen die Ionen (Atomrümpfe) zueinander recht nahe kommen, oder anschaulich gesprochen, sich berühren. Bei den

Metall	Atomvolumen	Ionenvolumen	ΣV_E	V_E	Tc
Li	12,6	1,5	11	11	—
Na	22,8	6,5	16	16	—
K	43,1	16,0	27	27	—
Rb	53,1	20.0	33	33	—
Cs	65,9	26,0	40	40	—
Cu	7,0	5,0	2	2	—
Ag	10,1	9,0	1	1	—
Au	10,1	—	—	—	—
Be	4,8	0	5	2,5	—
Mg	13,8	2	12	6	—
Ca	25,6	6,5	19	9,5	—
Sr	33,2	11.0	22	11,0	—
Ba	37,3	16	21	10,5	—
Zn	8,9	3	6	3	0,91
Cd	12,7	6	7	3,5	0,56
Hg	13,8	8	6	3	4,152
Al	9,9	0	10	3,3	1,20
Sc	(15)	2	13	4,3	—
Y	20,2	6	14	4,7	—
La	22,8	8	14	4,7	4,37
Ga	11,7	2	10	3,3	1,10
In	15,3	4	11	3,7	3,37
Tl	16,9	6	11	3,7	2,38
Ti	10,7	1	10	2,5	0,53
Zr	13,9	2,5	11,5	2,9	0,70
Hf	13,4	2,5	11	2,8	0,35
Th	19,7	6	14	3,5	1,39
Ge	13,5	1	12,5	3,1	—
Sn	16,0	2	14	3,5	3,73
Pb	17,9	5	13	3,2	7,22
V	8,2	—	8	1,6	5,1
Nb	10,8	—	11	2,2	8,0
Ta	10,9	—	11	2,2	4,4

Elementen dagegen, die typisch keine Supraleiter werden (z. B. Alkalien und Erdalkalien), sind diese Ionen weit voneinander entfernt in die Elektronenflüssigkeit eingebettet. Selbstverständlich gibt es auch Ausnahmen von dieser Regel, so werden z. B. von den Metallen mit nicht oder gerade nur abgeschlossenen *d*-Schalen nicht alle Supraleiter, was sie nach der hier gefundenen Regel sein sollten. Diese Ausnahmen haben jedoch recht plausible physikalische Ursachen. Ausserdem sind ja die Ionenradien garnicht gut definierte Grössen, und es ist auch oft fraglich, welche Elektronen schon zu den Leitungselektronen rechnen und welche noch zu den Atomrümpfen usw. Wir wollen jedoch auf diese Einzelheiten hier nicht näher eingehen.

§ 4

Tatsache ist jedoch, dass diese erwähnten Ionen (oder vielleicht richtiger gesagt Atomrümpfe) polarisierbar sind. Kommen sie also einander recht nahe, so können sie sich gegenseitig beeinflussen. Um zu sehen, was dann geschieht, wollen wir eine lineare Kette von solchen polarisierbaren Teilchen betrachten. Nehmen wir an, dass diese Kette schon polarisiert ist und bezeichnen das in einem Teilchen induzierte Moment mit P. Auf ein Glied der unendlich lang angenommenen Kette wird dann infolge der Polarisiertheit der übrigen Teilchen die elektrische Feldintensität

$$\mathfrak{E} = \Sigma E = 2\frac{2P}{r^3}\left(1 + \frac{1}{2^3} + \frac{1}{3^3} + \frac{1}{4^3} \cdots\right) = 4\frac{1{,}202}{r^3}P \tag{11}$$

einwirken. Bezeichnen wir ausserdem die Polarisierbarkeit eines Kettengliedes mit α, so haben wir, wenn kein äusseres Feld vorhanden ist,

$$P = \alpha\mathfrak{E}. \tag{12}$$

Unser aus (11) und (12) bestehendes Gleichungssystem hat (abgesehen von dem physikalisch uninteressanten Fall, dass die Koeffizienten in den zwei Gleichungen gerade einander gleich sind) nur die Lösungen $P = \mathfrak{E} = 0$ und $P = \mathfrak{E} = \infty$. Das erste Lösungssystem entspricht dem gewöhnlichen Fall, dass ohne ein äusseres Feld die Kette nicht polarisiert ist. Der zweite Fall entsteht dagegen dann, wenn die polarisierbaren Teilchen zueinander so nahe kommen, dass eine »Polarisationskatastrophe« eintritt. Selbstverständlich kann die Erscheinung nicht so krass werden, weil ja der lineare Zusammenhang (12) nur bis zu einer gewissen Feldintensität angenähert gültig sein kann. Das Auftreten der »Polarisationskatastrophe« drückt jedoch den physikalischen Sachverhalt aus, dass die Ionenladungen in diesem Fall zusammenfliessen. Es wurde jedoch schon von FRENKEL darauf aufmerksam gemacht, dass eine unendlich grosse Polarisierbarkeit mit einer unendlich grossen Leitfähigkeit äquivalent ist. Eigentlich sollte man in (11), weil die Atomrümpfe zueinander ganz nahe kommen, auch noch höhere Glieder berücksichtigen. Diese Frage wollen wir jedoch in einer weiteren Arbeit besprechen,

Wenn wir also tatsächlich nach dem vorher Gesagten annehmen, dass in den Supraleitern solche bis zur Sättigung polarisierte Ionenketten entstehen, so können wir auf die dieselben Gedankengänge anwenden, mit Hilfe deren wir bei der Besprechung der Atomketten der aromatischen Verbindungen gezeigt haben, dass dort zu dem absoluten Diamagnetismus und der Supraleitung ähnliche Erscheinungen auftreten. In unseren Polarisationsketten werden diese aber tatsächlich voll entwickelt sein. Die Atomketten in den kondensierten aromatischen Verbindungen werden jedoch durch starke Valenzkräfte zusammengehalten, welche bei normaler Temperatur die thermische Bewegung noch nicht zerreisen kann, unsere in den Supraleitern auftretenden

Polarisationsketten werden dagegen nur durch schwache Polarisationskräfte zusammengehalten und es ist deshalb leicht begreiflich, dass diese schon am Sprungpunkt der Supraleiter von der thermischen Bewegung zerstört werden. Berücksichtigen wir noch, dass wie wir dies schon erwähnt haben, die normale diamagnetische Suszeptibilität von einem cm^3 fester Materie die Grössenordnung -10^{-7} besitzt, die Suszeptibilität der absolut diamagnetischen Körper dagegen $\chi = -1/4\pi$ ist, so sehen wir unter Berücksichtigung der Formel (10), dass es zur Erklärung des Auftretens des absoluten Diamagnetismus genügt, wenn unsere Polarisationsketten auf einige hundert Atome ausgedehnt angenommen werden.

Zum Schluss wollen wir noch zeigen, wie einfach z. B. der Isotopeneffekt aus der neuen Theorie folgt. Die Wahrscheinlichkeit der thermischen Anregung am Sprungpunkt T_c beschreibt der Boltzmann-faktor

$$e^{-\frac{h\nu}{kT_c}}. \tag{13}$$

Andererseits ist die Schwingungsfrequenz eines Oszillators so klassisch wie auch quantenmechanisch

$$\nu = \frac{1}{2\pi}\sqrt{\frac{\overline{k^2}}{M}}. \tag{14}$$

Setzen wir jetzt (14) in (13) ein und berücksichtigen, dass bei den zu demselben Element gehörenden verschiedenen Isotopen am Sprungpunkt die Wahrscheinlichkeiten (der Zerreissung der Polarisationsketten) einander gleich sein müssen, so folgt

$$\sqrt{M} \cdot T_c = \text{konst}, \tag{15}$$

also

$$T_c \,\text{prop} \cdot M^{-1/2}. \tag{16}$$

LITERATUR

1. C. Benedicks, Ann. d. Phys. (5), **17,** 169, 1933; Zeitschr. f. Metallk., **25,** 197, 1933.
2. R. de L. Kronig, Zeitschr. f. Phys., **78,** 744, 1933; **80,** 203, 1933.
3. W. Heisenberg, Zeitschr. f. Naturf., **2***a*, 185, 1947; Ann. d. Phys. (6), **3,** 289, 1948.
4. H. London und F. London, Proc. Roy. Soc. A, **149,** 71, 1935; Physica, **2,** 341, 1935.
5. M. v. Laue, Theorie der Supraleitung, 2. Aufl. Springer Verlag. Berlin—Göttingen—Heidelberg, 1949.
6. J. Bardeen, L. N. Cooper und J. R. Schrieffer, Phys. Rev., **106,** 162, 1957; Phys. Rev., **108,** 1175, 1957; Vgl. auch C. J. Gorter, Progress in Low Temperature Physics, Vol. 3, North-Holland. Publ. Comp. Amsterdam, 1961. S. 170. Artikel von J. Bardeen und J. R. Schrieffer.
7. E. Maxwell, Phys. Rev., **78,** 477, 1950; C. A. Reynolds, B. Serin, W. H. Wright und L. B. Nesbitt, Phys. Rev., **78,** 487, 1950.

8. W. KLEMM, Magnetochemie, Akademische Verlagsges. Leipzig, 1936. S. 198; vgl. auch P. W. SELWOOD, Magnetochemistry, Interscience Publishers, New York, 1943. S. 66.
9. H. AKAMATU, H. INOKUCHI und T. HANDA, Nature, **168**, 520, 1951.
10. K. S. KRISHNAN und N. GANGULI, Proc. Roy. Soc. A, **177**, 168, 1941.
11. N. GANGULI und K. S. KRISHNAN, Nature, **144**, 667, 1939. Vgl. auch A. R. UBBELOHDE und F. A. LEWIS, Graphite and its Crystal Compounds, Clarendon Press, Oxford, 1960.
12. C. J. GORTER, Progress in Low Temperature Physics, Vol. 2, North-Holland Publ. Comp., Amsterdam, 1957. S. 138. Artikel von B. T. MATTHIAS.
13. W. BILTZ, Raumchemie der festen Stoffe, L. Voss, Leipzig, 1934. S. 207.
14. D. SHOENBERG, Superconductivity, University Press, Cambridge, 1952.
15. W. MEISSNER und G. SCHUBERT, Bayerische Akad. d. Wiss. Sitzungsber. der. math.-naturwiss. Klasse, München, 1943, S. 195.

О ПРОБЛЕМЕ АБСОЛЮТНОГО ДИАМАГНЕТИЗМА И СВЕРХПРОВОДИМОСТИ

Т. НАЙГЕБАЕР

Резюме

В рамках всей созданной до настоящего времени теории сверхпроводимости сделалась попытка для объяснения бесконечно большой проводимости. Свойства сверхпроводников вытеснять из своих внутренних частей магнитные трубные линии, то-есть явление абсолютного диамагнетизма, удалось вводить в теорию либо с помощью дальнейших гипотез, либо путём сложнейших математических выводов удалось доказать, или, по крайней мере, представлять вероятным, что разработанная теория сверхпроводимости способна объяснять явление абсолютного диамагнетизма. В противоположность этому данная работа непосредственно исходит из эффекта Мейсснера—Оксенфельда. Предметом первоначального исследования в ней является вопрос о возможности перехода между обычными диамагнитными явлениями и абсолютным диамагнетизмом сверхпроводников. С этой точки зрения особенным вниманием исследуется поведение конденсированных ароматических соединений и графита. Далее говорится об одном новом правиле распределения явлений сверхпроводимости в периодической системе элементов. Из этого правила и результатов, полученных в первой части статьи для диамагнетизма, вытекает новая теория для объяснения явлений абсолютного диамагнетизма и сверхпроводимости.

VERBUNDGRAPHENDARSTELLUNG DES S-OPERATORS BEIM QUANTENMECHANISCHEN MEHRTEILCHENPROBLEM

Von

W. MACKE und P. ZIESCHE

INSTITUT FÜR THEORETISCHE PHYSIK DER TECHNISCHEN UNIVERSITÄT, DRESDEN, DDR

(Eingegangen: 20. IX. 1963)

Der S-Operator — und damit auch der Grundzustand Φ, E und die Einteilchen-Ausbreitungsfunktion $G(x, x')$ — eines quantenmechanischen Mehrteilchensystems $H = H^0 + V$ mit quasistatisch ein- und ausgeschalteter Wechselwirkung V wird in einfacher Weise ganz allgemein durch nur zusammenhängende Graphen (Verbundgraphen) dargestellt.

Einleitung

Zur Behandlung quantenmechanischer Mehrteilchenprobleme wird oft der S-Matrix-Formalismus herangezogen [1—6]. Das System wird dabei beschrieben durch $H = H^0 + V$ und enthält einen Anteil H^0 wechselwirkungsfreier Teilchen und eine störende Teilchenwechselwirkung V. Letztere denkt man sich gewöhnlich im Zeitablauf während einer Zeitdauer τ quasistatisch ein- und nach Ablauf eines langen Zeitintervalls $T \gg \tau$ wieder ausgeschaltet. Zu Beginn des Einschaltens möge sich das System in einem Eigenzustand des ungestörten Anteils H^0 befinden. Falls dieser Eigenzustand von H^0 nicht entartet ist — eine Voraussetzung, die oft nur auf den Grundzustand Φ^0, E^0 zutrifft — und falls durch V nicht entsprechende Erhaltungssätze verletzt werden, entwickelt sich im Zeitablauf auf Grund der Schrödingergleichung aus dem ungestörten Grundzustand Φ^0, E^0 bei quasistatischen Einschalten der gestörte Grundzustand Φ, E. Der Zeitablauf wird einfach durch den sogenannten S-Operator beschrieben, der wiederum wie üblich durch eine alle möglichen virtuellen Prozesse wiedergebende Graphensumme übersichtlich abgekürzt wird. Da nun Φ und E sowie auch beliebige Erwartungswerte im Zustand Φ, insbesondere Dichtematrizen und Ausbreitungsfunktionen, durch den S-Operator bestimmt sind, werden auch alle diese interessierenden Grössen durch Graphen dargestellt. Im folgenden wird der S-Operator allgemein durch zusammenhängende Graphen (= Verbundgraphen) dargestellt, woraus sofort die bereits von GELL-MANN und LOW [7], HUBBARD [2] u.a. angegebene Verbundgraphendarstellung von Φ, E, aber auch beliebiger Erwartungswerte im Zustand Φ folgt.

1. Exponentialdarstellung von S

Der S-Operator eines beliebigen Systems $H = H^0 + V$ vermittelt bekanntlich in der Schrödingerdarstellung folgendermassen die Entwicklung eines beliebigen Anfangszustandes $\Psi(t)$:

$$\Psi(t) = U(t,t')\,\Psi(t'), \quad U(t,t') = e^{-\frac{i}{\hbar}H^0 t} S(t,t')\, e^{\frac{i}{\hbar}H^0 t'}. \tag{1}$$

Da $\Psi(t)$ im Zeitablauf die Schrödingergleichung erfüllt, ergibt sich für S die Darstellung als zeitgeordnetes Produkt

$$S(t,t') = P e^{-\frac{i}{\hbar}\int_{t'}^{t} dt''\, V(t'')}, \quad V(t) = e^{\frac{i}{\hbar}H_0 t} V e^{-\frac{i}{\hbar}H^0 t}. \tag{2}$$

Im Formalismus der Wellenquantelung hängt nun der S-Operator über die Teilchenwechselwirkung V von den Erzeugungs- und Vernichtungsoperatoren ψ^+, ψ bzw. nach deren Zerlegung $\psi^+ = \psi_+^+ + \psi^-$ und $\psi = \psi^+ + \psi^\pm$ von den Erzeugungs- und Vernichtungsoperatoren $\psi_\pm^+, \psi_\pm$ ab. Da für die neuen Vernichtungsoperatoren $\psi_\pm$ der ungestörte Grundzustand $\Phi^0 = >$ oder $<$ als Vakuum erscheint, also $\psi_\pm > = < \psi_\pm^+ = 0$ gilt, und da gerade die Anwendung von S auf Φ^0 interessiert, empfiehlt es sich, alle Vernichtungsoperatoren $\psi_\pm$ nach rechts zu tauschen, d. h. das sogenannte Normalprodukt herzustellen. Da die ψ_+ mit den $\psi_\pm^+$ nicht vertauschbar bzw. antivertauschbar sind, entstehen dabei alle möglichen Verknüpfungen der Wechselwirkungen, die gewöhnlich durch Graphen abgekürzt werden. So wird S aus einem zeitgeordneten Produkt (2) in eine Summe von Normalprodukten und damit in eine *Graphensumme* überführt (Wicksches Theorem [8]):

$$\begin{gathered} S = 1 + \Sigma\, G, \\ G = \int\int dx_1\, dx_1' \ldots dy_1\, dy_1' \ldots \psi_\pm^+(x_1)\ldots\psi_\pm^\pm(y_1)\ldots \\ \ldots f_\Gamma(x_1 x_1' \ldots) f_{\Gamma'}(y_1 y_1' \ldots) \ldots \psi_\pm(y_1')\ldots\psi_\pm(x_1')\ldots \\ NG = G, \quad NS = S. \end{gathered} \tag{3}$$

G besteht im allgemeinen aus mehreren voneinander völlig unabhängigen, miteinander nicht verknüpften, aber auch nicht weiter in einfachere Graphen zerlegbare *Verbundgraphen* Γ. Ein bestimmter Verbundgraph Γ liefert einschliesslich aller seiner nichtverknüpften Potenzen zum S-Operator den Beitrag

$$S_\Gamma = NS_\Gamma = N\sum_{n=0}^{\infty}\frac{\Gamma^n}{n!} = Ne^\Gamma. \tag{4}$$

Dabei verhindert die Normalprodukt-Vorschrift N Verknüpfungen der Γ miteinander. Der Faktor $n!$ bedeutet die Zahl der Permutationen der n gleichen

Verbundgraphen Γ und sorgt dafür, dass insgesamt jeder Graph genau einmal vorkommt. Für den gesamten S-Operator als Summe der nichtverknüpften Potenzprodukte aller Verbundgraphen ergibt sich damit

$$\boxed{S = NS = N\Pi S_\Gamma = N\Pi e^\Gamma = Ne^{\Sigma\Gamma}.} \tag{5}$$

(5) gibt die allgemeine Exponential- oder Verbundgraphen-darstellung des S-Operators wieder.

Bezeichnen die Γ^+ die Verbundgraphen, die zwar Erzeugungsoperatoren $\psi_\pm^+$ aber keine Vernichtungsoperatoren $\psi_\pm$ enthalten, die Γ^- die Verbundgraphen, die zwar Vernichtungsoperatoren $\psi_\pm$ aber keine Erzeugungsoperatoren $\psi_\pm^+$ enthalten, und die Γ_0 die Vakuumverbundgraphen, so liefert die Anwendung von S auf den Vakuumzustand

$$<S> = e^{<\Sigma\Gamma>} = e^{\Sigma\Gamma_0}, \quad S> = a^{\Sigma\Gamma_0+\Sigma\Gamma^+}, \quad <S = <e^{\Sigma\Gamma_0+\Sigma\Gamma^-}. \tag{6}$$

2. Verbundgraphendarstellung von Φ und E

Wählt man in (1,1) als Anfangszustand $\Psi(-\infty) = \Phi^0$ undschaltet die Wechselwirkung mit der in Abb. 1 wiedergegebenen Funktion $\xi(t)$ ein (und aus), so entwickelt sich im Zeitablauf aus dem ungestörten Grundzustand Φ^0, E^0 bei quasistatischen Einschalten $\tau \to \infty$ der gestörte Grundzustand Φ, E. Umgekehrt entwickelt sich beim Ausschalten wieder Φ^0, E^0 aus Φ, E. Dabei ist vorausgesetzt, dass diese Zustände nicht entartet sind und dass die Wechselwirkung die Gültigkeit entsprechender Erhaltungssätze garantiert die Übergänge in andere Zustände verhindern.

Für $T \gg \tau \to \infty$ gilt also

$$U\left(\pm\frac{\mathbf{T}}{2}, \quad \pm\infty\right)\Phi^0 \to \Phi, \quad (\Phi,\Phi) = (\Phi^0,\Phi^0) = 1, \tag{1}$$

bis auf Phasenfaktoren $\exp(\pm i\,\Delta E\,\tau/\hbar)$, die wegen $\tau \ll T$ gegenüber den interferierenden Faktoren $\exp(\pm i\,\Delta E\,T/\hbar)$ zu vernachlässigen sind. Im Zeitintervall zwischen $-T/2$ und $+T/2$ ist V voll eingeschaltet, also H nicht mehr zeitabhängig, so dass

$$U(t_1,t_2) = e^{-\frac{i}{\hbar}H(t_1-t_2)} \quad \text{für} \mid t_{1,2} \mid < \frac{\mathbf{T}}{2} \tag{2}$$

gilt. Da U gemäss (1,1) einfach mit S zusammenhängt, lassen sich Φ und E

durch S ausdrücken. Aus (1) und (2) folgt mit $U(0, \pm\infty) = U(0, \pm T/2) \cdot U(\pm T/2, \pm\infty)$ sofort

$$S(0, \pm\infty)\,\Phi^0 = \\ = U(0, \pm\infty)\, e^{-\frac{i}{h}E^0\left(\pm\frac{T}{2}\pm\tau\right)}\Phi^0 \to e^{-\frac{i}{\hbar}E^0\left(\pm\frac{T}{2}\right)} e^{\pm\frac{i}{\hbar}E\frac{T}{2}}\Phi = e^{\pm\frac{i}{\hbar}\Delta E\frac{T}{2}}\Phi. \quad (3)$$

Dabei bedeutet $\Delta E = E - E^0$ die Energieverschiebung des Grundzustands. Durch Linksmultiplikation mit Φ^0 lässt sich der Phasenfaktor eliminieren:

$$< S(0, \pm\infty) > \to e^{\pm\frac{i}{h}\Delta E\frac{T}{2}}(\Phi^0, \Phi), \quad (4)$$

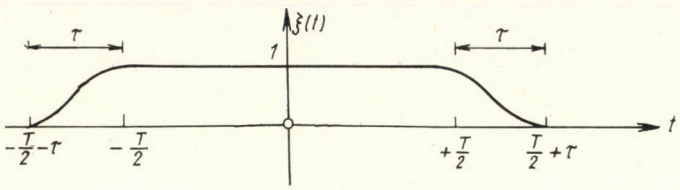


Abb. 1. Einschaltfunktion $\xi(t)$

so dass für Φ endgültig

$$\boxed{\frac{S(0, \pm\infty)}{< S(0, \pm\infty) >}\Phi^0 \to \frac{\Phi}{(\Phi^0, \Phi)} \equiv \Phi', \quad (\Phi^0, \Phi') = 1} \quad (5)$$

entsteht. Die Anwendung von $S \equiv S(\infty, -\infty)$ auf Φ^0 dagegen liefert

$$S\Phi^0 = e^{\frac{i}{h}H^0\left(\frac{T}{2}+\tau\right)} U e^{-\frac{i}{\hbar}\left(\frac{T}{2}+\tau\right)}\Phi^0 \to e^{\frac{i}{\hbar}E^0T} e^{-\frac{i}{\hbar}ET}\Phi^0 = e^{-\frac{i}{\hbar}\Delta ET}\Phi^0, \quad (6)$$

woraus sich für die Energieverschiebung ΔE

$$\boxed{< S > \to e^{-\frac{i}{\hbar}\Delta ET}} \quad (7)$$

ergibt. Ein Vergleich von (5) und (7) mit (1,6) zeigt: Die Energieverschiebung ΔE ist bis auf einen Faktor $-\hbar/iT$ durch die Vakuumverbundgraphen $\Sigma\Gamma_0$ gegeben:

$$\Delta E = -\frac{\hbar}{i}\frac{\Sigma\Gamma_0}{r}. \quad (8)$$

Weiter wird der Zustand Φ als Rechtsvektor nur durch die Verbundgraphen $\Sigma\Gamma^{+}$, die zwar $\psi_{\pm}^{+}$ aber keine $\psi_{\pm}$ enhalten, und der Zustand Φ als Linksvektor nur durch die Verbundgraphen $\Sigma\Gamma^{-}$, die zwar $\psi_{\pm}$, aber keine $\psi_{\pm}^{+}$ enhalten bestimmt, siehe auch GELL-MANN und LOW [7], SUCHER [3], RODBERG [4], HUBBARD [2].

3. Verbundgraphendarstellung von Erwartungswerten, Dichtematrizen und Ausbreitungsfunktionen

Den Erwartungswert einer beliebigen physikalischen Grösse A im Zustand $\Phi = \gg$ oder $\ll$ kann man mit (2.5) schreiben als

$$\boxed{\ll A \gg = \frac{<S(\infty, 0)\, AS(0, -\infty)>}{<S>} = <TAS><S>^{-1},\ S \equiv S(\infty, -\infty),} \tag{1}$$

wobei $S(t, t') = S^{+}(t, t')$ verwendet ist und T den Operator $A = A(0)$ in das Zeitintervall der S-Matrix hineinordnet. Falls A nur aus Einteilchenoperatoren $a(x', x) = \delta(t' - \varepsilon)\, a(\mathfrak{x}', \mathfrak{x})\, \delta(t)$ mit $\varepsilon \geqslant 0$ aufgebaut ist, lauten Operator A und Erwartungswert $\ll A \gg$ ausführlich

$$\begin{aligned} A &\equiv \iint d\mathfrak{x}'\, d\mathfrak{x}\, \psi^{+}(\mathfrak{x}')\, a(\mathfrak{x}', \mathfrak{x})\, \psi(\mathfrak{x}) = \iint dx'\, dx\, \psi^{+}(x')\, a(x', x)\, \psi(x), \\ \ll A \gg &= \mathrm{Sp}\, a\, \varrho = i\mathrm{Sp}\, a\, G, \end{aligned} \tag{2}$$

wobei *Einteilchendichtematrix* ϱ und *Einteilchenausbreitungsfunktion* G des Zustandes ähnlich (1) durch die S-Matrix ausgedrückt werden können:

$$\boxed{\begin{gathered} \varrho(\mathfrak{x}, \mathfrak{x}') = \ll \psi^{+}(\mathfrak{x}')\, \psi(\mathfrak{x}) \gg = <T\, \psi^{+}(x')\, \psi(x)\, S><S>^{-1},\ t' = t + 0 \\ G(x, x') = i \ll T\left(e^{\frac{i}{\hbar} Ht}\, \psi(\mathfrak{x})\, e^{-\frac{i}{\hbar} Ht}\right)\left(e^{\frac{i}{\hbar} Ht'}\, \psi^{+}(\mathfrak{x}')\, e^{-\frac{i}{\hbar} Ht'}\right) \gg = \\ = i <T\psi(x)\, \psi^{+}(x')\, S><S>^{-1}. \end{gathered}} \tag{3}$$

In (2) enthält die zweite Spurbildung auch die entsprechenden Zeitintegrationen. Bei fehlender Teilchenwechselwirkung, d. h. $V \to 0$, also $H \to H^{0}$, $\Phi \to \Phi^{0}$, $S \to 1$ entstehen jeweils die ungestörten Ausdrücke $\ll A \gg^{0} = <A>$, $\varrho^{0}(\mathfrak{x}, \mathfrak{x}')$, $G^{0}(x, x')$.

Mit der Verbundgraphendarstellung (1,5) der S-Matrix ist die Graphendarstellung von G, und damit auch von ϱ und $\ll A \gg$ leicht ablesbar:

$$G(x, x') = i <T\psi(x)\, \psi^{+}(x')\, N e^{\Sigma\Gamma}> e^{-\Sigma\Gamma_0} = i <T\psi(x)\, \psi^{+}(x')\, N(1 + \Sigma\Gamma_1)>. \tag{4}$$

Dabei liefern nach dem WICKschen Theorem nur solche Terme Γ_1 der Graphensumme $S <S>^{-1} = N \exp(\Sigma\Gamma - \Sigma\Gamma_0)$ Beiträge, die gerade ein Operatorenpaar $\psi^+ \psi$ enthalten, also alle Graphen, die durch Aufschneiden von Vakuumverbundgraphen entstehen:

$$\Gamma_1 = N \int\int dx_1\, dx_2\, \psi^+(x_1)\, i\, \Gamma_1(x_1, x_2)\, \psi(x_2). \tag{5}$$

(5) in (4) eingesetzt liefert die *Verbundgraphendarstellung von G* (siehe auch z. B. BOGOLJUBOW und SHIRKOW [9]):

$$\boxed{G(x, x') = G^0(x, x') + \int\int dx_1\, dx_2\, G^0(x, x_1)\, \Sigma\Gamma_1(x_1, x_2)\, G(x_2, x').} \tag{6}$$

Damit ist auch ϱ und $\ll A \gg$ allein durch Verbundgraphen dargestellt:

$$\begin{aligned} \varrho(\mathfrak{r}, \mathfrak{r}') &= \varrho^0(\mathfrak{r}, \mathfrak{r}') + \int\int dx_1\, dx_2\, G^0(\mathfrak{r}, t; x_1)\, \Sigma\Gamma(x_1, x_2)\, G^0(x_2; \mathfrak{r}', t) + \ldots, \\ \ll A \gg &= <A> + i\, \mathrm{Sp}\, a\, G^0\, \Sigma\Gamma_1\, G^0 + \ldots, \\ i\,\Gamma_1(x_1, x_2) &= -\,\delta(x_1 - x_2) \int dx_1'\, v(x_1, x_1')\,(-i)\, G^0(\mathfrak{r}_1', t_1'; \mathfrak{r}_1'; t_1' + 0) + \\ &\quad + v(x_1, x_2)\,(-i)\, G^0(x_1, x_2) + \ldots. \end{aligned} \tag{7}$$

Bis zur ersten Ordnung hat Γ_1 die angegebene Form.

Die Graphen der Einteilchen-Ausbreitungsfunktion hängen übrigens topologisch in einfacher Weise mit den Vakuumverbundgraphen zusammen. Würde in (3) nämlich nicht der Operator $T_{\psi\psi'^+}$ auf die S-Matrix wirken, so entstünden bei der Erwartungswertbildung die Vakuumgraphen $<S>$. Der Operator $T_{\psi\psi'^+}$ bewirkt nun, dass in dem aus Teilchenwechselwirkungen $v(x, x')$ und ungestörten Ausbreitungsfunktionen $G^0(x, x')$ aufgebauten $<S>$ der Reihe nach an allen möglichen Stellen je eine Ausbreitungsfunktion G^0 zerschnitten, d. h. gemäss $G^0(1, 2) = i <T_{\psi_1\psi_2^+}> \to i_{\psi_1\psi_2^+}$ ersetzt wird. Die Graphen der Einteilchen-Ausbreitungsfunktion entstehen aus den Vakuumverbundgraphen also durch sukzessives Zerschneiden der ungestörten Ausbreitungsfunktionen G^0, was durch die Gleichung

$$\frac{\delta \ln <S>}{\delta G^0(x_2, x_1)} = \Sigma\Gamma_1(x_1, x_2) \tag{8}$$

zum Ausdruck kommt.

Schliesslich sei noch eine weitere einfache topologische Eigenschaft der Vakuumverbundgraphen Γ_0 und der Einteilchengraphen Γ_1 erwähnt. Bekanntlich enthält jeder Graph Γ einen Symmetriefaktor $1/g(\Gamma)$. Dabei stellt $g(\Gamma)$ die Zahl von Permutationen der Integrationsvariablen (= Bezeichnungen innerer Punkte) dar, die den Graphen Γ völlig invariant lassen. Da nun die Graphen von $G(x, x')$ äussere Punkte besitzen, deren Bezeichnungen x, x'

nichtintegrierte Variable bedeuten, können diese Graphen keinerlei Symmetrie aufweisen, d. h. keine Permutation von Bezeichnungen innerer Punkte kann diese Graphen in sich spiegeln, so dass für die Symmetriefaktoren der Einteilchengraphen $g(\Gamma_1) = 1$ gilt. Daraus folgt in Verbindung mit (8), dass beim sukzessiven Zerschneiden eines Vakuumverbundgraphen Γ_0, der einen Symmetriefaktor $1/g(\Gamma_0)$ enthält, offenbar gerade an genau $g(\Gamma_0)$ gewissen Stellen ein bestimmter Graph Γ_1 entsteht. Da beim nichtrelativistischen Mehrteilchenproblem ein Vakuumverbundgraph Γ_0 der Ordnung $r(\Gamma_0)$ gerade $2r(\Gamma_0)$ Ausbreitungsfunktionen G_0 (= Zerschneidemöglichkeiten) besitzt, entstehen aus einem Γ_0 genau $n(\Gamma_0) = 2r(\Gamma_0)/g(\Gamma_0)$ verschiedene Einteilchengraphen $\Gamma_1, \Gamma_1', \Gamma_1'', \ldots, \Gamma^{(n(\Gamma_0))}$, aber jeder genau $g(\Gamma_0)$ mal:

$$n(\Gamma_0)\, g(\Gamma_0) = 2r(\Gamma_0)\,. \tag{9}$$

Daraus folgt anderseits für die Symmetriefaktoren der Vakuumverbundgraphen die Eigenschaft

$$n(\Gamma_0) \leqslant g(\Gamma_0) \leqslant 2r(\Gamma_0)\,. \tag{10}$$

Falls Γ_0 keinerlei Symmetrie aufweist, gilt $g(\Gamma_0) = 1$, falls Γ_0 die in der Ordnung $r(\Gamma_0)$ grösstmögliche Symmetrie aufweist, gilt $g(\Gamma_0) = 2r(\Gamma_0)$. Bei teilweiser Symmetrie liegt $g(\Gamma_0)$ zwischen 1 und 2 $r(\Gamma_0)$.

Die hier angegebene Verbundgraphendarstellung der S-Matrix, des Grundzustandes Φ, E sowie der Ausbreitungsfunktion $G(x, x')$ erleichtert konkrete störungstheoretische Untersuchungen.

LITERATUR

1. J. Goldstone, Proc. Roy. Soc., **239**, 267, 1957.
2. J. Hubbard, Proc. Roy. Soc. *A*, **240**, 539, 1957; **243**, 336, 1958; **244**, 199, 1958.
3. J. Sucher, Phys. Rev., **107**, 1448, 1957.
4. L. Rodberg, Phys. Rev., **110**, 277, 1958.
5. A. Klein und R. Prange, Phys. Rev., **112**, 994, 1958; **121**, 950 und 957, 1961.
6. F. D. Du Bois, Ann. Phys. (New York), **7**, 174, 1959; **8**, 24, 1959.
7. M. Gell—Mann und F. Low, Phys. Rev., **84**, 350, 1951.
8. G. C. Wick, Phys. Rev., **80**, 268, 1950.
9. N. N. Bogoljubow und D. V. Shirkow, »Einf. in die Quantumfeldtheorie«, Staatsverl. f. Techn.-Theor. Lit., Moskau, 1957.

СВЯЗНО-ГРАФИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ S-ОПЕРАТОРОВ В СЛУЧАЕ КВАНТОВОМЕХАНИЧЕСКОЙ ПРОБЛЕМЫ МНОГИХ ЧАСТИЦ

В. МАККЕ и П. ЦИШЕ

Резюме

S-оператор данной квантогомеханической системы многих частиц — вместе с тем и основное состояние Φ, E и одночастичная обобщённая функция $G(x, x')$ — $H = H^0 + V$ с квазистатическим включённым и выключенным взаимодействиями V представляется в общем случае простым методом через связых графиков.

A FIELD-THEORETICAL METHOD IN THE THEORY OF SUPERCONDUCTIVITY

By

A. ZAWADOWSKI

CENTRAL RESEARCH INSTITUTE FOR PHYSICS, BUDAPEST

and

G. PÓCSIK

INSTITUTE FOR THEORETICAL PHYSICS, ROLAND EÖTVÖS UNIVERSITY, BUDAPEST

(Presented by Z. Gyulai. — Received 20. IX. 1963)

The exact equations concerning the spectral functions of the simplest correlation functions appearing in the theory of superconductivity are derived. As an application it is shown that in the pole-approximation the equations lead to GORKOV's results.

§ 1. Introduction

One of the simplest and most convenient methods for the treatment of the properties of a system of fermions with attractive interaction has been given by GORKOV [1]. The essence of GORKOV's method is that on the basis of the equations of motion he establishes a coupled system of equations for the various one-particle Green's functions, which can be solved exactly if only the two-particle Green's functions are considered as the products of suitable one-particle Green's functions.

In the present paper a method is described which is based on the equations of motion and the commutation relations as is the method by GORKOV, but here use is made of integral relations concerning the spectral functions of the various correlation functions. These integral relations are exact ones and approximation need only be applied in the last step of the calculation. Namely, if we assume that only one stable excited state contributes to the spectral functions, we are led to GORKOV's results.

In the corresponding quantum field-theoretical case, the same method leads to a self-mass equation like the gap equation [2].

§ 2. Equations for the spectral functions

In the present paper the model characterized by the Hamiltonian

$$H = \int \left[- \left(\psi^+ (x) \frac{\Delta}{2m} \psi(x) \right) + \frac{g}{2} \left(\psi^+ (x) \left(\psi^+ (x) \psi(x) \right) \psi(x) \right) \right] d^3x \qquad (1)$$

is examined. The round brackets denote spin summations and $g < 0$. Eq. (1) is completed by the usual canonical commutation rules

$$\begin{aligned} &\{\psi_\alpha(x), \psi^+_\beta(x')\} = \delta_{\alpha\beta}\,\delta(x - x'), \\ &\{\psi_\alpha(x), \psi_\beta(x')\} = 0. \end{aligned} \qquad (x_0 = x'_0) \tag{2}$$

The Hamiltonian (1) provides the equations of motion

$$\begin{aligned} \left(i\frac{\partial}{\partial x_0} + \frac{\Delta}{2m}\right)\psi(x) &= g(\psi^+(x)\psi(x))\,\psi(x), \\ \left(i\frac{\partial}{\partial x_0} - \frac{\Delta}{2m}\right)\psi^+(x) &= -g\psi^+(x)\,(\psi^+(x)\psi(x)). \end{aligned} \tag{3}$$

In the following various correlation functions will be defined. First of all, we introduce two correlation functions of one-particle structure

$$K_{\alpha\beta}(x, x') = <N|\,\psi^+_\alpha(x)\psi_\beta(x')\,|N> = \delta_{\alpha\beta}K(x, x'), \tag{4.a}$$

$$K'_{\alpha\beta}(x, x') = <N|\,\psi_\beta(x')\,\psi^+_\alpha(x)\,|N> = \delta_{\alpha\beta}K'(x, x'), \tag{4.b}$$

where $|N>$ is the ground state of the system of fermions. Let the particle number N be even. All the calculations are confined to zero temperature. We must introduce also the correlation function

$$K^{(+)}_{\alpha\beta}(x, x') = <N|\,\{\psi^+_\alpha(x), \psi_\beta(x')\}\,|N> = \delta_{\alpha\beta}K^{(+)}(x, x'). \tag{5}$$

Besides, two anomalous correlation functions are defined by

$$\begin{aligned} F^{(+)}_{\alpha\beta}(x, x') &= <N-2|\,\{\psi_\alpha(x), \psi_\beta(x')\}\,|N>, \\ F_{1\alpha\beta}(x, x') &= <N-2|\,\psi_\alpha(x)\,\psi_\beta(x')\,|N>. \end{aligned} \tag{6}$$

The functions $F^{(+)}$ and F_1 can be written in the form

$$\begin{aligned} F^{(+)}_{\alpha\beta}(x, x') &= e^{-2iE_0x_0}\,\tilde{F}^{(+)}_{\alpha\beta}(x - x'), \\ F_{1\alpha\beta}(x, x') &= e^{-2iE_0x_0}\,\tilde{F}_{1\alpha\beta}(x - x'). \end{aligned} \tag{7}$$

The reason for this is that $|N>$ and $|N-2>$ are not the same states, but they belong to different energies

$$\begin{aligned} H\,|N> &= E_N\,|N>, \\ H\,|N-2> &= E_{N-2}\,|N-2>, \\ E_N - E_{N-2} &= 2E_0. \end{aligned} \tag{8}$$

E_0 is just the usual chemical potential at zero temperature.

Now, we derive an equation for $K^{(+)}$. For this purpose let us form

$$\left(i\frac{\partial}{\partial x_0} - \frac{\Delta}{2m}\right) K^{(+)}_{\alpha\beta}(x, x')$$

By using (3) we have

$$\left(i\frac{\partial}{\partial t} - \frac{\Delta}{2m}\right) K^{(+)}_{\alpha\beta}(x, x') + \\ + < N | \{\psi_\alpha{}^+(x)(\psi^+(x)\psi(x)), \psi_\beta(x')\} | N > = 0. \tag{9}$$

The anticommutator in (9) can be easily simplified for $x_0 = x_0'$, with regard to (2)

$$\left[\left(i\frac{\partial}{\partial t} - \frac{\Delta}{2m}\right) K^{(+)}(x, x')\right]_{t=0} = -g\delta(\underline{x} - \underline{x}')\frac{\varrho}{2}, \tag{10}$$

$$(t = x_0 - x_0')$$

where ϱ denotes the density of the particles.

In the same way we get for $F^{(+)}$

$$\left[\left(i\frac{\partial}{\partial x_0} + \frac{\Delta}{2m}\right) F^{(+)}_{\alpha\beta}(x, x')\right]_{t=0} = -g\delta(\underline{x} - \underline{x}')F_{1\alpha\beta}(x, x'). \tag{11}$$

It is very expedient to transcribe (10) for spectral functions,

$$\int\left(-E + \frac{\underline{p}^2}{2m}\right)(J(\underline{p},E) + J'(\underline{p},E))dE = -\frac{g\varrho}{2}. \tag{12}$$

Here $J(\underline{p}, E)$, $J'(\underline{p}, E)$ means the Fourier transform of the spectral function $J(\underline{x}, \underline{x}', E)$, $J'(\underline{x}, \underline{x}', E)$ of $K(x, x')$, $(K'(x, x')$ and for example

$$K(x, x') = \int J(\underline{x}, \underline{x}', E)e^{iEt}\,dE. \tag{13}$$

Eq. (11) can be written as follows

$$\int\left(2E_0 - E - \frac{\underline{p}^2}{2m}\right) L^{(+)}_{\alpha\beta}(\underline{p},E)\,dE = \tag{14}$$

$$= -\frac{g}{(2\pi)^3}\int L_{1\alpha\beta}(\underline{p}', E)\,d^3\,\underline{p}'\,dE,$$

where we have taken into account eq. (7). $L^{(+)}(L_1)$denotes the Fourier transform of $\tilde{F}^{(+)}$ $(\tilde{F}_1)$.

Eqs. (12) and (14) can be further simplified by using the commutation relations. Namely, from (2), (5) and (6) the conditions

$$\int (J(\underline{p}, E) + J'(\underline{p}, E))\, dE = 1 \tag{15}$$

and

$$\int L^{(+)}(\underline{p}, E)\, dE = 0 \tag{16}$$

follow. Therefore, instead of (12) and (14) we can write

$$\int E\,(J(\underline{p}, E) + J'(\underline{p}, E))\, dE = \frac{\underline{p}^2}{2m} + g\frac{\varrho}{2} \equiv \varepsilon_p \tag{17}$$

and

$$\int E L^{(+)}_{\alpha\beta}(\underline{p}, E)\, dE = \frac{g}{(2\pi)^3}\int L_{1\alpha\beta}(\underline{p}', E)\, d^3\underline{p}'\, dE. \tag{18}$$

ε_p expresses the renormalization of the free particle energy.

Eqs. (15)—(18) can form the basis of the calculations leading to the determination of the excitation spectrum. In the quantum field theory such equations lead to the mass theorem.

§ 3. Application

In order to get some information about the excitation spectrum we must make use of approximate spectral functions.

First of all we note the explicit form of the operator $\psi_\alpha(x)$

$$\psi_\alpha(\underline{x}, x_0) = \frac{1}{(2\pi)^{3/2}}\int a_\alpha(\underline{p}. x_0) e^{-i\underline{p}\underline{x}}\, d^3\underline{p}, \tag{19}$$

then the Fourier transforms of the correlation functions (4. a—b) are $<N|a_\alpha^+(-\underline{p}, x_0)\, a_\alpha(-\underline{p}, x_0')|N>$ and $<N|a_\alpha(\underline{p}, x_0')a_\alpha^+(\underline{p}, x_0)|N>$. Now, we assume that

$$<N\,|\,a_\alpha^+(-\underline{p}, x_0)a_\alpha(-\underline{p}, x_0')\,|\,N> = \tag{20}$$

$$= <N\,|\,a_\alpha^+(-\underline{p}, x_0)\,|\,N-1,\ \underline{p}><N-1, \underline{p}\,|\,a_\alpha(-\underline{p}, x_0')\,|\,N>,$$

where $|\,N-1, \underline{p}>$ is an exact energy eigenstate containing one-particle excitation. Similarly,

$$\begin{aligned} <N\,|\,a_\alpha(\underline{p}, x_0')\,|\,N+1, \underline{p}><N+1, \underline{p}\,|\,a_\alpha^+(\underline{p}, x_0)\,|\,N> = \\ = <N\,|\,a_\alpha(\underline{p}, x_0')\, a_\alpha^+(\underline{p}, x_0)\,|\,N>. \end{aligned} \tag{21}$$

For a large system the following relations hold:

$$H\,|\,N-1,\underline{p}>=(E_N-E_0+E_p)\,|\,N-1,\underline{p}>,$$
$$H\,|\,N+1,\underline{p}>=(E_N+E_0+E_p)\,|\,N+1,\underline{p}>. \tag{22}$$

In consequence of the time-dependence of the operators, the approximation used here consists in writing

$$J(\underline{p},E)=v_p^2\,\delta(E-E_0+E_p),$$
$$J'(\underline{p},E)=u_p^2\,\delta(E-E_0-E_p). \tag{23}$$

Here

$$v_p=|\,v\pm_{p\alpha}|\,,\ v_{p\alpha}=<N-1,\underline{p}\,|\,a_\alpha(-\underline{p},0)\,|\,N>,$$
$$u_p=|\,u\pm_{p\alpha}|\,,\ u_{p\alpha}=<N+1,\underline{p}\,|\,a_\alpha^+(\underline{p},0)\,|\,N>. \tag{24}$$

Furthermore, (15) gives

$$u_p^2+v_p^2=1. \tag{25}$$

Also in the case of L_1, $L^{(+)}$ we make use of the above-mentioned approximation. Thus, for example

$$L_{1\,\frac{1}{2}-\frac{1}{2}}(\underline{p},E)=\tilde{u}^*_{p\frac{1}{2}}v_{p-\frac{1}{2}}\,\delta(E-E_0+E_p), \tag{26}$$

where

$$\tilde{u}_{p\frac{1}{2}}=<N-1,\underline{p}\,|\,a^+{}_{\frac{1}{2}}(\underline{p},0)\,|\,N-2>. \tag{27}$$

For a large system $|\,\tilde{u}_{p\frac{1}{2}}\,|=u_p$.

From the consideration of the time-dependence and (16) it is clear that

$$L^{(+)}_{\frac{1}{2}\,-\frac{1}{2}}(p,E)=\tilde{u}^*_{p\frac{1}{2}}v_{p-\frac{1}{2}}\big(\delta(E-E_0+E_p)-\delta(E-E_0-E_p)\big). \tag{28}$$

We see that the continuous parts of the spectral functions are neglected and only the contribution corresponding to the one-particle excitation is maintained.

Our equations contain definite statements regarding the parameters u_p, E_p. Eqs. (23), (25) and (17) give

$$u_p^2=\frac{1}{2}\left(1+\frac{\Sigma_p}{E_p}\right),$$
$$\Sigma_p=E_p-E_0. \tag{29}$$

On the other hand (18), (26) and (28) lead to the result

$$2E_p \tilde{u}^*_{p\frac{1}{2}} v_{p-\frac{1}{2}} = -\frac{g}{(2\pi)^3} \int \tilde{u}^*_{p'\frac{1}{2}} v_{p'-\frac{1}{2}} d^3\underline{p}' \, dE \equiv \Delta. \tag{30}$$

Here Δ is independent of $\underline{p}$, hence the relative phase of $\tilde{u}_{p\frac{1}{2}}$ and $v_{p-\frac{1}{2}}$ is also independent of $\underline{p}$. It follows that

$$2E_p u_p v_p = \frac{-g}{(2\pi)^3} \int u_{p'} v_{p'} d^3\underline{p}' = |\Delta|. \tag{31}$$

Furthermore, the excitation energy is

$$E_p = \sqrt{\Sigma_p^2 + |\Delta|^2} \tag{32}$$

and for the energy gap we get

$$1 = \frac{-g}{2(2\pi)^3} \int \frac{d^3\underline{p}}{\sqrt{|\Delta|^2 + \Sigma_p^2}} \, . \tag{33}$$

These results agree with GORKOV's. Hereby it is shown that in the pole-approximation the exact equations of the various simple spectral functions lead to the usual superconducting solutions.

The study of other problems is in progress.

REFERENCES

1. L. P. GORKOV, JETP, **34**, 735, 1958.
2. G. PÓCSIK, Nucl. Phys., **49**, 1963; Acta Phys. Hung., **17**, 103, 1964.

МЕТОД ТЕОРИИ ПОЛЯ В ТЕОРИИ СВЕРХПРОВОДИМОСТИ

А. ЗАВАДОВСКИ и ДЬ. ПОЧИК

Резюме

Выводятся точные уравнения спектральных функций простейших корреляционных функций, появляющихся в теории сверхпроводимости. В качестве их применения показывается, что в полюсном приближении эти уравнения приводят к результатам Горкова.

PAIRING CORRELATION IN THE NUCLEAR SURFACE LAYER

By

P. SZÉPFALUSY

INSTITUTE FOR THEORETICAL PHYSICS, ROLAND EÖTVÖS UNIVERSITY, BUDAPEST

(Presented by Z. Gyulai — Received 20. IX. 1963)

A nonuniform system of fermions in superfluid state is treated starting from the equations for the one-particle Green's function in the framework of the generalized Hartree-Fock method taking into account pairing correlation. Equations for determining the self-consistent pairing potential and expressions for the phase space density of tractable form are deduced with the aid of an approximation method which to the lowest order gives formulas of similar form to those of infinite systems, but the one-particle effective potential, and therefore all the other quantities too depend on the position. In that approximation the phase space density goes over into that of the Thomas-Fermi model if the pairing correlation is neglected.

By taking into account first- and second-order terms we derived inhomogeneity corrections both to the phase space density and the equations determining the pair correlation. The presence of the pair correlation plays an intrinsic role in the evaluation of the correction terms to the phase space density from the Green's function. Thus the pairing correlation may have considerable effect on the density distribution.

1. Introduction

Since the theory of superfluidity of fermion systems was proposed [1] it has widely been applied to nuclear problems by many authors. It was found that in the shell model of nuclei the effects of the residual pairing interaction are very important to the understanding of many of the nuclear properties [2]. On the other hand, using realistic nucleon-nucleon forces, the pairing correlation was studied for nuclear matter, too. BOGOLIUBOV [3] and COOPER, MILLS and SESSLER [4] treated a model Hamiltonian assuming interaction only between pairs of particles with equal and opposite momentum and taking into account the effect of the other interactions by a momentum-dependent single-particle potential which determines the independent-particle motion in the normal state of the system. The energy gap in nuclear matter as a function of density was calculated by EMERY and SESSLER [5]. They found that the gap is of considerable magnitude only at densities characteristic of the nuclear surface layer. Thus using the Thomas-Fermi gas model, one can estimate the magnitude of the pairing correlation for finite nuclei. The Thomas-Fermi approximation, however, is a poor one in the surface region.

The aim of the present paper is to derive inhomogeneity corrections to the equations of the Thomas-Fermi gas model in the presence of pairing

correlation. Thus also improved equations taking into account gradient terms, are obtained for determining the pairing correlation.

We start from the Green's function formulation of the many-particle problem which is well suited to the examination of correlation effects. In Section 2 we define the model and give the equations to be treated. Section 3 contains the zeroth order approximation: the Thomas-Fermi gas model generalized to include the pair correlation. The higher-order corrections are calculated in the next section. In the last section the results are discussed.

2. Formulation

The generalization of the Hartree-Fock approximation was developed by BOGOLIUBOV [6] considering the general form of the quasi-particle transformation. Investigations by VALATIN [7] have given deeper insight into the connection between the original and the generalized Hartree-Fock method. In the framework of the Green's function formalism the most convenient method to take into account the pairing correlation was pointed out by GORKOV [8] introducing in addition to the conventional one particle Green's function

$$G(1,1') = -i < T\,\psi(1)\,\psi^+(1') > , \tag{1}$$

the anomalous Green's function

$$F(1,1') = -i < T\,\psi^+(1)\,\psi^+(1') > \tag{2}$$

too. The numerical arguments denote the space coordinates, spin and time variables. (The isotopic spin is not considered here but it can be taken into account in a straightforward manner). The operators ψ and ψ^+ are the annihilation and creation operators, respectively, expressed in the Heisenberg representation. T indicates the time-ordered product.

In the framework of the generalized Hartree-Fock method the equations for G and F are as follows:

$$i\hbar \frac{\partial}{\partial t_1} G(1,1') = \hbar\,\delta(1-1') + \int d(2)\,[v(1,2)\,G(2,1') + \\ + \mu(1,2)\,F(2,1')], \tag{3}$$

$$i\hbar \frac{\partial}{\partial t_1} F(1,1') = -\int d(2)[\mu^*(1,2)G(2,1') + v^*(1,2)\,F(2,1')]. \tag{4}$$

The self-consistent energy v is defined by

$$v(1,1') = -\frac{\hbar^2}{2m} \Delta_1^2\,\delta(1-1') + \int d(2)\,d(2')\,[V(1,2;1',2') - \\ - V(1,2,2',1')]\,h(2,2') - \lambda$$

and the self-consistent pairing potential μ by

$$\mu(1,1') = \int d(2)\, d(2')\, V(1,1';2,2')\, \chi(2,2'), \tag{5}$$

where

$$h(1',1) = <\psi^{+}(1)\,\psi(1')>,$$
$$\chi(1',1) = <\psi(1)\,\psi(1')>.$$

The expectation values are formed with respect to the vacuum state of the general quasi-particle transformation. λ is the chemical potential. The energy gap can be calculated with the aid of v and μ [7].

The nucleon-nucleon potential has a hard core which would make the self-consistent potential energy term in v infinite. To get a finite expression we should have to go beyond the generalized Hartree-Fock approximation taking into account at least further two-body correlations. For the sake of clarity we want to avoid the complications involved in the determination of the self-consistent potential. Therefore this will be replaced by some model field of force chosen in such a way that one may hope that it will to some extent approximate the self—consistent potential. We assume this model potential to be spin independent:

$$v(1,1') = v(r_1, r_1')\,\delta_{s_1 s_1'}\,\delta(t_1 - t_1') =$$
$$= -\frac{\hbar^2}{2m}\nabla_1^2\,\delta(1-1') + U(r_1, r_1')\,\delta_{s_1 s_1'}\,\delta(t_1 - t_1') - \lambda.$$

We further simplify the problem by taking into account pair correlation only between particles of opposite spin. We get from eqs. (3), (4) and (5):

$$i\hbar\frac{\partial}{\partial t_1}G(\mathbf{r}_1, \mathbf{r}_1';\, t_1 - t_1';\ \uparrow\uparrow) = \hbar\,\delta(\mathbf{r}_1 - \mathbf{r}_1')\,\delta(t_1 - t_1') +$$
$$+ \int d\mathbf{r}_2\,[v(\mathbf{r}_1, \mathbf{r}_2)\,G(\mathbf{r}_2, \mathbf{r}_1';\, t_1 - t_1';\, \uparrow\uparrow) + \tag{6}$$
$$+ \mu(\mathbf{r}_1, \mathbf{r}_2;\ \uparrow\downarrow)\,F(\mathbf{r}_2, \mathbf{r}_1';\, t_1 - t_1';\ \downarrow\uparrow)],$$

$$i\hbar\frac{\partial}{\partial t_1}F(\mathbf{r}_1, \mathbf{r}_1';\, t_1 - t_1';\ \downarrow\uparrow) =$$
$$= -\int d\mathbf{r}_2\,[\mu^{*}(\mathbf{r}_1, \mathbf{r}_2;\ \downarrow\uparrow)\,G(\mathbf{r}_2, \mathbf{r}_1';\, t_1 - t_1';\ \uparrow\uparrow) + \tag{7}$$
$$+ v(\mathbf{r}_1, \mathbf{r}_2)\,F(\mathbf{r}_2, \mathbf{r}_1';\, t_1 - t_1';\ \downarrow\uparrow)],$$

$$\mu(\mathbf{r}_1, \mathbf{r}_1';\ \uparrow\downarrow) = \sum_s V(\mathbf{r}_1 - \mathbf{r}_1';\ \uparrow\downarrow, s, -s)\,\chi(\mathbf{r}_1, \mathbf{r}_1'; s, -s), \tag{8}$$

where

$$\chi(\mathbf{r}_1, \mathbf{r}_1'; s_1, s_1') = [\chi(1,1')]_{t_1 = t_1'}. \tag{9}$$

We have assumed a local, instantaneous two-body potential. It is apparent that the expression for the pair potential μ is meaningful even for a potential with hard core.

We transform our equations to the mixed position-momentum representation by taking the Fourier transform with respect to $\mathbf{r}_1 - \mathbf{r}_1'$ and $t_1 - t_1'$ keeping $\mathbf{r} = 1/2\,(\mathbf{r}_1 + \mathbf{r}_1')$ fixed. (None of the functions occurring here depends on $t_1 + t_1'$). We shall use the operator Θ defined by BARAFF and BOROWITZ [9] as follows. If

$$K_1(\mathbf{r}_1\, t_1, \mathbf{r}_2\, t_2) = \int d\mathbf{r}_3\, dt_3\, K_2(\mathbf{r}_1\, t_1, \mathbf{r}_3\, t_3)\, K_3(\mathbf{r}_3\, t_3, \mathbf{r}_2\, t_2),$$

then

$$K_1(\mathbf{r}, \mathbf{p}, \omega) = \Theta[K_2(\mathbf{r}, \mathbf{p}, \omega)\, K_3(\mathbf{r}, \mathbf{p}, \omega)],$$

where

$$K_j(\mathbf{r}, \mathbf{p}, \omega) = \int d(\mathbf{r}_1 - \mathbf{r}_2)\, d(t_1 - t_2)\, K_j(\mathbf{r}_1\, t_1, \mathbf{r}_2\, t_2) \cdot \cdot \exp\left\{-i[\mathbf{p}(\mathbf{r}_1 - \mathbf{r}_2) - \omega\,(t_1 - t_2)]\right\}$$

and Θ is a differential operator of infinite order[1]:

$$\Theta[K_1\, K_2] = \lim_{\mathbf{r}' \to \mathbf{r};\ \mathbf{p}' \to \mathbf{p}} \exp\left[\frac{i\hbar}{2} \sum_{xyz} \left(\frac{\partial}{\partial x}\frac{\partial}{\partial p_x'} - \frac{\partial}{\partial p_x}\frac{\partial}{\partial x'}\right)\right] \cdot \cdot K_1(\mathbf{r}, \mathbf{p}, \omega)\, K_2(\mathbf{r}', \mathbf{p}', \omega). \tag{10}$$

The summation over x, y, z means that x, y, z are to replace each other cyclically. The transformed forms of eqs. (6), (7) are

$$\omega\, G(\mathbf{r}, \mathbf{p}, \omega) - \Theta[v(\mathbf{r}, \mathbf{p})\, G(\mathbf{r}, \mathbf{p}, \omega)] - \Theta[\mu(\mathbf{r}, \mathbf{p})\, F(\mathbf{r}, \mathbf{p}, \omega)] = \hbar, \tag{11}$$

$$\omega\, F(\mathbf{r}, \mathbf{p}, \omega) + \Theta[v(\mathbf{r}, \mathbf{p})\, F(\mathbf{r}, \mathbf{p}, \omega)] - \Theta[\mu^*(\mathbf{r}, \mathbf{p})\, G(\mathbf{r}, \mathbf{p}, \omega)] = 0, \tag{12}$$

where the spin indices are suppressed, i.e.:

$$G(\mathbf{r}, \mathbf{p}, \omega) \equiv G(\mathbf{r}, \mathbf{p}, \omega; \uparrow\uparrow),$$

$$F(\mathbf{r}, \mathbf{p}, \omega) \equiv F(\mathbf{r}, \mathbf{p}, \omega; \downarrow\uparrow),$$

$$\mu(\mathbf{r}, \mathbf{p}) \equiv \mu(\mathbf{r}, \mathbf{p}; \uparrow\downarrow).$$

Furthermore

$$v(\mathbf{r}, \mathbf{p}) = \frac{p^2}{2m} + U(\mathbf{r}, \mathbf{p}) - \lambda.$$

[1] This form of Θ can be found in Appendix B of BARAFF's paper [10].

We have used the fact, that

$$\mu(\mathbf{r}_1 s_1, \mathbf{r}_2 s_2) = -\mu(\mathbf{r}_2 s_2, \mathbf{r}_1 s_1),$$

whence

$$\mu(\mathbf{r}, -\mathbf{p}; \downarrow\uparrow) = -\mu(\mathbf{r}, \mathbf{p}; \uparrow\downarrow).$$

From the definition (1) of G we obtain for the density of particles with a given direction of the spin

$$n(\mathbf{r}) \equiv n(\mathbf{r}, \uparrow) = -iG(\mathbf{r}t\uparrow, \mathbf{r}t^+\uparrow) = \int \frac{d\mathbf{p}}{(2\pi\hbar)^3} h(\mathbf{r}, \mathbf{p}), \tag{13}$$

where $h(\mathbf{r}, \mathbf{p}) \equiv h(\mathbf{r}, \mathbf{p}, \uparrow\uparrow)$ and $h(\mathbf{r}, \mathbf{p}, s_1 s_1')$ is the Fourier transform (with respect to $\mathbf{r}_1 - \mathbf{r}_1'$) of

$$h(\mathbf{r}_1 s_1, \mathbf{r}_1' s_1') = [h(1, 1')]_{t_1 = t_1'},$$

i.e.

$$h(\mathbf{r}, \mathbf{p}) = -i \int \frac{d\omega}{2\pi\hbar} G(\mathbf{r}, \mathbf{p}, \omega) e^{i\omega 0^+}. \tag{14}$$

$h(\mathbf{r}, \mathbf{p})$ can be imagined as a joint position-momentum probability distribution function which it becomes in the quasi-classical limit.

According to the definitions (2) and (9) the Fourier transform of $\chi(\mathbf{r}_1, \mathbf{r}_1'; \uparrow\downarrow)$ with respect to $\mathbf{r}_1 - \mathbf{r}_1'$ becomes

$$\chi(\mathbf{r}, \mathbf{p}) \equiv \chi(\mathbf{r}, \mathbf{p}; \uparrow\downarrow) = \left[\int \frac{d\omega}{2\pi\hbar i} F(\mathbf{r}, \mathbf{p}, \omega) e^{i\omega 0^+} \right]^*. \tag{15}$$

3. The generalized Thomas—Fermi approximation

As it was pointed out by BARAFF and BOROWITZ [9] the operator Θ involves powers of $\hbar$ in series of terms the relative size of which depends on the importance of the uncertainty principle. The physically significant parameter exhibited by Θ is the fractional variation of the functions on which Θ operates within a cell of area $\hbar$ in the phase plane. We define the order of a term in Θ by the power of $\hbar$ it contains.

The zeroth order term consists of the product of the quantities on which Θ operates. We first neglect all the other terms of Θ except the one of zeroth order. In this case we obtain from eqs. (11) and (12)

$$G_0 = \hbar \frac{u_0^2}{\omega - E_0 + i\delta} + \hbar \frac{|v_0|^2}{\omega + E_0 - i\delta},$$

$$F_0 = \hbar u_0 v_0^* \left[\frac{1}{\omega - E_0 + i\delta} - \frac{1}{\omega + E_0 - i\delta} \right],$$

where

$$E_0 = \sqrt{\nu^2 + |\mu_0|^2}\,. \tag{16}$$

u_0 is real, v_0 is complex and

$$u_0^2 = \frac{1}{2}\left(1 + \frac{\nu}{E_0}\right), \tag{17}$$

$$|v_0|^2 = \frac{1}{2}\left(1 - \frac{\nu}{E_0}\right), \tag{18}$$

$$u_0 v_0^* = \frac{\mu_0^*}{2E_0}\,. \tag{19}$$

χ can be determined by substituting (17) into eq. (15). Making use of eq. (19) we obtain

$$\chi_0(\mathbf{r}, \mathbf{p}) = -\frac{\mu_0(\mathbf{r}, \mathbf{p})}{2E_0(\mathbf{r}, \mathbf{p})}\,, \tag{20}$$

which together with eq. (8), writing $\mu = \mu_0$, determines μ_0. After substituting G_0 into eq. (14) we get

$$h_0(\mathbf{r}, \mathbf{p}) = |v_0(\mathbf{r}, \mathbf{p})|^2,$$

which gives for the density

$$n_0(\mathbf{r}) = \int \frac{d\mathbf{p}}{(2\pi\hbar)^3} |v_0(\mathbf{r}, \mathbf{p})|^2.$$

If there is no pairing, i.e. $\mu_0 = 0$ we find

$$E_0 = \lambda - \frac{p^2}{2m} - U(\mathbf{r}, \mathbf{p}), \text{ if } |\mathbf{p}| \leq p_F,$$

$$= \frac{p^2}{2m} + U(\mathbf{r}, \mathbf{p}) - \lambda, \text{ if } |\mathbf{p}| > p_F,$$

where the Fermi momentum $p_F(\mathbf{r})$ is the momentum of the highest filled state defined by

$$\frac{p_F^2}{2m} + U(\mathbf{r}, p_F) = \lambda.$$

The density is then given by

$$n_0 = \frac{1}{6\pi^2\hbar^3} p_F^3.$$

These are the equations of the Thomas-Fermi gas model [11] for nucleons moving in a field with the potential $U(\mathbf{r}, \mathbf{p})$. Thus one can refer to the above

approximation where the pair correlation is taken into account as the generalized Thomas-Fermi approximation.

4. Higher-order approximations

It will prove useful to define $G^{(0)}$ and $F^{(0)}$ in a form similar to that of G_0 and F_0, respectively, but to replace μ_0 by a μ calculated in the approximation in question. Accordingly, we define E, u, v by equations similar to eqs. (16), (17), (18), (19) and replace μ_0 by μ everywhere[1]. We write

$$G = G^{(0)} + G^{(1)}, \tag{21}$$

$$F = F^{(0)} + F^{(1)}, \tag{22}$$

and rearrange the eqs. (11) and (12) as follows:

$$G^{(1)} = \frac{\mathrm{Q}(\omega + \nu) + P\,\mu}{\omega^2 - E^2}, \tag{23}$$

$$F^{(1)} = \frac{P(\omega - \nu) + Q\,\mu^*}{\omega^2 - E^2}, \tag{24}$$

with

$$Q = \Theta^{(1)}[\nu, G^{(0)} + G^{(1)}] + \Theta^{(1)}[\mu, F^{(0)} + F^{(1)}],$$

$$P = -\Theta^{(1)}[\nu, F^{(0)} + F^{(1)}] + \Theta^{(1)}[\mu^*, G^{(0)} + G^{(1)}],$$

where

$$\Theta^{(1)}[K_1 K_2] = \Theta[K_1 K_2] - K_1 K_2.$$

The first-order term of $\Theta^{(1)}$ according to eq. (10) composes the Poisson bracket of the functions on which it operates:

$$\Theta_1^{(1)}[K_1 K_2] = \frac{1}{2} i[K_1, K_2],$$

where

$$[K_1, K_2] = \nabla_{\mathbf{r}} K_1 \nabla_{\mathbf{p}} K_2 - \nabla_{\mathbf{p}} K_1 \nabla_{\mathbf{r}} K_2.$$

The first correction to the Thomas-Fermi approximation consists in taking into account $\Theta_1^{(1)}$ and neglecting $G^{(1)}$ and $F^{(1)}$ on the right-hand sides of eqs. (23) and (24) they being considered as corrections to $G^{(0)}$ and $F^{(0)}$, respectively.

[1] In a certain approximation all these quantities have the same index as the μ appearing in the appropriate expressions (e.g. $G_0^{(0)} = G_0$).

Thus we get

$$G_1^{(1)} = \frac{Q_1(\omega + \nu) + P_1 \mu_1}{\omega^2 - E_1^2},$$

$$F_1^{(1)} = \frac{P_1(\omega - \nu) + Q_1 \mu_1^*}{\omega^2 - E_1^2},$$

where

$$Q_1 = \Theta_1^{(1)}[\nu, G_1^{(0)}] + \Theta_1^{(1)}[\mu_1, F_1^{(0)}],$$

$$P_1 = -\Theta_1^{(1)}[\nu, F_1^{(0)}] + \Theta_1^{(1)}[\mu_1^*, G_1^{(0)}].$$

A straightforward calculation yields:

$$\frac{1}{\hbar^2} G_1^{(1)} = \frac{a_1}{\omega - E_1 + i\delta} + \frac{a_2}{(\omega - E_1 + i\delta)^2} + \\ + \frac{b_1}{\omega + E_1 - i\delta} + \frac{b_2}{(\omega + E_1 - i\delta)^2},$$

where

$$a_1 = -b_1 = -\frac{i}{8E_1^3}\{\nu[\mu_1, \mu_1^*] + \mu_1^*[\nu, \mu_1] + \mu_1[\mu_1^*, \nu]\},$$

$$a_2 = E_1 b_1 + \frac{i}{8E_1}[\mu_1, \mu_1^*],$$

$$b_2 = E_1 b_1 - \frac{i}{8E_1}[\mu_1, \mu_1^*]$$

and

$$\frac{1}{\hbar_2} F_1^{(1)} = -\frac{i}{4E_1}[\nu, \mu_1^*]\left\{\frac{1}{(\omega - E_1 + i\delta)^2} - \frac{1}{(\omega + E_1 - i\delta)^2}\right\}.$$

Each a_i, b_i is real and is equal to zero if μ_1 is real.

By substitution into eq. (14) the phase space density becomes

$$h_1(\mathbf{r}, \mathbf{p}) = |v_1|^2 + \hbar b_1.$$

It can be seen from eq. (15) that $F_1^{(1)}$ gives no contribution to χ (i.e. $\chi_1 = -\mu_1/2E_1$). Therefore we have $\mu_1 = \mu_0$ (i.e. $E_1 = E_0$, $v_1 = v_0$, $G_1^{(0)} = G_0$, $F_1^{(0)} = F_0$ in the preceding equations).

It is an important feature of $G_1^{(1)}$ and $F_1^{(1)}$ that each of their terms contains derivatives of μ_1 as a factor.

For an infinite system it is usually assumed that μ is a slowly varying function of $\mathbf{p}$ in the vicinity of the Fermi surface where it differs from zero (this assumption leads to the linearization of the gap equation). One may accept the same assumption for finite systems neglecting $\nabla_{\mathbf{p}} \mu$.

As for the gradients $\nabla_r \nu$ and $\nabla_r \mu$, they are nonzero only at the nuclear surface layer. μ changes there from a value of approximately zero to some MeV, so one can expect that those correction terms will be of greater importance which contain the gradient of ν only. Such terms come from higher order terms of $\Theta^{(1)}$. To get the lowest-order ones we take into account the next term in $\Theta^{(1)}$ neglecting $G_1^{(1)}$ and $F_1^{(1)}$ in the calculation of Q and P. Moreover, the calculation is simplified by neglecting curvature effects and considering the case of a plane nuclear surface so as to reduce the variation (with respect to $\mathbf{r}$) to a one-dimensional problem. Accordingly the corrections to $G^{(0)}$ and $F^{(0)}$ are written as

$$G_2^{(1)} = \frac{Q_2(\omega + \nu) + P_2 \mu_2}{\omega^2 - E_2^2},$$

$$F_2^{(1)} = \frac{P_2(\omega - \nu) + Q_2 \mu_2}{\omega^2 - E_2^2}.$$

with

$$Q_2 = \Theta_2^{(1)}[\nu G_2^{(0)}] = -\hbar^2 \frac{1}{8}\left(\frac{d^2\nu}{dx^2}\frac{dG_2^{(0)}}{dp_x^2} + \frac{d^2\nu}{dp_x^2}\frac{d^2 G_2^{(0)}}{dx^2}\right),$$

$$P_2 = -\Theta_2^{(1)}[\nu F_2^{(0)}] = \hbar^2 \frac{1}{8}\left(\frac{d^2\nu}{dx^2}\frac{d^2 F_2^{(0)}}{dp_x^2} + \frac{d^2\nu}{dp_x^2}\frac{d^2 F_2^{(0)}}{dx^2}\right),$$

where the expression (10) of Θ was used. Let us write ν in the effective mass approximation:

$$\nu = \frac{p^2}{2m^*} + U(x) - \lambda.$$

Performing the differentiations in Q_2 and P_2 and regarding μ as a constant we get

$$\frac{1}{\hbar^3} G_2^{(1)} = \sum_{j=1}^{4}\left\{\frac{A_j}{(\omega - E_2 + i\delta)^j} + \frac{B_j}{(\omega + E_2 - i\delta)^j}\right\}, \tag{25}$$

where

$$A_1 = -B_1 = \frac{|\mu_2|^2}{16 E_2^5}\left[\frac{3\nu}{m^*}\frac{d^2\nu}{dx^2} + \left(1 - 5\frac{\nu^2}{E_2^2}\right) W\right],$$

with

$$W = \left(\frac{p_x}{m^*}\right)^2 \frac{d^2\nu}{dx^2} + \frac{1}{m^*}\left(\frac{d\nu}{dx}\right)^2$$

and

$$\frac{1}{\hbar^3} F_2^{(1)} = \sum_{j=1}^{4}\left\{\frac{C_j}{(\omega - E_2 + i\delta)^j} + \frac{D_j}{(\omega + E_2 - i\delta)^j}\right\}, \tag{26}$$

where

$$C_1 = -D_1 = -\frac{\mu_2^*}{16E_3^2}\left[\left(1-3\frac{v^2}{E_2^2}\right)\frac{1}{m^*}\frac{d^2v}{dx^2} - \frac{v}{E_2^2}\left(3-5\frac{v^2}{E_2^2}\right)W\right].$$

The other C's and D's will not be needed. Substituting eqs. (21), (22), (25) and (26) into eqs. (14) and (15) we get for the phase space density and for the pair field, respectively

$$h_2(x,\mathbf{p}) = \frac{1}{2}\left(1-\frac{v}{E_2}\right) - \hbar^2\frac{|\mu_2|^2}{16E_2^5}\left[\frac{3v}{m^*}\frac{d^2v}{dx^2} + \left(1-5\frac{v^2}{E_2^2}\right)W\right] \tag{27}$$

$$\chi_2(x,\mathbf{p}) = -\frac{\mu_2}{2E_2}\left[1+\hbar^2\frac{1}{8E_2^2}\left\{\left(1-3\frac{v^2}{E_2^2}\right)\frac{1}{m^*}\frac{d^2v}{dx^2} - \frac{v}{E_2^2}\left(3-5\frac{v^2}{E_2^2}\right)W\right\}\right] \tag{28}$$

and μ_2 is determined by the eqs. (8) (with $\mu = \mu_2$) and (28). They always have the trivial solution $\mu_2 = 0$ describing the normal state. In that case $F = 0$ and

$$G_2^{(1)N} = -\frac{1}{4}\frac{1}{m^*}\frac{d^2v}{dx^2}\frac{\hbar^3}{(\omega - v + i\delta\,\sigma(v))^3} - \\ -\frac{1}{4}\left\{\left(\frac{p_x}{m^*}\right)^2\frac{d^2v}{dx^2} + \frac{1}{m^*}\left(\frac{dv}{dx}\right)^2\right\}\frac{\hbar^3}{(\omega - v + i\delta\,\sigma(v))^4}, \tag{29}$$

where $\sigma(v)$ is the sign of v. Taking into account our assumptions (model potential, effective mass approximation, one-dimensional case) this result is in agreement with that of BARAFF and BOROWITZ [9][1]. It has been shown by them that the phase space density has singularities of first and second derivatives of the delta function at the Fermi surface. In contrast the phase space density has no singularities in case of the superfluid solution.

5. Discussion

The eq. (8) (with $\mu = \mu_0$) and the eq. (20) of the generalized Thomas-Fermi approximation are of the same form as those for infinite systems but the quantities here depend on **r**. This dependence can also be replaced by a dependence on the density. Thus the calculation of EMERY and SESSLER [5] for nuclear matter of various densities is equivalent to the solution of these equations.

The evaluation of the pairing potential from our eqs. (8) (with $\mu = \mu_2$) and (28) is much more complicated and has not been done yet. However, an estimate of the correction term was made showing that this correction may

[1] They found their first nonvanishing correction term to the Thomas-Fermi density to be the same as the one found by KOMPANEETS and PAVLOVSZKII [12] and KIRZHNITS [13].

considerably alter the magnitude of the pair correlation calculated in the zeroth order approximation in the surface layer. We note that in the normal case, i.e. when $\mu_2 = 0$ our correction term in the expression (27) of the phase space density becomes zero. In that case, however, such terms which gave no contribution in case of the superfluid solution will yield according to eq. (29) the correction terms. Thus the pairing correlation may play an important role also in the calculation of the density distribution.

REFERENCES

1. J. Bardeen, L. H. Cooper and J. R. Schrieffer, Phys. Rev., **108**, 1175, 1957; N. N. Bogoliubov, V. V. Tolmachev and D. V. Shirkov, A new method in the theory of superconductivity, Publishing Department of the USSR Academy of Sciences, Moscow, 1958 and Consultants Bureau, New York, 1959; J. G. Valatin, Nuovo Cim., **7**, 843, 1958.
2. The first references are: A Bohr, B. R. Mottelson and D. Pines, Phys. Rev., **110**, 936, 1958; S. T. Beliaev, The Many-Body Problem (Les Houches, 1958), J. Wiley, New York, p. 377; B. R. Mottelson, The Many-Body Problem (Les Houches, 1958), J. Wiley, New York, p. 283.
3. N. N. Bogoliubov, ДАН, **119**, 52, 1958.
4. L. N. Cooper, R. L. Mills and A. M. Sessler, Phys. Rev., **114**, 1377, 1959.
5. V. J. Emery and A. M. Sessler, Phys. Rev., **119**, 248, 1960.
6. N. N. Bogoliubov, ДАН, **119**, 244, 1958; see further: S. V. Tyablikov, ДАН, **121**, 250, 1958; N. N. Bogoliubov, Усп. Физ.Наук, **67**, 549, 1959; N. N. Bogoliubov and V. G. Soloviev, ДАН, **124**, 1011, 1959; V. G. Soloviev, Doctoral Theses, Preprint, ОИЯИ, 1961.
7. J. G. Valatin, Phys. Rev., **122**, 1012, 1961.
8. L. P. Gorkov, ЖЭТФ, **34**, 735, 1958.
9. G. A. Baraff and S. Borowitz, Phys. Rev., **121**, 1704, 1961.
10. G. A. Baraff, Phys. Rev., **123**, 2087, 1961.
11. See for a review of the Thomas—Fermi model: P. Gombás, Statistische Behandlung des Atoms, Handbuch der Physik, Band XXXVI., Springer Verlag, 1956, p. 171.
12. A. S. Kompaneets and E. S. Pavlovskii, ЖЭТФ, **31**, 427, 1956.
13. D. A. Kirzhnits, ЖЭТФ, **32**, 115, 1957.

ПАРНАЯ КОРРЕЛЯЦИЯ В ПОВЕРХНОСТНОМ СЛОЕ ЯДРА

П. СЕПФАЛУШИ

Резюме

В работе, исходя из уравнений одночастичной функции Грина, в рамках обобщённого то-есть принимающего во внимание парную корреляцию метода Хартри—Фока, обсуждается неоднородная система ферми-частиц в состоянии сверх текучести. При помощи приближённого метода выводятся простые уравнения, определяющие самосогласованный парный потенциал, и выражения, относящиеся к плотности фазового пространства. Данный метод в наинизшем приближении даёт формулы, действительные в случае бесконечной системы. Отличительной чертой его является зависимость эффективного одночастичного потенциала (и всех других величин) от места. В этом приближении, пренебрегая парной корреляцией, для плотности фазового пространства получается тот же результат, что и в модели Томаса—Ферми.

Принимая во внимание члены первого и второго порядка, выводятся поправки неоднородности как к плотности фазового пространства, так и к уравнениям, определяющим парную корреляцию. Наличие парной корреляции играет значительную роль при вычислении коррекций плотности фазового пространства на основе функций Грина. Следовательно, парная, корреляция может оказать зналительное влияние на распределение плотности.

THE GROUND STATE ENERGY OF ODD FERMION SYSTEMS

By

J. NÉMETH

INSTITUTE FOR THEORETICAL PHYSICS, ROLAND EÖTVÖS UNIVERSITY, BUDAPEST
and
INSTITUTE OF NUCLEAR RESEARCH OF THE HUNGARIAN ACADEMY OF SCIENCES, DEBRECEN

(Presented by Z. Gyulai. — Received 20. IX. 1963)

The ground state energy of odd fermion systems is determined by the BCS approximation. An attempt is made to eliminate the error of the BCS method in the case of small particle number.

§ 1. Introduction

In the last years it became possible to explain superconductivity by supposing a pairing correlation among the electrons. This led to the idea of introducing the pairing correlation in the case of nuclei too [1]. With the help of this assumption it was possible to determine quantitatively numerous properties of nuclei and the results agreed with the experiments [2]. The difference in the energy spectra and in the ground state energies of even and odd nuclei became also understandable by the introduction of the pairing energy [3], but odd nuclei were not treated mathematically exactly: the ground state energy and excitation spectra of odd nuclei has not yet been determined with the BCS method [4], the reason for this being that the BCS approximation leads to good results only in the case of very large systems. The error of the method is of the order of $1/\Omega$ (Ω is the number of possible states), and the difference between even and odd nuclei may be of the same order. In the following we shall determine the ground state energy of an odd fermion system with the BCS approximation, we shall estimate the error of the result and we shall try to eliminate it as much as possible.

§ 2. The ground state of an even fermion system in the BCS approximation

In the following we shall deal for the sake of simplicity with a very simple nuclear model: we take no notice of the long-range nuclear forces and we shall only examine the effect of the pairing correlation. The model Hamiltonian will be the following:

$$H = \sum \varepsilon_\nu \left(a^+_{\nu+}\, a_{\nu+} + a^+_{\nu-}\, a_{\nu-}\right) - \sum' \sum V_{\nu_1 \nu_2}\, a^+_{\nu_1+}\, a^+_{\nu_1-}\, a_{\nu_2-}\, a_{\nu_2+}\,. \tag{1}$$

Here ε_ν is the kinetic energy in a ν state, and $V_{\nu_i\nu_k}$ is the matrix element of the pairing interaction. The simplest supposition for the shape of $V_{\nu_i\nu_k}$ is that $V_{\nu_i\nu_k} = V = \text{const.}$ if ν_i and ν_k are states in the last (open) shell,

$$V_{\nu_i\nu_k} = V \text{ if } \nu_i = \nu_k \,,$$

$V_{\nu_i\nu_k} = 0$ in other cases.

The ground state eigenfunction of the BCS theory is

$$\Phi_0 = \Pi_\nu(u_\nu + v_\nu a^+_{\nu+} a^+_{\nu-}) \,|\, 0 >, \tag{2}$$

where v_ν is characteristic for the filling up, and u_ν for the emptiness of the ν state. Because of the normality-conditions

$$u_\nu^2 + v_\nu^2 = 1 \,.$$

In the case of closed shells $v_\nu^2 = 1$, $u_\nu^2 = 0$ and in the case of empty shells $u_\nu^2 = 1$, $v_\nu^2 = 0$. Determining the energy with the help of this eigenfunction we get:

$$\begin{aligned} E' = <\Phi_0 \,|\, H \,|\, \Phi_0> = E_0 + 2 \sum_\nu \varepsilon_\nu v_\nu^2 - \\ - V \sum_\nu \sum_{\nu'} u_\nu v_\nu u_{\nu'} v_{\nu'} - V \sum_\nu v_\nu^4 , \end{aligned} \tag{3}$$

where E_0 is the energy of closed shells. The summation goes only over the states of the open shell. We may determine the value of v_ν by minimizing the energy, taking into account as a subsidiary condition the constancy of the particle number

$$\frac{d(E - \lambda n)}{dv_\nu} = 0, \tag{4}$$

where $E = E' - E_0$ and $n = \Sigma 2\, v_\nu^2$.

Let us introduce the expressions $\Delta = V\Sigma u_\nu v_\nu$, $\varepsilon = \lambda + V/2 + B + x_\nu$, where $B = (\varepsilon'' + \varepsilon')/2 - \lambda - V/2$ and x_ν changes between $-A = -(\varepsilon'' - \varepsilon')/2$ and $+A$(ε'' and ε' are the energy barriers of the open shell). So with the help of (4) we get the equation:

$$2(B + x_\nu) = \frac{\Delta(1 - 2v_\nu^2)}{u_\nu v_\nu} - V(1 - 2v_\nu^2). \tag{5}$$

The further calculations will be very much simplified, if we suppose that $\Delta \gg V$ and $\Delta^2 + B^2 = C^2 \gg 2\, x_\nu B + x_\nu^2$. The latter inequality means that x_ν is small compared to C.* Making the calculations to second-order approxim-

* This approximation is equal substantially to the strong interaction approximation of (2) ($\eta \leqslant 1$).

ation, it turns out that

$$v_\nu^2 = \frac{1}{2}\left[1 - \frac{B}{C} - \frac{x_\nu}{2C^3}(C^2 - B^2) - \frac{VB(C^2 - B^2)}{2C^4} + \right.$$

$$+ \frac{5}{2}\frac{V^2 B^3 (C^2 - B^2)}{C^7} - \frac{V^2 B^3 (C^2 - B^2)}{4C^5} + \qquad (6)$$

$$\left. + \frac{2VB^2(C^2 - B^2)}{C^6} - \frac{Vx_\nu (C^2 - B^2)}{2C^4} + \frac{3}{2}\frac{Bx_\nu^2 (C^2 - B^2)}{C^5}\right].$$

So as to determine the energy we need the values of $B(\lambda)$ and $C(\Delta)$, respectively. For this purpose we turn to integration instead of summation. For the sake of simplicity we shall suppose that the energy level density is constant in our system, so

$$\Sigma \sim \int_{\varepsilon'}^{\varepsilon''} d\varepsilon\, \varrho(\varepsilon) = \frac{\Omega}{\varepsilon'' - \varepsilon} \int_{-\frac{(\varepsilon''-\varepsilon')}{2}}^{\frac{\varepsilon''-\varepsilon'}{2}} dx.$$

We have now two equations for C and B, namely

$$n = \frac{\Omega}{A}\int_{-A}^{A} v^2\, dx\,, \qquad (7)$$

$$\Delta = \frac{\Omega V}{2A}\int_{-A}^{A} uv\, dx, \qquad (8)$$

where

$$A = \frac{\varepsilon'' - \varepsilon'}{2}\,.$$

From these we get in second-order approximation:

$$C = \frac{V\Omega}{2}\left[1 - \frac{1}{\Omega}\left(1 - \frac{n}{\Omega}\right)^2 + \frac{1}{2\Omega^2}\left(1 - \frac{n}{\Omega}\right)^2\left(\frac{2n}{\Omega} - \frac{n^2}{\Omega^2}\right) - \right.$$

$$\left. - \frac{-2V^2}{3V^2\Omega^2}\left(1 - 3\left(1 - \frac{n}{\Omega}\right)^2\right)\right],$$

$$B = \frac{V\Omega}{2}\left[1 - \frac{n}{\Omega} - \frac{1}{\Omega}\left(1 - \frac{n}{\Omega}\right) + \frac{4}{3}\frac{A^2}{V\Omega}\left(1 - \frac{n}{\Omega}\right)\right].$$

Now we can turn to the determination of the energy:

$$E = \frac{\Omega}{4A}\int_{-A}^{A} dx(\varepsilon'' + \varepsilon' + 2x)\,2v^2 - \frac{\Delta^2}{V} - \frac{V\Omega}{2A}\int_{-A}^{A} dx\, v^4 .$$

After integration substituting the values of C and B we get the following

$$E = \frac{n}{2}(\varepsilon'' + \varepsilon') - \frac{V\Omega n}{2}\left(1 - \frac{n}{2\Omega}\right) - \frac{V}{2}n - \\ - \frac{1}{6}(\varepsilon'' - \varepsilon')^2 \frac{n}{\Omega}\left(1 - \frac{n}{2\Omega}\right) + \frac{V}{2}n\left(1 - \frac{n}{2\Omega}\right). \tag{9}$$

Comparing our results with the formula (68) of (2) given for the case $\eta \leqslant 1$, we can see that they differ only in the last two terms. These expressions have appeared here because the terms with single summation were taken into account. In (2) these terms were neglected. $-V/2n$ is just the contribution corresponding to the interaction of particles with opposite angular momentum. The other term comes from the error of the BCS approximation. We see that these terms are of the order $1/\Omega$ compared to the main terms.

§ 3. The corrections of the results of the BCS approximation in the case of an even system

As it is well known, the inexactness of the BCS model, appearing in the case of small particle number, comes from the fact that the wave function (2) used here is not an exact eigenfunction of the particle number operator. For this reason in the BCS model $<N^2>$ and $<N>^2$ are not equal to each other, in contrast to the real physical situation, where the wavefunction is a common eigenfunction of the energy and particle number operators. We may correct our results, arrived at by the BCS approximation, by adding an expression $-V/4(\Delta N)^2$ to the transformation, where

$$(\Delta N)^2 = <N^2> - <N>^2.$$

With this modification it can be shown that the energy expression in the special case of a model with constant kinetic energy, which case is exactly solvable, gives the exact energy value.

Performing the modification mentioned above in the expression (3) we get the following simple formula:

$$E = \sum 2v_\nu^2 (E_\nu - \lambda - V/2) - V \sum' \sum' u_\nu v_\nu u_{\nu'} v_{\nu'} . \tag{10}$$

Minimizing this, the value of v_ν^2 turns out to be

$$v_\nu^2 = \frac{1}{2}\left[1 - \frac{B}{C} - \frac{C^2 - B^2}{C^3} x_\nu + \frac{3}{2}\frac{B x_\nu^2 (C^2 - B^2)}{C^5}\right].$$

Solving the equations (6) and (7) we get

$$C = \frac{V\Omega}{2}\left[1 - \frac{2}{3}\frac{A^2}{V^2\Omega^2}\left(1 - 3\left(1 - \frac{n}{\Omega}\right)^2\right)\right],$$

$$B = \frac{V\Omega}{2}\left(1 - \frac{n}{\Omega}\right)\left(1 + \frac{4}{3}\frac{A^2}{V\Omega}\right)$$

and for the energy

$$E = \frac{n}{2}(\varepsilon'' + \varepsilon') - \frac{V\Omega n}{2}\left(1 - \frac{n}{2\Omega}\right) - \frac{2}{3}\frac{A^2}{V}\frac{n}{\Omega}\left(1 - \frac{n}{2\Omega}\right) - \frac{V}{2}n. \quad (11)$$

It is easily seen that the difference between (9) and (11) is only the term $\frac{V}{2} n(1 - n/2\,\Omega)$. In the case of constant kinetic energy the exact expression is (5):

$$W = -\frac{V}{2} n\Omega\left(1 - \frac{n}{2\Omega} + \frac{1}{\Omega}\right). \quad (12)$$

Now we can see that if $\varepsilon'' = \varepsilon' = \varepsilon$, (11) becomes equal to (12), so really we have got back the exact results.

§ 4. The ground state of an odd fermion system in the BCS approximation

After dealing with even systems let us turn to the examination of odd ones. The Hamiltonian will again be given by (1), but the ground state wavefunction will differ from (2). In the sense of the BCS model we may introduce the following wavefunction:

$$\overline{\Psi}_0 = \Pi_\nu (u_\nu + r_\nu a_{\nu+}^+ + v_\nu a_{\nu+}^+ a_{\nu-}^+)\,|0>, \quad (13)$$

where r_ν^2 is the probability that the ν state is filled with a single particle. (13) is not exactly an odd particle number wave function, beside odd terms there will be even terms too, but with the two subsidiary conditions

$$\Sigma(2v_\nu^2 + r_\nu^2) = n \quad (14)$$

and

$$\Sigma r_\nu^2 = 1 \quad (15)$$

we can achieve that the most characteristic part of the wave function will be an odd one. Calculating the energy we get

$$E_{\text{odd}} = \sum (r_\nu^2 + 2v_\nu^2)\,\varepsilon_\nu - V \sum_\nu \sum_{\nu'} u_\nu v_\nu u'_\nu v'_\nu - \\ - V \sum v_\nu^2 + V \sum u_\nu^2 v_\nu^2 . \tag{16}$$

From the normalization of the wave function we get the following relation

$$u_\nu^2 + v_\nu^2 + r_\nu^2 = 1 .$$

Let us eliminate r_ν^2 from the expression to be minimized

$$E = \sum (1 - u_\nu^2 - v_\nu^2)(\varepsilon_\nu - \lambda) - V \sum_\nu \sum_{\nu'} u_\nu v_\nu u_{\nu'} v_{\nu'} - \\ - V \sum v_\nu^2 + V \sum u_\nu^2 v_\nu^2 - \left(\alpha + \frac{V}{2}\right) \sum (1 - u_\nu^2 - v_\nu^2) .$$

Here $\alpha + V/2$ is the Lagrange multiplicator of the second subsidiary condition. For u_ν and v_ν we get the following expressions:

$$v_\nu^2 = \frac{\alpha - B - x_\nu}{V} \left[\frac{\Delta}{\sqrt{\alpha^2 - (B + x_\nu)^2}} - 1 \right], \tag{17}$$

$$u_\nu^2 = \frac{\alpha + B + x_\nu}{V} \left[\frac{\Delta}{\sqrt{\alpha^2 - (B + x_\nu)^2}} - 1 \right]. \tag{18}$$

B and x_ν mean here the same as above and Δ is given by

$$\Delta = V \Sigma u_\nu v_\nu . \tag{19}$$

We may see that this solution of the variational problem for u_ν and v_ν is correct only if $\alpha - B > |x_\nu|$. If the inequality is not fulfilled, the solution will differ from (17) and (18), as it will be seen later. If the inequality is fulfilled, (14), (15) and (19) give three equations for the determination of α, B and Δ. Taking into account (17) and (18), these equations will be the following:

$$\frac{V(\Omega - n)}{2} = \sum (x_\nu + B) \left[\frac{\Delta}{\sqrt{\alpha^2 - (B + x_\nu)^2}} - 1 \right], \tag{14a}$$

$$\frac{2\alpha\,\Omega + V(\Omega - 1)}{2\alpha} = \sum \frac{\Delta}{\sqrt{\alpha^2 - (B + x_\nu)^2}} , \tag{15a}$$

$$\Delta = \frac{1}{\Omega - 1} \sum \sqrt{\alpha^2 - (B + x_\nu)^2} . \tag{19a}$$

Let us solve these equations. Introducing the expression $\alpha^2 - B^2 = C^2$ and expanding in a power series of x_ν we get

$$\Delta = \frac{\Omega}{\Omega - 1} C \left[1 - \frac{1}{6} \frac{A^2 \alpha^2}{C^4}\right],$$

$$\alpha = \frac{V(\Omega - 1)^2}{2\Omega} \left[1 - \frac{\Omega}{3} \frac{A^2 B^2}{C^4}\right],$$

$$B = \frac{V(\Omega - n)(\Omega - 1)}{2\Omega} \left[1 - \frac{\Omega}{3} \frac{A^2 \alpha^2}{C^4}\right],$$

With the help of these we can determine the energy:

$$\begin{aligned} E_{\text{odd}} = \sum (1 + w_\nu^2 - u_\nu^2) x_\nu + \frac{n(\varepsilon'' + \varepsilon')}{2} - \frac{\Delta^2}{V} - \frac{V}{2}(n-1) + \\ + V \sum u_\nu^2 v_\nu^2 = \frac{n}{2}(\varepsilon'' + \varepsilon') - \frac{V}{2}(n-1) - \\ - \frac{V}{2} n\Omega \left(1 - \frac{n}{2\Omega}\right) + \frac{V}{2} n \left(1 - \frac{n}{2\Omega}\right) - \\ - \frac{(\varepsilon'' - \varepsilon')^2}{12V} \frac{(\Omega - n)^2}{\Omega^2 - (\Omega - n)^2} + \frac{V}{2}\Omega - \frac{3}{4}V. \end{aligned} \tag{20}$$

The energy difference of the even and odd system will be the following

$$\delta^{(1)} = E_{\text{odd}} - E_{\text{even}} = \frac{V}{2}\Omega - \frac{V}{4} + \frac{(\varepsilon'' - \varepsilon')^2}{12V} \left[1 - 2\left(1 - \frac{n}{\Omega}\right)^2 - \left(1 - \frac{n}{\Omega}\right)^4\right]. \tag{21}$$

We may interpret our result easily. The first two terms occur because the single particle does not take part in the pairing correlation and the numbers of the possible states which can be occupied by pairs is reduced by one (the terms proportional to $(\varepsilon'' - \varepsilon')^2$ are not so easily understandable).

The probability that a state will be filled with a single particle is the following:

$$r_\nu^2 = 1 - u_\nu^2 - v_\nu^2 = \frac{1}{\Omega} - \Omega \frac{B x_\nu}{C^2} + \frac{A^2 \Omega}{6} \left(\frac{\alpha^2 + 2B^2}{C^4}\right).$$

We see that r_ν has almost a smooth distribution.

If the inequality

$$\alpha - B > 2A$$

is not fulfilled, the variational problem has only a trivial solution, namely $r_\nu = 0$ if $r \neq \nu_0$ and $r_{\nu_0} = 1$. In this case the wave function is really an odd one, namely

$$\Phi_0 = \prod_{\nu \neq \nu_0} (u_\nu + v_\nu a^+_{\nu+} a^+_{\nu-}) a^+_{\nu_0 \mathrm{r}} | 0 > . \tag{13a}$$

Calculating the energy with the help of (13*a*), we get

$$\begin{aligned} E'_{\mathrm{odd}} = & \Sigma 2 v_\nu^2 \varepsilon_\nu + \varepsilon_0 (1 - 2v_0^2) - V \Sigma\Sigma u_\nu v_\nu u'_\nu v'_\nu - \\ & - V \Sigma v_\nu^2 + 2V u_0 v_0 \Sigma u_\nu v_\nu - V v_0^2 (1 - 2v_0^2). \end{aligned} \tag{22}$$

The calculations proceed in just the same way as in the case of even systems, only now

$$\Delta = V[\Sigma u_\nu v_\nu - u_0 v_0],$$

$$n = \Sigma 2 v_\nu^2 + (1 - 2v_0^2).$$

Solving the equations we get for the energy

$$\begin{aligned} E'_{\mathrm{odd}} = & \frac{n}{2}(\varepsilon'' + \varepsilon') - \frac{V \Omega n}{2}\left(1 - \frac{n}{2\Omega}\right) - \frac{V}{2}(n-1) + \\ & + \frac{Vn}{2}\left(1 - \frac{n}{2\Omega}\right) + \frac{V}{2}\frac{n}{\Omega}\left(1 - \frac{n}{2\Omega}\right) + \frac{V}{2}\Omega - \\ & - \frac{3}{4}V - \frac{1}{6}(\varepsilon'' - \varepsilon')^2 \frac{n}{\Omega}\left(1 - \frac{n}{2\Omega}\right) + \frac{x_0}{2}\left(1 - \frac{n}{\Omega}\right). \end{aligned} \tag{23}$$

This has its minimum value if the last term is

$$- \frac{(\varepsilon'' - \varepsilon')}{4}\left|1 - \frac{n}{\Omega}\right| .$$

We may now determine the energy difference of even and odd systems. Taking into account (9) and (23) we get

$$\delta^{(2)} = E'_{\mathrm{odd}} - E_{\mathrm{even}} = \frac{V}{2}\Omega - \frac{V}{4} + \frac{V}{2}\frac{n}{\Omega}\left(1 - \frac{n}{2\Omega}\right). \tag{24}$$

Comparing (21) and (24) we see that their difference depends on the relation of V and $\varepsilon'' - \varepsilon'$ and on the numbers of particles in the open shell. If n is very small, (24) will be larger than (21), in this case the single particle will be in a well-determined state, while for large n and V the smooth distribution will occur.

§ 5. The correction of the BCS approximation in the case of an odd fermion system

If we correct the results of the BCS approximation in the way discussed in § 2 the energy will be the following:

$$E = \sum (r_\nu^2 + 2v_\nu^2)\,\varepsilon_\nu - V \sum\sum u_\nu v_\nu u_{\nu'} v_{\nu'} - \\ - V \sum v_\nu^2 - \frac{V}{4} \sum (1 - u_\nu^2 - u_\nu^2)\,(u_\nu^2 + v_\nu^2). \tag{25}$$

It is easily seen that in this way we do not get exact results even in the simplest case. Namely, if the kinetic energy is equal to zero, from the relations $\Sigma r^2 = 1$ and $\Sigma 2\, v^2 = n - 1$ we get:

$$v^2 = \frac{n-1}{2\Omega}, \quad u^2 = 1 - \frac{n+1}{2\Omega}, \quad r^2 = \frac{1}{\Omega}.$$

The energy in the corrected BCS approximation will be the following:

$$E = -\frac{V}{2}(n-1)\left(\Omega - 1 - \frac{n-1}{2} + 1\right) - \frac{V}{4}\left(1 - \frac{1}{\Omega}\right). \tag{26}$$

The exact energy in the case of zero kinetic energy is the following:

$$W = -\frac{V(\Omega-1)(n-1)}{2}\left(1 - \frac{n-1}{2(\Omega-1)}\right) - \frac{V}{2}(n-1). \tag{27}$$

(26) differs from (27) but the error in the original BCS approximation is $V/2(n-1)(1-n+1/2\Omega)$ and now it is only $V/4(1-1/\Omega)$, so it seems to be better to calculate with the corrected expression.

Using the method discussed in § 4 taking into account the subsidiary conditions (14), (15) we get

$$u_\nu^2 = \frac{\alpha + B + x_\nu}{\alpha V}\left[(\alpha + B + x_\nu)\left\{\frac{\Delta}{\sqrt{\alpha^2 - (B + x_\nu)^2}} - 1\right\} + \frac{V}{4}\right],$$

$$v_\nu^2 = \frac{\alpha - B - x_\nu}{\alpha V}\left[(\alpha + B + x_\nu)\left\{\frac{\Delta}{\sqrt{\alpha^2 - (B + x_\nu)^2}} - 1\right\} + \frac{V}{4}\right].$$

The equations (14), (15), (19) will be the following:

$$\alpha\Delta = \sum \Delta(\alpha + B + x_\nu) - \sum \sqrt{\alpha^2 - (B + x_\nu)^2}\left(\alpha + B + x_\nu - \frac{V}{4}\right),$$

$$\frac{\alpha V(\Omega - n)}{2} = \sum (B + x_\nu)\left[(\alpha + B + x_\nu)\left(\frac{\Delta}{\sqrt{\alpha^2 - (B + x_\nu)^2}} - 1\right) + \frac{V}{4}\right],$$

$$\frac{V}{4}(\Omega - 2) = \sum (\alpha + B + x_\nu)\left[\frac{\Delta}{\sqrt{\alpha^2 - (B + x_\nu)^2}} - 1\right].$$

Solving again these equations, we get for the energy

$$E = \frac{n}{2}(\varepsilon'' + \varepsilon') - \frac{V}{2} n\Omega \left(1 - \frac{n}{2\Omega}\right) - \frac{V}{2}(n-1) + \\ + \frac{V}{2}(\Omega - 1) + (\varepsilon'' - \varepsilon')^2(\ldots).$$

[We neglect the terms proportional to $(\varepsilon'' - \varepsilon')^2$ as they are not important compared to the error of the method.] The energy difference of an even and odd system in the corrected method will thus be

$$\delta^{(3)} = \frac{V}{2}\Omega + (\varepsilon'' - \varepsilon')^2(\ldots).$$

If we examine now the results obtained for the value of the pairing energy δ we may say that although the errors of the BCS method and the pairing energy are of the same order of magnitude, in first approximation the error is the same for even and odd systems and so the main terms of the pairing energy are correct. With two different Hamiltonians we have obtained the same result. This indicates that even with the non-exact BCS method we obtain rather good results for the pairing energy, and it seems worth while to determine the level density of odd nuclei with the same method. Comparing the level density relation of odd and even nuclei with the experimental results, it is possible to check the basic idea of our very simple model.

Appendix

If we want to examine a pure odd system, we have to use the following wave function

$$\Phi_0 = \sum_{\nu'} r'_{\nu} a^+_{\nu' +} \prod_{\nu' \neq \nu} (u_{\nu} + v_{\nu} a^+_{\nu +} a^+_{\nu -}) \,|\, 0), \tag{A1}$$

with the subsidiary conditions

$$u_{\nu}^2 + v_{\nu}^2 = 1,$$

$$\Sigma r_{\nu}^2 = 1, \tag{A2}$$

$$\Sigma v_{\nu}^2 (1 - r_{\nu}^2) = n - 1. \tag{A3}$$

Here r_{ν} is again the probability that the ν state is filled with a single nucleon. Determining with (A1) the energy in the BCS approximation we get

$$E = \Sigma 2\varepsilon_{\nu} v_{\nu}^2 (1 - r_{\nu}^2) + \Sigma \varepsilon_{\nu} r_{\nu}^2 - V \Sigma\Sigma u_{\nu} v_{\nu} u'_{\nu} v'^2_{\nu} (1 - \\ - r_{\nu}^2 - r'^2_{\nu}) + V \Sigma u_{\nu}^2 v_{\nu}^2 (1 - 2r_{\nu}^2) - V \Sigma v_{\nu}^2 (1 - r_{\nu}^2). \tag{A4}$$

In the case of constant kinetic energy

$$r_\nu^2 = 1/\Omega, \quad v_\nu^2 = \frac{n-1}{2(\Omega-1)}, \quad u_\nu^2 = \left(1 - \frac{n+1}{2(\Omega-13)}\right)$$
$$E = \varepsilon n - \frac{V}{2}(n-1)\left(\Omega - 1 - \frac{n-1}{2} + \frac{n-1}{2(\Omega-1)}\right) \tag{A5}$$

so the error is exactly as much as in the case of § 4. But if we improve again the BCS approximation in the way discussed in § 3, we get

$$(\Delta N)^2 = -\Sigma\Sigma v_\nu^2 v_\nu'^2 r_\nu^2 r_\nu'^2 - \Sigma v_\nu^4(1-2r_\nu^2) + \Sigma v_\nu^2(1-r_\nu^2).$$
$$E' = E - \frac{V}{4}(\Delta N)^2 = \Sigma 2\varepsilon_\nu v_\nu^2(1-r_\nu^2) + \Sigma\varepsilon_\nu r_\nu^2 - \tag{A6}$$
$$- V\Sigma\Sigma u_\nu v_\nu u_\nu' v_\nu'(1-r_\nu^2-r_\nu'^2) - V\Sigma v_\nu^2 +$$
$$+ V\Sigma\Sigma v_\nu' v_\nu'^2 r_\nu^2 r_\nu'^2,$$

and in the case of constant kinetic energy

$$E' = \varepsilon n - \frac{V}{2}(n-1)\left[\Omega - 1 - \frac{n-1}{2} + 1\right], \tag{A7}$$

which is just the exact energy result. So it is probable that we can get the best results for the odd system using the (A1) wave function and the (A7) energy expressions, but the variation process will then be very complicated so that we did not use it for the first approximation.

REFERENCES

1. A. Bohr, B. R. Mottelson and D. Pines, Phys. Rev., **110,** 936, 1958.
2. S. T. Belyaev, Mat. Fys. Medd. Dan. Vid. Selsk., **31,** no 11, 1959.
3. S. T. Belyaev, Les Houches, 1958.
4. J. Bardeen, L. N. Cooper and J. R. Schrieffer, Phys. Rev., **108,** 1175, 1957.

ОСНОВНОЕ СОСТОЯНИЕ ЭНЕРГИИ НЕЧЕТНЫХ СИСТЕМ ФЕРМИ—ЧАСТИЦ

Й. НЕЙМЕТ

Р е з ю м е

BCS-аппроксимацией определяется основное энергетическое состояние нечетных систем Ферми-частиц. Сделалась попытка для ликвидации ошибки в метода BCS в случае небольшого числа частиц.

FINAL STATE $n-n$ INTERACTION IN THE THREE-PARTICLE PHOTODISINTEGRATION OF TRITON

By

G. GYÖRGYI and P. HRASKÓ

CENTRAL RESEARCH INSTITUTE OF PHYSICS OF THE HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST

(Presented by Z. Gyulai. — Received 20. IX. 1963)

The cross-section of the $\gamma + t \to n + n + p$ reaction has been calculated, taking into account the interaction of the two neutrons in the final state. The proton spectrum is found to be sensitively dependent on the $^1S_0$ neutron-neutron scattering length in the neighbourhood of its upper limit.

A task of fundamental importance in nuclear physics is to investigate the forces acting in two-nucleon systems. Since there is no evidence for the existence of a dineutron, and neutron targets are not available at present, the problem of n — n forces cannot be approached in the same direct way as that of $n-p$ and $p-p$ forces. In a recent elegant experiment Jugoslav workers [1] chose the

$$n + d \to n + n + p \tag{1}$$

reaction to investigate the influence of final state interaction of two neutrons on the energy distribution of protons. Calculations described in the present paper concern the three-particle photodisintegration of triton,

$$\gamma + t \to n + n + p \tag{2}$$

a reaction leading to final products identical with those of reaction (1). In our opinion reaction (2) deserves attention for the following reasons. In the work quoted above calculations concerning reaction (1) were performed using a kind of Born approximation in which the primary interactions between the incident neutron and nucleons in the deuteron are treated as the perturbing potential. There is no *a priori* justification to do so for (1) is a strong nuclear process and, indeed, there is a discrepancy in the absolute value of the cross section. On the other hand, it is entirely justified to use perturbation methods in the case of reaction (2) as this is an electromagnetic process. It should be mentioned that in the final state a square-well interaction between the two neutrons is taken into account exactly both in work [1] and, following the same method, in our calculations. Another point of interest is that in [1] the interaction between the proton and either neutron in the final state has been neglected. There are certain qualitative arguments according to which this approximation should be expected to be good for the high-energy

part of the proton spectrum. One does not know, however, how large the error introduced in this way is. As far as this approximation is concerned, reaction (2) has the advantage that in the electric dipole transition which is to be expected to predominate, the proton emerges in a p state relative to the center of mass of the two neutrons, while in (1) emission of s state protons occurs. We cannot give, either, the magnitude of the error involved in this approximation, we believe, however, that it is rather more justified to use it in the case of reaction (2) than in the case of reaction (1). On the other hand, we think that it is of importance to investigate the role of the interactions neglected.* In this paper, however, no such attempt is made. The first who considered reaction (2) were M. VERDE [2], as well as J. C. GUNN and J. IRVING [3]. Calculations were performed using oscillator wave functions for the triton and free waves for the product particles. In [3] it is argued that in view of the greater phase space available the three-particle disintegration mode might be expected to predominate over the two-particle mode. The necessity of an allowance for distortion of the outgoing waves was stressed by the authors quoted, but in view of the difficulties their treatment was confined to free outgoing waves. Final state interaction formalism suggested by K. M. WATSON [4] and used by the Jugoslav group in their analysis of reaction (1), however, seemed to us to offer a possibility to fill this gap in former treatments. In our calculation, too, an oscillator wave function is used for the triton. We are aware of the limitations arising from the asymptotic behaviour of such a function which does not correspond to reality. On the other hand the calculations which are made possible by the use of this simple function enable us to form a clear physical picture of the process and give an idea about the character of the effect to be expected. Probably for a comparison with the experiment calculations will be needed which use more realistic triton wave functions. Experimental data on (2), which, to our knowledge, are not available at present, would be highly desirable.

To calculate the differential cross section $d\sigma$ of the electric dipole absorption process (1) giving rise to protons of energies in the interval $(E_p, E_p + dE_p)$, we start from the expression

$$d\sigma = \frac{(2\pi)^2 \omega}{c} |(\Phi_f, D_p \Phi_i)|^2 \varrho(E, E_p) dE_p, \tag{3}$$

where ω is the frequency of the incident photon, D_p is the component of electric dipole moment of the proton along the direction of photon polarization, $\varrho(E, E_p)dEdE_p$ is the number of final states in the intervals $(E, E + dE)$ and $(E_p, E_p + dE_p)$ of total and proton energies, Φ_i and Φ_f are wave functions

* This point has been particularly emphasised by Professor N. AUSTERN in discussions with him.

of the initial and final states, respectively; c is the velocity of light. The (normalized) triton wave function is assumed to be of the form

$$\Phi_i = \left(\frac{4\alpha^2}{3\pi^2}\right)^{3/4} e^{-\frac{1}{3}\alpha[(\mathbf{r}_1-\mathbf{r}_2)^2+(\mathbf{r}_2-\mathbf{r}_3)^2+(\mathbf{r}_3-\mathbf{r}_1)^2]} \tag{4}$$

which can be written alternatively as

$$\Phi_i = \nu_i(\mathbf{r})\,\pi_i(\mathbf{r}'), \tag{5}$$

$$\nu_i(\mathbf{r}) = \left(\frac{\alpha}{\pi}\right)^{3/4} e^{-\frac{\alpha}{2}\mathbf{r}^2}, \qquad \pi_i(\mathbf{r}') = \left(\frac{4\alpha}{3\pi}\right)^{3/4} e^{-\frac{2\alpha}{3}\mathbf{r}'^2}, \tag{6}$$

ν_i and π_i being normalized functions. Here r is a vector joining the two neutrons, while r' is the position vector of the proton with respect to the center of mass of the two neutrons. In our assumption (4) the magnetic dipole transition is forbidden. The final state is written as follows:

$$\Phi_f = \nu_f(\mathbf{r})\,\pi_f(\mathbf{r}'), \tag{7}$$

$$\nu_f(\mathbf{r}) = \left(\frac{1}{2\pi R}\right)^{1/2} \cdot e^{i\delta} \cdot k \cdot \frac{\sin(kd+\delta)}{\sin Kd} \cdot \frac{\sin Kr}{kr} \qquad (r \leqslant d), \tag{8}$$

$$= \left(\frac{1}{2\pi R}\right)^{1/2} \cdot e^{i\delta} \cdot k \cdot \frac{\sin(kr+\delta)}{kr} \qquad (r > d), \tag{9}$$

$$\pi_f(\mathbf{r}') = \left(\frac{2}{R'}\right)^{1/2} \cdot k' \cdot \left[\frac{\sin k'r'}{(k'r')^2} - \frac{\cos k'r'}{k'r'}\right] \cdot Y_{1m}(\vartheta', \varphi'). \tag{10}$$

As already mentioned above, in the final state a square-well interaction of the two neutrons is taken into account exactly. The radius of this well is d, while its depth V_0 is related to the inner and outer wave numbers through

$$K^2 = K_0^2 + k^2, \qquad K_0^2 = MV_0/\hbar^2, \tag{11}$$

where M denotes the nucleon mass. Substitute now the wave functions (5) and (7) into (3) and use the expression

$$\varrho(E, E_p) = \frac{RR'}{2\pi^2\hbar^2} \cdot \frac{M}{\sqrt{E_p(2E-3E_p)}} \tag{12}$$

for the density of final spherical wave states; here R and R' denote radii of normalization spheres in $\mathbf{r}$ and $\mathbf{r}'$ spaces, respectively. The formula obtained in this way for the differential cross section $\sigma' \equiv d\sigma/dE_p$ differs from the value σ_0' valid for $V_0 = 0$ in a relatively simple factor $F(E; E_p)$:

$$\frac{d\sigma}{dE_p} = F(E; E_p)\,\frac{d\sigma_0}{dE_p}, \tag{13}$$

$$d\sigma_0 = \frac{e^2}{\hbar c}\sqrt{24}\,\frac{\pi}{a}\,(\varepsilon+\varepsilon_T)\,\varepsilon_p^{3/2}\left(\varepsilon-\frac{3}{2}\,\varepsilon_p\right)^{1/2} e^{-\varepsilon}\,d\varepsilon_p, \tag{14}$$

$$\varepsilon = \frac{ME}{\hbar^2 a}, \qquad \varepsilon_p = \frac{ME_p}{\hbar^2 a}, \qquad \varepsilon_T = \frac{ME_T}{\hbar^2 a}; \tag{15}$$

here E_r means the triton binding energy. Factor $F(E;\,E_p)$ represents the ratio of the squared neutron overlap integrals with and without final state interaction:

$$F(E;\,E_p) = \left|\frac{I_n}{I_n^0}\right|^2, \tag{16}$$

$$I_n = \int v_f(\mathbf{r})\,v_i(\mathbf{r})\,dv, \qquad I_n^0 = \int v_f^0(\mathbf{r})\,v_i(\mathbf{r})\,dv. \tag{17}$$

Here v_f^0 denotes a free two-neutron function which is obtained from (9) putting $\delta = d = 0$. The overlap integrals are:

$$I_n^0 = N\int_0^\infty r e^{-\frac{a}{2}r^2}\sin kr\,dr = \frac{N}{4\pi}\left(\frac{2\pi}{a}\right)^{3/2} k e^{-\frac{k^2}{2a}}, \tag{18}$$

$$I_n = I_n^0 + N\left\{\int_0^a\left(\frac{[1+ikf(k)]\sin kd + kf(k)\cos kd}{\sin Kd}\cdot\sin Kr - \right.\right.$$

$$\left.\left. - \sin kr\right) e^{-\frac{a}{2}r^2} r\,dr + kf(k)\int_a^\infty e^{ikr} e^{-\frac{a}{2}r^2} r\,dr\right\}$$

$$\left(N = 4\pi\left(\frac{1}{2\pi R}\right)^{1/2}\left(\frac{a}{\pi}\right)^{3/4}\right). \tag{19}$$

The ${}^1S_0$ $n-n$ scattering amplitude $f(k)$ has been introduced here instead of the phase shift δ, the connection between them being expressed by

$$f(k) = \frac{1}{k}\,e^{i\delta}\sin\delta\,. \tag{20}$$

For the latter quantities the effective-range approximation has been used:

$$k\,\mathrm{ctg}\,\delta = -\frac{1}{a} + \frac{r_0}{2}k^2, \tag{21}$$

$$f(k) = \frac{1}{-\frac{1}{a} + \frac{r_0}{2}k^2 - ik}\,. \tag{22}$$

The factor F has been evaluated numerically for several cases.

In Figs. 1, 2 and 3 results for $\sigma' \equiv d\sigma/dE_p$ are shown and, in order to exhibit the influence of final state interaction, compared with the corresponding curves for $\sigma_0' \equiv d\sigma_0/dE_p$, calculated by using free-wave final states. The

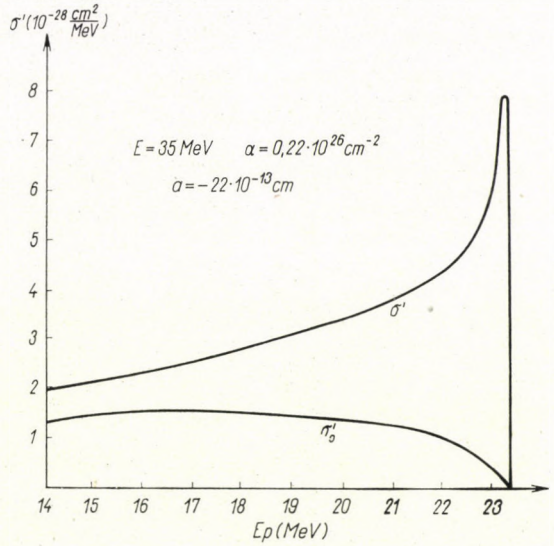


Fig. 1

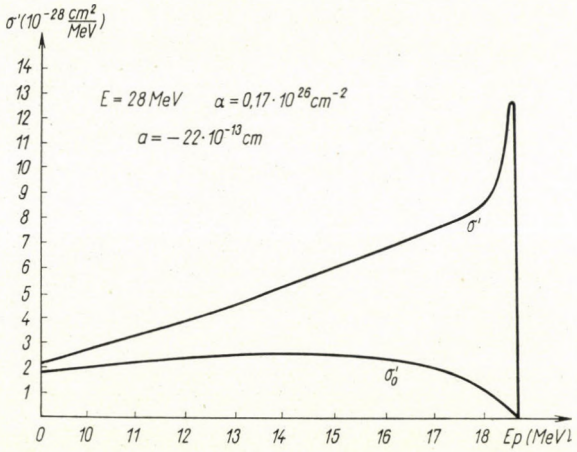


Fig. 2

value of the parameter a corresponding to Fig. 1 was obtained by identifying the difference in the binding energies of H^3 and He^3 with the He^3 Coulomb energy. In Figs. 2 and 3 two other values were used which were also considered in the work of GUNN and IRVING. Results are not sensitive to the value of the effective range; the figure $r_0 = 2.4 \cdot 10^{-13}$ cm was adopted throughout the whole calculation. Values of the total energy E (equal to the gamma energy

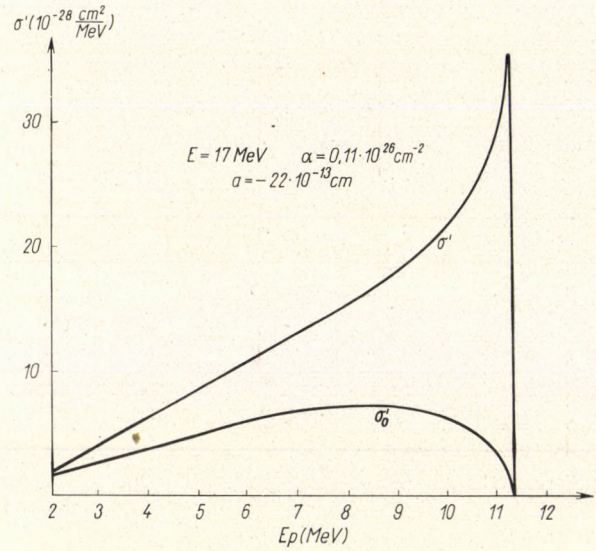


Fig. 3

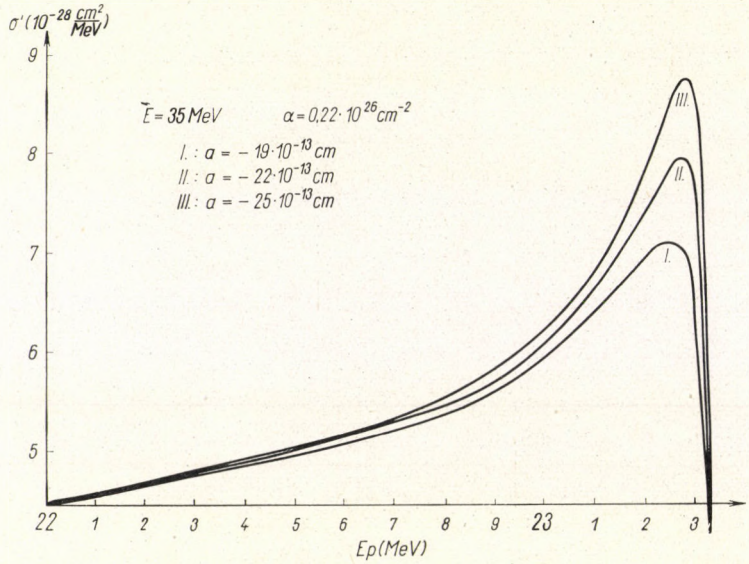


Fig. 4

minus the triton binding energy) were chosen in the regions of maximum total cross section for the respective parameter values. The sharp maxima in σ' obtained next to the high-energy limit of the proton spectrum are very sensitive to the value of the n — n scattering length. To exhibit this the high-energy part of the σ' curves is shown in Figs. 4, 5 and 6 for three different values of the scattering length. Comparing this part of the calculated spectrum with

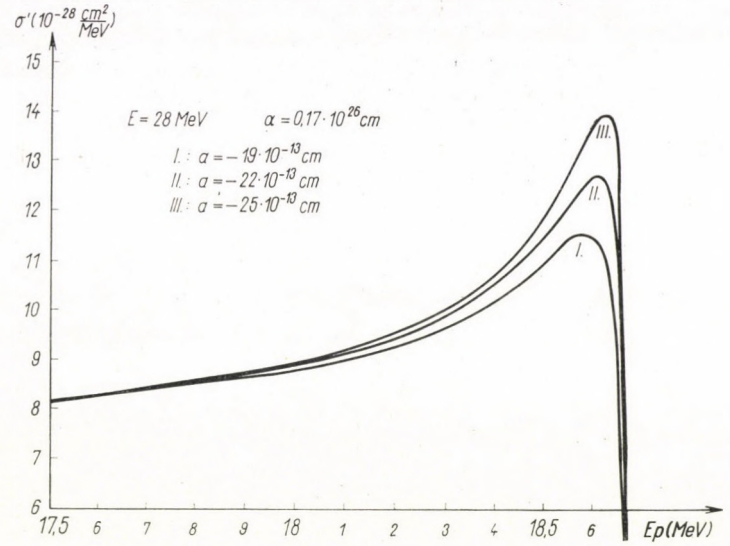


Fig. 5

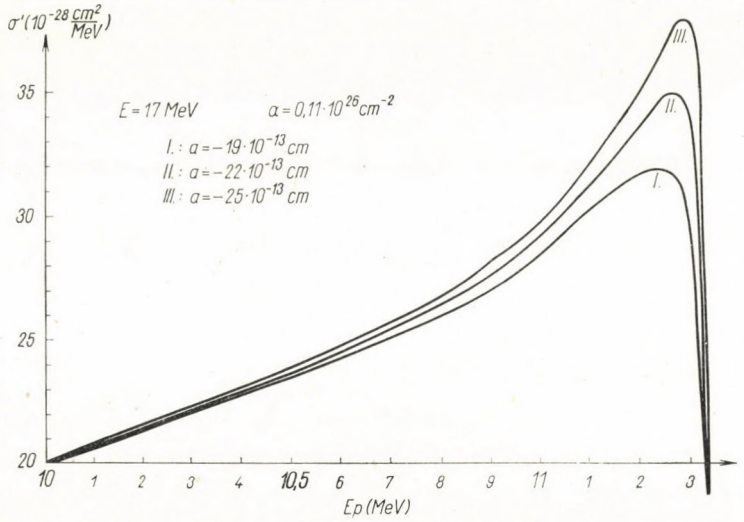


Fig. 6

the experiment an attempt might be made to determine the $^1S_0$ $n - n$ scattering length. In Fig. 7 the influence of the final state interaction on the energy dependence of the total cross section $\sigma(E) = \int_0^{\frac{2}{3}E} \sigma' dE_p$ is exhibited. The position of the maximum of curve $\sigma(E)$ is strongly dependent on a, which means that experimental data concerning this point can provide information on the value of this parameter.

One of the authors (G. Gy.) is indebted to Professors N. Austern and A. deShalit for valuable remarks. Thanks are due to Mr. G. Németh for performing numerical computations.

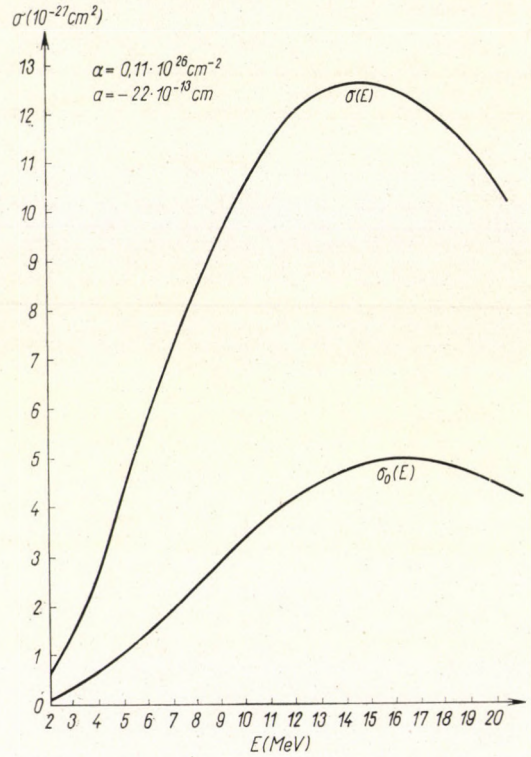


Fig. 7

LITERATURE

1. K. Ilakovac, L. G. Kuo, M. Petravić and I. Šlaus, Proc. Rutherford Jubilee Internat. Conf., Manchester 1961, p. 157; Phys. Rev., **124**, 1923, 1961.
2. M. Verde, Helvetica Physica Acta, **23**, 453, 1950.
3. J. C. Gunn and J. Irving, Phil. Mag., **42**, 1353, 1951.
4. K. M. Watson, Phys. Rev., **88**, 1163, 1952.

ВЗАИМОДЕЙСТВИЕ *n*—*n* В КОНЕЧНОМ СОСТОЯНИИ В ТРЁХЧАСТИЧНОМ ФОТОРАСПАДЕ ТРИТОНА

Г. ДЬЕРДИ и П. ХРАШКО

Резюме

Работа ознакомливает читателя с определением эффективного сечения $\gamma + t \rightarrow$ $\rightarrow n + n + p$ реакции, принимая во внимание взаимодействие двух нейтронов в конечном состоянии. Результаты указывают на то, что спектр протона в соседстве верхнего предела чувствительно зависит от значения длины нейтрон-нейтронного рассеяния $^1S_0$.

DIE STATISTISCHE THEORIE DES ATOMKERNS

V. TEIL

Von

P. GOMBÁS und D. KISDI

PHYSIKALISCHES INSTITUT DER UNIVERSITÄT FÜR TECHNISCHE WISSENSCHAFTEN, BUDAPEST

und

FORSCHUNGSGRUPPE FÜR THEORETISCHE PHYSIK DER UNGARISCHEN AKADEMIE DER WISSENSCHAFTEN, BUDAPEST

(Eingegangen 20. IX. 1963)

Die Dichteverteilungen der Nukleonen werden für die in den vorangehenden Teilen dieser Arbeit entwickelten statistischen Kernmodelle für verschiedene Wechselwirkungsenergien und zwar für eine YUKAWAsche, eine exponentielle und eine GAUSS-sche Wechselwirkung der Nukleonen diskutiert. Ausserdem werden für diese verschiedenen Wechselwirkungen der Nukleonen für einige Kerne die elektrostatischen equivalenten Kernradien, die Dicke der Oberflächenschicht, sowie die »half-way«-Radien berechnet und miteinander verglichen. Es zeigt sich, dass die mit der exponentiellen oder GAUSS-schen Wechselwirkung berechneten Daten und die diesen zugrunde liegenden Dichteverteilungen der Nukleonen im Falle einer exponentiellen oder GAUSS-schen Wechselwirkung zwischen den Nukleonen bedeutend besser mit der Erfahrung übereinstimmen als im Falle der YUKAWAschen Wechselwirkung.

Das Ziel der vorliegenden Arbeit ist die Dichteverteilung der Nukleonen sowie einige mit der Kerngrösse und Kernform im Zusammenhang stehende Daten und zwar den elektrostatischen equivalenten Radius, die Dicke der Oberflächenschicht und den »half-way«-Radius für die in den Teilen I bis IV[1] entwickelten statistischen Kernmodelle zu untersuchen. Die Bezeichnungen sind hierbei dieselben wie die in den ebengenannten vorangehenden Arbeiten.

Es sei betont, dass das zugrunde liegende statistische Modell nur eine sehr grobe Näherung darzustellen vermag, denn in diesem wurden für die Wechselwirkungskräfte zwischen den Nukleonen einfache MAJORANAsche Kräfte angesetzt. Die Berechnungen wurden in den früheren Arbeiten für die folgenden Typen der Wechselwirkungsenergie zweier Nukleonen durchgeführt

1. für die YUKAWAsche Wechselwirkung

$$J_1(r) = -\varepsilon_1 \frac{e^{-r/r_0}}{r/r_0} \; ; \tag{1}$$

2. für die exponentielle Wechselwirkung

$$J_2(r) = -\varepsilon_2 \, e^{-r/r_0} \; ; \tag{2}$$

3. für die GAUSS-sche Wechselwirkung

$$J_3(r) = -\varepsilon_3 \, e^{-(r/r_0)^2} , \tag{3}$$

[1] P. GOMBÁS, Acta Phys. Hung. **1**, 329, 1952.
P. GOMBÁS, Acta Phys. Hung. **2**, 223, 1952.
P. GOMBÁS, E. MÁGORI, B. MOLNÁR, É. SZABÓ, Acta Phys. Hung. **4**, 267, 1954.
P. GOMBÁS, P. SZÉPFALUSY, E. MÁGORI, Acta Phys. Hung. **7**, 251, 1957.

wo wir die Reichweite r_0 der Kernkräfte mit der COMPTON-Wellenlänge der π-Mesonen gleichsetzten: $r_0 = 1{,}355 \cdot 10^{-13}$ cm und die Kopplungskonstante $\varepsilon_i (i = 1, 2, 3)$ so festsetzten, dass die berechneten Kernenergien mit den gemessenen am besten übereinstimmen. Hieraus ergab sich

$$\varepsilon_1 = 71{,}28\,\text{MeV}, \quad \varepsilon_2 = 91{,}17\,\text{MeV}, \quad \varepsilon_3 = 104{,}27\,\text{MeV} \tag{4}$$

Für die Neutronendichte ϱ_n und Protonendichte ϱ_p wurden in den vorangehenden Arbeiten in zweiter Näherung die Ansätze gemacht

$$\varrho_n = \varrho_{n0}\, e^{-a^2 r^2} \left(1 + \frac{1}{3}\,\gamma a^2 r^2\right)^3, \tag{5}$$

$$\varrho_p = \varrho_{p0}\, e^{-a^2 r^2} \left(1 + \frac{1}{3}\,\gamma a^2 r^2\right)^3, \tag{6}$$

wo ϱ_{n0} und ϱ_{p0} Normierungskonstanten sind und a sowie γ Variationsparameter bezeichnen, die aus der Minimumsforderung der Energie bestimmt wurden. Zur besseren Übersicht sind die Parameterwerte für einige Kerne in den Tabellen 1, 2 und 3 zusammengestellt.

Tabelle 1

Werte von a und γ für die YUKAWAsche Wechselwirkung. a in 10^{13} cm^{-1} Einheiten

A	Z	a	γ
16	8	0,607	0,00
40	19	0,576	0,0025
80	37	0,507	0,11
120	54	0,481	0,29
200	85	0,428	0,45
240	99	0,403	0,49

Tabelle 2

Werte von a und γ für die exponentielle Wechselwirkung. a in 10^{13} cm^{-1} Einheiten

A	Z	a	γ
16	8	0,500	0,54
80	36	0,362	1,4
200	82	0,271	1,8

Tabelle 3

Werte von a und γ für die GAUSS-sche Wechselwirkung. a in $10^{13}\,\mathrm{cm}^{-1}$ Einheiten

A	Z	a	γ
16	8	0,645	0,43
80	36	0,493	1,4
200	83	0,378	1,9

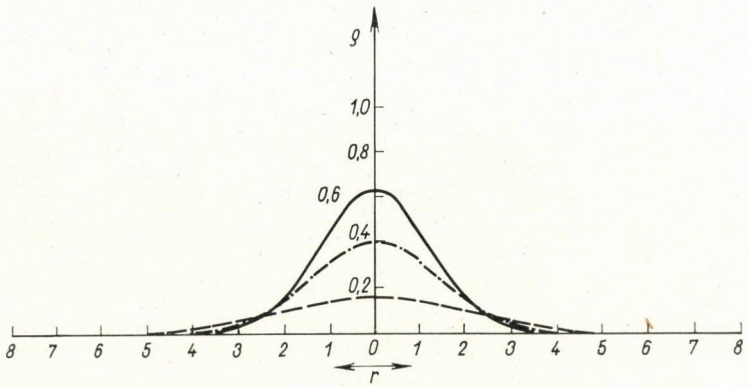


Fig. 1. Verlauf der Nukleonendichte $\varrho = \varrho_n + \varrho_p$ für die Isobaren tiefster Energie mit der Massenzahl $A = 16$. r in 10^{-13} cm, ϱ in $10^{39}\,\mathrm{cm}^{-3}$ Einheiten.
——— mit der YUKAWAschen Wechselwirkung J_1 berechnet,
— — — mit der exponentiellen Wechselwirkung J_2 berechnet,
—·—·— mit der GAUSS-schen Wechselwirkung J_3 berechnet.

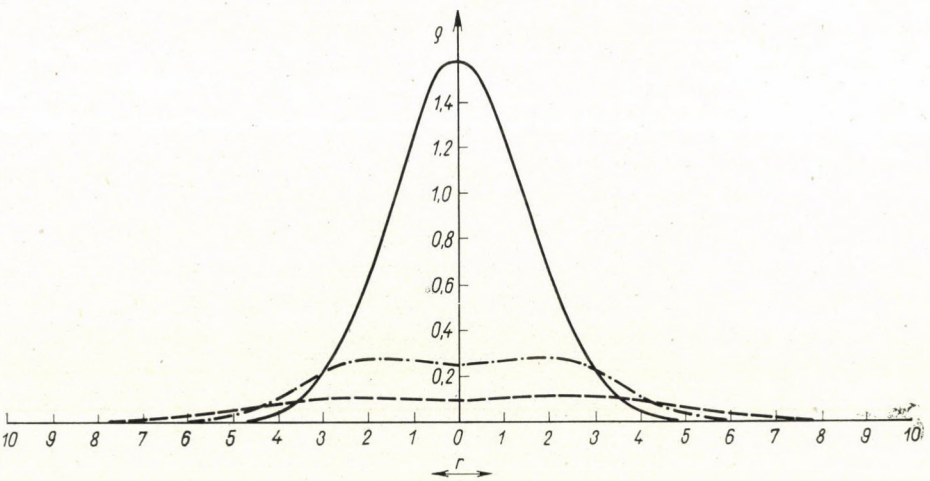


Fig. 2. Verlauf der Nukleonendichte $\varrho = \varrho_n + \varrho_p$ für die Isobaren tiefster Energie mit der Massenzahl $A = 80$. r in 10^{-13} cm, ϱ in $10^{39}\,\mathrm{cm}^{-3}$ Einheiten.
——— mit der YUKAWAschen Wechselwirkung J_1 berechnet,
— — — mit der exponentiellen Wechselwirkung J_2 berechnet,
—·—·— mit der GAUSS-schen Wechselwirkung J_3 berechnet.

Die Gesamtdichte $\varrho = \varrho_n + \varrho_p$ der Nukleonen ist für einige Kerne für die verschiedenen drei Wechselwirkungen in den Figuren 1, 2 und 3 dargestellt.

Die in den Figuren 1 bis 3 dargestellten Dichteverteilungen der Nukleonen können mangels empirischer Daten mit der Erfahrung nicht unmittelbar verglichen werden. Man kann jedoch auf Grund der Dichteverteilungen der Nukleonen einige Parameter sowie den elektrostatischen equivalenten Kern-

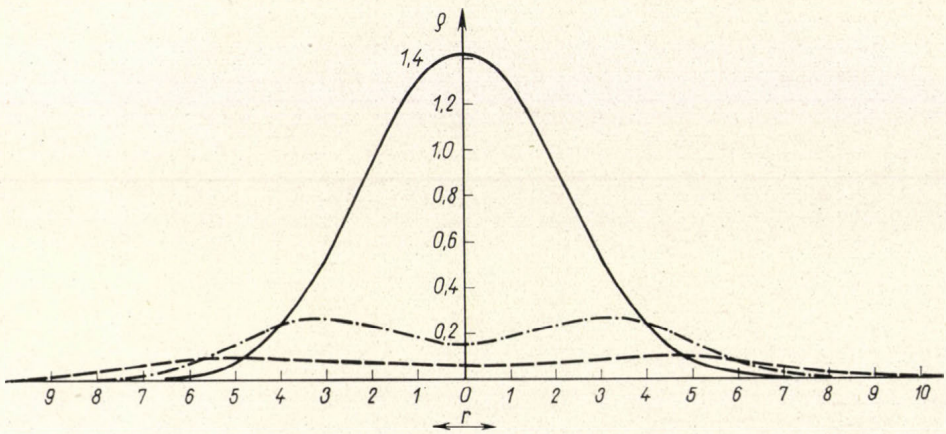


Fig. 3. Verlauf der Nukleonendichte $\varrho = \varrho_n + \varrho_p$ für die Isobaren tiefster Energie mit der Massenzahl $A = 200$. r in 10^{-13} cm, ϱ in 10^{39} cm^{-3} Einheiten.
——— mit der YUKAWAschen Wechselwirkung J_1 berechnet,
— — — mit der exponentiellen Wechselwirkung J_2 berechnet,
—·—·— mit der GAUSS-schen Wechselwirkung J_3 berechnet.

radius, die Dicke der Oberflächenschicht, bzw. den sogenannten »half-way«-Radius des Kerns, weiterhin den maximalen Wert der Nukleonendichte einfach berechnen und mit den entsprechenden empirischen Werten vergleichen, wodurch in Bezug auf die theoretischen Dichteverteilungen Schlüsse gezogen werden können.

Wir wollen nun diese Parameter mit unseren Dichteverteilungen der Reihe nach berechnen.

Der elektrostatische equivalente Kernradius ist folgendermassen definiert

$$R_e = \left(\frac{5}{3Z} \int \varrho_p r^2 \, dv \right)^{1/2} . \tag{7}$$

Mit der statistischen Dichteverteilung (6) ergibt sich hierfür

$$R_e = \left(\frac{5P_e(\gamma)}{2P_0(\gamma)} \right)^{1/2} \frac{1}{a} , \tag{8}$$

wo $P_e(\gamma)$ und $P_0(\gamma)$ die folgenden Ausdrücke bezeichnen

$$P_e(\gamma) = 1 + \frac{5}{2}\gamma + \frac{35}{12}\gamma^2 + \frac{35}{24}\gamma^3 , \tag{9}$$

$$P_0(\gamma) = 1 + \frac{3}{2}\gamma + \frac{5}{4}\gamma^2 + \frac{35}{72}\gamma^3 . \tag{10}$$

Unter der Dicke R_s der Oberflächenschicht des Kerns versteht man die Breite derjenigen Schicht, innerhalb welcher die Nukleonendichte von 90% auf 10% desjenigen Wertes abfällt, den die Nukleonendichte im Kerninneren aufweist. Der »half-way«-Radius R_c ist diejenige Entfernung vom Kernmittelpunkt, bei welcher die Nukleonendichte auf 50% des Wertes im Kerninneren abfällt.

Mit den statistischen Dichteverteilungen (5) und (6) der Nukleonen erhält man für R_c und R_s die folgenden Ausdrücke

$$R_c = \sqrt{3\varkappa_{1/2}}\,\frac{1}{a} , \tag{11}$$

$$R_s = \sqrt{3}\,(\sqrt{\varkappa_{1/10}} - \sqrt{\varkappa_{9/10}})\,\frac{1}{a} , \tag{12}$$

wo $\varkappa_n$ ($n = 1/2, 1/10, 9/10$) die Wurzeln der folgenden transcendenten Gleichung sind

$$(1 + \gamma\varkappa_n)\, e^{-\varkappa_n} = n^{1/3} . \tag{13}$$

Für den maximalen Wert der Nukleonendichte ergibt sich mit den statistischen Verteilungen (5) und (6) der Ausdruck

$$\varrho_{\max} = \frac{a^3\, S(\gamma)}{\pi^{3/2}\, P_0(\gamma)}\, A, \tag{14}$$

wo

$$S(\gamma) = \begin{cases} 1 & \text{für } \gamma \leqslant 1 \\ \gamma^3 \exp\left(-3\,\dfrac{\gamma - 1}{\gamma}\right) & \text{für } \gamma \geqslant 1 \end{cases} \tag{15}$$

ist.

Die mit diesen Formeln berechneten Werte von R_e, R_c, R_s und $\varrho_{\max}$ haben wir in den Figuren 4—7 graphisch dargestellt; zum Vergleich sind auch die halbempirischen Werte von ELTON[2] angegeben. Bei allen vier Grössen lässt

[2] L. R. B. ELTON, Nuclear Sizes, S. 31, Oxford University Press, London, 1961.

sich ein wesentlicher Unterschied zwischen den auf Grund der Wechselwirkung J_1 einerseits und den auf Grund der Wechselwirkungen J_2 und J_3 andererseits berechneten Werten feststellen und zwar fallen die halbempirischen Werte in allen vier Fällen zwischen den mit den Wechselwirkungen J_2 und J_3 berechneten Werten, während die auf Grund der YUKAWAschen Wechselwirkung J_1

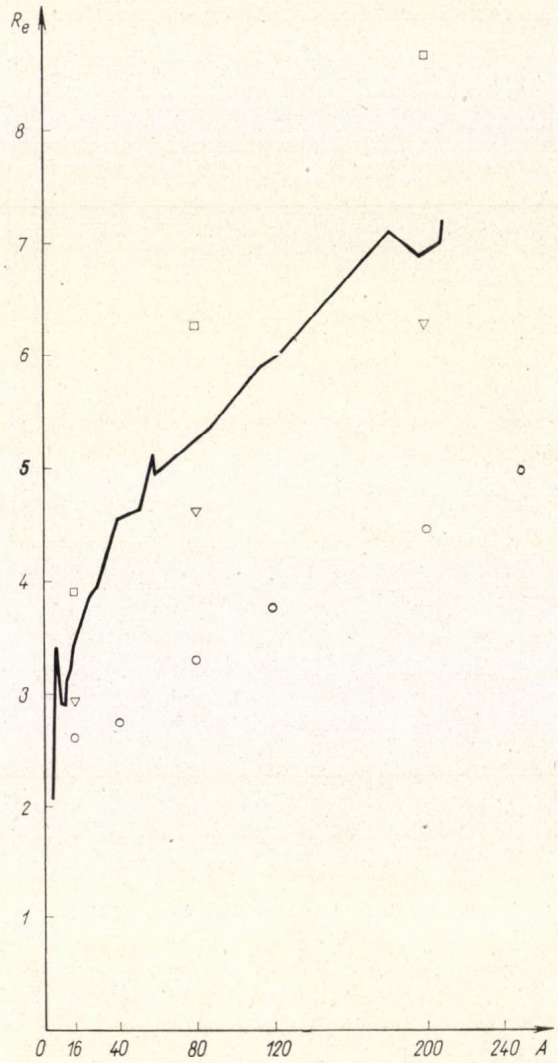


Fig. 4. R_e als Funktion von A in 10^{-13} cm Einheiten.
○ mit der YUKAWAschen Wechselwirkung J_1 berechnet,
□ mit der exponentiellen Wechselwirkung J_2 berechnet,
▽ mit der GAUSS-schen Wechselwirkung J_3 berechnet.
— halbempirische Daten von ELTON.

berechneten Werte von den halbempirischen stark abweichen. Besonders augenfällig ist dies bei ϱ_{max}, man vgl. Fig. 7.

Der grosse Unterschied zwischen den mit der YUKAWAschen Wechselwirkung J_1 einerseits und den Wechselwirkungen J_2 und J_3 andererseits berechneten obengenannten Parameterwerten, sowie den entsprechenden Nukleonendichten ist darauf zurückzuführen, dass die Wechselwirkung J_1 bei $r = 0$

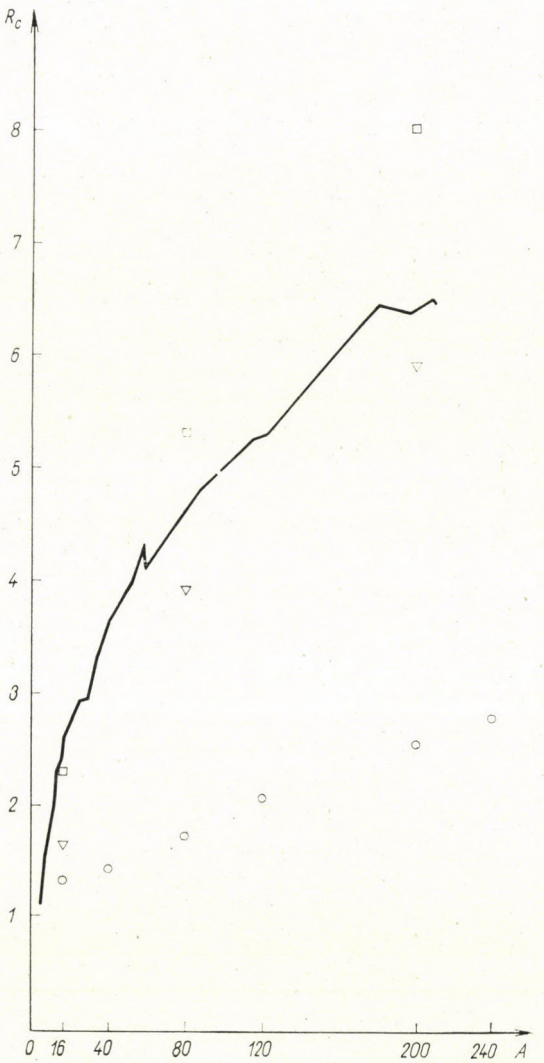


Fig. 5. R_c als Funktion von A in 10^{-13} cm Einheiten.
○ mit der YUKAWAschen Wechselwirkung J_1 berechnet,
□ mit der exponentiellen Wechselwirkung J_2 berechnet,
▽ mit der GAUSS-schen Wechselwirkung J_3 berechnet,
— halbempirische Daten von ELTON.

eine Singularität von der Ordnung $1/r$ aufweist, während bei den anderen beiden Wechselwirkungen dies nicht der Fall ist. In diesem vereinfachten statistischen Modell, in welchem zwischen den Nukleonen nur Kräfte vom MAJORANAschen Typ in Betracht gezogen wurden, ergibt sich also mit den sin-

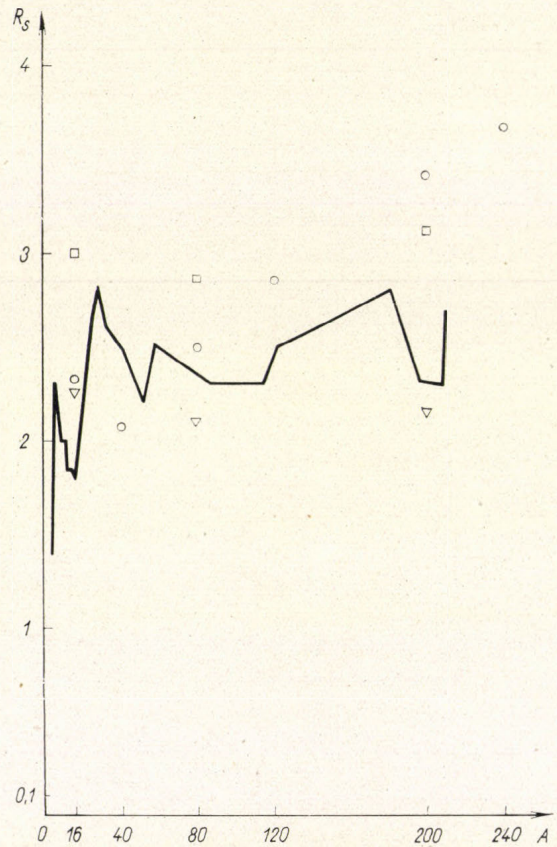


Fig. 6. R_S als Funktion von A in 10^{-13} cm Einheiten.
○ mit der YUKAWAschen Wechselwirkung J_1 berechnet,
□ mit der exponentiellen Wechselwirkung J_2 berechnet,
▽ mit der GAUSS-schen Wechselwirkung J_3 berechnet,
— halbempirische Daten von ELTON.

gularitätsfreien Wechselwirkungen J_2 und J_3 eine bedeutend bessere Übereinstimmung mit der Erfahrung als mit J_1.

Zwischen den mit den Wechselwirkungen J_2 und J_3 berechneten Daten bestehen keine grossen Unterschiede. Eine genauere Untersuchung der in den Figuren 4 bis 7 dargestellten Werte zeigt jedoch, dass die mit der Wechselwirkung J_3 berechneten Daten mit den halbempirischen Werten etwas besser übereinstimmen als die, die man auf Grund der Wechselwirkung J_2 erhält.

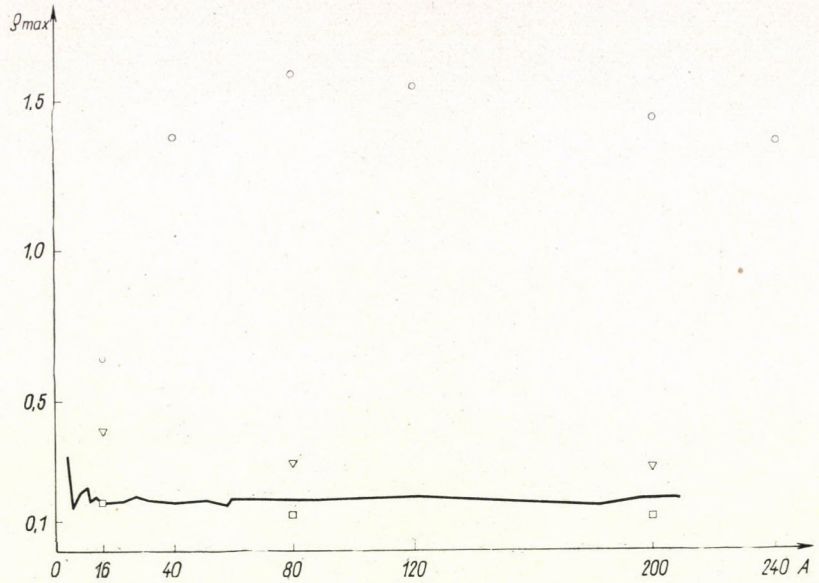


Fig. 7. $\varrho_{\max}$ als Funktion von A in 10^{39} cm^{-3} Einheiten.
○ mit der YUKAWAschen Wechselwirkung J_1 berechnet,
□ mit der exponentiellen Wechselwirkung J_2 berechnet,
▽ mit der GUASS-schen Wechselwirkung J_3 berechnet,
— halbempirische Daten von ELTON.

СТАТИСТИЧЕСКАЯ ТЕОРИЯ АТОМНОГО ЯДРА. ЧАСТЬ V.

П. ГОМБАШ и Д. КИШДИ

Р е з ю м е

В работе дискутируется распределение плотности нуклонов в ядерной модели, разработанной авторами в предыдущих частях данной серии работ для различных энергий взаимодействия между нуклонами, а именно для взаимодействия Юкава, электромагнитного и Гаусса. Кроме этого, для этих различных взаимодействий нуклонов в случае одного ядра определяются и сравниваются между собой электростатический эквивалентный радиус ядра, толщина поверхностного слоя, а также и радиус «полпути». Оказывается, что вычисленные данные при помощи экспоненциального и гауссовского взаимодействий, в основе которых лежит распределение плотности нуклонов в случае взаимодействий Гаусса и экспоненциального, значительно лучше согласуются с результатами эксперимента, чем в случае взаимодействия Юкава.

INDEX

Acta Phys. Hung. Tom. XVII. Fasc. 1—2. Budapest, 28. II. 1964

Printed in Hungary

A kiadásért felel az Akadémiai Kiadó igazgatója Műszaki szerkesztő: Farkas Sándor

A kézirat nyomdába érkezett: 1963. XII. 16. — Terjedelem: 23,75 (A/5) ív, 32 ábra

60.58124 Akadémiai Nyomda, Budapest — Felelős vezető: Bernát György

The *Acta Physica* publish papers on physics, in English, German, French and Russian.
The *Acta Physica* appear in parts of varying size, making up volumes.
Manuscripts should be addressed to:

Acta Physica, Budapest 502, Postafiók 24.

Correspondence with the editors and publishers should be sent to the same address.
The rate of subscription to the *Acta Physica* is 110 forints a volume. Orders may be placed with "Kultúra" Foreign Trade Company for Books and Newspapers (Budapest I., Fő u. 32. Account No. 43-790-057-181) or with representatives abroad.

Les *Acta Physica* paraissent en français, allemand, anglais et russe et publient des travaux du domaine de la physique.
Les *Acta Physica* sont publiés sous forme de fascicules qui seront réunis en volumes.
On est prié d'envoyer les manuscrits destinés à la rédaction à l'adresse suivante:

Acta Physica, Budapest 502, Postafiók 24.

Toute correspondance doit être envoyée à cette même andresse.
Le prix de l'abonnement est de 110 forints par volume.
On peut s'abonner à l'Entreprise du Commerce Extérieur de Livres et Journaux «Kultúra» (Budapest I., Fő u. 32. — Compte-couraut No. 43-790-057-181) ou à l'étranger chez tout les représentants ou dépositaires.

«*Acta Physica*» публикют трактаты из области физических наук на русском, немецком, английском и французском языках.
«*Acta Physica*» выходят отдельными выпусками разного объема. Несколько выпусков составляют один том.
Предназначенные для публикации рукописи следует направлять по адресу:

Acta Physica, Budapest 502, Postafiók 24.

По этому же адресу направлять всякую корреспонденцию для редакции и администрации.
Подписная цена «*Acta Physica*» — 110 форинтов за том. Заказы принимает предприятие по внешней торговле книг и газет «*Kultúra*» (Budapest I.,) Fő u. 32. Текущий счет: № 43-790-057-181) или его заграничные представительства и уполномоченные.

68,— Ft

Index: 26.022

ACTA PHYSICA

ACADEMIAE SCIENTIARUM HUNGARICAE

TOMUS XVII FASCICULUS 3

AKADÉMIAI KIADÓ, BUDAPEST
1964

ACTA PHYS. HUNG.

ACTA PHYSICA

A MAGYAR TUDOMÁNYOS AKADÉMIA FIZIKAI KÖZLEMÉNYEI

SZERKESZTŐSÉG ÉS KIADÓHIVATAL: BUDAPEST V., ALKOTMÁNY UTCA 21.

Az *Acta Physica* német, angol, francia és orosz nyelven közöl értekezéseket a fizika tárgyköréből.

Az *Acta Physica* változó terjedelmű füzetekben jelenik meg: több füzet alkot egy kötetet.

A közlésre szánt kéziratok a következő címre küldendők:

Acta Physica, Budapest 502, Postafiók 24.

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Az *Acta Physica* előfizetési ára kötetenként belföldre 80 forint, külföldre 110 forint. Megrendelhető a belföld számára az Akadémiai Kiadónál (Budapest V., Alkotmány utca 21. Bankszámla 05-915-111-46), a külföld számára pedig a „Kultúra" Könyv- és Hírlap Külkereskedelmi Vállalatnál (Budapest I., Fő u. 32. Bankszámla 43-790-057-181 sz.), vagy annak külföldi képviseleteinél és bizományosainál.

Die *Acta Physica* veröffentlichen Abhandlungen aus dem Bereiche der Physik in deutscher, englischer, französischer und russischer Sprache.

Die *Acta Physica* erscheinen in Heften wechselnden Umfanges. Mehrere Hefte bilden einen Band.

Die zur Veröffentlichung bestimmten Manuskripte sind an folgende Adresse zu richten:

Acta Physica, Budapest 502, Postafiók 24.

An die gleiche Anschrift ist auch jede für die Redaktion und den Verlag bestimmte Korrespondenz zu senden.

Abonnementspreis pro Band: 110 Forint. Bestellbar bei dem Buch- und Zeitungs-Aussenhandels-Unternehmen »Kultúra« (Budapest I., Fő u. 32. Bankkonto Nr. 43-790-057-181) oder bei seinen Auslandsvertretungen und Kommissionären.

THE INFLUENCE OF THE AMBIENT TEMPERATURE ON THE MOVING STRIATION PROCESSES OF LOW PRESSURE DISCHARGES

By

G. LAKATOS and J. BITÓ

INDUSTRIAL RESEARCH INSTITUTE FOR TELECOMMUNICATION TECHNIQUE, BUDAPEST

(Presented by G. Szigeti — Received 30. V. 1963)

The information available so far in respect of moving and stationary striation processes is presented by the authors. The temperature dependence in the range of 20—40° C of the more important characteristics of moving striations was investigated in low-pressure Hg—A direct current discharges. In these investigations the whole discharge tube was surrounded by a stabilized water jacket.

The dependence of the striation frequency on the ambient temperature is given and reference is made to the part played by the external and internal parameters in the development of the striation. With the help of an equivalent circuit the influence of the external parameters on the development of oscillations is discussed theoretically.

1. Introduction

The different gas and vapour discharges are generally accompanied by a number of oscillation effects. Under the simplest conditions, even in the case of direct current discharges, several discharge anisotropies and/or instabilities arising in the equilibrium of the plasma exist that may be the cause of the generation of oscillations.

Outstanding among these oscillations is the so-called plasma oscillation, which is to be considered as the most regular one, and which has been shown to exist already by LANGMUIR and TONKS and treated theoretically by them [1]. These oscillations are characteristic of the given discharge plasmas, they accompany every discharge and their existence can be demonstrated by the Lecher pair of wires in the positive column of the discharge. Their frequency is given by the following relation [1]:

$$n_e = \left[\frac{N_1 \cdot e^1}{\pi \cdot m_e}\right]^{1/2}, \tag{1}$$

where n_e is the frequency, N_e the concentration of the electrons, e the charge of the electrons, m_e the mass of the electrons.

The very same relationship may be applied to the case of ions as well, when the corresponding quantities are substituted. The frequency of electron oscillations in case of an electron concentration of 10^{10} electron/cm^3 and on the basis of the relationship (1) is found to have a value around 10^4 cps. The

frequency of ion oscillations is much smaller than this, on account of the bigger mass and the concentration of the ions which has to be assumed to be smaller than the electron concentration in every case.

In addition to these typical plasma oscillations (electron and ion oscillations) another process may be mentioned that will also accompany regularly the direct current discharges. This process is known under the name of "striation". We distinguish two forms of it [2], the stationary and the moving striation.

With respect to its frequency the striation can be investigated by simpler means than the so-called plasma oscillation, since in this case the frequency is found to be in the range of 10^2—10^4 cps. The results connected with this and which have been reported in the literature, have been described in detail in a previous work [2] by one of the authors.

In the technical journals many authors deal with the striation and search for its origin by varying the experimental conditions. In this respect no opinion that would have been unanimously accepted has as yet been formed. Some are searching for the anisotropy bringing about the process of striation — which as a rule is accompanied by light fluctuation and which can be well shown to occur in the positive column — in the processes taking place in the neighbourhood of the electrodes, others in the variation in time of the basic processes) (generation, diffusion, concentration fluctuation of metastable atoms, recombination, ionization, etc.) of the positive column.

Further, many authors study the parameters which influence these striations [2, 3, 4] and attempt to come thereby to conclusions regarding their physical nature.

The striation — which may both extend axially in the positive column and develop in a stationary way — may be considered also as a space charge wave, the development of which may be influenced by factors both inside and outside the discharge [2, 3, 4], Thus among others in the case of the Penning gas (e.g. with the mercury—argon gas blend) it will presumably be possible to influence the process of striation also by the variation of the external parameter, the ambient temperature, through the variation of the mercury vapour pressure and thereby through that of the concentration of the neutral atoms. Namely, it is generally known that in mercury vapour discharges the concentration of the mercury atoms in the discharge is determined, through the partial pressure resulting from the saturated vapour pressure of the mercury, by the wall temperature of the discharge vessel.

The object of the present investigations is to study in the case of the Penning gas the influence of the ambient temperature on the moving striation of the direct current mercury discharge, as well as to interpret the dependence of the striation on external and internal parameters by utilizing the equivalent circuit.

So far no tests are known in the available reference material that – dealing with the external parameters of the moving striation — would examine the part played by the temperature in particular. This may be explained by the character of the mercury vapour discharge in this type of direct current discharge.

The oscillation phenomenon will be described here following the considerations of KAPZOW [1] and use will be made of his computations based on an equivalent circuit.

2. Measuring method, test conditions

In their measurements the authors have employed a measuring method which was outlined in one of their earlier works [3]. The investigation of the frequency, node, light maxima of the moving striation has been effected by the photocell method, where it has been possible to observe oscilloscopically the signals of the photo-cell amplified with the help of an amplifying stage. Through the use of the oscilloscope it has been possible to obtain the wave length and the frequency (by the Lissajoux curve method), of the light intensity of the strata developing in the course of striation as well as to investigate this along the length of the discharge tube, as function of location.

The discharge has been maintained by a stabilized direct current voltage supply source and the discharge current has been limited by symmetrically arranged ohmic adapters.

The length of the discharge tube of a wall thickness of 1 mm, was 1200 mm, its diameter 38 mm, and after proper vacuum technical treatment it was charged with 3 mmHg pressure argon gas and mercury of about 60 mg weight. In order to ensure the stability of the pressure of mercury vapour the whole discharge tube was placed in a water jacket through which water was flowing all the time, at a well-defined velocity. By the use of an ultra-thermostate the temperature of the water used for thermostation could be adjusted to the desired value with a precision of $\pm 0{,}02^\circ$ C.

In the course of the experiments no external magnetic field has been employed.

The tests took place after 20 minutes operation of the discharge tube during which its performance could be considered to have been stationary at the respective stabilized temperatures of 20, 25 and 40° C set in advance.

3. Results

Fig. 1 shows the tube voltage V_t and the discharge current characteristics I_t of the discharge tube recorded at a stabilized wall temperature of

25° C. It may be seen that at low current intensities the characteristics is very steep and of a negative character.

The dependence of the frequency of the moving striation on the discharge current at a stabilized wall temperature of 25°C is shown in Fig. 2. The measurements have been effected in the 5—200 mA current range. The current scale plotted on the abscissa of Fig. 2 is logarithmic.

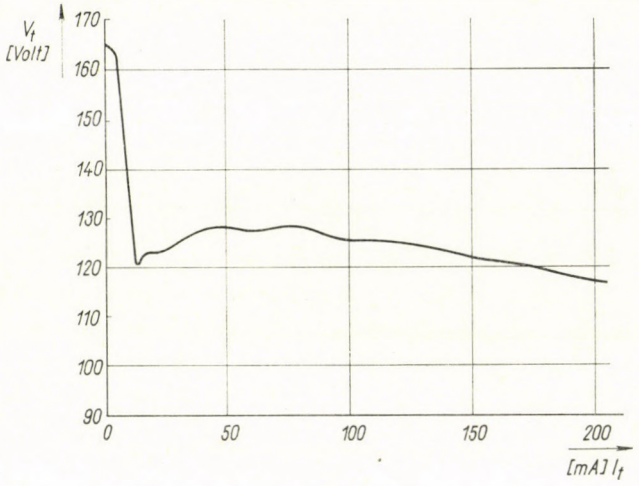


Fig. 1. Discharge current (I_t) — tube voltage (V_t) characteristics of the discharge tube, record at a stabilized wall temperature of 25° C

In the current sector between 5—20 mA the applied external limiting resistance influenced the frequency [4].

The frequency breaking point appearing when the discharge current has a value of 120 mA could be found at all the three tested temperatures (20, 25, 40° C) at the very same value of the current.

The current dependence of the striation frequency follows the character of the curve shown in Fig. 2 which refers to a temperature of 25° C, also in the case of the other two temperatures investigated. The temperature dependence of the frequency at the individual discharge currents is shown in Fig. 3. The temperature dependence in case of currents lower than 120 mA is of a diminishing character; with increasing wall temperature the frequency of the striations will be reduced.

At 120 mA (where the breaking of the frequency current characteristics shown in Fig. 2 takes place) the character of the temperature dependence shown in Fig. 3 already changes; from then onwards at higher currents the frequency of the striation grows with the increasing wall temperature.

As already pointed out above, the maintenance of the wall temperature at a constant value means keeping the value of concentration of the mercury

atoms participating decisively in the discharge fixed. Namely on account of their higher generation and ionization potential the argon atoms are represented to a lesser extent in the basic processes of the stationary discharge in a state of equilibrium than are the mercury atoms.

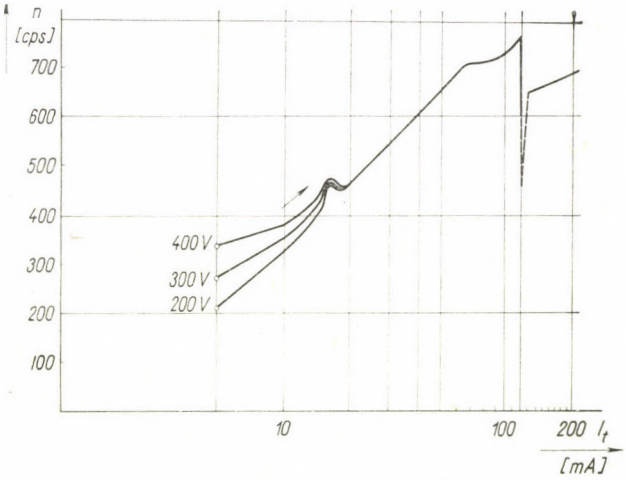


Fig. 2. Dependence of frequency of the moving striation on discharge current

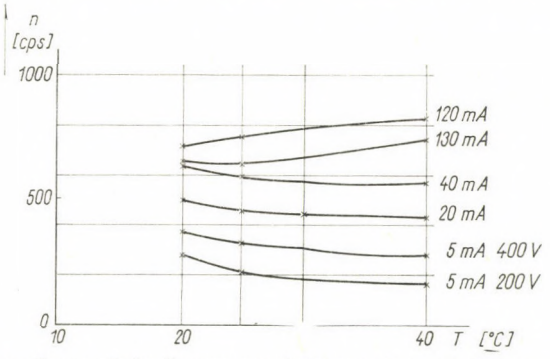


Fig. 3. Temperature-dependence of the frequency of the moving striation at various discharge currents

The saturated vapour pressure and the number of the mercury atoms at the individual temperatures investigated are shown in Table I [7].

Table I

	20° C	25° C	40° C
Vapour pressure of mercury atoms 0^{-3} mm Hg	1,201	1,85	6,07
Concentration of mercury atoms, atom/cm³	$3{,}95 \cdot 10^{13}$	$5{,}94 \cdot 10^{13}$	$20 \cdot 10^{13}$

The well-known rapid increase of the concentration of mercury atoms is striking in the range between the temperatures of 25 and 40° C.

4. Discussion

Before starting the discussion of the results obtained it is worth while to pay attention to certain typical sections of the current-voltage characteristics shown in Fig. 1. At low currents the characteristics have a negative slope, while at currents higher than 20 mA they are in some sections of a positive, in some sections of a negative type.

The character of the curve representing the current dependence of the striation frequency shown in Fig. 2 can be interpreted through the specific features of the characteristics described above, as well as through the influence of the internal parameters of the discharge. In this connection the authors have expressed their opinion in an earlier lecture [4] in which they pointed out that in the current range extending up to 20 mA, the external parameters — mainly the value of the limiting resistance — exercise a decisive influence on the frequency of striation while at currents above 20 mA the frequency of striation is mostly determined by the internal parameters. The authors have also pointed out in [4] that the theoretical description of the influence of the external parameters may be effected by the method of the equivalent circuit introduced by KAPZOW [1]. This will be done further on.

In each curve representing the dependence of the frequencies on current at the three temperatures given above (Fig. 2) one breaking point may be found at 120 mA. Here the frequency decreases by 200—300 cps. For currents higher than 120 mA up to 200 mA the curve increases monotonically from the original frequency value to the value corresponding to the breaking point, in a near linear way. The temperature dependence of this frequency break-down is demonstrated in Fig. 3 by the curve indicated by the parameter value of 120 mA. The theoretical relationships explaining the influence of the temperature as shown in Fig. 3 may be established on the basis of further investigations only.

What can be established in any case without further investigations is that the influence of the ambient temperature may be connected with the change of the concentration of the neutral mercury atoms according to Table I. A consequence of this may be for instance that at current intensities of less than 120 mA, the frequency of the striation diminishes with increasing mercury atom concentration.

The fact that with the increase of the concentration of mercury atoms the direction of the change of the frequency will be reversed at 120 mA, admits of the conclusion that on account of the higher energy losses belonging

to the higher current intensity, over and above this current intensity the change in quality of the discharge has to be traced back to the rate of the heat dissipation. This might affect among others the course of the temperature dependence of the striation frequency.

In the development of the moving striation the influence of the temperature should be considered as an external factor. As has been mentioned earlier [4—6] among the external parameters the resistance, capacity and induction

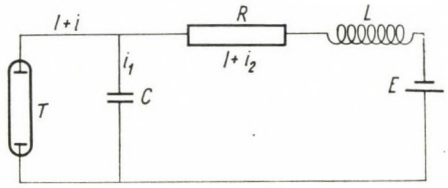


Fig. 4a. Electric circuit of the discharge tube
C — overall capacity of the circuit, R — the overall ohmic and induction resistance, E — the direct current voltage maintaining the discharge

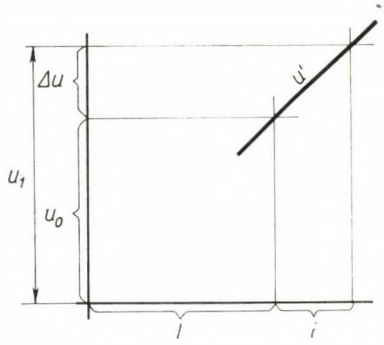


Fig. 4b. Section of the discharge tube characteristics

of the external electric circuit play a very important part in the single sectors of the investigated current range.

The method for the discussion of this has been given by KAPZOW [1] who derived a second-order differential equation for an equivalent circuit and on hand of its solution discussed the conditions of oscillation.

Accordingly, in the present investigations the electric circuit of the discharge may be characterized by the equivalent circuit shown in Fig. 4, where C is the overall capacity of the water jacket, enveloping the discharge circuit and the tube and influencing the discharge, R is the overall resistance of the

external circuit, L is the total induction supplied by these resistances and other circuit elements, E is the voltage of the direct current voltage supply source maintaining the discharge.

In the oscillation-free equilibrium of the discharge the following relationship can be written down for the discharge circuit shown in Fig. 4:

$$E = I \cdot R + U_0, \tag{2}$$

where U_0 is the tube voltage of the discharge at the current I.

Should the current of the discharge change to the value $I + i$ through the starting of some process of oscillation, then corresponding to the discharge characteristics of slope U, shown in Fig. 4, the voltage of the tube will change to the value of U_1.

On account of the existing oscillations the current flowing through the condenser C will be i_1 and also the current flowing through the equivalent chain R L will be changed to some value i_2.

The following current circuit relation may be written down, similarly to the equ. (2):

$$E = U_1 + R[I + i_2] + L\frac{d_{i2}}{dt}, \tag{3}$$

since

$$U_1 = U_0 + \Delta U \tag{4}$$

and

$$\Delta U = U' - i,$$

as well as

$$U_1 - U_0 = \Delta U = \frac{1}{C}\int i_1 \, dt, \tag{6}$$

and from this

$$CU' \cdot i = \int i_1 \, dt,$$

$$CU' \frac{di}{dt} = i_1. \tag{8}$$

Substituting this into equ. (3) and taking into consideration that

$$i_2 = i + i_1 \tag{9}$$

the following relationships are obtained:

$$E = U_0 + U' i + RI + Ri + Ri_1 + L\frac{di}{dt} + L\frac{di_1}{dt}, \tag{10}$$

$$E = U_0 + U' i + RI + Ri + RCU'\frac{di}{dt} + L\frac{di}{dt} + \underbrace{L\frac{di_1}{dt}}_{\| \; LCU'\frac{d^2 i}{dt^2}}. \tag{11}$$

By re-arranging equs. (10) and (11) the following differential equation is obtained:

$$LCU'\frac{d^2 i}{dt^2} + (RCU' + L)\frac{di}{dt} + (R + U')\, i + \underbrace{RI + U_0 - E}_{=0} = 0, \tag{12}$$

$$\frac{d^2 i}{dt_2} + \frac{RCU' + L}{LCU'}\frac{di}{dt} + \frac{R + U'}{LCU'}\, i = 0. \tag{13}$$

Looking for the solution of (13) in the form

$$i = A_1 \exp(r_1 t) + A_2 \exp(-r_2 t), \tag{14}$$

where A_1, A_2 are constants, the following expression containing a square root is obtained for r:

$$r = -\frac{1}{2}\frac{RCU' + L}{LCU'} \pm \sqrt{\frac{1}{4}\left[\frac{RCU' + L}{LCU'}\right]^2 - \frac{R + U'}{LCU'}}, \tag{15}$$

which may be brought into a clearer form

$$r = -\frac{1}{2}\left[\frac{R}{L} + \frac{1}{CU'}\right] \pm \sqrt{\frac{1}{4}\left[\frac{R}{L} + \frac{1}{CU'}\right]^2 - \frac{1}{LC}\left(\frac{R}{U'} + 1\right)}. \tag{16}$$

a) In case the quantity under the root is positive the current fluctuation is periodic, the r_1, r_2 are real. In this case the condition of stability is that

$$r_1 < 0, \tag{17}$$

$$r_2 < 0. \tag{18}$$

This may be realized if first of all

$$\frac{R}{L} + \frac{1}{CU'} > 0, \tag{19}$$

and secondly if the absolute value of the term under the square root is smaller than the square of the first term. Otherwise also a positive r would be possible. Therefore in order to fulfil the conditions (17, 18) it is necessary that

$$\frac{1}{4}\left[\frac{R}{L}+\frac{1}{CU'}\right]^2-\frac{1}{LC}\left[\frac{R}{U'}+1\right]<\left\{\frac{1}{2}\left[\frac{R}{L}+\frac{1}{CU'}\right]\right\}^2 \tag{20}$$

that is that the inequality

$$\frac{1}{LC}\left[\frac{R}{U'}+1\right]>0 \tag{21}$$

exists. Since L and C are positive

$$\frac{R}{U'}+1>0. \tag{22}$$

b) In case the quantity under the root of equ. (16) is negative, the solutions can be found in the form

$$i=A_1\exp(-\delta t+j\omega t)+A_2\exp(-\delta t-j\omega t), \tag{23}$$

where A_1, A_2 are constants,

$$\delta=\frac{1}{2}\left[\frac{R}{L}+\frac{1}{CU'}\right], \tag{24}$$

$$jw=\left\{\frac{1}{4}\left[\frac{R}{L}+\frac{1}{CU'}\right]^2-\frac{1}{LC}\left[\frac{R}{U'}+1\right]\right\}^{1/2}. \tag{25}$$

The condition of stability is that $\delta>0$, therefore also the right-hand expression of (24) should be larger than zero, that is the inequality

$$\frac{1}{2}\left[\frac{R}{L}+\frac{1}{CU'}\right]>0 \tag{26}$$

should be fulfilled. Should both conditions (22) and (26)exist simultaneously, the external parameters will not cause instability in the discharge.

On the basis of the relationship (25) it is possible to determine the frequency oscillations at workpoint which may generate moving striations.

By substituting the values of the external circuit elements appearing in the relationship (25) and the slope of the characteristics belonging to the respective current value, a quantitative relationship is obtained, according to which the frequencies of the oscillations are in fact in the range between 100 and 1000 Hz.

For this computation the ohmic resistance, the capacity as well as the inductivity of the external electric circuit applied in the course of the present investigations, has been determined. The inductivity referred to resulted from the self-inductance of the ohmic resistance.

The capacity of the water jacket surrounding the discharge tube had to be calculated separately and it was found that beside the latter the other capacitive elements of the circuit could be neglected. Table II presents the parameters of the circuit belonging to the discharge current of 50 mA and the value of the slope (U') of the characteristics taken from Fig. 1.

Table II

R	C circuit	C water jacket		U'
3440 ohm	$14 \cdot 10^{-12}$ F	$2{,}05 \cdot 10^{-8}$ F	$12 \cdot 10^{-3}$ Hy	150 ohm

The value of the frequencies is strongly influenced by the value of U' and that of R, L changes to a small extent only, while C may be considered as a constant in this case (water jacket).

Depending on the slope and character of the characteristics, in the case of the various currents different frequency values will result from equation (25) which, however, as regards the order of magnitude, are in agreement with the frequencies found experimentally.

Summarizing, it appears from our investigations that the external parameters, and among them the temperature, together with the elements of the external electric circuit have considerable influence on the striation processes. Further investigations are needed to obtain more information on the basic phenomena taking place in the plasma.

REFERENCES

1. N. A. Kapzow, Elektrische Vorgänge in Gasen und im Vakuum, Deutscher Verlag der Wissenschaften, Berlin, 1955.
2. J. Bitó, Hung Phys. Journal, *X*, 303 1962.
3. G. Lakatos and J. Bitó, Acta Phys. Hung., **13,** 193, 1961.
4. G. Lakatos and J. Bitó, Lecture delivered at the II Czechoslovakian Electronic Conference, 1962. Czechosl. J. Phys., in the press.
5. G. Lakatos and J. Bitó, J. T. F., **32,** 902, 1962.
6. G. Lakatos and J. Bitó, Acta Phys. Hung., **13,** 245, 1961.
7. W. Uyterhoeven, Elektrische Gasentladungslampen, Springer, Berlin, 1938.

ВЛИЯНИЕ ТЕМПЕРАТУРЫ ОКРУЖАЮЩЕЙ СРЕДЫ НА ПРОЦЕССЫ ПОДВИЖНОГО СЛОЕОБРАЗОВАНИЯ РАЗРЯДОВ ПРИ НИЗКИХ ДАВЛЕНИЯХ

ДЬ. ЛАКАТОШ и Й. БИТО

Р е з ю м е

Авторы ознакомливают читателя со встречающимися до настоящего времени в литературе материалами по отношению подвижного и неподвижного слоеобразовательных процессов. В Hg—A разряде постоянного тока при низких давлениях исследуется немпературная зависимость главных параметров подвижного слоеобразования в области температур 20—40° C при стабилизированных экспериментальных условиях с водяной тубашкой. Даётся зависимость частоты слоеобразования от температуры окружающей рреды. Рассматривается роль внешних и внутренних параметров разряда при формиросании слоеобразования. Применяя эквивалентную схему, теоретически истолкуется лияние внешних параметров на формирование колебаний.

ANODE OSCILLATIONS OF DISCHARGES

By

J. Bitó

INDUSTRIAL RESEARCH INSTITUTE FOR TELECOMMUNICATION TECHNIQUE, BUDAPEST

(Presented by G. Szigeti — Received 10. X. 1963)

The author presents a survey of the available literature dealing with the experimental results obtained so far in connection with the oscillation effects appearing on the anodic side of discharges. Experiments were carried out by the author in the case of direct current mercury-argon discharges and the experimental methods and the measuring conditions are described. A special anode construction suggested by G. Szigeti was used and the observed oscillation phenomena were influenced externally through an electric circuit. Conclusions are drawn concerning the possibility of increasing the stability of the anode space.

1. Survey of the available literature

As is known [1—8] even the simplest discharge processes are generally accompanied by instabilities and oscillation effects of various character. In recent years research workers dealing with the investigation of gas discharges have paid close attention also to the oscillation effects which originate in the space around the electrodes (cathode, anode, probes, auxiliary electrodes) or in the electrodes themselves, [1—101]. To-day attention is focused mainly on the oscillation processes going on in the plasma of the positive column (in the presence of an external magnetic field and without it), the various aspects of application always being kept in mind. The oscillation effects appearing in the discharges may be influenced both through the variation of some of the parameters of the space around the electrodes and by that of some of the internal or external parameters of the discharge [54, 57, 58, 63, 64, 79, 80]. The development of any oscillation and its characteristisc is determined not only by a discharge anisotropy or other single factor [64], but in a manner that is not as yet completely understood, these factors in their totality exert an influence to a lesser or higher degree upon the oscillation phenomena of the discharge.

Accordingly the origin of the vibrational state and the striations, moving or stationary, in the positive column of the discharge cannot be decided upon unambiguously either [1, 12, 29, 32, 38, 40, 41, 48, 50—60].

Some authors [56] attempted to trace the origin of the instabilities back to the space-charging phenomena in the synthetized plasma. These investig-

ations based on some model, however, did not provide a sufficiently general picture.

Also the part played by the electrodes in the development of these vibrations has not been unambiguously assessed by the research workers [39—50, 54, 59].

It is evident, however, that oscillations may appear in both the cathode space [99—101] and the anode space, and both can be the source of periodic or aperiodic anisotropy.

These phenomena may be explained also theoretically when a Maxwell—Boltzmann or Fermi distribution is assumed — as this has been done also by KOVRIZSNIKH in the case of the electron-ion plasma [76] — by solving the dispersion equations. Optical and acoustic (electron and ion) oscillations appear in case the wave number vector has a small value, while in case it has a large value there are only acoustic oscillations. However, the frequencies to be calculated theoretically in this way form only a part of the frequency spectrum within which oscillations can be experimentally shown to occur. This admits of the conclusion that in addition to the factors considered here there are still others that influence the oscillations or that themselves can give rise to such periodic or aperiodic anisotropy in the space around the anode. As is known such factors may be the anode spot [66—68, 70—74, 78], the shape and construction of the anode [71], the arrangement of the anode as against the wall of the discharge vessel [73], the variation of the value of the anode fall distance [52] disturbances of various character arising in the energy equilibrium of the anode [70], periodic anode emission [52], etc.

Several research workers have connected also the formation of the moving layer with the phenomenon of the anode spot appearing on the anode [34, 60, 66, 68, 81—84]. Tests carried out with discharges without the positive column have shown [28] that as soon as one or more anode spots develop on the anode surface, oscillations will start, the existence of which may be demonstrated by the fluctuation of the light emitted by the discharge or by that of the discharge current. In a certain pressure range the anode spot will appear at a given current intensity and upon its appearance the value of the anode fall will be reduced [66, 61, 82]. With the increase of pressure the number of anode spots will increase and their dimensions will decrease as far as the negative space-charge region before the anode, where they will disappear altogether [66, 81, 82].

According to probe measurements [66, 81, 82] in some cases upon the appearance of the anode spot the anode fall is found to consist of two parts. The first part is the section of positive anode fall between the end of the positive column and the anode spot, while the second is the region of negative anode fall between the anode spot and the anode. According to KLARFELD and NERETINA [66] the generation of the anode oscillations appearing in the

sound frequency range takes place very intensively in the first section (between the positive column and the anode spot), and very much weaker in the second section.

ZAJTSEV [85] points out that also the striation phenomena of the discharges may originate in certain cases in anode instabilities, anode spots, anodic phenomena, and these features of the anode strongly influence the oscillations. Observations of a similar character have been made by others as well [50—53, 86—94].

NÖLLE ascribes the origin of the oscillations to the anode fall variations [69]. He was able to demonstrate the occurrence of typical Kipp oscillations. He traces the variations of the anode fall back to space-charge anisotropy [69]: if before the anode a large negative space-charge appears for an instant, the value of the anode fall will increase and consequently the electrons getting into the anode space will be considerably accelerated and will ionize once more in their progress towards the anode. Should a sufficiently large number of positive ions be formed by this process these will first reduce and then compensate the negative space charge. Consequently, the value of the anode fall will diminish and it may even become negative, while the anode will repulse the positive ions. This process may occur periodically its frequency depending on the particular conditions [92—94].

ROHNER, in one of his papers [52] suggests the following mechanism for the occurrence of oscillations in the case of anodes emitting ions. If the ions are emitted by the anode (e.g. if the ionization potential of the gas is smaller than the work function of the anode, as in the case of Cs steam and W anode [98], the electrons create relatively less ions on the anode side of the region of anode fall. Thereby the anode fall may diminish, namely in this way a smaller accelerating space is required for the creation of the smaller number of ions. It may happen that in the course of the discharge the emission of the anode changes (if it did not emit ions initially then for instance, on account of the strong heating up it may become capable of emission), and this in accordance with the above will bring about the changing of the anode fall, its disappearance and possibly its becoming negative. Should the anode fall be reduced because of the additional ions emitted by the anode, the energy of the electrons accelerated in the anode space will be reduced too and for this reason a smaller number of ionizations will occur, in consequence of which the original discharge equilibrium will be restored. After a certain time, however, the temperature of the anode will become lower because, owing to the reduced anode fall, the electrons incident on the surface of the anode have a smaller energy. Thereby the emission of the anode will diminish or even stop, which necessitates the increase of the anode fall to its initial value. This process is again of a periodic character, the period being strongly influenced by the heat capacity conditions of the anode [52].

In addition to the more important vibrational mechanisms discussed here briefly, the following factors have also been investigated: the influence of the anode-cathode distance upon the anode oscillations [29, 67]; the effect of the anode heating upon the oscillation frequency (the oscillation frequency increases with increasing anode temperature [2, 52]), the effect of the discharge tube wall [73] and the effect of the auxiliary anodes arranged by the side of the anode [71, 50, 53] as well as the part played by the dimensions and shape of the anode surface [70, 71].

The condition of oscillation established in this way in the anode space is propagated [65] as a modulation of the space charge wave [64], in the direction of the cathode, through the flow of the positive ions.

Investigations have also been made concerning the possibility of utilizing the oscillations of the anode space. OGAWA [35] by making use of certain favourable properties of the oscillations constructed a glimm discharge anode oscillatron tube. RUBTSINSKI and his collaborators [68] applied the process of oscillation going on on the small-surface anode for measuring vapour pressure by making use of the well-reproducible dependence of the amplitude on the vapour pressure of the discharge space.

In the present article the author reports the results of investigations carried out with auxiliary electrodes of different potentials and shapes and with special anode constructions for influencing the oscillations around the anode. The object of the experiments was the investigation of the stability conditions in the space around the anode, as well as the demonstration of the possibility of influencing — from the anode side — the properties of moving striations appearing in direct current discharges.

2. Experimental method

The block diagram of the electric setup applied in the investigations is shown in Fig. 1. The direct current was provided by the stabilized direct current voltage supply source SDC. The current of the discharge tube T was limited by the symmetrically inserted ohmic resistances R_1, R_2. The voltage of the discharge tube T could be read on the instrument V_t, its discharge current on the instrument I_t.

The electric circuit connected to the cathode K provided the external heating of the cathode. The heating current of the heating circuit to be read on the instrument I_h could be adjusted through the choice of the heating voltage V_h and the variation of R_k. The experiments were conducted partly with externally heated cathodes and partly with cathodes without external heating. In one part of the investigations the external heating was provided by a stabilized alternating current voltage source, in the other by a stabilized direct current voltage source, depending on the phenomenon to be investigated.

To the anode A a separate auxiliary electric circuit was connected with the help of which it was possible to influence the anode oscillations and the moving striations — which was the object of the investigations. The auxiliary circuit was fed by a stabilized direct current voltage power source APS which in its turn was fed by a separating transformer. The current flowing in the auxiliary circuit could be read on the instrument I_a, the voltage difference between the outlets on the instrument V_a. The current I_a flowing in the circuit

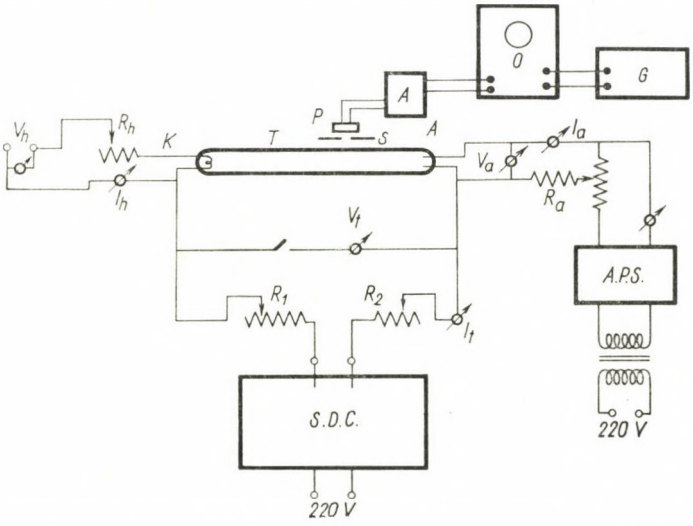


Fig. 1. Circuit diagram of the experimental arrangement used in the investigation of the anode oscillations

could be adjusted to the desired value through the variation of the stabilized voltage source APS which could be controlled continuously, and the resistance R_a.

Three traditional methods are known for the detection of the oscillations [12, 29, 64]: the rotary mirror, the rotary disc provided with slots of the proper number and dimensions, and the oscilloscopic methods (observation of current fluctuation, luminous flux fluctuation). In the present investigations the two latter methods have been employed. With the help of the method where a rotary disc is used lower frequency oscillations could be shown to occur. Series of discs of different shapes, dimensions and number of slots were constructed for the observations.

The oscilloscopic observation served for the detection of oscillations appearing in the discharge current or luminous flux. In the investigation of current oscillations the voltage difference appearing on the vertical input of oscilloscope O came from the terminals of the resistances R_a or R_2, while on the horizontal input the certifying frequency of the sound frequency generator G appeared.

The oscillations of the anode space and the positive column are as a rule also shown by the periodic character of the generation processes as well as by that of the light emission. The results of the previous investigations have shown [53, 57, 58, 64] that the frequency of the oscillations to be found in the discharge current generally agrees with the frequency of the light fluctuations. This empirical fact was again checked and accordingly in a part of the measurements only the frequency and the amplitude of the light fluctuation were recorded [12]. In the case of rather high frequencies the photocell P served for frequency measurements. It was illuminated through the slit S by light emerging from the discharge space to be investigated. The fluctuation of the photocell current was intensified by the amplifier A and thereafter the signal passed to the vertical input of the oscilloscope O. In such a manner it was possible to investigate the shape as well as the amplitude of the light fluctuation. The sound frequency generator G served for the determination of the frequency. The frequency of the oscillation was established by the Lissajoux curve method.

3. Experimental conditions

The measurements were carried out in an ambient temperature of $25 \pm 1°$ C. The length of the employed glass-walled discharge tube was 500 mm, the internal radius 36 mm and the wall thickness 1 mm. A wolfram double spiral provided with a coating to promote electron emission was used as the cathode of the discharge tube. On the two sides of the cathode at a distance of 3 mm auxiliary electrodes of thickness 0,2 mm, width 5 mm and lentgh 14 mm each and of potentials identical with that of the spiral were placed.

The anode of the tube consisted of two parts. In some of the tests, tubes were employed the anodes of which were formed by a cylinder in which a disc was placed perpendicularly to the anode-cathode direction, while the anode employed in the further measurements consisted of a cylinder in which a rod placed parallel to the axis of the cylinder and pointing towards the positive column of the discharge served as an auxiliary anode. These anode constructions are shown in Figs. 2/a and 2/b while their characteristic dimensions are shown in the corresponding section diagrams Figs. 2/c and 2/d. The material of the anode and that of the auxiliary anode was nickel. The wall thickness of the cylinder used as anode was 2 mm, the thickness of the nickel disc shown in Fig. 2/a was similarly 2 mm. The individual parts of the electrode were provided with separate outlets.

The discharge tube passed through the usual vacuum treatment, at the end of which it was filled with 3 mmHg pressure argon and some 60 mg of mercury. During the experiments the gas pressure and the kind of gas remained unaltered. The gas pressure of the discharge tubes could be adjusted to a pre-

cision of $\pm 0{,}05$ mmHg. The pressure of the mercury vapour was determined by the lowest wall temperature of the discharge tube, which, corresponding to the variations of the ambient temperature, could fluctuate within $\pm 1°$ C.

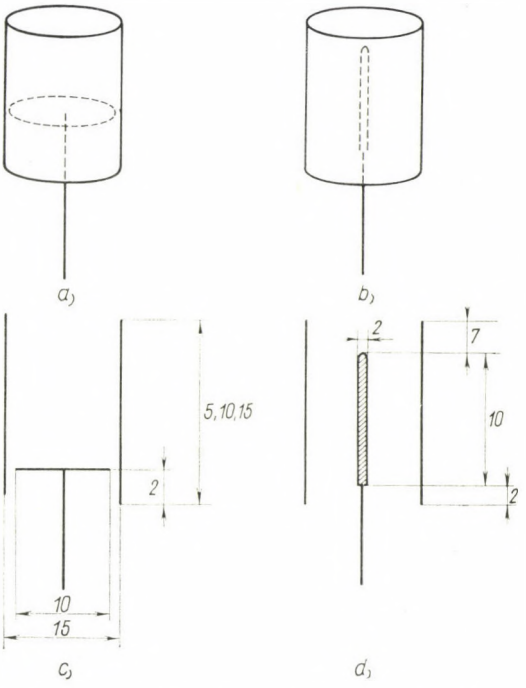


Fig. 2. The employed nickel anode constructions (a, b) and their more important dimensions (c, d)

Prior to starting the measurements the discharge tube was operated for 30 minutes under stabilized conditions. The discharge was started after proper cathode heating and the high frequency preionization of the discharge space.

4. Results

a) *Connection between the light fluctuations and the current fluctuations*

The oscilloscope test method described above served for ascertaining whether or not there is some connection between the characteristics (frequency and amplitude) of the light fluctuations appearing in the anode space and the characteristics of the oscillations appearing (for instance, when there is no voltage difference between the terminals of the resistance R_2) in the discharge circuit current. It was found that the frequencies of the light fluctuations and current fluctuations tally in this case too, they have more or less identical

harmonic oscillations and their amplitudes are proportional to each other. Thus it was possible to demonstrate under the pressent test conditions as well, the connection between the light fluctuation and the discharge current fluctuation, reference to which had been made earlier by several authors [53, 57, 58, 64]. No deeper theoretical interpretation will be offered in this paper. The result has been utilized in the course of the investigations as a possibility for double observation and control only.

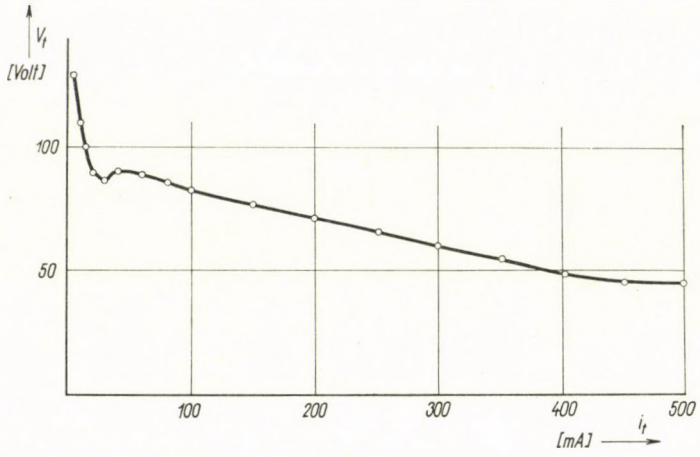


Fig. 3. Typical tube voltage-discharge current characteristics of the tested discharge tubes

A typical voltage-current characteristics of the discharge tubes investigated, which turned out to be the same for both anode constructions, may be seen in Fig. 3. At the lower discharge current the characteristics are very steep and have a typical inflexion point at 30 mA. After 40 mA they may be characterized by a straight section of nearly constant negative slope.

b) *Influence of the anode construction on the oscillation*

A comparison was made, at identical discharge and heating currents, of the influence exerted by the anode constructions *a* and *b*, respectively, shown in Fig. 2, on the space around the anode. The lower frequencies were determined by the above-mentioned rotary disc method, and the higher ones by the photo-cell method.

When the cylinder and the disc (or rod) placed inside the cylinder had identical potential, the positive column of the discharge extended with its initial width as far as the anode cylinder. There was no significant generation of light around and behind the cylinder. Under these conditions the frequency of the oscillations observed was found to amount to 1350 cps.

If the cylinder was not part of the electric circuit and only the disc or the rod acted as anode, then light generation took place also at the end of the tube behind the cylinder. Directly before the anode, oscillations of 1350 cps frequency could be shown in the case of the disc method (Fig. 2/a) and of 5400 cps frequency in the case of the rod anode (the constructions shown in Fig. 2/b). The frequency determinations were carried out at a discharge current of 400 mA in the case of a feed voltage of 200 V. A considerable difference appeared betwen the amplitudes as well. In the case when a rod had been arranged in the inside of the anode cylinder as the sole anode (Fig. 2/b) the amplitude of the vibrations was found to be the 5—6 fold of the amplitude observed in the case of the other construction (Fig. 2/a).

As has been mentioned in the Introduction similar effects had been found earlier and in this connection the conclusion had been drawn that the characteristics of the anode oscillations can also depend on the shape and dimensions of the anode [70, 71]. On this basis RUBTSINSKI and his collaborators [68] actually constructed a vapour pressure meter of suitable sensitiveness.

So far no explanations have been offered for the effects observed here. In the opinion of the author an explanation can possibly be found through the mechanism of the anode fall. COBINE in a discussion of the anode effects, observes that the surface of the anode may influence the dimensions of the anode fall as well. Should the size of the surface of the anode be smaller or larger than a critical value, anode fall shiftings may appear.

Starting from this and regarding the anode as a probe placed into the discharge space, the following explanation may be offered in respect of the observed phenomena, i.e. when oscillations of relatively high amplitudes and comparatively high frequencies appear in the case of small anode surfaces. In the case of a small anode surface — if its size is below a critical value — the anode will limit the electron current which may flow into the external circuit, no current of the intensity required by the circuit elements and by the discharge itself will be able to flow. To overcome this, according to COBINE [70], a space charge will develop by the side of the anode and this space charge will act as an auxiliary anode. Thus the pickup surface of the anode will increase accordingly. This increase, however, will be accompanied also by the increase of the anode fall. Since the concentration of the space charge near the anode will fluctuate according to the fluctuations of the discharge current, also the value of the anode fall will change with a similar periodicity. The energy of the electrons incident on the anode will similarly vary as function of the value of the anode fall, and it may happen that the electrons entering the anode space will acquire so much energy that they are able to ionize once more there. This will cause the increase of the discharge current, on account of which the additional space-charge concentration developing by the side of the anode

will again be reduced, and so will the value of the anode fall. Thereafter the process described here will be repeated.

Similar effects may appear if the size of the surface of the anode is larger than a critical value, in this case a positive space-charge layer will be formed close to the surface of the anode and a retarding potential corresponding to the space-charge will reduce the value of the anode fall, which then leads to the mechanism described formerly.

Therefore it may be stated that the oscillation of higher frequency, as well as of higher amplitude experienced in the case of the rod-shaped anode, is connected with the dimensions of the anode surface. With the reduction of the anode surface — along a certain surface region — the increase of the frequency and amplitude of oscillation may be expected. This may be assumed because of the negative space charge involving more uncertainty. To confirm the results obtained up to now more experiments will be required.

The preceding report may have shown how many different conditions of oscillations resulted in the anode space, in the case of two kinds of anode constructions employed here.

In the sectional drawing of the anode construction shown in Fig. 2/a, three casing height values are given at the side of the anode cylinder. Tests have been conducted under perfectly identical conditions, to find out what variations can be observed in the oscillations of the anode space in the case of these three heights. What resulted was that with the increase of the casing height the frequency of the oscillations did not change, fewer harmonic oscillations could be shown to occur and the amplitude of the oscillations diminished. Also in this case the anode cylinder was on a potential identical with that of the disc placed in it, the dimensions and the position of which were constant during these experiments. On the basis of the above findings one may suppose that the increase of the surface of the anode has not been of such a measure as to affect the change in the value of the anode fall. The surface of the anode could here have been of a size between the two critical surface sizes mentioned by COBINI [70]. In any case it may be seen, that the increase of the anode surface to such an extent acted in a stabilizing way upon the oscillations.

c) *The dependence of oscillation frequency and amplitude on the discharge current*

The dependence of the oscillation frequency of the anode space on the discharge current can be seen in Fig. 4. The results shown in the Figure have been obtained in the case of unheated cathode discharges, and with the anode construction outlined in Fig. 2/a. The oscillations could be shown to occur in the frequency range of 500—2500 cps. The oscillation frequency was growing with the increase of the discharge current. Findings of a similar character, obtained under different discharge conditions, in respect of the positive

column of the discharge, have been known from the relevant literature [79, 80]. The three curves shown belong to three different values of the feed voltage (200 V, 300 V, 400 V). The fact that also the oscillations of the anode space depend on the dimensions of the ohmic restriction elements of the external circuit, admits of the conclusion that there is a connection between these anode oscillations and the striation processes generally appearing in the positive column in the case of direct current discharges. This will be discussed in more detail further below.

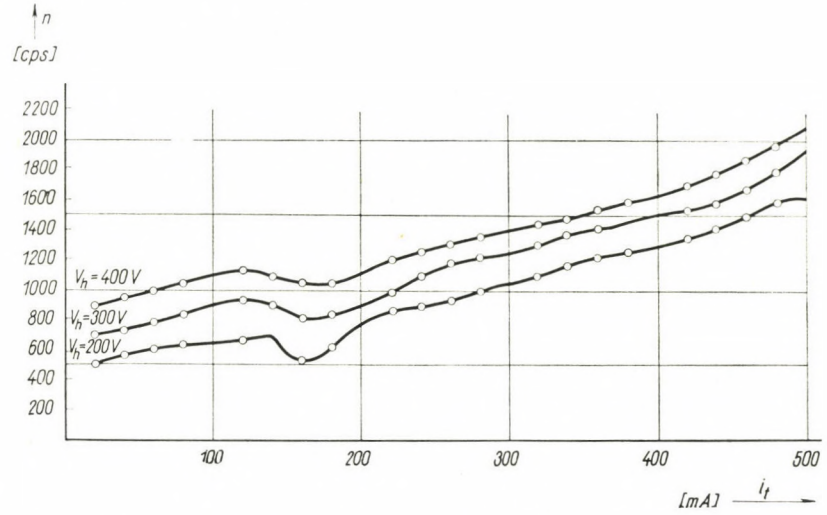


Fig. 4. The dependence of the frequency n of the anode oscillations on the discharge current i_t

The curves shown in Fig. 4 have a characteristic frequency breakdown at a discharge current of 160 mA which with the increase of the applied feed voltage (i.e. with the increase of the angle of the voltage-current characteristics and the limiting resistance) will more or less flatten out.

The current dependence of the amplitude A recorded under similar conditions, can be seen in Fig. 5. The highest amplitudes have been obtained when the experiments were carried out at a feed voltage of 200 V. Also these curves have a characteristic point at 160 mA discharge current. The amplitude of the oscillation is here maximum at least inside the investigated current range. Since the characteristics shown in Fig. 3 do not have a distinguished point in the region of the discharge current of 160 mA, the characteristic point of the curves shown in Figs. 4 and 5 appearing at this value of the current, can be explained neither by the peculiarities of the characteristics nor by the influence of the applied external resistance. From the maximum of the amplitude to be found at a discharge current of 160 mA one may infer psosible resonance effects connected with the peculiarities of the anode space, which

resonance effects are characteristic of the given anode conditions arising in the case of the particular value of the current density — and that of the anode fall — belonging to the discharge current of 160 mA.

When investigating the oscillations of the positive column several authors have studied the dependence of the frequency of the oscillations on the discharge current [32, 33, 64, 79]. They stated that also the frequency of the oscillations increases with the increase of the discharge current. At the

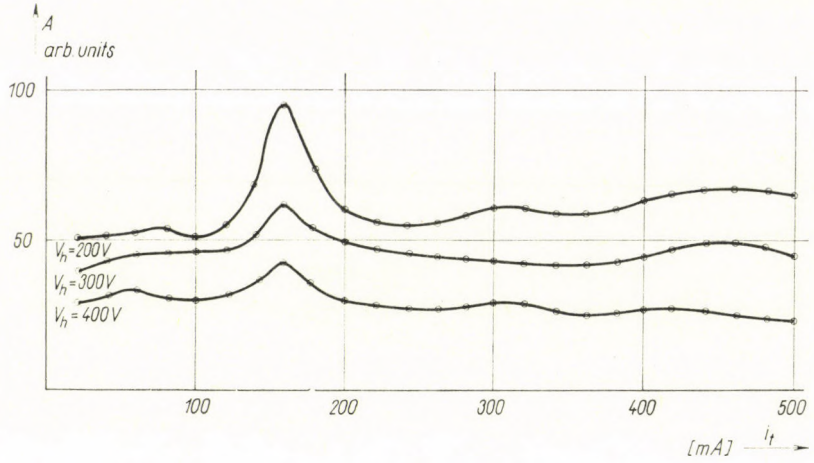


Fig. 5. The dependence of the amplitude A of the anode oscillations on the discharge current i_t measured in arbitrary units

same time in some of the experiments [79, 80] generally a characteristic breakdown point could be found on the frequency curve, which appeared at a certain discharge current, the position of this breakdown point being independent of the external parameters of the discharge. As was shown above, a current dependence of a perfectly similar character was found also in the case of the frequencies arising in the anode space and moreover the frequency breaking point there as well remained at a given current value independent of the parameters of the external current circuit. Here the dependence on temperature, as an external parameter, has not been investigated. In any case the similarity found between the current dependence of oscillations of the positive column with the current dependence of oscillations of the anode space, allows one to infer that there is some connection between them.

d) *The influence of the auxiliary electric circuit of the anode on the anode oscillations*

In the auxiliary electric circuit connected to the anode and reproduced in Fig. 1, the perturbing voltage was provided by the stabilized direct current power source APS. In these investigations the anode construction shown in

Fig. 2/a was employed, with a nickel cylinder of 10 cm casing height. In a part of the measurements the disc placed in the cylinder was on a higher potential than the anode cylinder. After the interchange of the poles of the power source APS the experiments were continued with the anode cylinder on a higher positive potential than the disc placed in it.

Otherwise these two series of measurements were conducted under perfectly identical conditions. The switching of polarity at a given perturbing

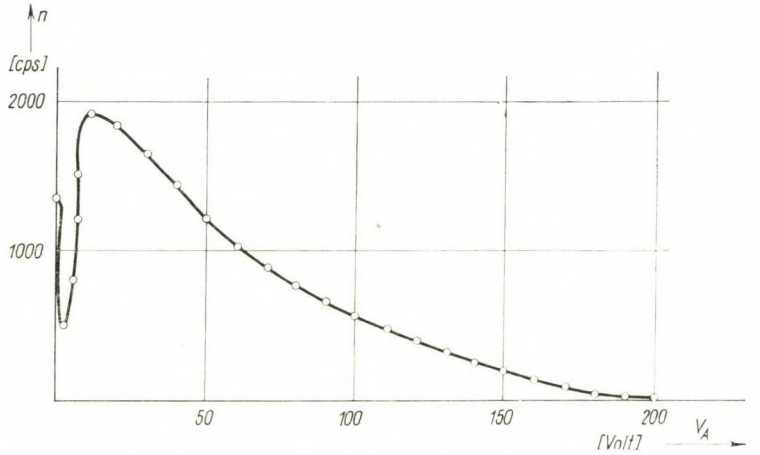


Fig. 6. The dependence of the frequency n of the anode oscillations on the direct current voltage V_A existing between the anode cylinder and the anode disc

voltage did not mean a change in the frequency of the appearing oscillations. In the case when the disc in the interior of the cylinder was on a higher potential, the number of harmonic oscillations increased and it was possible at the same time to show substantially lower frequencies as well, in the frequency range between 6—10 cps. Against this the amplitude of the oscillations increased twofold in the latter case.

In both cases the frequency of the anode oscillations varied in the same way with the perturbing voltage. The dependence of the frequency V_A on the perturbing voltage is shown in Fig. 6. The frequency reached its maximum height around 10 V, and with the further increase of the perturbing voltage it was reduced to a value around 10 cps. Then a luminosity fluctuation of regular period and which could be followed by the naked eye appeared at the end of the anode cylinder facing the positive column. All this indicates that with the help of the external electric circuit and with the anode construction used here it is possible to influence the frequency of the oscillations around the anode, and what is more, through the employment of a correspondingly high voltage (200 V) a very considerable reduction of the frequency (1900 cps — 10 cps) can be achieved.

The measurements have been effected at a discharge current of 500 mA, in the case of a feed voltage of 300 V. In the case when the external perturbing voltage was of zero value, a voltage difference of 10 V could be measured with the V_A static voltage measuring instrument shown in Fig. 1. The frequency of the anode oscillations amounted also in this case to 1900 cps. By the application of a perturbing voltage lower than 10 V a relatively stabilizing effect could be achieved, expressing itself in the local minimum of the curve plotted in Fig. 6 (frequency breakdown to 500 cps).

At other values of the perturbing voltage the vibration amplitude as compared to the amplitude belonging to the value $V_A = 0$ was reduced by some 50%.

e) *The influence of the shape of the anode upon the space around the anode*

The investigation of the influence exerted by the shape and the dimensions of the anode upon the space around the anode, without any external electric circuit on the anode side, was carried through at a discharge current of 400 mA and a feed voltage of 200 V. The measurements have been carried out with the electric setup shown in Fig. 1, and with the anode construction plotted in Fig. 2/a, by the method described above and under the known conditions.

In the course of the investigation first the anode cylinder was used as anode and electrically connected to the external electric circuit, then the anode disc placed inside the cylinder and finally the two jointly connected conductively. In the first two cases the anode disc and the anode cylinder, respectively, were not connected electrically, and then their outlets were in the air.

In all three cases the frequency of the anode oscillations was 1350 cps. When the cylinder acted alone as an anode or the cylinder and the disc placed in the cylinder were connected and jointly formed the anode, the positive column extended as far as the cylinder. No oscillation or light fluctuation could be shown in the space behind the cylinder. The shape of the anode space in this case can be seen on the photograph in Fig. 8. In this case the anode disc had been coated by glow light both on the side facing the positive column and on that facing the end of the tube, whereby it appeared as if both plates (the front and the back plate) of the anode which had been placed perpendicularly to the axis of the discharge would participate in the discharge. The glow light on the side of the cylinder facing the tube end fluctuated with an irregular frequency and these frequencies could be shown to exist also among the frequencies of the oscillations before the anode. The frequency of the glow light at the tube end varied up to 2—30 cps without showing any regularity.

As may be seen from a comparison of Figs. 7 and 8 the shape, dimensions and arrangement of the anode influenced in this case substantially and visibly

the character of the anode space. A situation similar to that shown on the photograph in Fig. 8 developed around the anode also when the rod shown in Fig. 3/b was placed inside the cylinder and acted as anode. In this case, how-

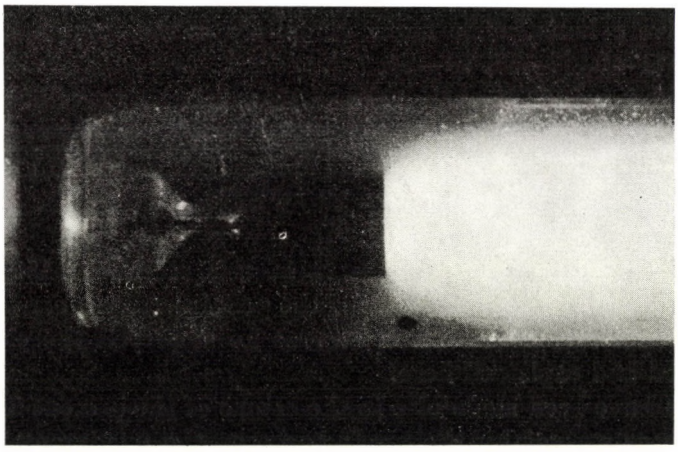


Fig. 7. Image of the anode space in the case when the cylinder acts as the anode, or when the cylinder and the disc are on the same potential and are connected to act as the anode

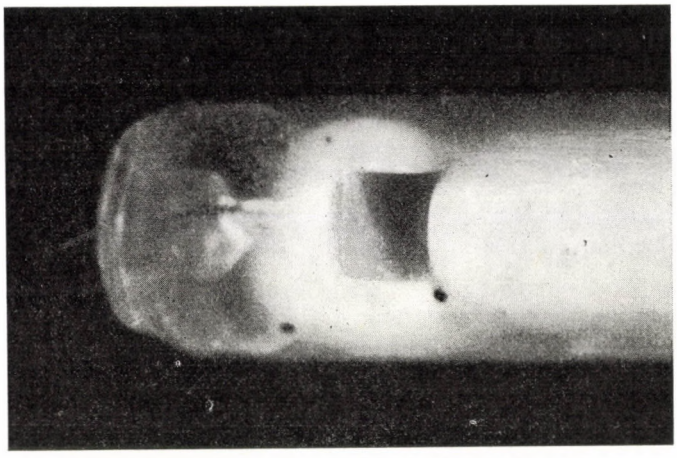


Fig. 8. Image of the anode space in the case when the disc inside the anode cylinder acts as the anode; the cylinder itself is not connected to the circuit

ever, the frequency of the oscillations developing at the tube end were higher, the oscillations became even more irregular and their amplitudes were about the four-fivefold of those found in the cases described above.

j) *The connection between the anode oscillations and the striations of the positive column*

As has been observed in the Introduction there are many who believe the cathode space [99—101] to be the source of the striation process appearing in the positive column, while many others are of the opinion that it is the anode space [1—9, 66—68, 70—74, 78].

In the opinion of the author the moving striation travelling in the positive column and consisting of space charges of different signs is influenced both by the cathode and by the anode spaces and the origin of these space-charge waves may also be traced back to the spaces before the corresponding electrodes.

SZIGETI carried out measurements in this connection [50] already in 1930 and showed that the oscillations of the positive column can be influenced from the anode side. In his measurements he employed the electrode construction used and explained here.

Studies have been made also in the course of the present investigations of the influence exerted by the anode space on the moving striation of the positive column.

Already during the description of the preceding experiments mention has been made of some similarities between the striations of the positive column and the oscillations of the anode space. In the relevant measurements one photo-cell detected the oscillations of the anode space, while another one showed the oscillations occurring in the middle of the positive column. The current fluctuations supplied by the photocells, amplified to the same degree were passed to the vertical and horizontal inputs of an oscilloscope, where frequency identification was effected by the Lissajoux curve method described above.

With the help of this measuring method in which two photo-cells were employed it was possible to demonstrate that the frequency of the anode oscillations modified and determined by the different anodic influences, agrees with the frequency of the wave of the moving striation travelling in the positive column towards the cathode.

In this way it has been proved that under the present conditions of investigation, the characteristics of one of the space-charge waves (positive space-charge wave) of the moving striation may be influenced from the anode side, through the anode oscillations. Also other oscillations were found in the case of the space-charging wave of moving striation, these repeated themselves irregularly or their amplitudes could be neglected as against the above-mentioned oscillations.

f) *Stability of the anode space*

As can be seen from the experimental results under the conditions described here it is expedient from the point of view of the stability of the anode space, to employ relatively large-surface anodes. The amplitudes of the appearing oscillations are smallest in this case.

The frequency of the anode oscillations and also their amplitude can be reduced through the application of the external anode circuit shown in Fig. 1 when a sufficiently high voltage (200 V) is applied. In this case the external anode circuit acts as a stabilizing circuit. A very interesting result was obtained showing that also in the case of an anode circuit voltage of a few Volts (2—3 Volts) very good stabilization can be achieved.

It can be seen that the shape of the anode considerably influences the oscillation conditions. Similar statements have been made in the literature for the case of anodes of a spiral shape.

The question of the stability of the anode space is important not only in order to ensure the conditions of local equilibrium but, as has been shown, the oscillations arising here will extend further in the positive column and may increase the energy losses through the influence they exert on the lamination processes.

5. Acknowledgement

The author wishes to express his gratitude to G. SZIGETI, Member of the Hungarian Academy of Sciences, for raising the subject and for his continuous guidance as well as to G. LAKATOS for his valuable remarks in connection with the experiments.

REFERENCES

1. W. PUPP, Phys. Z., **33,** 844, 1932.
2. W. PUPP, Phys. Z., **34,** 756, 1933.
3. M. STEENBECK, Müller-Pouillets Lehrbuch der Physik, IV, 3. Springer, Berlin, 1933.
4. I. LANGMUIR, Zeitschrift für Physik, **46,** 271, 1927.
5. L. VON ISSENDORF, Zeitschrift für Physik, **67,** 556, 1931.
6. T. TAKAMINE et al., Sci Papers I. P. C. R., **403,** 63, 1933.
7. T. TAKAMINE et al., Sci. Papers I. P. C. R., **403,** 26, 1933.
8. M. J. DRUYVESTEYN, Physica, **1,** 273, 1934.
9. N. A. KARELINA and B. N. KLARFELD, ISTF., **18,** 1235, 1948.
10. L. PEKAREK, Czechosl. J. Phys., **8,** 32, 1958.
11. W. PUPP, Z. f. techn. Physik, **7,** 257, 1934.
12. T. DONAHUE and G. H. DIEKE, Phys. Rev., **81,** 248, 1951.
13. A. A. ZAJTSEV, Vestn. Mosk. Univ. **10,** 1952.
14. G. W. FOX, Phys. Rev., **35,** 1066, 1930.
15. G. W. FOX, Phys. Rev., **37,** 815, 1931.
16. H. S. ROBERTSON, Phys. Rev., **105,** 368, 1957.
17. A. V. NEDOSPASOV, ZSTF., **28,** 173, 1, 1958.

18. B. N. KLARFELD, JETP., **22,** 66, 1952.
19. W. D. FARRIS, Proc. Phys. Soc., B**68,** 381, 1955.
20. W. DE GROOT, Physica, **8,** 23, 1928.
21. B. N. KLARFELD and V. SOBOLJEV, ZSTF., **17,** 319, 1947.
22. V. D. FARRIS, J. of Electronics, **1,** 1, 1955.
23. M. J. DRUYVESTEYN and F. PENNING, Rev. Mod. Phys., **12,** 2, 1940.
24. S. KOJIMA et al., J. Phys. Soc. Jap., **14,** 821, 1959.
25. S. KOJIMA et al., J. Phys. Soc. Jap., **11,** 1276, 1957.
26. H. J. MERRIL and H. W. WEBB, Phys. Rev., **55,** 1191, 1939.
27. D. H. LOONEY and S. C. BROWN, Phys. Rev., **93,** 965, 1954.
28. K. OGAWA, J. Phys. Soc. Jap. **14,** 1746, 1959.
29. N. L. OLESON and A. W. COOPER, Phys. Rev., **105,** 1411, 1957.
30. L. A. PARDUE and J. S. WEBB, Phys. Rev., **32,** 946, 1928.
31. N. R. LABRUN and E. K. BIGG, Proc. Phys. Soc., **65**B, 356, 1932.
32. H. YOSHIMOTO, J. Phys. Soc. Japan, **8,** 59, 1953.
33. H. YOSHIMOTO, J. Phys. Soc. Japan, **8,** 69, 1953.
34. M. J. DRUYVESTEYN and F. M. PENNING, Rev. Mod. Phys., **12,** 172, 1940.
35. K. OGAWA, J. Phys. Soc. Japan, **14,** 385, 1959.
36. J. D. COBINE and C. J. GALLAGHER, Phys. Rev. **70,** 113, 1946.
37. J. MICHEL, Wiss. Z. E. Moritz Arndt Univ. Greifswald, VIII, 295, 1958/59.
38. W. PUPP, Phys. Z., **36,** 61, 1935.
39. VAN GORKUM, Physica, **2,** 535, 1935.
T. DONAHUE, Phys. Rev., **82,** 571, 1951.
E. B. ARMSTRONG et al. Proc. Roy. Ir. Acad. (A) **54,** 291, 1951.
40. N. L. OLESON, Phys. Rev., **92,** 848, 1953.
41. S. WATANABE and N. L. OLESON, Phys. Rev., **99,** 1701, 1955.
42. H. S. ROBERTSON, Phys. Rev., **99,** 1662, 1955.
43. N. L. OLESON, Phys. Rev., **98,** 559, 1955.
44. V. D. FARRIS, J. Electronics, **1,** 60, 1955.
45. K. G. EMELEUS and N. R. DALY, Proc. Phys. Soc., B**69,** 114, 1956.
46. K. G. EMELEUS and N. R. DALY, Proc. Phys. Soc., B**69,** 433, 1956.
47. R. PILON, Phys. Rev., **107,** 2, 1957.
48. H. ACHTERBERG and J. MICHEL, Ann. Physik, **2,** 365, 1959.
49. P. J. HUTTON, Proc. 4th Int. Conf. Ion. Phen. in Gases, 1960. IB, 185, North-Holland Publishing Co.
50. G. SZIGETI and J. BITÓ, Acta Phys. Hung., **11,** 103, 1960.
51. J. R. M. COULTER et al., Physica, **24,** 828, 1958.
52. E. ROHNER, Appl. Sci. Res. Sec., B, **5,** 90, 1955.
53. J. BITÓ, Hung. Telecom. Techn., XI, 23, 1960.
54. L. PEKAREK and V. KREJCI, Preprint of Czech. J. Phys., 1960.
55. E. E. SALPETER, Phys. Rev., **120,** 1528, 1960.
56. A. L. EICHBAUM et al., J. Appl. Phys., **32,** 16, 1961.
57. G. LAKATOS and J. BITÓ, Acta Phys. Hung., **13,** 193, 1961.
58. G. LAKATOS and J. BITÓ, Acta Phys. Hung., **13,** 245, 1961.
59. K. WOJACZEK, Acta Phys. Hung., **11,** 35, 1960.
60. H. S. ROBERTSON and M. A. HAKEEM, Plasma Phys. Bull. of Dept. of Phys. Univ. Miami, **1—2,** 2, 1960.
61. H. YOSHIMOTO et al., J. Phys. Soc. Jap., **13,** 734, 1958.
62. A. V. NEDOSPASOV et al., JTP, **30,** 125, 1960.
63. G. LAKATOS and J. BITÓ, JTP, **32,** 902, 1962.
64. J. BITÓ, Hung. Phys. J., X, 303, 1962.
65. B. SAGGAU, Proc. 4th Int. Conf. Ion. Phen. in Gases, IIA, 281, 1960.
66. B. N. KLARFELD and N. A. NERETINA, JTP, **30,** 186, 1960.
67. J. R. M. COULTER, Physica, **26,** 949., 1960.
68. A. V. RUBSTINSKI et al., Radiotechn. i Elektronika, No 7, 1311, 1959.
69. E. NÖLLE, Technisch.-Wiss. Abhandl. Osram. Ges. **7,** 65, Springer Verlag, Berlin, 1958.
70. J. D. COBINE, Gaseous Conductors, Dover Publications, 1958.
71. N. A. KAPZOW, Elektrische Vorgänge in Gasen und in Vakuum, Deutscher Verlag der Wissenschaften, Berlin, 1955.
72. W. UYTERHOEVEN, Elektrische Entladungslampen, Springer Verlag, Berlin, 1938.
73. S. FLÜGGE, Handbuch der Physik, XXII, 422.
74. K. G. EMELEUS, The Conduction of Electricity through Gases, Methuens, New York, 1951.
75. N. A. NERETINA and B. N. KLARFELD. Rediotechn. i Elektronika **8,** 1301, 1959.

76. L. M. Kovrizsnikh, JETP, **37,** 1692, 1959.
77. H. L. Steele, Phys. Rev., **82,** 571, 1961.
78. J. R. M. Coulter et al, Proc. Phys. Soc., **77,** 476, 1961.
79. G. Lakatos and J. Bitó, Czech. J. Phys., (to be published).
80. G. Lakatos and J. Bitó, Acta Phys. Hung. (to be published).
81. B. N. Klarfeld and N. A. Neretina, JTP, **28,**, 296, 1958.
82. B. N. Klarfeld and N. A. Neretina, JTP, **29,** 15, 1959.
83. M. A. Hakeem and H. S. Robertson, J. Appl. Phys., **31,** 2063, 1960.
84. H. S. Robertson and M. A. Hakeem, Plasma Phys. Bull. Miami, **3,** 3, 1959.
85. A. A. Zajcev, Dokl. Akad. Nauk., **84,** 41, 1952.
86. A. W. M. Cooper, USA Naval Postgrad. School Calif., Priv. Comm.
87. E. V. Appleton and A. G. D. West, Phil. Mag., **45,** 879, 1923.
88. S. M. Rubens and J. E. Henderson, Phys. Rev., **58,** 446, 1940.
89. T. Takamine, T. Suga and A. Yanagihara, Sci. Papers IPCR Tokyo, **454,** 69, 1933.
90. M. Hoyaux and P. Gans, Bull. Ac. Roy. Soc. Belg., **42,** 1050, 1956.
91. W. Finkelburg, Techn. Phys. in Einzeldarst., **6,** 42, 1948.
92. I. Forsythe and A. Adams, Fluorescent and Other Disch. Lamps, J. Wiley, New York 78, 1948.
93. C. J. Warren, Ill. Eng., **47,** 37, 1952.
94. A. von Engel, Phil. Mag., **32,** 417, 1941.
95. W. Brunhart et al, Bull. SEV., **44,** 617, 1953.
96. H. J. Bürgin et al., Diplomaarbeit, E. T. H. 1953.
97. H. L. Steele, Ill. Eng., **49,** 349, 1954.
98. A. von Engel, Ionized Gases, Clarendon Press, Oxford, 1955.
99. P. Chapnik, JETP, **6,** 1497, 1958.
100. L. Pekarek, Czechl. J. Phys., **8,** 32, 1958.
101. L. B. Loeb, Phys. Rev., **76,** 255, 1949.

АНОДНЫЕ КОЛЕБАНИЯ РАЗРЯДОВ

Й. БИТО

Резюме

Автором статьи даётся литературный обзор о результатах исследования колебательных явлений, появляющихся в анодной части разрядов. Излагаются условия измерений и метод исследования, применённый автором. Опыты проводились при ртутно-аргонном разряде постоянного тока. При опытах применялась специальная анодная конструкция, предложенная Сигети; воздействие на наблюдаемые колебательные явления осуществляется извне электрической цепью. Сделается вывод по отношению возможности увеличения стабильности анодного пространства.

DETERMINATION OF WAVE FUNCTIONS OF MOLECULAR SYSTEMS BY THE METHOD OF THE MOMENTA I.

By

T. SZONDY

RESEARCH GROUP FOR THEORETICAL PHYSICS OF THE HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST

(Presented by A. Kónya. — Received 13. XI. 1963)

It is possible to determine an approximate wave function of a molecular system by the requirement that the wave function should be invariant under certain kinds of improvements. This possibility is used for developing a method for the determination of molecular wave functions. The method makes possible *a)* to avoid a considerable part of the difficulties of integration of the usual quantumchemical calculations, *b)* to obtain in some cases qualitative information about the accuracy of the wave function in certain regions of the configurational space, *c)* to find relations between different variational methods used in quantumchemical calculations. It is remarkable, that almost all the variational methods applied so far to problems of quantum chemistry — e. g. the method of energy variation and the method of local energies — can be formulated also in terms of the present method.

1. Introductory remarks

A variational method for the determination of molecular wave functions will be considered [1]. In order to obtain a clear picture of the advantages as well as the shortcomings of the method, let us first recall the main steps in the determination of a molecular wave function by some usual variational method of quantum chemistry e.g. the method of energy variation (MEV) [2].

1. The choice of the variational function. As no rigorous method is known for the estimation of the (potential) accuracy of a given variational function, the choice must be based *a)* on a — generally qualitative — mathematical investigation of the problem, *b)* on the results of exactly solvable model problems [3], *c)* on the results of similar approximate calculations already carried out for related systems, *d)* on experimental information about the considered system or related ones, *e)* on the actual computational possibilities.

2. The determination of the values of the variational parameters.

3. The checking of the results. As no non-empirical method is known so far, which would be practically useful for this purpose in the case of systems that are not extremely simple, generally the best that can be done is to compare the calculated values of some carefully chosen properties of the system with the empirical data. Although in general no rigorous conclusions can be drawn from these data concerning the accuracy of other physical quantities [4],

they can give a rather reliable picture of the accuracy of the obtained approximate wave function.

In this paper we shall be concerned with step 2. In particular, our aim will be to reduce the restrictions on the choice of the variational function which are forced upon us by difficulties of integration.

2. The method of the momenta

Let H be the electronic Hamiltonian of the considered system, ψ an eigenfunction of H belonging to the eigenvalue E, and let $\varphi(a)$ be a variational wave function involving a set of variational parameters $a_0, a_1, \ldots, a_m$ briefly denoted by a. It will be assumed, that there exists some domain D in the space of the a_i's with the property, that if $a \in D$, $\varphi(a)$ can be regarded as a good approximation to ψ, and our aim will be to find such an $a \in D$. Our considerations will be restricted to real Hamiltonians and wave functions [5] and it will be assumed, that the set a includes also the normalization factor of φ.

Let $w_1, w_2, \ldots, w_n$ $(n \geq m)$ be a set of linearly independent functions of the (spatial and spin) coordinates of the electrons of the system satisfying the same symmetry, differentiability and boundary conditions as φ. Let us consider the (normalized) wave function

$$\chi(a, \beta) = (1 + \beta_0)\, \varphi(a) + \sum_{i=1}^{n} \beta_i w_i \tag{1}$$

at some fixed values of the a_i 's as a function of the set of variational parameters $\beta_0, \beta_1, \ldots, \beta_n$ briefly denoted by β. Determining the values of the β_i 's by MEV, we obtain the set of equations

$$< \varphi(a) \,|H - \varepsilon|\, \varphi(a) > (1 + \beta_0) + \sum_{i=1}^{n} < \varphi(a) \,|H - \varepsilon|\, w_i > \beta_i = 0, \tag{2a}$$

$$< w_j \,|H - \varepsilon|\, \varphi(a) > (1 + \beta_0) + \sum_{i=1}^{n} < w_j \,|\, H - \varepsilon|\, w_i > \beta_i = 0, \tag{2b}$$

$$(\mathrm{j} = 1, 2, \ldots, n),$$

ε denoting that root of the secular equation of (2), which belongs to the state ψ. We assume, that for the a's to be considered, such a root exists.

Evidently the β_i 's satisfying (2) depend on the values of the a_i 's. It can be expected, that the better an approximation is $\varphi(a)$ to ψ, the less can it be improved further, i.e. the smaller will be the absolute values of the β_i 's satisfying (2). It seems thus to be reasonable to determine

the "best" values of the a_i 's by the requirement that the improvement due to the function $\sum_{i=1}^{n} \beta_i w_i$ should be as small as possible.

In general, the smaller the absolute values of the integrals $<w_i|H - \varepsilon|\varphi(a)>$, the smaller will be the absolute values of the β_i 's determined by (2). More rigorous statements about the values of the β_i 's can be made in general only by a careful investigation of eq. (2). No such problem arises, however, if it can be achieved by a suitable choice of the a_i 's that $\varphi(a)$ satisfies the equations

$$<w_i|H - \varepsilon|\varphi(a)> = 0, \tag{3a}$$

$$<\varphi(a)|\varphi(a)> = 1, \tag{3b}$$

$$(i = 1, 2, \ldots, n).$$

In this case we obtain namely $\beta_0 = \beta_1 = \ldots = \beta_n = 0$ thus $\varphi(a)$ cannot be improved within the framework of MEV by adding to it *any* linear combination of the form $\beta_0 \varphi(a) + \sum_{i=1}^{n} \beta_i w_i$. It should be noted that if the w_i 's are chosen according to the principles to be discussed later, the compatibility of eqs. (3) is generally automatically ensured for the case $m = n$.

We shall refer to the integrals $<w_i|H - \varepsilon|\varphi>$ as *momenta* of $(H - \varepsilon)\varphi$ with the *weight functions* w_i, and we shall regard (3) as the basic equation of a variational method for the determination of molecular wave functions, called briefly *the method of the momenta* (MM).

A possible generalization of (3) for the case $n > m$, when the equations (3) are generally contradictory, is

$$\sum_{i=1}^{n} c_i |<w_i|H - \varepsilon|\varphi(a)>|^2 = \min, \tag{4a}$$

$$<\varphi(a)|\varphi(a)> = 1, \tag{4b}$$

the c_i 's denoting some suitably chosen positive weight factors. Eqs. (4) will not be considered in detail in this paper [6].

An important property of MM follows directly from the way it has been introduced. If a wave function cannot be improved by the introduction of some new variational parameters, this can provide information about its accuracy in certain regions of the configurational space. Let us assume, for example, that the set $w_1, w_2, \ldots, w_n$ contains a subset $w_{i_1}, w_{i_2}, \ldots, w_{i_k}$ with the property that $\sum_{j=1}^{k} \beta_{i_j} w_{i_j}$ is highly flexible in some domain of the configurational space, and has very small values outside it. As $\varphi(a)$ has the property that it cannot be

improved by adding to it any linear combination $\sum_{j=1}^{k} \beta_{i_j} w_{i_j}$ it can be expected, that it is itself a particularly good approximation to ψ in the considered domain. There may even exist the possibility that by the suitable choice of the weight functions we may — to some extent — prescribe the regions in which the accuracy of $\varphi(a)$ should be particularly high. Evidently such a "guided approximation" (PREUSS [7]) will be successful only if $\varphi(a)$ contains parameters, which make possible a particularly good approximation to ψ in the considered region. More generally, it seems to be necessary in any application of MM to ensure that the regions of high flexibility of $\varphi(a)$ and $\sum_{i=1}^{n} \beta_i w_i$ should at least roughly coincide. The principles for the construction of weight functions to be considered in the next sections generally automatically satisfy this requirement.

Let us assume now that we have a set of weight functions $\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_n$, and we have some good reason to expect that they lead to a wave function $\varphi(\bar{a})$ which is a good approximation to the exact one in those regions of the configurational space which are important for our actual purposes. The weight functions $\bar{w}_i$ are, however, of an inconvenient mathematical form which — because of the practical difficulties involved — prevents the computation of the integrals $<\bar{w}_i|H - \varepsilon|\varphi>$. If similar difficulties are met in the case of other variational methods of quantum chemistry, generally the best that can be done is to make neglections in either H or φ. In the first case this means that we study a model problem instead of the actual one, in the second case we must content ourselves with a rougher approximation to the wave function. In the case of MM, on the other hand, it is often possible to restrict the necessary simplifications to the weight functions, which have no direct physical meaning, but which play only a rather qualitative role. Replacing the $\bar{w}_i$'s by some mathematically more convenient approximations $\tilde{w}_i \approx \bar{w}_i$ we obtain the wave function $\varphi(\tilde{a})$ instead of $\varphi(\bar{a})$. As $\varphi(\tilde{a})$ must be insensitive, by definition, to improvements similar to those of $\varphi(\bar{a})$, it may be expected, that $\varphi(\tilde{a})$ will differ only slightly from $\varphi(\bar{a})$. (It should be remembered, that both $\varphi(\bar{a})$ and $\varphi(\tilde{a})$ are the exact solutions of an ordinary MEV approximation to ψ, carried out with two very similar variational wave functions, respectively.)*

This relative freedom in the choice of the weight functions is one of the most important features of MM in its practical use. In the next section relations of MM to other variational methods will be discussed, and it will become possible to find highly reliable sets of $\bar{w}_i$'s and in this way a considerable part of the arbitrariness in the choice of the weight functions can be eliminated. These

* Note added in proof. The conditions under which this expectation is justified will be investigated in more detail in the next part of this paper. If the weight functions are no approximations to some highly reliable $\bar{w}_i$'s but are "chosen at random", the reliability of MM may be compared with that of the method of local energies.

$\overline{w}_i$'s are, however, generally difficult to handle, and by replacing them by more convenient approximations $\widetilde{w}_i$, we may obtain an approximate solution to the original variational problem.

Evidently (3) and (4) generally do not loose their sense if the weight functions do not satisfy all the differentiability and boundary conditions required of a wave function. E.g. as $(H - \varepsilon)\varphi$ vanishes exponentially in infinity, moderate changes in the asymptotic behaviour of the weight functions generally do not modify the roots of (3) and (4) considerably. This is an important circumstance, as just those types of weight functions, that offer the most considerable computational advantages, have an incorrect asymptotic behaviour. This shows also that MM is a more general method than MEV. In the next section we shall see that MEV can be regarded (not only in the case of linear parameters) as an important special case of MM.

So far we have disregarded a problem of practical importance. The determination of the value of ε occurring in (3) and (4) requires namely the computation of just those complicated integrals that we wanted to avoid. However, the smaller the absolute values of the integrals $<w_i|\varphi>$, the less the roots of (3) and (4) depend on ε. Consequently, we have to prefer weight functions, which — depending themselves on a — satisfy the condition

$$<w_i(a)|\varphi(a)> = 0. \tag{5}$$

It can be readily seen that all our previous considerations remain valid also for weight functions depending on a, and that (5) does not restrict the generality of MM. In the following (5) will turn out to be quite a natural requirement, automatically satisfied by the most important type of $\overline{w}_i$'s.

Evidently if (5) is not exactly satisfied, ε most be replaced by some approximate value, e.g. the empirical value of the energy of the system.

3. Relations between MM and other variational methods of quantum chemistry

Many of the variational methods used in quantumchemical calculations can be formulated also in terms of MM. A few examples of this kind will be considered in this section. At the same time these relations provide highly reliable sets of $\overline{w}_i$'s.

1. The method of energy variation. MEV determines the values of the variational parameters from the condition

$$\varepsilon(a) = \frac{<\varphi(a)|H|\varphi(a)>}{<\varphi(a)|\varphi(a)>} = \text{stationary}. \tag{6}$$

Differentiating (6) with respect to a_i, and taking into account that φ is real and H is self-adjoint, we obtain

$$< \partial\varphi(a)/\partial a_i \mid H - \varepsilon \mid \varphi(a)> = 0 \,, \qquad (7)$$

$$(i = 0, 1, \ldots, m.)$$

If $\varphi(a)$ is normalized

$$<\varphi(a) \mid \varphi(a)> = 1 \,, \qquad (8)$$

one of the parameters of the set $a_0, a_1, \ldots, a_m$, say a_0, can be expressed by the others and thus we can drop one equation from (7), say the 0th. From the comparison of (7) with (3a) it follows immediately that

$$w_i = \partial\varphi/\partial a_i \,, \qquad (9a)$$

$$(i = 1, 2, \ldots, m) \,,$$

$$\varepsilon = \varepsilon(a) \,. \qquad (9b)$$

Differentiation of (8) with respect to a_i leads to

$$<\partial\varphi(a)/\partial a_i \mid \varphi(a)> = 0 \,, \qquad (10)$$

i.e. (5) is automatically satisfied by the weight functions (9a).

The weight functions (9a) seem to be particularly suitable to serve as $\bar{w}_i$'s.

2. The PREUSS method. This method determines the wave function from the condition

$$<f(H - \varepsilon)\,\varphi(a) \mid f(H - \varepsilon)\,\varphi(a)> = \min \,, \qquad (11a)$$

$$<\varphi(a) \mid \varphi(a)> = 1 \,, \qquad (11b)$$

where ε denotes the value $\varepsilon(a)$ satisfying (6) and f is some suitably chosen weight function (not to be confused with the w_i's) depending on the electron coordinates.

Differentiating (11a) with respect to a_i we obtain similarly to the former case

$$w_i = f^2(H - \varepsilon)\partial\varphi/\partial a_i \,. \qquad (12)$$

3. The local energy method. Let x stand for the spatial and spin coordinates of all the electrons of the system, and let x_i be some fixed

value of x. The basic equation of the local energy method can be written in the form

$$\sum_{i=1}^{n} c_i \,|\, < \tilde{\delta}(x - x_i) \,|\, H - \varepsilon |\, \varphi(a) > |^2 = \min, \tag{13a}$$

$$<\varphi(a) \,|\, \varphi(a)> = 1\,, \tag{13b}$$

the c_i's denoting positive weight factors, $\tilde{\delta}$ denoting Dirac δ-functions in the spatial and Kronecker δ's in the spin coordinates, while ε denotes some approximation to the energy of the system. It follows immediately that

$$w_i = \tilde{\delta}(x - x_i)\,. \tag{14}$$

4. The method based on the equations

$$<\varphi(a) - \psi \,|\, \varphi(a) - \psi> = \min\,, \tag{15a}$$

$$<\varphi(a) \,|\, \varphi(a)> = 1\,. \tag{15b}$$

(This method is closely related to the method of the natural spin orbits [8] and also to the methods of A. DALGARNO, J. T. LEWIS and C. SCHWARTZ [9]). Differentiating (15a) and (15b) with respect to a_i, we obtain

$$<\partial\varphi/\partial a_i |\, \psi> = 0\,, \tag{16a}$$

$$<\partial\varphi/\partial a_i |\, \varphi> = 0\,. \tag{16b}$$

From (16) it follows after some computation that

$$<\partial\varphi/\partial a_i |\, 1 - P \,|\, \varphi(a)> = 0\,, \tag{17}$$

$1 - P$ being the operator projecting off the state ψ. Making use of the fact, that $1 - P$ is self-adjoint, we can rewrite (17) and obtain the equation

$$<(H - \varepsilon)^{-1}\,(1 - P)\,\partial\varphi/\partial a_i |\, H - \varepsilon \,|\, \varphi(a)> = 0\,, \tag{18}$$

ε denoting any approximation to E, of which we assume for the sake of simplicity, that it does not coincide with any eigenvalue $E' \neq E$ of H. Comparing (18) with (3a) we find that (15) is equivalent to an MM approximation, the weight functions being solutions of the equation

$$(H - \varepsilon)\, w_i = (1 - P)\,\partial\varphi/\partial a_i\,. \tag{19}$$

4. Some aspects concerning the construction of the $\widetilde{w}_i$'s

In this section some types of functions will be considered, that seem to be particularly suitable to serve as building elements of the $\widetilde{w}_i$'s. It will be tacitly assumed throughout the section — if not explicitly indicated otherwise — that φ is built from determinants of single-particle spin-orbitals, the space-dependent parts of which are linear combinations of Slater functions centered at arbitrary fixed points, in general at the nuclei. If not indicated otherwise the wave function is assumed *not* to include HYLLERAAS-type correlation factors.

The following types of functions seem to have particular advantages as building elements of the $\widetilde{w}_i$'s:

1. Polynomials of the cartesian coordinates of the electrons with spin-dependent coefficients. All the integrals occurring in the momenta can be obtained in this case in the form of rather simple closed analytical expressions, the most complicated terms being two-centre Coulomb integrals. According to WEIERSTRASS' theorem, there is no limitation to the accuracy, with which the $\widetilde{w}_i$'s can approximate to the $\overline{w}_i$'s, naturally apart from their asymptotical behaviour. However, — assuming the same degree of approximation — the smaller the spatial extension of the system the simpler the expressions for the $\widetilde{w}_i$'s. Consequently this type of weight functions seems to be particularly advantageous in the case of not very large "compact" molecules, as methane for example.

2. Polynomials and Gaussian functions of the cartesian coordinates of the electrons with spin-dependent coefficients. Expanding the Slater functions into series of Gaussian functions [10], the computation of the momenta requires — besides elementary integrations — only the computation of one and two-dimensional numerical quadratures resp. cubatures. This is by far less tedious than the computation of three and four-centre integrals between Slater orbitals [11].

3. $\tilde{\delta}$-functions of the electron coordinates as defined in 3) of the previous section. This special case of MM is equivalent to the method of local energies, and is consequently limited to systems involving a very low number of electrons.

4. Let us consider now wave functions consisting of antisymmetrized products of two-electron orbitals [12] involving HYLLERAAS-type correlation factors between electrons belonging to the same orbital, but no such factors between electrons belonging to different orbitals. MEV leads in this case to integrals involving the coordinates of all the electrons inseparably. In the case of MM, on the other hand, if the weight functions are built from products of single-particle functions, the most difficult integrals occurring in the momenta involve the coordinates of four electrons.

Evidently also other types of weight functions (e.g. plane waves) may have — in certain cases — considerable computational advantages.

Finally some general remarks: MM seems to be applicable also to time-dependent problems. Calculations are in preparation on the inelastic scattering of hydrogen atoms by approximately solving the time-dependent Schrödinger equation for a time-dependent variational wave function.

Probably semiempirical corrections, similar to those proposed by A. MUKHERJI and M. KARPLUS [13] for the case of MEV will be highly useful also in the case of MM.

5. Numerical results

The values of the variational parameters of simple variational wave functions for the ground state of the helium atom and the hydrogen molecule have been computed by MM. Both the wave and the weight functions used in the calculations are extremly simple (the latter can be regarded as "chosen at random") and the results are only meant as illustrations. The results of more elaborate calculations will appear in subsequent papers.

The variational functions used in our calculations were

$$\varphi = a_0 \exp[-a_1(r_1 + r_2)] \tag{20}$$

for the helium atom, and

$$\varphi = a_0 \left\{\exp[-a_1(r_{a1} + r_{b2})] + \exp[-a_1(r_{b1} + r_{a2})]\right\} \tag{21}$$

for the hydrogen molecule. The nuclei of the hydrogen molecule were assumed to be at the fixed distance of 1,40 at. u. (the experimental value of the equilibrium distance).

a_0 is determined by the normalization condition, and consequently one weight function was needed for each computation. The computations were carried out separately with the weight functions

$$w_1^{(1)} = \text{const} \tag{22a}$$

and

$$w_1^{(2)} = 1 + c(r_1^2 + r_2^2) \tag{22b}$$

for the helium atom and with the weight functions

$$w_1^{(1)} = \text{const} \tag{23a}$$

and

$$w_1^{(2)} = 1 + c(r_{a1}^2 + r_{b2}^2) + c(r_{b1}^2 + r_{a2}^2) \tag{23b}$$

for the hydrogen molecule. The value of the parameter c has been determined in both cases by the condition (5)

$$< w_1^{(2)} \mid \varphi > = 0. \tag{24}$$

In the cases $w_1^{(1)} = \text{const}$, when it is evidently impossible to satisfy the orthogonality condition, the calculations were carried out separately by first taking ε equal to the empirical value and then to the value determined by (6). The results of the calculations, together with the results of the corresponding MEV calculations are listed in the Table.

Table

System	Weight function	ε (at. u.)	a_1 (at. u.)
He	$w_1^{(1)}$	−2,904*	1,72
	$w_1^{(1)}$	−2,848**	1,69
	$w_1^{(2)}$	—	1,63
	Determined by MEV		1,69
H_2	$w_1^{(1)}$	−1,175*	1,28
	$w_1^{(1)}$	−1,138**	1,25
	$w_1^{(2)}$	—	1,31
	Determined by MEV		1,17

* The empirical value.
** The value (6).

Note. The reader's attention is drawn to the paper of D. M. SCHRADER and S. PRAGER [14], where the basic problem of the work presented here, the difficulties of integration in quantum chemistry, is attacked from a completely different and very attractive starting point. It seems, however, to be too early to make any comparison of the usefulness of these two methods.

Interesting results concerning the problem of "guided approximations" are contained in the papers of VETCHINKIN [15]. These results are of a considerable interest also for the problems discussed in this paper.

Acknowledgement. The author wishes to express his best thanks to his wife for many helpful discussions.

REFERENCES

1. Similar methods have already been used in mathematical physics to solve complicated differential and integral equations. Cf. e.g. L. COLLATZ, Numerische und graphische Methoden. (Encyclopedia of Physics, Springer Vlg. Berlin—Göttingen—Heidelberg, 1955. Vol. II, Sect. 61, pp. 438—9.)
2. Concerning questions about other variational methods of quantum chemistry the reader is referred to the paper by H. PREUSS in Fortschritte der Physik, **10,** 271, 1962. This paper contains a detailed discussion of the problems considered in this section.
3. E.g. non-interacting electrons in some suitably chosen external potential.
4. It is shown in [2] that the energy of the system is no exception to this statement.
5. MM can be generalized also to complex Hamiltonians and wave functions, but some practical problems require further investigation. It is not clear, for example, whether it is better to solve directly the equation.

$$H\psi - E\psi = 0$$

or to solve the set of simultaneous equations

$$H_1 \psi_1 - H_2 \psi_2 - E \psi_1 = 0 ,$$

$$H_1 \psi_2 + H_2 \psi_1 - E\psi_2 = 0 ,$$

H_1 and H_2, ψ_1 and ψ_2 denoting the real and the imaginary part of H and ψ, respectively. As regards a justification for the use of real wave functions, the reader is referred to the paper by P. GOMBÁS in Acta Phys. Hung., **3,** 105, 1953, in particular to footnote 5. It should be noted, that if H is real and E is non-degenerate, ψ *must* be real. Were namely Re ψ and Im ψ linearly independent, E would be degenerate.
6. An equation of this type has already been proposed in the paper by T. SZONDY in Z. Naturforschg., **17a,** 552, 1962.
7. The idea of such a "guided approximation" has been raised in the important papers by H. PREUSS in Z. Naturforschg., **13a,** 439, 1958; **16a,** 598, 1961.
8. Cf. e.g. P. O. LÖWDIN and H. SCHULL, Phys. Rev., **101,** 1730, 1956.
9. A. DALGARNO and J. T. LEWIS, Proc. Roy. Soc. (London) **A233,** 70, 1956;
C. SCHWARTZ, Annals of Physics, **6,** 156, 1959; **6,** 170, 1959; **6,** 178; 1959.
10. S. F. BOYS, Proc. Roy. Soc. (Lond.) **A200,** 542, 1950.
11. R. KIKUCHI, J. Chem. Phys., **22,** 148, 1954.
Cf. also M. KARPLUS and I. SHAVITT, J. Chem. Phys., **36,** 550, 1962.
12. A detailed discussion of the uses of this type of wave functions together with many other references can be found in the paper by E. KAPUY, Acta Phys. Hung., **15,** 341, 1963.
13. A. MUKHERJI and M. KARPLUS, J. Chem. Phys., **38,** 44, 1963.
14. D. M. SCHRADER and S. PRAGER, J. Chem. Phys., **37,** 1456, 1962.
15. *С. И. Ветчинкин,* Доклады Акад. Наук СССР, **147,** 1328, 1962; Ж. Оптика и Спектроскопия **14,** 314, 1963; **15,** 291, 1963.

ОПРЕДЕЛЕНИЕ ВОЛНОВОЙ ФУНКЦИИ МОЛЕКУЛЯРНЫХ СИСТЕМ МЕТОДОМ МОМЕНТОВ

Т. СОНДИ

Р е з ю м е

В принципе имеется возможность для определения волновой функции молекулярной системы с соблюдением требования, согласно которому волновая функция при проведении исправлений определённого рода должна быть инвариантной. Использованием этой возможности в работе выводится метод для определения волновой функции молекулярной системы. Исследованный метод делает возможным а) избегать значительную часть трудностей интегрирования, имеющих место при применении обычных квантовохимических приближённых методов, б) в некоторых случаях получить качественные сведения о степени точности волновой функции в определённых областях конфигурационного пространства, в) находить связь между различными вариационными методами, применёнными при квантовохимических вычислениях.

ON THE NUCLEATION IN DIFFUSION CRYSTAL GROWING

By

E. LENDVAY

RESEARCH INSTITUTE FOR TECHNICAL PHYSICS OF THE HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST

(Presented by G. Szigeti — Received: 2. I. 1964)

Practically insoluble crystals can be grown from solutions by the method of diffusion ion transport. In this process first of all the concentrations of the different ions must reach values that are equal or higher than that of the solubility product. When this has happened nucleation and growing begins. By the variation of the diffusion circumstances (transport rate, ion sources, geometry of the growing system, etc.) one can regulate the initial period and the coordinates resp. the rate of the solid phase separation. The mentioned factors are of decisive importance in the morphology of separating solids. This paper, starting from the time-dependence of concentration distributions, investigates the role of the factors influencing the diffusion transport and nucleation.

1. Introduction

Most of the crystal growing methods are based on the supersaturation principle, and differ from each other only in the methods by which supersaturation is produced. Generally, if the crystals are insoluble, the growing methods become very difficult. In these methods high temperature or pressure or both are used. (growing from vapour phase, melts, hydrothermal growing, etc.). This renders the technology difficult and raises the probability of disagreeable impurities being built into the crystals.

Many years ago JOHNSTON and FRENELIUS prepared some crystals by a method in which the formation of insoluble crystals was simpler than in the usual cases, mentioned above [1, 2]. The principle of this diffusion crystal growing is the chemical precipitation with regulated rate. The properties of the separated solid phase strongly depend on the rate of mass separation. During the usual chemical precipitation this process is too fast, therefore the precipitate has an amorphous or microcrystalline structure. If the rate of precipitation is regulated by the diffusion ion transport and only low supersaturation is produced in the system, there is a possibility that a macrocrystalline solid phase is being formed.

No work has been published dealing with this method in detail. In the present paper the diffusion nucleation is investigated as a first step towards a general crystal growing method. Monocrystals (sulfides, carbonates, Mn-phtalate, etc.) were prepared by the diffusion method and in this way the

theoretical statements were confirmed by the real processes. The detailed description of the methods and the monocrystals themselves does not fall within the scope of this paper. We hope to return to this in due course.

2. The diffusion nucleation

We investigate the formation of insoluble crystalline phases. Assuming that the Law of Mass Action can be applied to this problem, we base our treatment on the Solubility Product Principle.* If this proves to be necessary, in concentrated solutions we shall take into account also the influence of foreign ions. As in the diffusion system (reaction volume) the ions are generated separately, the influence of the foreign ions can be of great importance in some cases in that region of the system, where nucleation begins. If the solubility of the formed solid is extremely low, this effect is negligible because the ion strength around the nucleus is also small. At higher solubilities, however, the effect of foreign ions and the activity must be considered.

At the beginning of the growing process the crystals or solutions of the raw materials are put into determined parts of the system. The remaining part of the system is occupied by the pure solvent. At the beginning of the diffusion processes no part of the system contains both ions of the precipitated compound, therefore nucleation in any part of the system is impossible. As in the whole system the diffusion processes endeavour to balance the concentration differences, a time-dependent concentration product is originated everywhere in the system. If the concentration product reaches the solubility product in a certain part of the system, nucleation begins. As we shall see later, the dimensions of the volume in which the process is started depend on the experimental circumstances.

The fundamental differential equation which determines the mass transport resp. the factors of the concentration product belonging to the corresponding element of the volume in time is

$$\left(\frac{\partial C}{\partial t}\right)_x = \operatorname{div}(D \operatorname{grad} C). \tag{1}$$

In our cases the medium is generally isotropic and the diffusion coefficient is constant.

The treatment of the three-dimensional cases even if the system is of a rather simple geometry is very difficult, therefore our studies will be confined to the treatment of cases in which the one-dimensional solution is sufficient. For the majority of such cases it is characteristic that the ion sources

* In very dilute solutions the activity of an ion is practically equal to its concentration therefore we may substitute concentrations for activities.

are planar. The planes of the ion sources are parallel and the active surfaces are equal or the perpendicular cross-section of the system. If the cross-section of the growing system is constant, the concentration is constant in time in the cases mentioned i.e. when the planes are parallel to the ion source surfaces. Many problems of radial flow in a sphere can also be deduced immediately from the corresponding linear problems, when the following substitution is introduced:

$$U = Cr\,, \tag{2}$$

where C is the concentration at a point and r the radius vector pointing to that point. With the above substitution we obtain instead of eq. (1)

$$\left(\frac{\partial U}{\partial t}\right) = D\left(\frac{\partial^2 U}{\partial r^2}\right). \tag{3}$$

Eq. (3) is the well-known Fick Law for linear diffusion with U, t and r, therefore this three-dimensional problem can be reduced to a linear one.

2.1. *Solid-liquid systems in case of uniform media*

In the simple case of diffusion growing the crystals of the raw materials are in equilibrium with the solvent in definite parts of the system. This means that the ion sources are the saturated solutions of suitable compounds. Until the solid phases of the raw materials are present, the continuous dissolution of the solids ensures the constant concentrations of the ion sources. In these cases the initial distribution of all types of ions is restricted to a finite region and we have an initial state defined by

$$\begin{aligned} c_A &= c_A^0 \quad x = 0\,, \qquad \text{for all } t > 0 \\ c_B &= c_B^0 \quad x^1 = 0\,, \end{aligned}$$

and

$$\begin{aligned} c_A &= 0 \quad x > 0\,, \qquad \text{at} \quad t = 0\,, \\ c_B &= 0 \quad x^1 > 0\,, \end{aligned}$$

where $x^1 = f(x)$ is a transformed coordinate. If the ion sources are at the end points of a linear system with a length of L, it is possible to choose the transformation $f(x) = L - x = x^1$. The solution of eq. (1) with the mentioned initial conditions for an ion component A is [3]

$$C_A(x, t) = \frac{1}{2}\, C_A^0 \operatorname{erfc} \frac{x}{2(D_A t)^{1/2}}\,. \tag{4}$$

The same procedure may be applied to the component B and the solution for B is

$$C_B(x,t) = \frac{1}{2} C_B^0 \operatorname{erfc} \frac{L-x}{2(D_B t)^{1/2}}, \tag{5}$$

where C_A^0 and C_B^0 are the constant (saturation) concentrations of the ion sources, D_A resp. D_B are the diffusion coefficients in the medium. Determining the saturation concentrations in $\frac{1}{2} C_i^0$, the factor of 0,5 can be neglected, as in the presence of solid raw materials for practically all t

$$x < 0 \qquad C_A = C_A^0 \quad \text{and} \quad x > L \quad C_B = C_B^0.$$

The foregoing solutions are corresponding to the infinite medium. Because of the low solubility of the forming solid phase, by a suitable choice of the length of the reaction volume it can be achieved that the concentration of opposite ions in the neighbourhood of the ion sources in the nucleation sphere is negligible. In this case we can neglect the reflection of the ion flow from the boundaries and by equs. (4) and (5) the concentration distributions, resp. the concentration product can be described in the whole system. The concentration product, $K(x, t)$, should, therefore, be represented as a function of x and t by an equation of the type

$$K(x,t) = (C_A^0)^m (C_B^0)^n \left(\operatorname{erfc} \frac{x}{2(D_A t)^{1/2}}\right)^m \left(\operatorname{erfc} \frac{L-x}{2(D_B t)^{1/2}}\right)^n \tag{6}$$

because the concentration product in general is

$$K(x,t) = (C_A)^m (C_B)^n. \tag{7}$$

The value of $K(x, t)$ for all x is at $t = \infty$

$$K(x,\infty) = (C_A^0)^m (C_B^0)^n \tag{8}$$

corresponding to the saturation of the medium. In practice this case is never realized because of the chemical reaction between ions A and B and the separation of solid $A_m B_n$ from the system.

As we now investigate the processes only in the nucleation stage, the mentioned phase separation does not disturb our treatment at all. At $t = 0$ the value of $K(x, 0)$ for all x is

$$K(x,0) = 0\,, \qquad 0 < x < L.$$

If the concentration of the sources and the diffusion coefficients are equal

($C_A^0 = C_B^0$ and $D_A = D_B$) the function $K(x, t)$ has a maximum in the interval $0 < x < L$. In a somewhat simple, but usual case $m = n = 1$, and the concentration product function is symmetrical. The concentration product at successive times and under the conditions mentioned is shown in Fig. 1. The number on the curves represent the values of $2(Dt)^{1/2}$.

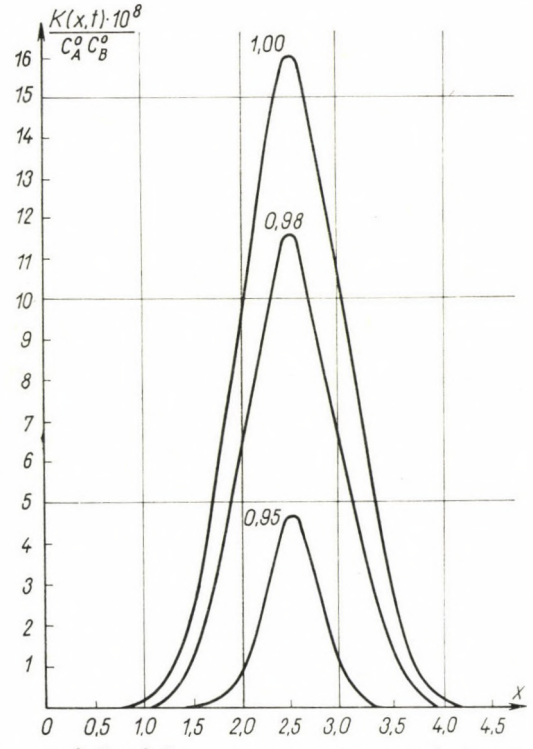


Fig. 1. The curves $K(x, t)/(C_A^0)^m\ (C_B^0)^n$ in the range $0 > x > L$. Numbers on curves are the values of $2(Dt)^{1/2}$. The value of L is 5,0 cm, $D_A = D_B$ and $m = n = 1$

We are now going to investigate another phenomenon, namely the time-dependence of the maximum value of $K(x, t)$. This function can be seen in Fig. 2. It appears from this Figure that the increase of the maximum is extremely fast already in the middle range of the solubility of the formed crystals (the solubility product is about $10^{-6} - 10^{-10}\ g/L$). At lower solubilities this increase will be much faster than in the cases mentioned. If the value of the solubility product is above 10^{-5}, the increase of the maximum of $K(x, t)$ as a function of time will be slower. The change of the maximum value of $K(x, t)$ results in that in the average solubility range the initial moment of nucleation is the same for different compounds even if their solubility products differ by 4—5 orders of magnitude. If the solubility product is above 10^{-5} already small solubility differences cause a large time difference between the initial

nucleation of different compounds. This phenomenon is very important from the point of view of segregation processes in diffusion systems. For the illustration of the described statements some data relating to different compounds are listed in Table I. All data refer to 5 cm reaction volume lengths, equal

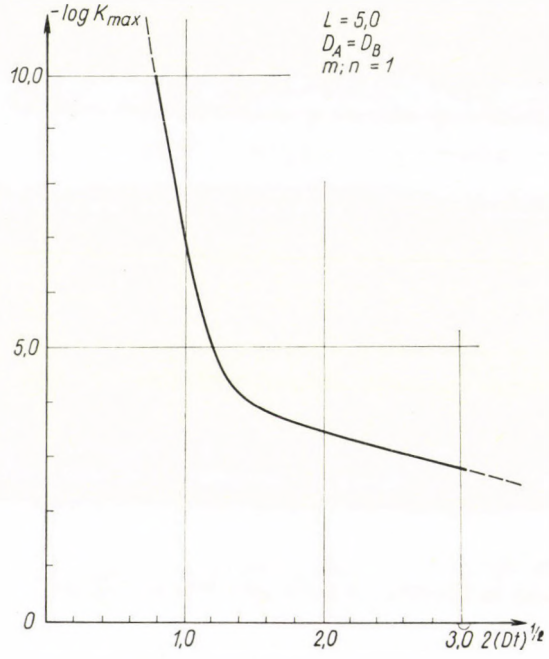


Fig. 2. The time-dependence of the maximum of $K(x, t)$ in the system characterized by the data of Fig. 1

Table I

Compound	CuS	ZnS	FeS	$BaSO_4$	$CaCO_3$	$CaSO_4$	$CaCrO_4$
Solubility product [4]	$8{,}0 \cdot 10^{-47}$	$7{,}0 \cdot 10^{-26}$	$3{,}7 \cdot 10^{-19}$	$1{,}0 \cdot 10^{-10}$	$1{,}0 \cdot 10^{-8}$	$6{,}1 \cdot 10^{-5}$	$2{,}3 \cdot 10^{-2}$
Initial time of nucleation t_k (day)	0,025	0,060	0,067	0,016	0,21	0,33	1,56

ion-source concentrations (the molarity of all ion-sources is the same) and diffusion coefficients ($D_A = D_B$). As the saturation concentrations and diffusion coefficients can be changed in a wide range by the suitable choice of the raw materials and media, the case mentioned is not specific. It can be seen from the data in Table I that while the solubility products of ZnS and

CuS differ by about 21 orders of magnitude, the initial nucleation times are very similar to each other. Naturally the theoretically determined values of t_k are only approximations. Because of the slower increase of the maximum of $K(x, t)$ in the region of higher solubility (solubility product $> 10^{-10}$) e.g. in the case of $BaSO_4$ and $CaSO_4$, already a difference of 4 orders of magnitude raises the difference in the initial nucleation times from 0,16 to 0,21 days. According to the curve of Fig. 2 in the region of high solubility (solubility product $> 10^{-4}$) a small change in the solubility causes a strong separation in the initial nucleation times. In this region the processes are very similar to the fractional crystallization and when the solubility is above 10^{-1} the components crystallize from the solution at different separated times. The indefinite value of t_k in this range is rather a consequence of the appearance of supersaturation than the result of the approximative character of our model. This supersaturation effect, of course, practically never appears in the low solubility range.

From the time dependence of the maximum value of $K(x, t)$ also other conclusions may be drawn. If the nucleation process begins already in the region of extremely low concentration, the slope of the concentration product curve near the maximum is small. This is the simple result of the high value of the argument of the erfc function at small dimensions of growing space. As a consequence of the flat character of the maximum, the volume in which the nucleation takes place, is rapidly expanding in the system. Because of this the probability that macrocrystals will grow decreases.

The detailed study of the process will not be discussed here. It is easy to understand that the dimensions of the reaction volume are also of great importance in the mentioned processes.

The latter effect plays a role chiefly in that it changes the rate of mass separation, therefore it is of importance in the growing stage, too.

In the examples mentioned we investigated special cases. As the character of $K(x, t)$ in a general case is determined by eq. (6) let us examine this function in detail. By the differentiation of this expression with respect to x, we have

$$\left(\frac{\partial K}{\partial x}\right)_t = (C_A^0)^m (C_B^0)^n \frac{\partial}{\partial x} \left\{ \left[\operatorname{erfc} \frac{x}{2(D_A t)^{1/2}} \right]^m \left[\operatorname{erfc} \frac{L-x}{2(D_B t)^{1/2}} \right]^m \right\} = 0, \quad (9)$$

Rearrangement of eq. (9) leads to

$$\frac{m}{n} \left(\frac{D_A}{D_B}\right)^{1/2} \frac{\exp\left[-\dfrac{(L-x)^2}{4D_B t}\right]}{\exp\left[-\dfrac{x^2}{4D_A t}\right]} = \frac{\operatorname{erfc} \dfrac{x}{2(D_A t)^{1/2}}}{\operatorname{erfc} \dfrac{L-x}{2(D_B t)^{1/2}}} . \quad (10)$$

As a fundamental assumption was that the solubility of $A_m B_n$ is low, at the onset of nucleation the argument of the erfc function must be higher than unit. The *erfc* (z) function at $z \gg 1$ can be expanded in a series, therefore

$$\operatorname{erfc}(z) = \frac{1}{\pi^{1/2} z} \exp(-z^2) \left(1 - \frac{1}{2z^2} + \frac{1.3}{(2z^2)^2} - \frac{1.3.5}{(2z^2)^4} + \ldots\right). \quad (11)$$

Substituting this result into eq. (10) the exponentials can be eliminated, and the following simple result is obtained:

$$\frac{m}{n}\left(\frac{D_B}{D_A}\right)^{1/2} = \frac{L-x}{x}\left(\frac{D_A}{D_B}\right)^{1/2}. \quad (12)$$

In a general case from eq. (12) one obtains for the coordinate of the maximum of $K(x, t)$

$$x_m = \frac{L}{1 + \frac{mD_B}{nD_A}}. \quad (13)$$

From eq. (13) it may be seen that in the cases of certain ion sources the maximum value of $K(x, t)$ is independent of the time. Its value is determined by the dimensions of the reaction volume, the diffusion coefficients of certain ions and the stoichiometric data of the precipitated compound (m and n). Naturally, m and n cannot be changed, therefore in a given system the value of x_m can be changed only by choosing the diffusion coefficients suitably.

The diffusion coefficient depends on the medium and on the compounds which were used as ion sources. Let there be AC the chemical composition of one of the compounds of the ion sources and DB that of the other. From the reaction

$$(A^+ + C^-) + (D^+ + D^-) = AB + D^+ C^-$$

it appears as if D^+ and C^- would play no role in the reaction, and remain unchanged in the solution. In real cases they are not indifferent, and in some cases their influence on the nucleation is very important. This foreign ion effect appears not only in the influence of ions on the ion strength, but in the value of the diffusion coefficients, too, i.e. the values of the diffusion coefficients in all cases belong to certain compounds, therefore D_A and D_B are characteristic of the compounds of the ion sources and not of the single ions. This means that the diffusion coefficient of the reacting ions (D_A and D_B) can be changed within a wide interval by the variation of the so-called indifferent ions (in our example C^- and D^+). In case of the crystalline precipitation of ZnS we may use $ZnCl_2$, $ZnSO_4$ or Zn-acetate as ion source. In the latter case the

diffusion coefficient of Zn^{++} ions is about a tenth of the diffusibility of Zn^{++} with Cl^- anions because of the small mobility of CH_3COO^- ions. Similarly, the diffusion coefficient of S^{--} ions can also be changed by changing the cation of the soluble sulfide. This means that the indifferent ions of the ion sources by changing the solubility of the precipitated compound (A_m B_n) influence the initial period of nucleation and furthermore, as they determine the diffusion coefficients in a certain medium, they regulate the coordinates

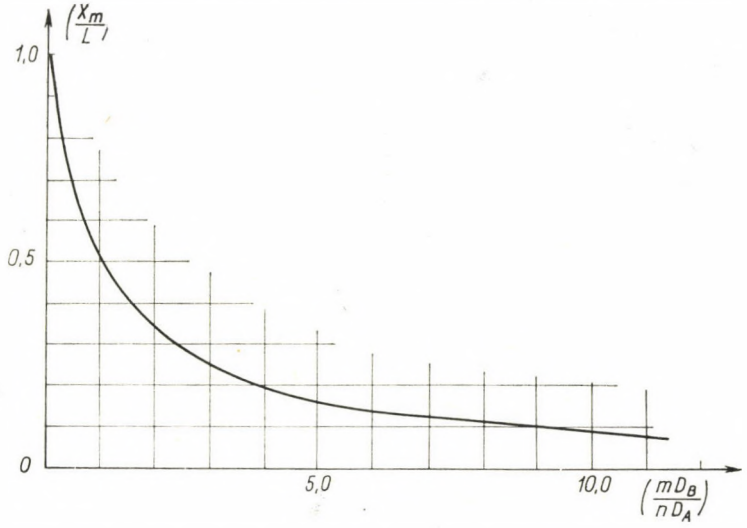


Fig. 3. The time-independent coordinate of the maximum of $K(x, t)$ as function of mD_B/nD_A in linear systems

of the initial nucleation, too. Both effects have a strong influence on the properties of the precipitated, crystalline phase $A_m B_n$.

From what has been said it follows that the local volume of the initial nucleation in a system can be changed by changing the length of L. This makes focusing of the nucleation possible. By choosing the dimensions of the growing system, the media and compounds of the ion sources in a way that the function $K(x, t)$ has a suitable, sharp maximum, in principle we can limit the initial nucleation to an extremely small volume or point in the system, and further this point of initial nucleation may be interchanged for another without restriction in a given range of the interval $0 < x < L$. If, in the notation of eq. (13), we denote the point of initial nucleation by x_m, the value of it shows a hyperbolic decay as a function of mD_B/nD_A. (See Fig. 3). The two characteristic limits of the function are

$$\lim_{\frac{mD_B}{nD_A} \to 0} x_m = L \qquad \text{and} \qquad \lim_{\frac{mD_B}{nD_A} \to \infty} x_m = 0. \tag{14}$$

From the above relations it is clear that if there exists an essential difference between the diffusion coefficients, the maximum of $K(x, t)$ moves in the direction of the ion sources where there is a slower diffusion ion flow. Above a certain limit (1 : 8) this effect is disadvantageous as the separation of the solid phase takes place in the growing system on the active surface of the ion source itself. The crystalline phase, separated on the ion source is generally an unpermeable crystal layer that practically stops the process entirely in short time.

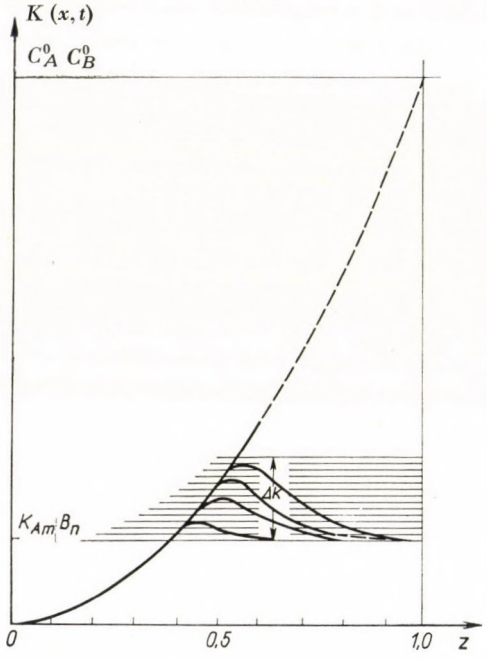


Fig. 4. The function $K(x, t)$ at $X = X_m$ in its dependence on diffusion and precipitation in case $m = n = 1$. The striated part of the figure corresponds to the uncertainty of the solue bility product in real systems. The two limits, $K_{\text{amorphous}}$ and K_s are the concentration products for amorphous phase in the presence of strong electrolytes and for macrocrystals in pur-solvent resp.

In the simple case described, from eq. (13) with $m = n = 1$ and $D_A = D_B$ the coordinate of the maximum of x_m is found to be 0,5 L. This value corresponds to the curves of Fig. 1. Substituting the time-independent value of the maximum from eq. (6), we have

$$K(x,t) = kz^2 , \tag{15}$$

where $$k = C_A^0 C_B^0 \qquad \text{and} \qquad z = \operatorname{erfc} \frac{x_m}{2(Dt)^{1/2}} . \tag{16}$$

As z is an erfc function, it has a lower and upper limit az $t = 0$, $z = 0$, and $t = \infty$, $z = 1{,}0$ for all values of L and D. The function given in (15) is represented by the curve in Fig 4. If in the system no reaction takes place, the value of the concentration product rises along a parabola and at $t = \infty$ it reaches the maximum value, $c_A^0 c_B^0$. Let us denote the solubility product by K_s. As at a certain temperature the solid phase is in thermodynamic equilibrium with its own dissoluted ions at the concentrations determined by K_s, the concentration product curve is folloved closely by the parabola only up to the value of K_s. If the concentration product is equal to the solubility product, $K(x, t) = K_s$, the curve turns into a straight line parallel to the x-axis as a consequence of mass separation from the system.

If the solubility of the separated solid falls into the middle range of the solubility product ($K_s < 10^{-10}$) the value of K_s is influenced by the concentration of foreign ions. In these cases — in addition to the effects caused by the ion strength — a supersaturation can also be originated, therefore the value of K_s is not well defined. In most cases the initial nucleation does not begin at the moment when $K(x, t)$ reaches the exact value of the solubility product. The value mentioned is only the lower limit for the initial nucleation. This can easily be understood because this value corresponds rather to the solubility of macrocrystals than to that of the nucleus. The curve $K(x, t)$ can also exceed the value of K_s of the macrocrystals in the cases, when the separated solid phase is not crystalline. It is well known that the solubility of amorphous or freshly precipitated materials is always higher than the solubility of macrocrystals in pure, ionfree media. In Fig. 4 the striated part of the figure corresponds to the mentioned uncertainty. It can be seen that the lower limit of the initial nucleation period is K_s, the upper limit is the solubility of amorphous phases in the presence of strong electrolytes. The difference between the limits is

$$\Delta K = K_{\text{amorphous}} - K_s, \tag{17}$$

where $K_{\text{amorphous}}$ and K_s correspond to the mentioned solubilities and solubility products, resp. If the diffusion process requires a longer time, the real solubility product is determined by the ageing of the separated solid phase. In these cases the curve $K(x, t)$ possesses a maximum before the linear section of the curve begins. Above K_s the increasing part of the curve is caused by the varying solubility of the nucleus, the decreasing part is formed in consequence of ageing processes. In cases of compounds with a higher solubility the ion strength also plays an important role. It is easy to understand that in diffusion systems an inverse connection can be assumed between the crystalline state and the foreign ion effect. The crystalline nucleus can be formed only when the ion transport is slow, therefore the concentration of indifferent ions around the nucleus is small. In case of rapid ion flow the formed nucleus

is rather amorphous than crystalline, furthermore, strong accumulation and adsorption of foreign ions takes place. These processes cause an increase in the ion strength and raise the solubility of the precipitated phases. The third effect which has an influence on the shape of the real function $K(x, t)$ is the dependence of the activities of the ion source compounds on the concentration. This dependence is different for different compounds. As the result of the mentioned three effects the maximum value of $K(x, t)$ in the striated interval of Fig. 4 may appear at different values of z, according to the transport rate and ion pairs. The three effects always appear together, therefore the maximum value of $K(x, t)$ can be varied inside the interval ΔK by varying the experimental circumstances.

In the case treated the initial time of mass separation can also be determined approximately from eq. (15). This, of course, refers only to the lower limit of the solubility product. Let us denote by t_K the initial time of nucleation and by z_K the corresponding erfc function. As

$$z_K = \left(\frac{K_S}{k}\right)^{1/2}, \tag{18}$$

it follows that z_K changes monotonically with K_s. If z_K has a large value, t_K in the argument of erfc(z) is also large. As

$$\operatorname{erfc} \frac{x_m}{2(Dt_k)^{1/2}} = \left(\frac{K_S}{k}\right)^{1/2} \tag{19}$$

substituting x_m from eq. (13) into eq. (19) we have

$$\operatorname{erfc} \frac{L}{4(Dt_k)^{1/2}} = \left(\frac{K_S}{k}\right)^{1/2}. \tag{20}$$

For $k = 1$ we obtain the data of Table I. In Fig. 5 the connection of initial nucleation time and the solubility product (K_s) can be seen. The numbers on the curves belong to the different values of $L/4D^{1/2}$.

The above calculation naturally does not allow for the effect of mixing, large density differences, etc. If these are present, the nucleation begins earlier than given by the calculated values.

A rather complicated equation in terms of the diffusion distribution and solubility product can be obtained when treating a general case. Substituting the coordinate of the maximum of $K(x, t)$ into eq. (6) we have

$$K(x, t)_{x=x_m} = (C_A^0)^m (C_B^0)^n \left\{\operatorname{erfc} nA \left(\frac{D_A}{t}\right)^{1/2}\right\}^m \left\{\operatorname{erfc} mA \left(\frac{D_B}{t}\right)^{1/2}\right\}^n, \tag{21}$$

where

$$A = \frac{L}{2(nD_A + mD_B)} .$$

If the nucleation begins at low concentrations ($K_s < 10^{-10}$) an approximate formula can be derived from eq. (11) from which the initial nucleation time

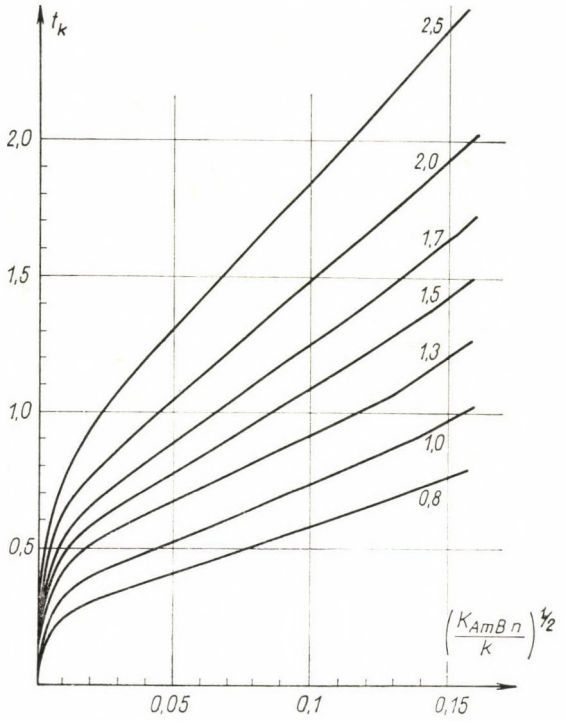


Fig. 5. The initial nucleation time (t_K) as a function of $(K_s/k)^{1/2}$ at $D_A = D_B$ and $m = n = 1$

(t_K) can be determined for a general case. This is

$$K(x,t)_{x=x_m} = c' \, t^{\frac{m+n}{2}} \exp\left(-\frac{B}{t}\right), \tag{22}$$

where c' and B contain the constants of the raw materials

$$c' = \frac{(nD_A + mD_B)^{m+n} (C_A^0)^m (C_B^0)^n}{n^m m^n D_A^{m/2} D_B^{n/2} L^{m+n}} \tag{23}$$

and

$$B = \frac{mn L^2}{4(nD_A + m D_B)} . \tag{24}$$

For the examined simple case $D_A = D_B$ and $m = n = 1$ and we have

$$K(x, t)_{x=x_m} \approx \frac{4C_A^0 C_B^0 D}{L^2} t \exp\left(-\frac{L^2}{8Dt}\right). \tag{25}$$

For very large times the exponential in eq. (25) may be replaced by unit, and the slope of the resulting straight line is determined by the diffusion coefficient and the geometrical characteristics of the growing system.

2.2. *Liquid-liquid systems in case of uniform media*

The previous solutions correspond to systems containing solids of the raw materials. Similar, but more complicated calculations can be carried out if the concentration of the ion sources continuously decreases during the growing processes. These cases correspond to ion sources containing the solutions of the raw materials in a limited volume without solid phase. The continuous decrease of the concentration results from the diffusion transport and the separation of the forming compound, $A_m B_n$.

Similarly as in paragraph 2.1, let us assume that because of the low solubility of the precipitated compound it is not necessary to take into account the reflexions of the ion flow at the boundaries. In this case the initial conditions for the solution of eq. (1) for one of the components are

$$\text{in } -h < x < h \qquad \text{interval } C(x, t) = C_A^0 , \quad \text{at } t = 0 .$$

$$\text{in } x < |\, h \,| \qquad \text{interval } C(x, t) = 0 .$$

For cases of extended sources in infinite or semi-infinite media, where the the diffusing substance initially occupies the region $-h < x < h$ the solution of the problem is [3] in the case of uniform concentration C_0:

$$C(x, t) = \frac{1}{2} C_0 \left[erf \frac{h - x}{2(Dt)^{1/2}} + \operatorname{erf} \frac{h + x}{2(Dt)^{1/2}} \right]. \tag{26}$$

The expression is symmetrical with respect to $x = 0$, therefore the system can be cut in half by a plane $x = 0$ without affecting the distribution.

Using the described coordinate transformation and the relation

$$\operatorname{erf}(-x) = -\operatorname{erf}(x),$$

the concentration product can be described as a function of the following type:

$$K(x,t) = \frac{(C_A^0)^m (C_B^0)^n}{2^{m+n}} \left\{ \operatorname{erf} \frac{x + h_A}{2(D_A t)^{1/2}} - \operatorname{erf} \frac{x - h_A}{2(D_A t)^{1/2}} \right\}^m \times \\ \times \left\{ \operatorname{erf} \frac{L + h_B - x}{2(D_B t)^{1/2}} - \operatorname{erf} \frac{L - h_B + x}{2(D_B t)^{1/2}} \right\}^n . \tag{27}$$

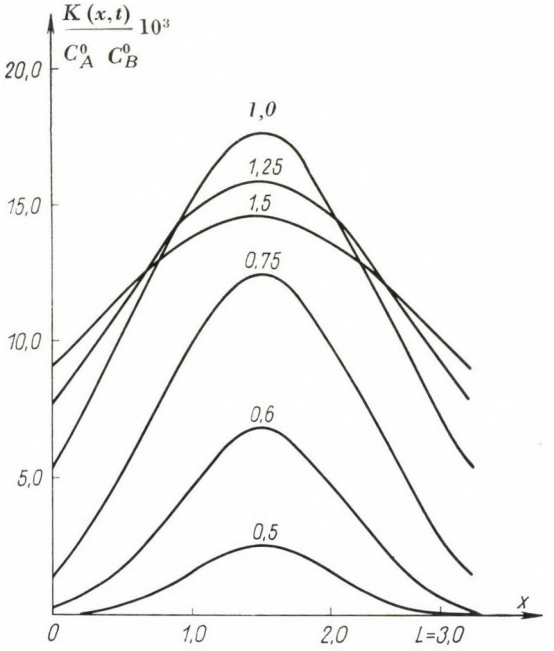


Fig. 6. The concentration product function for a system containing ion sources with time-dependent concentration. The numbers on the curves refer to the values of $2(Dt)^{1/2}$ at $D_A = D_B$ and $m = n = 1$. The length of the reaction volume is 3,0 cm

Eq. (27) is more complicated than eq. (6). For a simple case $h_A = h_B$, $m = n = 1$ and $D_A = D_B$. The derivative of eq. (27) in case of $\partial K/\partial x = 0$ is

$$\frac{\exp(-\alpha^2) - \exp(-\beta^2)}{\exp(-\gamma^2) - \exp(-\delta^2)} = \frac{\operatorname{erf}\alpha - \operatorname{erf}\beta}{\operatorname{erf}\gamma - \operatorname{erf}\delta} , \tag{28}$$

where

$$\alpha = \frac{x_m + h}{2(Dt)^{1/2}} ; \qquad \beta = \frac{x_m - h}{2(Dt)^{1/2}} ;$$

$$\gamma = \frac{L + h - x_m}{2(Dt)^{1/2}} \quad \text{and} \quad \delta = \frac{L - h + x_m}{2(Dt)^{1/2}} .$$

Eq. (28) is satisfied for $x_m = L/2$, namely after the substitution of this value both sides of the equation will be unit. This means that the maximum of $K(x, t)$ similarly to the solid-liquid systems is independent of time and that it moves towards the ion sources in cases of unequal D_A and D_B.

In Fig. 6 a series of solutions of eq. (27) can be seen. It is clear from the figure that contrary to the solid-liquid example discussed earlier, the value of $K(x, t)$ $x = x_m$ does not show a monotonic tendency in its change. Because of the extended monotonically decreasing concentration of the ion sources

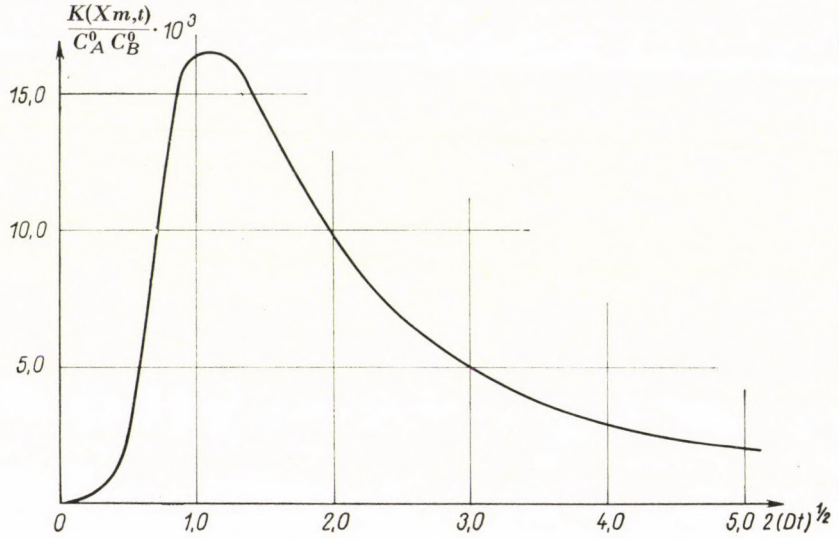


Fig. 7. $K(x, t)/k$ at $x = x_m$ as a function of $2(Dt)^{1/2}$ at $D_A = D_B$ and $m = n = 1$ in case of time-dependent ion sources

the curve $K(x, t)_{x=x_m}$ at first increases and after passing through a maximum shows a decreasing tendency. We have

$$\lim_{t\to 0} K(x, t)_{x=x_m} = \lim_{t\to\infty} K(x, t)_{x=x_m} = 0. \tag{29}$$

$K(x_m, t)$ as a function of $2(Dt)^{1/2}$ is illustrated in Fig. 7. For the examined special case using the solubility product, the value of t_K, too, can be determined. Substituting x_m into eq. (27) we get a quadratic equation similar to eq. (15):

$$K(x, t)_{x=x_m} = k' (z')^2 \tag{30}$$

where

$$k' = \frac{k}{4} = \frac{C_A^0 C_B^0}{4}$$

and

$$z' = \operatorname{erf} \frac{L + 2h}{2(Dt)^{1/2}} - \operatorname{erf} \frac{L - 2h}{2(Dt)^{1/2}}$$

From the fact that at $t = t_K$

$$z' = \left(\frac{4K_S}{k} \right)^{1/2} \tag{31}$$

the connection between t_K and the solubility product can be determined. From curves, especially from $K_c - t^{1/2}$ it can be seen that $K(x,t)$ reaches twice the solubility product. This of course can be realized only in cases when the reaction volume compared to the ion sources is an infinite or semi-infinite one. In this case from eq. (30) one can determine approximately the initial and final moment of the mass separation. This also means that by the variation of h and L the rate of the separation of the crystalline phase can be varied.

2.3. *Systems with two immiscible liquid phases*

Another question which we investigate is the analysis of systems containing two media which are insoluble in each other. If we have ion sources with constant concentrations (saturated solutions in the presence of solid raw materials) and one of the media belongs only to the ion sources, the concentrations of the reacting ions will be different, but constant at both sides of the boundaries. The concentrations mentioned above are determined by the distribution equilibrium between the liquid phases:

$$\frac{C_A^a}{C_B^b} = \varkappa . \tag{32}$$

In eq. (32) indices a and b refer to the different solvents and $\varkappa$ to the distribution coefficient. In the mathematical treatment it must be taken into account that the diffusion coefficients have a discontinuity at the boundaries and for these the coordinates $x = 0$ resp. $x = L$ were chosen.

In the medium "a" ($x > 0$) the diffusion coefficient is D_{1A} and in the medium "b" it is d_{2A} ($x < 0$), the following conditions are satisfied at the beginning of the diffusion process for the component A

$$x < 0 \quad \text{at} \quad t = 0 , \quad C_{1A} = C_A^0 ,$$

$$x > 0 \quad \text{at} \quad t = 0 , \quad C_{2A} = 0 .$$

First of all the mentioned distribution appears at the boundaries. After this if we assume that at the interface there is no accumulation of diffusion substance, we have

$$D_{1A}\left(\frac{\partial C_{1A}}{\partial x}\right)_{x=0} = D_{2A}\left(\frac{\partial C_{2A}}{\partial x}\right)_{x=0} .$$

This equation corresponds to eq. (1). Its solution under the conditions mentioned here in a semi-finite or infinite medium is given [3] by:

$$C_{2A}(x,t) = \frac{\varkappa_A C_A^0}{1+\varkappa_A\left(\frac{D_{2A}}{D_{1A}}\right)^{1/2}} \operatorname{erfc}\frac{x}{2(D_{2A}t)^{1/2}} . \tag{33}$$

In the region $x < 0$ the concentration distribution is uniform and equal to the saturation concentration of the raw material (ion source).

The time-dependent distribution of the component B can be described by a similar expression, therefore the concentration product is

$$K(x,t) = c''\left(\operatorname{erfc}\frac{x}{2(D_{2A}t)^{1/2}}\right)^m\left(\operatorname{erfc}\frac{L-x}{2(D_{2B}t)^{1/2}}\right)^n. \tag{34}$$

This equation is quite similar to eq. (6) and differs from it only in the constant factor, because

$$c'' = \frac{(\varkappa_A C_A^0)^m(\varkappa_B C_B^0)^n}{\left\{1+\varkappa_A\left(\frac{D_{2A}}{D_{1A}}\right)^{1/2}\right\}^m\left\{1+\varkappa_B\left(\frac{D_{2B}}{D_{1B}}\right)^{1/3}\right\}^n} . \tag{35}$$

Generally, this means that the problem can be reduced to those in paragraph 2.1.

Similar solutions can be gained if only one of the ion sources contains an immiscible solvent in the growing system. Naturally, in this case the expression for the constant factor is simpler than that given by eq. (35). If the concentration of one or both of the ion sources is changing, the suitable solutions are similar to those given in 2.2.

2.4. *Liquid-gas systems*

All the foregoing remarks refer to the solid-liquid systems. Apart from these the liquid-gas and the solid-liquid-gas systems resp. should be mentioned, too. In these systems one of the reacting components is in the liquid phase and the other is in the gaseous phase or it is the gas itself. E.g. in the preparation of carbonates or sulfides the anion source can be CO_2 and H_2S resp. as a gas. The partial pressure of the gases determines the concentration of the ion source. According to Henry's general law and using the activity coefficient of γ_x we have

$$\gamma_x X_i^T = \frac{f_i}{k''}, \tag{36}$$

where X_i^T is the mol fraction of solved gas at saturation, f_i its fugacity and k'' is constant. The partial pressure of the solved gas must be low during the growing process because the rate of nucleation must also be low. Similarly to the case of dilute solutions the fugacities f_i can be substituted by partial pressures, or by choosing k'' suitably; it can be expressed in arbitrary concentration units. If the mentioned condition is satisfied the gas behaves as a perfect gas and we have

$$p_i = RTc_i \tag{37}$$

and Henry's Law can be described by the following expression:

$$\frac{C_i^{\text{sol}}}{C_i^{\text{gas}}} = Q_i, \tag{38}$$

where Q_i is the Ostwald absorption coefficient for the i-th component of the gas.

The simplest case occurs when the ion source in the liquid phase corresponds to the first type of ion source discussed with a constant concentration (see 2.1). If we treat the gaseous phase as an infinite medium (the reacting compound is e.g. the CO_2 of the air) the whole problem is reduced to that treated in 2.1 (see eqs. (6) and (34). Namely, in this case the concentrations of both ion sources are constant in time. The concentration of the cation source is determined by the solubility of the suitable raw material, the concentration of the anion source by Henry's Law.

These conditions often are not fulfilled because the reaction volume is finite and C_B^0 generally has a small value, therefore we must take into account a reflection of the ion flow directed towards the solution interface. If we want to eliminate the reflection effects, the value of C_i^{gas} must be raised. Generally this can only be carried out by using a closed growing apparatus. If the gaseous phase is finite, at the gas-liquid interface the concentration of reacting gas continuously decreases because of its dissolution. This decrease in concentration causes a diffusion directed towards the boundary and this means that the problem is mathematically quite similar to those investigated in 2.3.

As the diffusion coefficients measured in gas and liquid essentially differ from each other it is possible to apply eqs. (34) and (33) to our problem. The solution for infinite gas phase is simpler than for the mentioned liquid-liquid problem, because if

$$\frac{D_{\text{liq}}}{D_{\text{gas}}} \ll 1{,}0 \quad \text{and} \quad Q < 1{,}0$$

the terms $Q\left(\frac{D_{\text{liq}}}{D_{\text{gas}}}\right)^{1/2}$ in eq. (34) are negligible.

After the substitution of suitable values for the distribution of the component originated from the gas phase into eq. (33), we have

$$C_{\text{liq}}(x,t) = QC_B^0 \operatorname{erfc} \frac{x}{2(D_{\text{liq}}\, t)^{1/2}} \tag{39}$$

and the concentration product function can be constructed by multiplication of certain distributions of the types mentioned. Generally, the quotient of

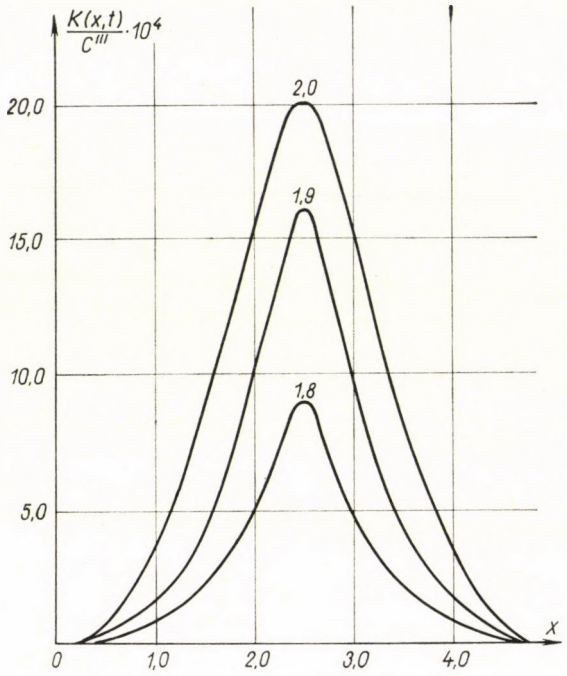


Fig. 8. The concentration product curves in the range of $0 < x < L$ for systems containing two immiscible liquids and a gas phase at $D_{2A} = D_B$ and $m = n = 1{,}0$

the diffusion coefficients in liquid and gas is about 10^{-5}. Practically this means that in case the gas volume is large enough, the concentration distribution in the gaseous phase remains unchanged for all t. Therefore the solution of the problem corresponds to eq. (6) and (34), resp., and differs from them in the constant factor. In some cases this factor is similar to C^{11}, but it contains Ostwald's coefficient instead of $\varkappa$.

If we combine systems where the concentration of one ion source is time dependent (e.g. one of the ion sources is a finite volume gas and the other a saturated solution with solid raw material) we get solutions for the concentration product which may be expressed by functions which were described above. If one of the ion sources corresponds to eq. (6) and the other to eq. (26) the expression describing the ion distribution is the following:

$$K(x,t) = \frac{(C_A^0)^m (QC_B^0)^n}{2^n} \left\{\operatorname{erfc} \frac{x}{2(D_A t)^{1/2}}\right\}^m \times \times \left\{\operatorname{erf} \frac{L+h-x}{2(D_B t)^{1/2}} - \operatorname{erf} \frac{L-h+x}{2(D_B t)^{1/2}}\right\}^n \tag{40}$$

where L is the total height of the system, h is twice as high as the gas-cylinder and D_A and D_B are the diffusion coefficients.

In case of systems containing immiscible liquid media similar expressions can be gained for the concentration product even if the concentration of one of the ion sources is changing in time. In these expressions we cannot neglect the term $(D_2/D_1)^{1/2}$ because $D_2 \sim D_1$. For $m = n = 1$ and the type of concentration product function is the following:

$$K(x,t) = \frac{\varkappa C_A^0 C_B^0}{1 + \varkappa \left(\frac{D_{2A}}{D_{1A}}\right)^{1/2}} \left\{\operatorname{erf} \frac{h+x}{2(D_B t)^{1/2}} + \operatorname{erf} \frac{h-x}{2(D_B t)^{1/2}}\right\} \operatorname{erfc} \frac{L-x}{2(D_A t)^{1/2}} \tag{41}$$

The curves shown in Fig. 8 correspond to eq. (41). The numbers on the curves refer to the different values of $2(Dt)^{1/2}$, the constant C^{111} is the constant factor before the erf and erfc functions in eq. (41).

3. The reflection of the ion flow

All mentioned solutions of the differential equation governing diffusion correspond to semi-infinite or infinite media. The foregoing considerations are strictly correct in the first stage of nucleation. If the solubility of the forming compound is high or if there is a large difference between the rates of oppositely directed ion flows, the problem is complicated by the reflections of the ion flows. The reflection appears at the boundaries and entirely changes the concentration distribution. This explains why after some time the concentration product function shows different, secondary time-dependent maxima or minima.

In this case the condition that c should tend towards zero if $x \to \infty$ must be changed for the condition that through the boundaries there is no flow. The condition to be satisfied at the impermeable boundary is

$$\frac{\partial C_A}{\partial x} = 0 \text{ at } x = L \text{ for component } A,$$

and

$$\frac{\partial C_B}{\partial x} = 0 \text{ at } x = 0 \text{ for component } B.$$

This condition is satisfied if we reflect the concentration distributions at the boundaries, and superpose them on the original distribution. Mathematically the resulting resolution is given by an infinite series, which converges quite rapidly. As the differential equation is linear, the superposition is allowed and the resulting infinite series is the solution of the problem.

For a detailed study it is advantageous to use the Laplace transform of the expression. As

$$\bar{c}(x, t) = \int_0^\infty \exp(-pt)\, C(x, t)\, dt \tag{42}$$

the transformed form of eq. (1) is

$$\frac{d^2 \bar{c}}{dx^2} = q^2 \bar{c} \text{ in the interval } 0 < x < L, \tag{43}$$

where $q^2 = P/D$ and L is the length of the growing system. Firstly we investigate the case of ion sources with constant concentrations (see 2.1). As it was described in 2.1 conditions are

$$c = c_0, \qquad x = 0, \qquad t \geqslant 0,$$

$$c = 0, \qquad x > 0, \qquad t = 0.$$

In finite systems with a length L and when there is reflection of the ion flows the mentioned conditions must be completed with the condition of impermeable boundaries described above. In this case the conditions for the solution of the transformed differential equation for the component A are

$$\frac{d\bar{c}_A}{dx} = 0, \qquad \text{at} \quad x = L,$$

$$\bar{c}_A = \frac{1}{p} C_A^0 \quad \text{at} \quad x = 0$$

and the exact solution is

$$\bar{c}_A = \frac{C_A^0}{p} \frac{\cosh q_A (L - x)}{\cosh q_A L}. \tag{44}$$

A similar solution can be gained for the component B:

$$\bar{c}_B = \frac{C_B^0}{p} \frac{\cosh q_B x}{\cosh q_B L}. \tag{45}$$

The hyperbolic functions can be expressed in terms of negative exponentials and expanded in a series according to the binominal theorem.

$$\bar{c}_B = \frac{C_B^0}{p} \{\exp[-q_B\,(L-x)] + \exp[-q_B\,(L+x)]\} \times \\ \times \sum_{j=0}^{\infty}{}' (-)^j \exp(-2j\,q_B\,L)\,. \tag{46}$$

After suitable multiplication we have

$$\bar{c}_B = \frac{C_B^0}{p} \left\{ \sum_{j=0}^{\infty}{}' (-)^j \exp[-q_B[(2j+1)L - x]\right\} + \\ + \sum_{j=0}^{\infty}{}' (-)^j \exp\{-q_B\,[(2j+1)L + x]\}\Big\}. \tag{47}$$

As

$$\frac{\exp(-qx)}{p} = \operatorname{erfc}\frac{x}{2(Dt)^{1/2}} \tag{48}$$

from eq. (47) we get the solution of eq. (1) for a finite system satisfying the conditions described above. This is

$$C_B(x,t) = C_B^0 \left\{ \sum_{j=0}^{\infty}{}' (-)^j \operatorname{erfc}\frac{(2j+1)L - x}{2(D_B t)^{1/2}} + \right. \\ \left. + \sum_{j=0}^{\infty}{}' (-)^j \operatorname{erfc}\frac{(2j+1)L + x}{2(D_B t)^{1/2}} \right\}. \tag{49}$$

The solution for the component A is similar

$$C_A(x,t) = C_A^0 \left\{ \sum_{j=0}^{\infty}{}' (-)^j \operatorname{erfc}\frac{(2j+1)L - (L-x)}{2(D_A t)^{1/2}} + \right. \\ \left. + \sum_{j=0}^{\infty}{}' (-)^j \operatorname{erfc}\frac{(2j+1)L + (L-x)}{2(D_A t)^{1/2}} \right\}. \tag{50}$$

With eqs. (49) and (50) the concentration product can be given. Both functions converge quite rapidly at small values of Dt/L^2.

If $Dt/L^2 < 2$, it is sufficient to pay attention to the first two terms of the infinite series. For $Dt/L^2=1$ the third term of eq. (49) is only 0,0008, for $Dt/L^2 = 0{,}25$ already the second term is only 0,0001.

In Fig. 9 the concentration product constructed from eqs. (49) and (50) can be seen for $D_A = D_B$. All curves are normalised to the $L/2$ value in order to make comparison possible. It can be seen that the reflexion effects at first cause the broadening of the curves and then a strong rise of the curves can be observed at the boundaries. Because of this rapid increase at the end points

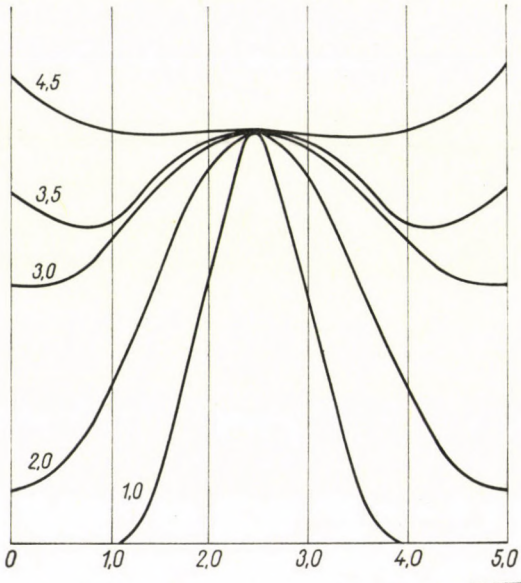


Fig. 9. The concentration product curves in case of reflection of the ion flows. All curves are normalized to the $L/2$ value. The number on the curves are the values of $2(Dt)^{1/2}$ at $D_A = D_B$ and $m = n = 1$, The length of the reaction volume is 5,0 cm

for large times the maximum curve changes to the minimum curve. From the Figure it is evident that there is a continuous transition between the curves in time. During this transition, some concentration product curves corresponding to $x = \text{const.}$ can be found. If the solubility product falls within these ranges of the curves, after the separation on the ion sources the nucleation in the whole system begins suddenly. This is favourable for the formation of homodisperse systems, in contrast to cases of compounds where the solubility product is about 10^{-10}.

A similar treatment can be applied to ion sources when the concentration is time-dependent. The solution corresponding to eq. (26) and taking reflection into consideration is then [3]

$$C_A(x,t) = \frac{1}{2} C_A^0 \sum_{j=0}^{\infty}{}' \left\{ \operatorname{erf} \frac{h + 2jL - x}{2(D_A t)^{1/2}} + \operatorname{erf} \frac{h - 2jL + x}{2(D_A t)^{1/2}} \right\}. \qquad (51)$$

For the component B the solution is

$$C_B(x,t) = \frac{1}{2} C_B^0 \sum_{j=0}^{\infty}{}' \left\{ \operatorname{erf} \frac{h + 2jL - (L-x)}{2(D_A t)^{1/2}} + \operatorname{erf} \frac{h - 2jL + (L-x)}{2(D_A t)^{1/2}} \right\}.$$

Eqs. (51) and (52) similarly to eqs. (49) and (50) converge rapidly, therefore in the first period of diffusion growing the first terms of the series describe

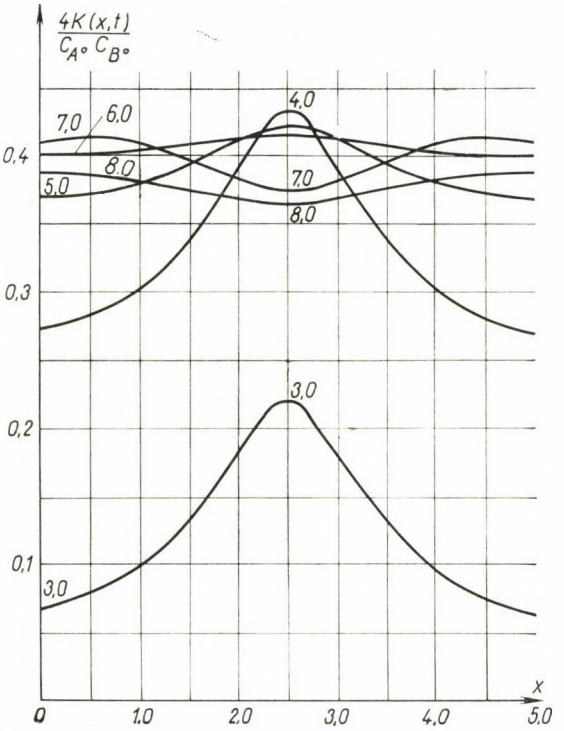


Fig. 10. The curves $K(x, t)/k^1$ in case of reflection of the ion flows and time-dependent ion sources. The numbers on the curves refer to the values of the $2(Dt)^{1/2}$. All the other data correspond to the data of Fig. 9

the process exactly. The function $K(x, t)$ constructed from the mentioned functions can be seen in Fig. 10. The concentration product curves show the distributions at given values of $2(Dt)^{1/2}$ with $D_A = D_B$. The shape of $K(x, t)$ is similar to that in the foregoing case of reflection distribution, but for the time-dependent increase of the decreasing concentration of the ion sources at $x = 0$ and $x = l$ in time.

Among the concentration product curves mentioned we find some with extensive horizontal sections as in the foregoing reflection distributions. This means that in such systems the extended nucleation and formation of homodisperse particle-size distributions is also possible.

As we have mentioned, the reflexion problems are important in the cases, when the value of the solubility product is above 10^{-3}. As has already been described in the linear case the mathematical description is difficult. In three dimensions the reflection distributions are more complicated. As the problem occurs very rarely in the diffusion growing, we do not examine it in detail. It is sufficient to note that by the superposition of erf and erfc functions in principle all the cases can be solved.

Concluding, we should like to remark that the problem treated is only the first, but very important stage of the growing process. Without the knowledge of the factors influencing nucleation the problem of growing processes cannot be examined.

REFERENCES

1. J. Johnston, J. Am. Chem. Soc., **36**, 16, 1914.
2. W. C. Frenelius and K. D. Detling, J. Chem. Educ., **11**, 176, 1934.
3. J. Crank, The Mathematics of Diffusion, Clarendon Press, Oxford, 1956.
4. H. Rauscher et al., Chemische Tabellen und Rechentafeln, VEB Deutscher Verlag für Grundstoffindustrie, Leipzig, 1961.

ВОЗНИКНОВЕНИЕ КРИСТАЛЛИЧЕСКИХ ОЧАГОВ ПРИ ДИФФУЗИОННОМ ВЫРАЩИВАНИИ КРИСТАЛЛОВ

Э. ЛЕНДВАИ

Р е з ю м е

Из растворов диффузионным переносом вещества можно выращивать и практически нерастворимые кристаллы. Как первый этап диффузионного процесса должны осуществляться такие распределения концентрации, произведение которых достигает или превышает произведение растворимости. При достижении произведения растворимости начинается выделение твёрдой фазы. При помощи диффузионного переноса, применением различных ионных источников и выбором соответствующей геометрии системы имеется возможность для регулировки времени начала, области и скорости образования очага. Все три фактора решительным образом влияют на морфологические свойства выделяющейся фазы. В работе, исходя из временной зависимости распределений концентрации, исследуется роль параметров, оказывающих до начала образования очагов влияние на диффузионный перенос вещества.

A RE-DETERMINATION OF THE HALF-LIFE OF RUBIDIUM-87

By

A. Kovách

INSTITUTE OF NUCLEAR RESEARCH OF THE HUNGARIAN ACADEMY OF SCIENCES, DEBRECEN

(Presented by A. Szalay. — Received 16. II. 1964)

A determination of the half-life of rubidium-87 was made by liquid scintillation techniques, using natural rubidium. For the half-life a value of $T_{1/2} = (4{,}77 \pm 0{,}10).10^{10}$ years was obtained.

Introduction

The so-called rubidium-strontium method has become of great importance in absolute geochronology during the past years. The need for increased accuracy of absolute age data, however, requires a more precise knowledge of the decay constant of rubidium-87 than has been obtained up to now.

There are essentially two methods suggested for the determination of the half-life of rubidium-87 differing from each other in basic principles. The decay constant can be determined either by the direct measurement of the specific activity of rubidium-87 (or that of the natural element) itself, or on the other hand by determining the amount of radiogenic strontium-87 formed by the decay of rubidium-87 in minerals of known absolute age. The value of the decay constant, supplying the measured parent/daughter ratio through a known time interval can be determined in this case with the aid of the exponential law of radioactive decay.

The Rb^{87}—Sr^{87} decay is known to be a non-unique third-forbidden 3^+ beta transition. Although the maximal energy of the beta spectrum is rather high, about 275 keV [1], because of the highly forbidden spectrum shape most of the emitted beta particles exhibit low energies, the average energy of the spectrum being less than about 40 keV.

Investigations concerning the shape of the spectrum and supplying uniform results [1], [2], [3] show a steady rise of the spectrum with decreasing energy. The investigations quoted above give reliable information about the spectrum shape down to about 6—8 keV, but there is no reason to suppose that it would exhibit a behaviour differing from the general picture below this limit.

The determination of the half-life of rubidium-87 was the subject of several investigations — but there is considerable disagreement in the values

obtained [1]—[29], these ranging from $4{,}3 \cdot 10^{10}$ years to about $6{,}5 \cdot 10^{10}$ years, if two older and much higher values [6], [7] are omitted. Recent investigations carried out on minerals of known absolute age [4], [5] indicate a value of about $5 \cdot 10^{10}$ years to be the most probable. There remains, however, the question whether the basic age values do indeed represent the real age of the minerals investigated, or whether the value above gives only a "geochemical" half-life, supplying concordant age results as compared to other age estimation methods. Although a careful selection of the material to be used evidently precedes investigations of this kind, the error of the final half-life value will always depend on the uncertainties of the original age data.

The first measurements concerning the determination of the specific activity of rubidium by physical methods have been carried out mainly by gas counters. However, owing to the low specific activity and to the low average energy of the beta particles, great difficulties arose in preparing the sources, and lack of information about the shape of the spectrum made the corrections for absorption and self-absorption losses inevitably uncertain.

Recently, scintillation methods became prevalent in the study of the specific activity of rubidium. Much lower energies can be measured by this method and at the same time the geometrical efficiency is also near to 100% because the source is incorporated into the detector itself.

In Table 1 available literature data on the half-life of Rb^{87} are summarized quoting the name of the author(s), year of publication and the method employed. Complete data on the source material are given in the references at the end of this paper.

From the data included in Table 1 it may be seen that even results obtained in the last years by employing accurate methods are in considerable disagreement, the scatter of values being by far larger than the experimental errors given. It should be noted however, that at the present time the value given by FLYNN and GLENDENIN [1], obtained by liquid scintillation techniques and equal to $4{,}7 \cdot 10^{10}$ years can be accepted as the most probable one, although this value needs still further confirmation.

Experimental methods

In the measurements reported here the determination of the specific activity of rubidium has been carried out by means of liquid scintillation methods and by employing an Rb-doped NE 220 (made by Nuclear Enterprises Ltd., England) anthracene + POPOP + dioxane liquid scintillator. In order to diminish the quenching effect, rubidium was built into the scintillator solution in the form of Rb-octoate (Rb salt formed with 2-ethylcaproic acid) which was prepared as follows:

Table 1

Literature data on the half-life of rubidium-87

Half-life in units of 10^{10} years	Method	Author(s) and year of publication
7,5	Measurement of Sr/Rb ratio in minerals of known age	O. HAHN—M. ROTHENBACH, 1919 [6]
12,0	Proportional counter	W. MÜHLHOFF, 1930 [7]
4,4	Cloud chamber	G. ORBAN, 1931 [8]
6,3	Measurement of Sr^{87}/Rb^{87} ratio in minerals of known absolute age	F. STRASSMANN—E. WALLING, 1938 [9]
7,5	Measurement of Sr/Rb ratio in minerals of known absolute age	P. H. CHAUDHURY, 1942 [10]
$5{,}80 \pm 1{,}00$	GM counter	S. EKLUND, 1946 [11]
$6{,}50 \pm 0{,}60$	2π GM counter	O. HAXEL—F. G. HOUTERMANS—M. KEMMERICH, 1948 [12]
$6{,}00 \pm 0{,}60$	2π GM counter	M. KEMMERICH, 1949 [13]
$6{,}15 \pm 0{,}30$	High pressure proportional counter	S. C. CURRAN—D. DIXON—H. W. WILSON. 1951 [14] [15]
$5{,}90 \pm 0{,}30$	RbI(Tl) scint. crystal	G. M. LEWIS, 1952 [16]
$6{,}23 \pm 0{,}30$	Gas counter. Sample enriched in Rb^{87}	M. H. MCGREGOR—M. L.WIEDENBECK, 1952 [17] [18]
4,8	4π GM counter	I. GEESE-BÄNISCH—E. HUSTER—W. WALCHER, 1952 [19]
$4{,}30 + 0{,}30 - 0{,}20$	4π GM counter	I. GEESE-BÄNISCH—E. HUSTER, 1954 [20] [21] [22]
$6{,}10 \pm 0{,}20$	GM counter	J. FLINTA—S. EKLUND, 1954 [23]
$5{,}00 \pm 0{,}20$	Measurement of Sr^{87}/Rb^{87} ratio in minerals of known absolute age determined by U/Pb methods	L. T. ALDRICH — et al. 1956 [4] [24]
$4{,}60 \pm 0{,}50$	Measurement of Sr^{87}/Rb^{87} ratio in minerals of known absolute age determined by the K/A method	K. FRITZE—F. STRASSMANN, 1956 [25]
$4{,}88 \pm 0{,}20$	4π GM counter	W. F. LIBBY, 1957 [26]
$4{,}70 \pm 0{,}05$	Rb-doped liquid scintillator	K. F. FLYNN—L. E. GLENDENIN, 1961 [27] [1]
$5{,}00 \pm 0{,}20$	Measurement of Sr^{87}/Rb^{87} ratio in minerals of known absolute age determined by K/A method	G. B. OVTSHINNIKOVA, 1960 [5]
$5{,}82 \pm 0{,}10$	RbI(Tl) scint. crystal	E. EGELKRAUT—H. LEUTZ, 1961 [2]
$5{,}53 \pm 0{,}10$	Rb-doped NaI(Tl) scint. crystal	G. B. BEARD—W. H. KELLY, 1961 [3]
$5{,}25 \pm 0{,}10$	4π proprtional counter	A. MCNAIR—H. W. WILSON, 1961 [28]
$5{,}80 \pm 0{,}12$	RbI(Tl) scint. crystal	H. LEUTZ—H. WENNINGER—K. ZIEGLER, 1962 [29]

An anion exchange column was made of Dowex 1×4; 50—100 mesh anion exchange resin, which was treated with 10% NaOH solution and thus converted into alkaline form. Through the column, in order to produce rubidium hydroxide a solution of RbCl (Merck, reagent grade) was led. The rubidium hydroxide solution obtained at the bottom of the column and containing no Cl^- ions at all was evaporated under vacuum to a minimal volume, and then contacted with about 10% excess of 2-ethylcaproic (octoic) acid (Fluka, reagent grade). The Rb-octoate solution was then evaporated with absolute ethanol under vacuum several times to remove the residual water from the preparation. The Rb-octoate sample was obtained in the form of a very viscous syrup, which was dissolved in a small amount of NE 220 liquid scintillator to get a free flowing liquid, as our previous experiments have shown, that Rb-octoate is well soluble in dioxane. The stock solution obtained in this way was then used for the measurements after further dilution with the liquid scintillator.

The activities of the samples were measured in an IDL 2022 Liquid Measuring Head, originally constructed for tritium and carbon-14 determinations. The used measuring head contains two RCA 6292 photoelectric multipliers mounted into a lead shielding and "viewing" from both sides a flat glass cuvette of 6 mm inner thickness, 28 mm inner diameter and of about 1 mm wall thickness. The volume of the cuvette was found to be 3 ccm.

Signals from the two multipliers were summed by a simple adding circuit. The summed signals were led through a cathode follower and after about 1 microsec. delay to an IDL 652 Wide Band Amplifier. Amplified signals were fed into a Nuclear Data multichannel analyser. Summing the pulses from both sides resulted in an improvement in the signal-to-noise ratio, as pulses originating from scintillations in the solution and thus appearing on both multipliers have got doubled amplitudes, while noise pulses appear only with single amplitudes — except for the spurious coincidences.

Uniform coincidence pulses obtained on an IDL 2032 Coincidence Control Unit were used after appropriate pulse shaping to operate the gate circuit of the multichannel analyser. By the use of this coincidence gating the background due to photomultiplier noise was eliminated — except of course for the spurious coincidences, and the total background was reduced to background originating from radioactive impurities and cosmic ray effects.

The block scheme of the measuring system is demonstrated in Fig. 1.

The determination of the actual rubidium-87 content of the samples was made subsequent to the activity measurements with the aid of stable isotope dilution techniques. The total amount of each sample solution was recovered and evaporated to dryness, following this the organic constituents were destroyed by perchloric and nitric acids on a water steam bath. Residual perchloric acid was removed by moderate heating on a sand bath, and rubidium

was converted into a chloride form by multiple evaporation with HCl. Previous investigations have shown that the procedure above does not cause a noticeable loss of rubidium.

The RbCl solution was evaporated to dryness, picked up again with some water and diluted to a volume depending on the estimated rubidium content. The determination of the exact Rb^{87} content was made by the stable isotope dilution method, using an MI 1305 mass spectrometer with a single

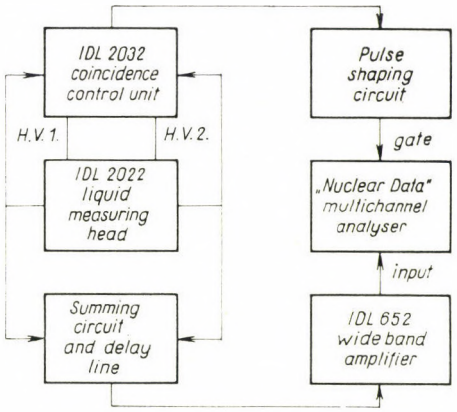


Fig. 1. Block scheme of the apparatus used in the activity measurements

filament surface ionization source. Molar abundance of Rb^{87} in the enriched sample proved to be 95,6%, the Rb^{87} concentration of the enriched spike solution was $9{,}72 \cdot 10^{-6}$ g Rb^{87}/ccm. The rubidium used for our activity measurements proved to be of normal isotopic composition. Relative error in the quantitative Rb^{87} determinations was less than 2%.

Experimental results

The investigations described here were not undertaken in an effort to obtain precise spectrum shapes, by applying a multichannel analyser only integral spectra offering good possibilities to extrapolate in the low energy region had to be achieved. Measurements were carried out at three different concentration levels, obtained by diluting the stock solution in about 1 : 1, 1 : 2 and 1 : 5 proportions, but the actual Rb^{87} contents were determined in each case independently.

The dependence of the quenching effect on the rubidium octoate content of the scintillator solution was investigated. As it is rather difficult to evaluate the beta spectrum near the maximal energy and to obtain the exact end-point

ihannel number the transformation of the spectra in a Fermi—Curie plot s needed; to evaluate the relative quenching effect an external radiation ource was used.

A Cs^{137} source of about 10 mC source strength was placed on to the outer urface of the lead shielding of about 6 cm average thickness and containing

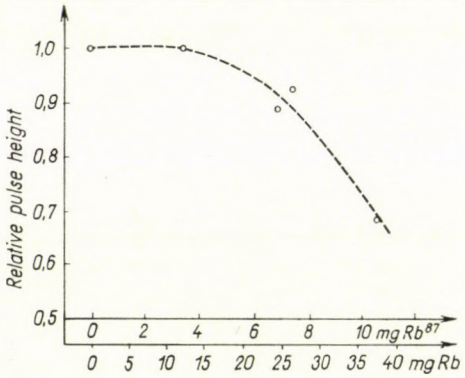


Fig. 2. Relative amplitude heights obtained at various rubidium concentration levels. Rb-concentrations are expressed in the actual Rb^{87} contents in 3 ccm liquid scientillator. On a second axis the corresponding amount of rubidium is shown

the cuvette with the Rb-doped liquid scintillator. Because of the known properties of liquid scintillators no peaks due to photoeffect did appear in the gamma-spectra, but the abrupt front of the Compton edge was very easy to evaluate. End-point channel values with respect to that obtained with undoped scintillator yielded a convenient measure of the quenching effect. Relative amplitude heights obtained in this way and related to the amplitude heights measured with undoped NE 220 liquid scintillator are shown in Fig. 2 as function of the Rb^{87} concentration (proportional to the actual Rb concentration as shown on a second axis) in 3 ccm scintillator solution.

The experimental curve given in Fig. 2 enabled us to determine the background spectra applicable to each concentration level as well. A basic background spectrum was taken with an undoped NE 220 liquid scintillator, and this spectrum was quenched to various levels according to the actual Rb^{87} content of the sample, using the quenching factor determined from the curve in Fig. 2. Careful experiments have shown, that the Rb-octoate solution did not contain any impurity which would cause the elevation of the total background values with respect to that obtained with an undoped scintillator. By mass spectrometrical methods the potassium contamination in the rubidium sample proved to be about 0,5% and owing to the fact that potassium contains the beta-active K^{40} isotope only to an extent of 0,0118%, no corrections to this impurity were found to be necessary. The total back-

ground values — even in the lowest concentration range investigated — were not higher than about 10% of the total counting rates.

The whole beta spectrum was measured in each case. A typical beta spectrum obtained is shown in Fig. 3*a*, where the number of pulses pro minute in each channel is plotted against the channel number. Fig. 3*b* shows the Fermi-Curie plot of the same spectrum, where also representative points of FK-plots obtained by FLYNN and GLENDENIN [1], EGELKRAUT and LEUTZ [2] and BEARY and KELLY [3] are included for comparison, normalized at the 80 keV point to our spectrum. It may be seen, that the shape of the spectrum measured by us corresponds to the spectrum shapes obtained by the authors quoted, giving the best coincidence with that of FLYNN and GLENDENIN [1].

Determination of the total counting rates was made by extrapolation to zero energy, using integral spectra deduced from the differential ones obtained with the multichannel analyser.

Our equipment enabled us to obtain reasonable results down to about 15 keV. This value is somewhat higher than, for example, that of FLYNN and GLENDENIN [1] but the obtained spectra were very suitable for the extrapolation through the low energy region. Extrapolation to the total counting rate was made by using the first six points of the integral spectra over 15 keV, by second order interpolation. Higher order interpolation was not necessary, because the second order correction already did contribute to the total counting rate by less than 0,6% in each case.

Extrapolation to zero beta energy was made with regard to the fact, that — indicated by investigations not referred to here — the intrinsic response of the scintillator needs about 2 keV minimal energy for the absorbed particles to produce any measurable pulse, i.e. to produce at least one photoelectron on the photocathodes of the multipliers.

Results of our measurements are summarized in Table 2 and Fig. 4, where extrapolated total net counting rates are plotted versus measured amounts of Rb^{87} loaded into the liquid scintillator solution.

Table 2

Rb^{87} contents of samples and extrapolated total net counting rates

No.	Measured amount of Rb^{87} used for doping the scintillator solution	Extrapolated total net counting rate, background deduced
1.	$3{,}50 \pm 0{,}08$ mg Rb^{87}	$664{,}4 \pm 5$ d/min
2.	$6{,}87 \pm 0{,}12$ mg Rb^{87}	$1342{,}7 \pm 40$ d/min
3.	$7{,}38 \pm 0{,}10$ mg Rb^{87}	$1408{,}3 \pm 15$ d/min
4.	$10{,}60 \pm 0{,}15$ mg Rb^{87}	$2026{,}9 \pm 10$ d/min

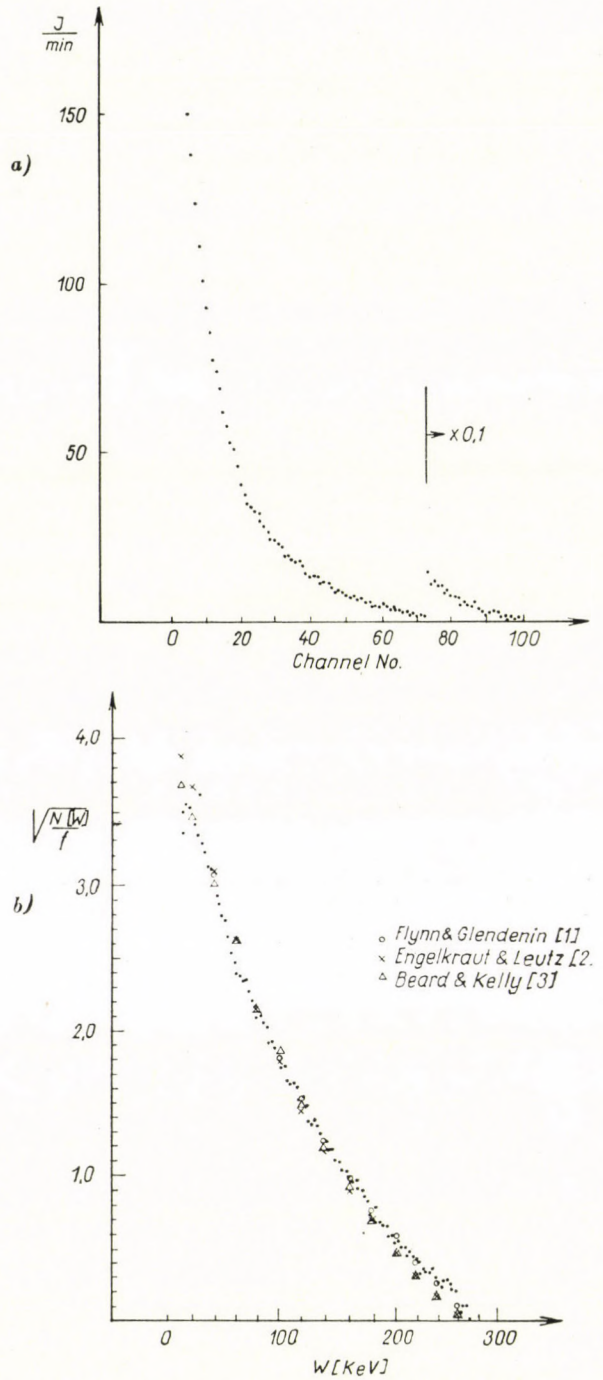


Fig. 3. a) Typical differential beta spectrum obtained with the apparatus shown in Fig. 1 *b)* Fermi-Curie plot of the same spectrum as shown under *a)*. Representative points of spectra obtained by other authors [1] [2] [3] normalized to our spectrum at the 80 keV point are also shown

Measured points on Fig. 4 can be well interpolated by a straight line, the slope of which is yielding the specific activity of rubidium-87.

Linear interpolation was made using a least squares fit of the points, taking into account errors originating from both coordinate values. A value of 191,6 $\pm$ 4,1 d/min · mg Rb^{87} was obtained for the specific activity of

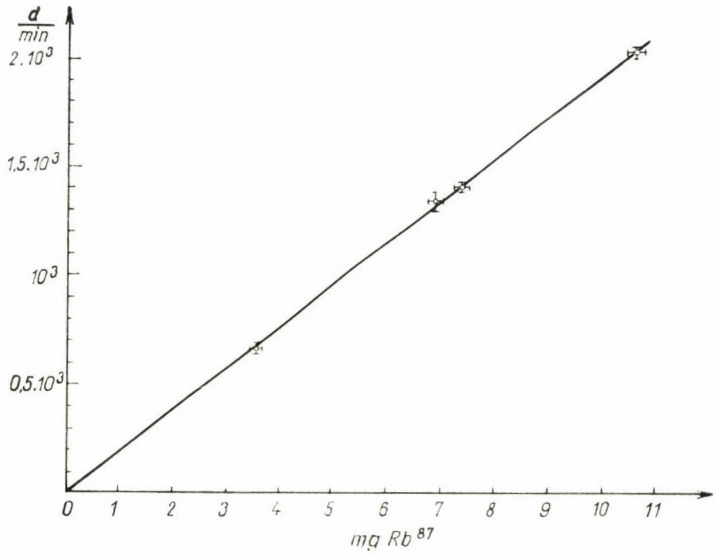


Fig. 4. Extrapolated total net counting rates as function of the actual Rb^{87} content of the scintillator solution

rubidium-87 from the least squares evaluation of the measurement data, corresponding to a half life of $(4{,}77 \pm 0{,}10) \cdot 10^{10}$ years and to a decay constant of $(1{,}45 \pm 0{,}03) \cdot 10^{-11}$ years^{-1}.

Discussion

The value obtained for the half life of rubidium-87 confirms the value given by FLYNN and GLENDENIN [1] within the limits of error, although it lies by about 1,5% higher. One should note, however, that errors given in this report are statistical errors only, and do not include those effects that might cause a systematic underestimation of the total counting rates, i.e. the overestimation of the half life, and which are certainly existing. Although FLYNN and GLENDENIN [1] have shown that liquid scintillators do exhibit a near to 100% counting efficiency even at very low energies, it is not improbable that some losses in the total counting rates could occur due to decreasing counting efficiency in the low energy region. It should be mentioned here, that the counting efficiency of the measuring system was checked with a calibrated tritium sample.

Another source of possible errors lies in the fact that total counting rates could be achieved by an extrapolation method only, crossing a region where no information is available on the shape of the spectrum, not even a theoretical one. The validity of the extrapolation is thus only probable, and supported by the monotonous increase in the differential spectrum with decreasing energy.

As going down to lower energies the differential spectrum gets always steeper, we can assume that if the spectrum differs from the general picture predicted on the basis of the known energy ranges, it can be only steeper than the assumed shape. This means, that if the real value of the specific activity differs from the value determined by the extrapolation method, it can be only higher that that, and correspondingly the true half life can be only lower than the one determined. In this respect our result can also be considered like an upper limit for the half life.

Errors originating from an uncorrect determination of background values are largely compensated by the method of evaluation, although background measurements using the inactive Rb^{85} isotope would be highly desirable.

Acknowledgements

The measurements dealt with in this report were made in the autumn of 1963 at the Institute of Physics of the Helsinki University during the author's stay in Finland which was supported by the Hungarian Academy of Sciences. Thanks are due to Professor M. NURMIA who kindly offered all facilities necessary to carry out the measurements reported here. His encouraging interest in this work is highly appreciated. The kind help of Mr. fil. lis. K. VALLI is warmly acknowledged.

Mass spectrometric measurements completing the reported investigations have been carried out at the Institute of Nuclear Research of the Hung. Acad. Sci. The kind interest of the Director of this Institute, Professor A. SZALAY is also highly appreciated.

REFERENCES

1. K. F. FLYNN and L. E. GLENDENIN, Phys. Rev., **116,** 744, 1959.
2. K. EGELKRAUT and H. LEUTZ, Z. Phys., **161,** 13, 1961.
3. G. B. BEARD and W. H. KELLY, Nucl. Phys., **28,** 570, 1961.
4. L. T. ALDRICH et al., Phys. Rev., **103,** 1045, 1956.
5. G. V. OVTSHINNIKOVA, Geokhimiya, No. 5, 392, 1960.
6. O. HAHN and M. ROTHENBACH, Phys. Z., **20,** 194, 1919.
7. W. MÜHLHOFF, Ann. Phys., **1,** 205, 1930.
8. G. ORBAN, Sb. d. Akad. Wiss. Wien, Math.-Nat. Kl(IIa), **140,** 121, 1931.
9. F. STRASSMANN and F. WALLING, Ber. Deut. Chem. Ges., **71,** 1, 1938.
10. P. H. CHAUDHURY, Proc. Natl. Inst. Sci. India, **8,** 45, 1942.
11. S. EKLUND, Ark. Mat. Astr. Fys. **33A,** 60, 1946.

12. O. Haxel, F. G. Houtermans and M. Kemmerich, Phys. Rev., **74,** 1886, 1948.
13. M. Kemmerich, Phys. Z. **126,** 399, 1949.
14. S. C. Curran, D. Dixon and H. W. Wilson, Phys. Rev., **84,** 151, 1951.
15. S. C. Curran, D. Dixon and H. W. Wilson, Phil. Mag., **43,** 82, 1952.
16. G. M. Lewis, Phil. Mag., **43,** 82, 1952.
17. M. H. McGregor and M. L. Wiedenbeck, Phys. Rev., **86,** 420, 1952.
18. M. H. McGregor and M. L. Wiedenbeck, Phys. Rev., **94,** 138, 1954.
19. I. Geese-Bänisch, E. Huster and W. Walcher, Naturwiss., **38,** 379, 1952.
20. I. Geese-Bänisch and E. Huster, Naturwiss., **41,** 495, 1954.
21. I. Geese-Bänisch, Z. Phys., **142,** 565, 1955.
22. E. Huster, *in* Nuclear Processes in Geological Settings. NAS—NRC Publ. No. 400 Washington, 1956, p. 195.
23. J. Flinta and S. Eklund, Ark. Fys., **7,** 401, 1954.
24. L. T. Aldrich et al., Bull. Amer. Phys. Soc., **1,** 31, 1956.
25. K. Fritze and F. Strassmann, Z. Naturforsch., **11a,** 277, 1956.
26. W. F. Libby, Anal. Chem., **29,** 1566, 1957.
27. L. E. Glendenin, Ann. N. Y. Acad. Sci., **91,** 156, 1961.
28. A. McNair and H. W. Wilson, Phil. Mag., **6,** 563, 1961.
29. M. Leutz, H. Wenninger and K. Ziegler, Z. Phys., **169,** 409, 1962.

ПОВТОРНОЕ ОПРЕДЕЛЕНИЕ ПЕРИОДА ПОЛУРАСПАДА РУБИДИЯ-87

А. КОВАЧ

Резюме

Период полураспада рубидия-87 был определен методом жидкой сцинтилляционной техники, с использованием рубидия естественного изотопного состава. По нашим данным период полураспада составляет $(4{,}77 \pm 0{,}10) \cdot 10^{10}$ лет.

A FAST NEUTRON FLIGHT TIME SPECTROMETER UTILIZING THE RECOIL ALPHA PARTICLES

By

A. ÁDÁM, GABRIELLA PÁLLA and P. QUITTNER

CENTRAL RESEARCH INSTITUTE OF PHYSICS, BUDAPEST

(Presented by L. Jánossy — Received 18. II. 1964)

The fast neutron spectrometer built for the investigation of nuclear reactions induced by the 14,7 MeV energy neutrons produced by a $H^3(d, n)He^4$ source is described. The zero time in the time-of-flight measurement is determined by the recoil alpha particle. This method is compared with another conventional procedure in fast neutron spectroscopy, where a pulsed neutron source is used. The usual data and — in order to illustrate the spectrometer characteristics — measured flight time spectra of the $C^{12}(n, n')$ neutrons are given.

Introduction

In neutron scattering experiments an extensively used method for determining the energies involved is the measurement of the flight times of the neutrons. If the neutrons are generated by the frequently used H^3 $(d, n)He^4$ reaction, the recoil alpha particles can be utilized for determining the zero time [1]. An alternative method is the use of a pulsed neutron source [2]. In the present paper these two methods will be compared and an apparatus, utilizing the associated alpha particle described.

Comparison of the two methods

The main advantage of the associated particle method lies in that it does not require any device for collimating or attenuating the neutron beam. The direction of the emitted neutrons is definitely related to that of the recoil alpha particle, thus the solid angle relating to the alpha particles equally defines the direction of the neutrons. The neutrons flying outside this solid angle may cause but random coincidences and contribute in this way only to the background. To be able to dispense with the need of collimating the neutrons is a considerable asset, since an efficient collimation or attenuation at such high energies is known to be very difficult. The only drawback of the associated particle method is that the neutron yield available for analysis is limited by the number of random coincidences. In the case of pulsed neutron sources, on the other hand, there are no other restrictions on the intensity which may

be applied than the performance of the apparatus, i.e. the maximum attainable yield and the possible detector load. Yet, when using the latter method, systematic coincidences may be caused by neutrons flying in random directions. Thus, the accelerating assembly, the support of the scattering target, etc. are likely to contribute significantly to the background. Therefore, the detectors must be surrounded by collimators which increases the weight and the size of the setup and imposes restrictions upon the positioning of the detectors e. g. in angular distribution measurements.

For a quantitative comparison of the two methods the relative errors obtained in identical experiments over the same measuring time will be considered. The following simplifying assumptions and approximations will be used: The solid angles of the alpha detector and the scattering target with respect to the neutron source are assumed to be equal (Ω_a); the scattering target is placed close to the neutron source so as to have the solid angle of the neutron detector (Ω_n) with respect to the scattering target equal to that to the neutron source; the energy dependence of the efficiency (ε_n) of the neutron detector is left out of consideration.

If one applies the associated particle method, the number of systematic coincidences will be given by

$$N_s^a = \frac{1}{4\pi} N_0 \Omega_a \Omega_n p \varepsilon_n, \tag{1}$$

while the number of background i. e. random coincidences is

$$N_b^a = \frac{2}{(4\pi)^2} N_0^2 \Omega_a \Omega_n \varepsilon_n \tau. \tag{2}$$

N_0 is the neutron yield for the associated particle method, p is the probability that in the target the reaction in question takes place and that the neutron leaves in the given direction at unit solid angle its energy being within an energy interval corresponding to the selected channel, while τ is the time interval corresponding to a channel in the pulse height analyser.

Using a pulsed neutron source, we have

$$N_s^p = \frac{1}{4\pi} \nu n \Omega_a \Omega_n p \varepsilon_n, \tag{3}$$

$$N_b^p = \frac{1}{2\pi} \nu^2 n a \Omega_n \varepsilon_n \tau, \tag{4}$$

where ν is the repetition frequency of the neutron bursts, n is the number of neutrons per burst, a is the attenuation. Here we have assumed that the neutrons after passing the attenuator arrive at the detector uniformly distributed in time.

In the case of the Poisson distribution, the relative error is given by

$$\delta = \frac{\sqrt{N_s + 2N_b}}{N_s}. \tag{5}$$

Using the associated particle method the relative error is seen to decrease with increasing intensity and to approach its minimum value $\delta^a_{\min} = \frac{\sqrt{2N^a_b}}{N^a_s}$, however, at high intensities it changes but slowly. The intensity, however, should not be chosen too high, since then the number of systematic coincidences is obtained as the difference between high counts and in this case increased stability is required. Therefore the condition $N^a_b = N^a_s$ represents the most convenient choice for the intensity in the case of the associated particle method. In this case we obtain

$$N_0 = \frac{2\pi p}{\tau} \qquad \text{and} \qquad \delta^a = \sqrt{3/2\, \delta^a_{\min}}. \tag{6}$$

Making use of the above considerations, the ratio of the relative errors can be expressed as

$$\frac{\delta^a}{\delta^p} = \frac{3}{2\pi} \sqrt{\frac{\nu n\, \Omega_a \tau}{\Omega_a p + 4\nu a \tau}}. \tag{7}$$

We apply now formula (7) to two cases: the 4,45 MeV excitation level in carbon and the maximum, that is about 1 MeV energy level, in the bismuth evaporation spectrum using the data from [3] and [4] as well as those of the present experimental apparatus. i.e. $\Omega_a = 8{,}5 \cdot 10^{-2}$, $\nu = 10^7$, $n = 5$, $\tau = 7 \cdot 10^{-10}$ sec and $a = 1{,}4 \cdot 10^{-2}$. The value of p depending on the material to be analysed on the chosen solid angle and on the energy. In the first case at an angle of 45° with a 10 cm long carbon target 10^{-2} was found which resulted in $\delta^a/\delta^p \sim 1$. In the second case there was $p_{\beta i} \sim 4 \cdot 10^{-4}$ and $\delta^a/\delta^p \sim 3$. In the first case the two methods are found to be equivalent, in the second one the pulsed method seems to be preferable. This comparison, however, does not reveal any significant difference. The associated-particle method appears to be simple, moreover it does not require the use of a collimator, a particularly attractive feature if two detectors have to be used as in the investigation of $(n, 2n)$ reactions.

The pulsed and the alpha methods have also been compared by RETHMEIER et al. [5]. They infer from their experiments that the alpha method is highly preferable to the pulsed source method. Their preference, however, seems not to be fully justified, since in their comparison the authors used the optimum

intensity of the associated particle method also in the pulsed source experiments, though the latter permit the use of much higher intensities.

A method combining the advantages of the above two methods has been worked out by the authors. It is described elsewhere [6].

Description of the apparatus

The deuterium beam feeding the H^3 (d, n) He^4 neutron source is obtained from a 200 keV Cockcroft—Walton type particle accelerator. The atomic ion component is brought to the target after magnetic deflection (Fig. 1). For

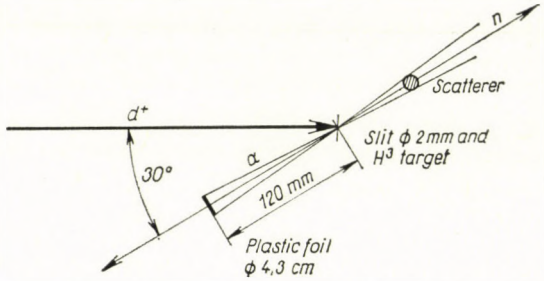


Fig. 1. Target arrangement

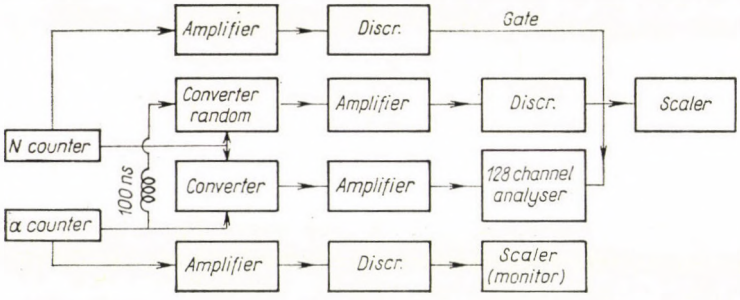


Fig. 2. Block diagram of the measuring apparatus

accurate definition of the alpha solid angle a 2 mm diameter aperture is applied in front of the target. Thus only a 2 mm diameter surface area of the H^3 target is actually exposed to the deuterium beam, though for better exploitation of the target it can be turned under vacuum to vary the area of exposure. The target is pasted to the accelerator tube by araldit. To minimize the background, the neutrons have to cross only a 0,1 mm thick molybdenum backing within the solid angle determined by the alpha particle. The alpha particles, leaving the target at 150° to the deuterium beam, are detected by a plastic scintillator, 4,3 cm in diameter and 0,1 mm thick, mounted on a RCA 6810/A

photomultiplier. The light is shut off by a vacuum deposited aluminium on plastic foil. In order to increase the possible counting rate, cathode follower stabilization is applied [7]. The neutrons are detected by another plastic scintillator, 50 mm in diameter and 230 mm long, coupled also to a RCA 6810/A multiplier.

The block diagram of the counting apparatus is shown in Fig. 2. The flight times are measured with the use of a 100 mμ sec time-to-pulse height converter and a 128 channel pulse height analyser. The time resolution and the linearity of the system have been checked against the flight time spectrum of the 14,7 MeV energy primary neutrons (Fig. 3). The neutron detector efficiency has been experimentally determined [8]. The threshold sensitivity

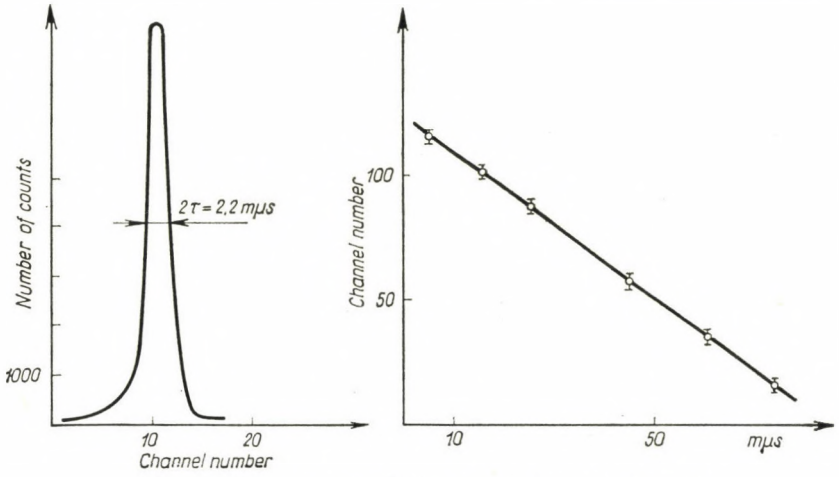


Fig. 3. Curves to check time resolution and linearity

of the apparatus, as measured with the use of the known pulse height spectra of the gamma radiations from Na^{22} and Co^{60}, was found to be below 30 keV electron energy.

The pulse height spectrum of the recoil alpha particles is to be seen in Fig. 4. The alpha signals seem to emerge quite well from the background. The large amplitude signals appearing in the spectrum are due to the 2,6 MeV protons from the $D(d, p)H^3$ reaction. The proton yield increases with increasing deuterium buildup in the target and with increasing target wear. If the target can be kept sufficiently clean, no proton signals will appear in the alpha pulse height spectrum. The deuterium contamination of the target manifests itself usually only after 8 hours of operation. The absolute value of the neutron yield is checked by means of a differential discriminator monitor.

In Fig. 5 the number of neutron-alpha coincidences, characteristic for the determination of the solid angle is plotted as a function of the angle.

To illustrate the energy resolution of the spectrometer the shapes of the flight time spectra measured on neutrons scattered in carbon at 45° are shown for flight paths both of 1 and 2 m (Fig. 6). For 1 m the 9,6 MeV scattering peak of the neutrons is well apparent and also the neutrons corresponding to

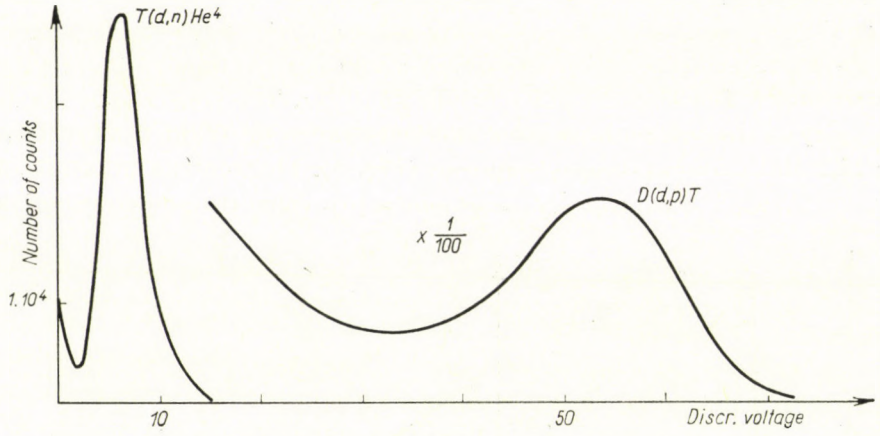


Fig. 4. Pulse height distribution of alpha particles

the 7,7 MeV level can be seen, though for the latter the statistics are rather poor. For 2 m flight path the 10 MeV neutrons scattered from the first level can be well distinguished from the 13,9 MeV neutrons due to elastic scattering.

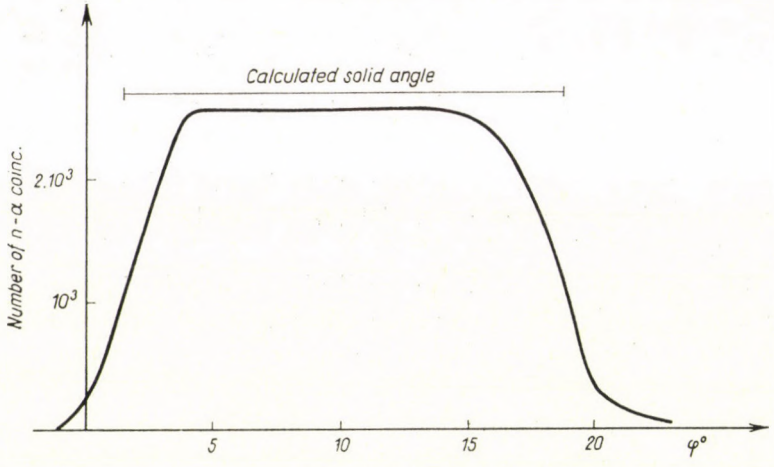


Fig. 5. Alpha solid angle

The authors are indebted to E. PÁSZTOR and to the members of their team for the excellent operation of the measuring apparatus and the accelerator.

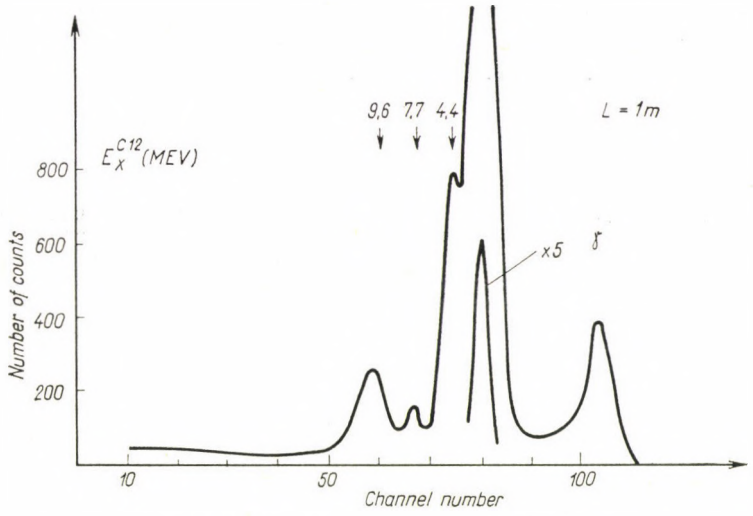


Fig. 6a. C^{12} (n, n') spectrum for $L = 1$ m

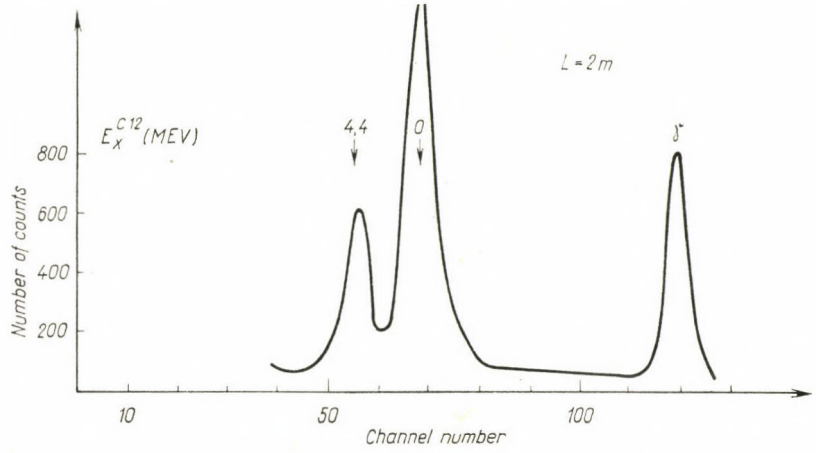


Fig. 6b. C^{12} (n, n') spectrum for $L = 2$ m

REFERENCES

1. G. V. O'Neil, Phys. Rev., **95,** 1235, 1954.
2. L. Crenberg, and J. S. Lewin, Phys. Rev., **103**, 343, 1956.
3. A. Ádám, G. Pálla, and P. Quittner, KFKI Közl., **11,** 197, 1963.
4. A. Neszmélyi and G. Pálla, KFKI Közl., **8,** 275, 1960.
5. Rethmeier et al., Nucl. Instr. and Meth., **17,** 273, 1962.
6. A. Ádám, G. Pálla, E. Pásztor and P. Quittner, Nucl. Instr. and Method. **25**, 365, 1964
7. J. V. Kane, Rev. Sci. Instr., **28,** 582, 1957.
8. A. Ádám, G. Pálla and P. Quittner, KFKI Közl., (to be published).

БЫСТРОНЕЙТРОННЫЙ СПЕКТРОМЕТР ВРЕМЕНИ ПРОЛЁТА ИСПОЛЬЗУЮЩИЙ ОТБРОШЕННЫЕ НАЗАД α-ЧАСТИЦЫ

А. АДАМ, Г. ПАЛЛА и П. КВИТТНЕР

Р е з ю м е

Описывается быстронейтронный спектрометр, сконструированный для исследования ядерных реакций, вызываемых нейтронами с энергией 14,7 МеВ и полученных прядерной реакции $H^3(d, n)$ He^4. При измерении времени пролёта нулевая точка измерения времени определяется отброшенной назад α-частицей. Данный метод сравнивается с другим приемом, — пульсацией источника нейтронов, — часто применяемым в спектроскопии быстрых нейтронов. Для характеристики спектрометра наряду с техническими данными приводится и спектр пролётного времени нейтронов $C^{12}(n, n')$.

THE RENORMALIZABLE VECTOR BOSON THEORY OF WEAK INTERACTION

By

A. FRENKEL and P. HRASKÓ

CENTRAL RESEARCH INSTITUTE OF PHYSICS OF THE HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST

(Presented by L. Jánossy. — Received 4. V. 1964)

A renormalizable vector boson theory of weak interaction can be constructed on the basis of the boson propagator $\frac{i}{(2\pi)^4}\left(g_{\mu\varrho}-\frac{k_\mu k_\varrho}{k^2}\right)\frac{1}{M_v^2-k^2}$. The main consequences of such a theory in allowed nuclear β-decay are as follows:

1. The ratio of the O^{14} and neutron half-lives gives for the ratio $|\lambda|$ of the axial vector and vector coupling constants the value $1 \leq |\lambda| \leq 1{,}1$.
2. This theory gives a somewhat better fit for the unpolarized neutron β-spectrum than the Fermi theory.
3. For the H^3 nucleus the theory reproduces the allowed shape of the spectrum as well as does the Fermi theory.

If further experimental tests should confirm the validity of the theory in nuclear β-decay and in μ-meson decay, we would have a completely renormalizable theory of elementary particles.

In the intermediate boson theory of weak interaction one generally takes the boson propagator in the form

$$\frac{i}{(2\pi)^4}\left(g_{\mu\varrho}-\frac{k_\mu k_\varrho}{M_v^2}\right)\frac{1}{M_v^2-k^2}. \tag{1}$$

This propagator has the merit that it leads to the four—fermion theory in all cases when $M_v \gg |k_\mu|$ and further that it resolves the problem of the "unitarity catastrophe" of the high-energy weak scattering processes. This is true, however, only to the first non-vanishing order of the perturbation expansion. Higher-order terms cannot be calculated, because the propagator (1) leads to a non-renormalizable interaction and therefore a theory based on (1) is not much better than the conventional non-renormalizable Fermi theory.

The possibility of the construction of a renormalizable vector boson theory (r.v.b. theory in the following) with the boson propagator

$$\frac{i}{(2\pi)^4}\left(g_{\mu\varrho}-\frac{k_\mu k_\varrho}{k^2}\right)\frac{1}{M_v^2-k^2} \tag{2}$$

has been emphasized by BIALYNICKI—BIRULA [1]. The same form of the vector boson propagator has also been mentioned by SCHWEBER in [2]. The problem of the compatibility of such a propagator with the basic principles of the

field theory has not been investigated in detail by these authors. In spite of this lack in mathematical rigor as to the foundations of the r.v.b. theory, we started calculations on the basis of the propagator (2), because the success or the failure of the r.v.b. theory in fitting experimental data seems to be more important than its connection with the contestable axioms of present-day field theory.

For $M_v \gg |k_\mu|$ (2) gives $\frac{i}{(2\pi)^4}\left(g_{\mu\varrho} - \frac{k_\mu k_\varrho}{k^2}\right)\frac{1}{M_v^2}$ and one immediately sees that serious departures from the Fermi theory may occur. In part I of this paper the predictions of the r.v.b. theory for the energy distribution of the β-particle in allowed nuclear β-decay are compared with the experiment. In part II some theoretical considerations and suggestions are summarized.

I. Let us write the interaction Lagrangian for the nuclear β-decay in the form

$$L(x) = d{:}\psi_n(x)\frac{(1+i\lambda\gamma^5)}{2}\gamma^\mu\psi_p(x)B_\mu(x){:} + \text{h.c.} +$$

$$+ d:\psi_e(x)\frac{(1+i\gamma^5)}{2}\gamma^\mu\psi_\nu(x)B_\mu(x):+\text{h.c.}$$

Here ψ_n, ψ_p, ψ_e, ψ_ν stand for the neutron, proton, electron and neutrino fields, respectively. $B_\mu(x)$ is the μ-th component of the vector-boson operator. d is the dimensionless ($\hbar = c = 1$) coupling constant of the weak interaction, h.c. stands for the Hermitian conjugate. Then straightforward lowest-order calculation leads to the following result for the energy distribution $W(x)$ of the β-particle, emitted in allowed nuclear β-decay:*

$$W(x) = \frac{d^4}{M_v^4}\frac{m_e^4}{16\pi^3}\left[|<1>|^2 G_F(x,a) + \lambda^2 |<\sigma>|^2 G_{GT}(x,a)\right],$$

$$G_F(x,a) = (a-x)^2 x\sqrt{x^2-1} - \frac{a^2(a-x)\sqrt{x^2-1}}{2[1+4a(a-x)]} -$$

$$- \frac{3a^2-4ax}{8}\ln\left\{1 + \frac{4(a-x)\sqrt{x^2-1}}{1+2(a-x)(x-\sqrt{x^2-1})}\right\}, \qquad (3)$$

$$G_{GT}(x,a) = (a-x)^2 x\sqrt{x^2-1} + \frac{(a-x)\sqrt{x^2-1}}{6}\left(1 - \frac{a^2}{1+4a(a-x)}\right) -$$

$$- \frac{1}{24}(1-4ax+3a^2)\ln\{\quad\}.$$

* The condition $M_v \gg |k_\mu|$ has been maintained and the nucleus has been treated non-relativistically.

In (3) m_e stands for the electron mass, $\langle 1 \rangle$ and $\langle \sigma \rangle$ for the Fermi and Gamow—Teller nuclear matrix element, respectively. x stands for the total energy of the outcoming β-particle, while a denotes the endpoint energy. Both x and a are dimensionless and measured in units of m_e, so that $1 \leq x \leq a$. The mean lifetime τ is connect ed with $W(x)$ by the formula

$$\frac{1}{\tau} = m_e \int_1^a W(x)\, F(x, \pm Z)\, dx, \tag{4}$$

or, going over from the $\hbar = c = 1$ system to the CGS system

$$\frac{1}{\tau} = \frac{m_e c^2}{\hbar} \int_1^a W(x)\, F(x, \pm Z)\, dx \cdot \sec^{-1}. \tag{5}$$

In (4) and (5) the Coulomb correction function $F(x, \pm Z)$ has been introduced.

It can be seen, that the first term in the functions G_F and G_{GT} is the well-known Fermi-spectrum for allowed β-decay. The new terms arise from the term $k_\mu k_\varrho / k^2$ of the propagator (2).

Let us now turn to the experimental test of the r.v.b. theory.

1. Determination of the ratio d^2/M_v^2

The half-life $T_{1/2}^{O^{14}} = \tau^{O^{14}} \ln 2$ of the O^{14} nucleus according to [3] equals (71,34 $\pm$ 0,08) sec. Inserting this value into (5) and taking into account that $|\langle 1 \rangle|^2 = 2$, $|\langle \sigma \rangle|^2 = 0$, $J_i = J_f = 0$ for O^{14}, we find

$$\frac{d^2}{M_v^2} = 4{,}84 \cdot 10^{-49}\, \mathrm{erg} \cdot \mathrm{cm}^3. \tag{6}$$

This corresponds to a Fermi coupling constant

$$f = \frac{1}{\sqrt{8}} \frac{d^2}{M_v^2} = 1{,}71 \cdot 10^{-49}\, \mathrm{erg} \cdot \mathrm{cm}^3$$

instead of $f = 1{,}4 \cdot 10^{-49}$ erg $\cdot$ cm^3 resulting from the Fermi theory. A comparison between the two constants is, however, of little value in our case, since in the r.v.b. theory the energy dependence of the spectrum is modified as compared to the Fermi theory.

2. Determination of $|\lambda|$

This parameter has been determined, as is done usually, from the ratio $T_{1/2}^{0^{14}} : T_{1/2}^{n}$. The O^{14} half-life has been taken from [3]. For the neutron half-life there are two data, reported in [4]. The early measurement of ROBSON et al. gives for the neutron lifetime. $T_{1/2}^{n} = (12{,}8 \pm 2{,}5)$ min., while the measurement of SOSNOVSKY et al. resulted in $(1{,}17 \pm 0{,}3)$ min. Inserting these data (and values $|\langle 1 \rangle|^2 = 1$, $|\langle \sigma \rangle| = 3$, $J_i = J_f = \frac{1}{2}$ for the neutron) one gets

$$|\lambda| = 1{,}02 \quad \text{for} \quad T_{1/2}^{n} = 12{,}8 \quad \text{min.},$$

$$|\lambda| = 1{,}08 \quad \text{for} \quad T_{1/2}^{n} = 11{,}7 \quad \text{min.}$$

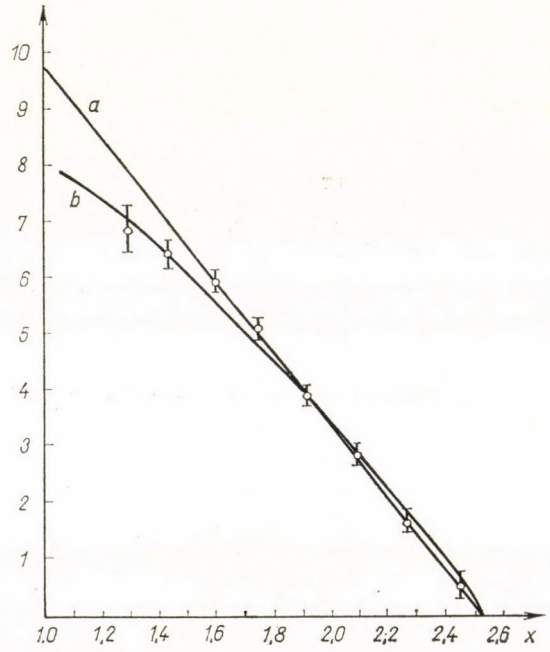


Fig. 1. Experimental and theoretical Kurie-plots for the neutron β-decay.
a) Fermi theory.
b) r.v.b. theory

Both these values of $|\lambda|$ differ significantly from the value of $\lambda = 1{,}25$ given by the Fermi theory. Their deviations from $|\lambda| = 1$ may be caused by secondary unknown factors in the structure of the O^{14} nucleus, or by the electromagnetic interaction of the intermediate boson with the nucleus (see the end of part I). Therefore in the following we accept the value $\lambda = 1$, i.e. the equality of the vector and axial vector coupling constants. We should also like to point out

that as shown by explicit numerical calculations, the *shape* of the β-spectra are rather insensitive to the variation of the value of λ even in the r.v.b. theory (in the Fermi theory the shape of the allowed transition spectra does not depend at all on λ). For example the neutron spectrum was also calculated with $|\lambda| = 1{,}25$, and practically the same shape was obtained as with $|\lambda| = 1$.

3. The β-spectrum of the neutron decay

The best experimental data, published in the literature, are apparently those of ROBSON from 1951, reported in [4]. We reproduced in Fig. 1 the experimental Kurie-plot and the straight line of the Fermi theory as given by ROBSON, and also our theoretical curve. It may be seen that the r.v.b. theory gives a somewhat better agreement with the measured data. The errors of the experimental data are however so large, that no clear-cut conclusion can be drawn. A sufficiently precise measurement of the β-spectrum of the neutron may possibly decide between the thoeries.

4. The β-spectrum of the H^3 nucleus

One of the outstanding successes of the Fermi theory is its excellent agreement with the experimental data for the H^3 decay. To test the r.v.b. theory, we have calculated the functions $G_F(x, a)$, $G_{GT}(x, a)$ for many x-values up to 6 figures and constructed the Kurie plot. The deviations from a straight line are so small that it is difficult to show this graphically. Instead, in Table 1 we give the the fluctuations of the ratio $W(x)/a - x$. One finds that these fluctuations are of the order of $^1/_{1000}$ for nearby points, and of $^1/_{100}$ for distant ones. From this one has to conclude that within the experimental errors both theories are equally acceptable for the description of the H^3 decay.

The mathematical reason for this coincidence of the two theories may be understood from the following properties of the functions $G_F(x, a)$, $G_{GT}(x, a)$:

a) for $x \leqq a$, i.e. at the high-energy end of the spectrum, in the expansion in powers of $a - x$ of these functions many cancellations occur and the first non-vanishing term in both of them is proportional to $(a - x)^2$. So the energy dependence at the high-energy end is the same as in the Fermi theory.

b) for $1 \leqq x$, i.e. at the low-energy end of the spectrum, the expansion of G_F and G_{GT} in powers of $\sqrt{x^2-1}$ gives a similar result: the energy dependence of the first non-vanishing term of both G_F and G_{GT} agrees here also with the energy dependence of the Fermi distribution.

c) It is therefore not surprising, that when the end-point energy a is small enough, i.e. when the two conditions $1 \leqq x$ and $x \leqq a$ are both fulfilled

the whole β-spectrum will show the properties of a Fermi spectrum. More precisely, one can rigorously show, that if

$$4a(a-x) < 1 \qquad \text{for all} \qquad 1 \leq x \leq a,$$

which is equivalent to the condition

$$a < \frac{1+\sqrt{2}}{2} \approx 1{,}2$$

(this condition is largely satisfied for H^3, where $a = 1{,}036$), one finds for the first non-vanishing term of the functions:

$$G_F(x,a) = (a-x)^2 \sqrt{x^2-1}\,[2a-x]\,(a^2-1) + \ldots$$

$$G_{GT}(x,a) = (a-x)^2 \sqrt{x^2-1}\left[x + \frac{1}{3}(a^2-1)(2a-x)\right] + \ldots \qquad (7)$$

The Kurie-plot as given by these terms is

$$\left[\frac{|<1>|^2 G_F + \lambda^2 |<\sigma>|^2 G_{GT}}{x\sqrt{x^2-1}}\right]^{1/2} =$$

$$= \left[(a^2-1)\frac{2a-x}{x}|<1>|^2 + \right.$$

$$\left. + \lambda^2 |<\sigma>|^2 \left[1 + \frac{1}{3}(a^2-1)\frac{2a-x}{x}\right] + \ldots\right]^{1/2} (a-x)$$

and it can easily be seen that the x-dependence of the radical is indeed very weak for e.g. $a = 1{,}036$.

It is, however, difficult to evaluate the precise effect of the omitted terms, and this is why we calculated the H^3 spectrum also numerically.

5. General behaviour of the allowed β-spectra

Explicit numerical computations of the functions $G(x, a)$ lead to the following conclusions for the spectral shapes of the allowed transitions:

a) For small end-point energy the spectrum is of the allowed shape in case of both Fermi- and Gamow—Teller transitions, as we have seen above for H^3.

b) For large end-point energy only the pure Gamow—Teller transitions lead to nearly the allowed shape, while the Fermi transitions deviate from it

especially at the low-energy end of the spectrum. (To illustrate this statement the Kurie-plots of the functions $G_F(x, a)$ and $G_{GT}(x, a)$ for three different values of a are shown in Fig. 2.)

Since the deviations from the allowed shape are mainly due to the presence of the Fermi matrix element, a sufficiently precise measurement of

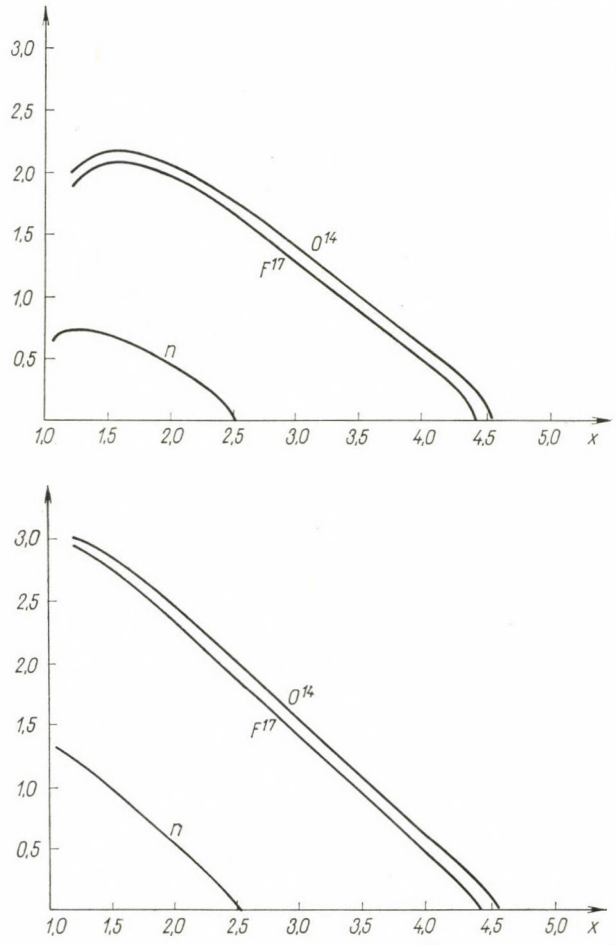


Fig. 2. a) The curve $G_F(x, a)$ for $a^n = 2.527$, $a^{F17} = 4{,}420$ and $a^{O14} = 4{,}552$
b) The curve $G_{GT}(x, a)$ for $a^n = 2{,}527$, $a^{F17} = 4{,}420$ and $a^{O14} = 4{,}552$

allowed β-spectra with pure Fermi transition, or even of a mixed transition (as in the case of the neutron) may be decisive in checking the r.v.b. theory.

The theoretical lifetimes calculated on the basis of the r.v.b. theory are also more or less different from those given by the Fermi theory. In most cases these discrepancies are of the same order of magnitude as the experimental errors and the uncertainties arising from insufficient knowledge of the nuclear

structure. A considerable effect may arise, however, for β-decay when $a \ll 1{,}2$ and $|\langle\sigma\rangle|^2 \approx 0$. In this case the formula (7) holds and we obtain with $|\lambda| = 1$

$$W_{\text{r. v. b.}}(x) = \left\{ |<1>|^2 (a^2 - 1) \left(\frac{2a}{x} - 1 \right) + \right.$$

$$\left. + |<\sigma>|^2 \left[1 + \frac{1}{3} (a^2 - 1) \left(\frac{2a}{x} - 1 \right) \right] \right\} (a - x)^2 x \sqrt{x^2 - 1},$$

while the Fermi theory results in

$$W_F(x) = \{ |<1>|^2 + |<\sigma>|^2 \} (a - x)^2 x \sqrt{x^2 - 1}.$$

Now for a pure Fermi transition we have

$$W_{\text{r. v. b.}}(x) = |<1>|^2 (a^2 - 1) \left(\frac{2a}{x} - 1 \right) (a - x)^2 x \sqrt{x^2 - 1} \ll$$

$$\ll |<1>|^2 (a - x)^2 x \sqrt{x^2 - 1} = W_F(x)$$

and the lifetime predicted by the r.v.b. theory turns out to be larger by several orders of magnitude than the lifetime given by the Fermi theory. For transitions where $|\langle\sigma\rangle|^2$ is not small enough no such a large difference between the two theories arises.

Table 1

The fluctuations in the slope of the Kurie-plot as given by the r. v. b. theory for H^3

x	$W(x)/_{a-x}$	x	$W(x)/_{a-x}$
1,0025	1,7168	1,0200	1,7445
1,0050	1,7207	1,0225	1,7489
1,0075	1,7246	1,0250	1,7535
1,0100	1,7284	1,0275	1,7573
1,0125	1,7322	1,0300	1,7625
1,0150	1,7364	1,0325	1,7667
1,0175	1,7407	1,0350	1,7714

It must be emphasized, that when comparing the r.v.b. theory with the experiment, the Coulomb corrections have to be handled with special care. The point is that in the r.v.b. theory one should in principle, take into account (in addition to the usual Coulomb correction for the β-particle) the interaction of the intermediate boson with the Coulomb field of the nucleus. The need

for such a correction is indicated by the β-spectrum of the F^{17} nucleus. This is one of the well-known cases where the Fermi theory fails to describe the experimental spectra [5—7]. If only the usual Coulomb correction is applied, the r.v.b. theory does not give better agreement with the experiment than the Fermi theory. It is possible, however, that by taking into account the electromagnetic interaction of the intermediate boson agreement with the experiment will improve.

While for the shape of the spectra the Coulomb interaction of the boson may be important, the total area under the curve $W(x)$ is presumably not seriously changed by it. That is why the values of a^2/M_v^2 and $|\lambda|$ calculated from the O^{14} spectrum with the usual Coulomb correction only, seem to be reliable. We should like to point out that a small change in the numerical values of these constants does not affect the main conclusions reported in this paper.

II. Let us now mention some of the theoretical consequences of the r.v.b. theory of the weak interaction.

a) If further experimental tests should confirm the validity of the r.v.b. theory in nuclear β-decay and in μ-meson decay, we would have a completely renormalizable theory of elementary particles, where the three basic interactions (weak, electromagnetic, strong) all would be of the fermion—fermion—boson type.

b) The connection between the weak and the electromagnetic interactions becomes particularly close. One may, indeed, say that there exists in nature a vector boson described by propagator (2), which in its neutral state participates with coupling constant $e = \sqrt{\frac{4\pi}{137}}$ and with mass $M_v = 0$ in processes where parity is conserved, and in its electrically charged states with coupling constant d and mass $M_v \gg m_e$ in processes where parity is not conserved. Such a picture may perhaps be considered as a relaization of the degenerated NAMBU boson [8], which also has one state with $m = 0$ and another with $m > 0$, and the very existence of which is also connected with the degeneracy of the vacuum.

c) We may now investigate the possibility of $d = e$. Taking into account the experimental values $e = \sqrt{\frac{4\pi}{137}}$ and $\frac{d^2}{M_v^2} = 4{,}84 \cdot 10^{-49}$ erg $\cdot$ cm^3 (the latter found in the present paper for the charged vector boson), we get for the mass of this boson

$$M_v \approx 47 M_N \approx 45 \text{ GeV}.$$

Such a boson might be found at present only in cosmic-ray nuclear emulsion plates. It is plausible that this heavy boson is highly unstable and if it decays

into many particles, it may imitate a jet with an unexpectedly high energy release, where, however, the characteristic two-cone structure will be missing.

We should like to stress that the identity of d with e is not at all necessary for the validity of the r.v.b. theory, it is simply a possibility, which may or may not be realized.

*

The authors are aware of the fact that the results reported in part I of this paper are far from providing a full experimental verification of the r.v.b. theory. But the results obtained so far seem to be promising, and moreover certain aspects of the new theory are attractive. Thus it appears that further work along these lines may be of value. The authors are deeply indebted to Drs. G. Bozóki, E. Fenyves, L. I. Lapidus and G. Marx for many interesting discussions on the subject. Thanks are also due to Mrs. T. Gál, who has carried out the tedious numerical computations.

REFERENCES

1. I. Bialynicki-Birula, Journal of Mathematical Physics, **3,** 1094, 1962.
2. S. S. Schweber, An Introduction to Relativistic Quantum Field Theory, p. 604. Row, Petterson and Co., New-York, 1961.
3. D. L. Hendrie and J. B. Gerhard, Phys. Rev., **121,** 846, 1961.
4. Handbuch der Physik, XLI/2, 1962. p. 76. ref. [1].
5. С. С. Василев, Л. Я. Шасталов, ЖЭТФ, **36,** 317, 1959. Изв. Ак. Наук. СССР, сер. физ., 22, 788, 1958.
6. C. Wong, Phys. Rev., **95,** 765, 1954.
7. R. Wallace and J. A. Welch, Phys. Rev., **117,** 1297, 1960.
8. Y. Nambu and G. Jona-Lasino, Phys. Rev., **122,** 345, 1961.

ПЕРЕНОРМИРУЕМАЯ ТЕОРИЯ СЛАБЫХ ВЗАИМОДЕЙСТВИЙ

А. ФРЕНКЕЛЬ и П. ХРАШКО

Р е з ю м е

Перенормируемая теория слабых взаимодействий с промежуточным векторным бозоном может быть построена, если функцию распространения бозона взять в виде $\frac{i}{(2\pi)^4}(g_{\mu\varrho} - k_\mu k_\varrho/k^2)\frac{1}{M_v^2 - k^2}$. Главные следствия такой теории в области разрешенного ядерного β — распада следующие:

1. Отношение времен полураспада O^{14} и нейтрона дает для отношения $|\lambda|$ аксиал-векторной и векторной постоянной взаимодействия значение $1 \leq |\lambda| \leq 1{,}1$.
2. Перенормируемая теория согласуется с экспериментальным β — спектром неполяризованного нейтрона несколько лучше, чем теория Ферми (фиг. 1).
3. Для H^3 распада теория согласуется с разрешенным видом β — спектра так же хорошо, как теория Ферми.

Если дальнейшая экспериментальная проверка подтвердит состоятельность теории для ядерного β — распада и для распада μ — мезона, то вся теория элементарных частиц станет перенормируемой.

HÖHERE NÄHERUNGEN DES STATISTISCHEN ATOMMODELLS IN WELCHEM DIE ELEKTRONEN NACH DER HAUPTQUANTENZAHL GRUPPIERT SIND II

Von

P. GOMBÁS und T. SZONDY

PHYSIKALISCHES INSTITUT DER UNIVERSITÄT FÜR TECHNISCHE WISSENSCHAFTEN, BUDAPEST
und
FORSCHUNGSGRUPPE FÜR THEORETISCHE PHYSIK DER UNGARISCHEN AKADEMIE DER WISSENSCHAFTEN, BUDAPEST

(Eingegangen 7. V. 1964)

Für das statistische Atommodell, in welchem die Elektronen nach der Hauptquantenzahl gruppiert sind, wird eine höhere Näherung entwickelt, indem für das Paulische Besetzungsverbot-Potential, das dem Paulischen Besetzungsverbot der von den Rumpfelektronen vollbesetzten Quantenzuständen Rechnung trägt, ein genauerer Ausdruck verwendet wird. Die Berechnungen der Elektronenverteilung für die Atome Ne und Ar, sowie für das Rb^+-Ion führen zu Resultaten, die mit den Hartreeschen merklich besser übereinstimmen als die in den vorangehenden Arbeiten erhaltenen, was sich besonders bei den diamagnetischen Suszeptibilitäten stark bemerkbar macht. In der Energie der Atome Ne und Ar und in der des Rb^+-Ions ergibt sich eine geringe Erhöhung.

Das statistische Atommodell, in welchem die Elektronen nach der Hauptquantenzahl gruppiert sind, wurde in einigen vorangehenden Arbeiten ausgearbeitet.[1] In diesem Modell werden die Elektronen in den einzelnen Elektronenschalen von einander gesondert statistisch behandelt. In bezug auf ein Elektron in einem Quantenzustand mit der Nebenquantenzahl l und der Hauptquantenzahl n (d.h. in der n-ten Elektronenschale) wurde das Paulische Besetzungsverbot der inneren (energetisch tieferen) Elektronenzustände mit der Nebenquantenzahl l durch das nicht-klassische Abstossungspotential, das sogenannte Besetzungsverbot-Potential[2]

$$G_l^{(n)} = -\frac{\pi^2}{8(2l+1)^2}\, e a_0\, D_l^2 - \frac{1}{4}\, e a_0 \frac{1}{r^2} \tag{1}$$

ersetzt, wo D_l die Summe der radialen Dichten der Elektronen mit der Nebenquantenzahl l in den Elektronenschalen mit der Hauptquantenzahl 1, 2, . . ., $n-1$ bezeichnet; weiterhin ist r die Entfernung vom Kern, e die positive Elementarladung und a_0 der erste Bohrsche Wasserstoffradius.

Aus den Potentialen $G_l^{(n)}$ wird für die global behandelten Elektronen der n-ten Elektronenschale das mittlere Besetzungsverbot-Potential $G^{(n)}$ fol-

[1] P. GOMBÁS u. K. LADÁNYI, Acta Phys. Hung. **5**, 313, 1955; Acta Phys. Hung. **7**, 255, 1957; Acta Phys. Hung. **7**, 263, 1957; Acta Phys. Hung. **8**, 301, 1958; Zeitschrift für Physik, **158**, 261, 1960.

[2] P. GOMBÁS, Acta Phys. Hung. **1**, 285, 1952. Man vgl. auch P. GOMBÁS, Handbuch der Physik, Bd. **36**/2, S. 168 ff. Berlin-Göttingen-Heidelberg, Springer, 1956.

gendermassen definiert

$$G^{(n)} = \frac{1}{N_n} \sum_l n_l G_l^{(n)}, \tag{2}$$

wo n_l die Anzahl derjenigen Elektronen der n-ten Elektronenschale bezeichnet, die in einem Quantenzustand mit der Nebenquantenzahl l gebunden sind und N_n die gesamte Elektronenzahl der n-ten Elektronenschale ist.

Die Elektronendichte der n-ten Elektronenschale haben wir in der Form

$$\varrho_n = A_n r^{2(\varkappa_n - 1)} e^{-2\lambda_n r} \tag{3}$$

angesetzt, wo A_n einen Normierungsfaktor und $\varkappa_n$ sowie λ_n Variationsparameter bezeichnen,[3] die aus dem Energieminimum des Atoms bestimmt werden. Bezüglich weiterer Einzelheiten verweisen wir auf die eingangs zitierten vorangehenden Arbeiten.

In der vorliegenden Arbeit wird für das Besetzungsverbot-Potential $G_l^{(n)}$ ein inzwischen hergeleiteter[4] genauerer Ausdruck herangezogen, der in bezug auf ein Elektron mit der Nebenquantenzahl l und der Hauptquantenzahl n folgende Gestalt hat

$$G_l^{(n)} = -\frac{\pi^2}{8(2l+1)^2} ea_0 (D_l^2 + 2D_l P_l) - \frac{1}{4} ea_0 \frac{1}{r^2}, \tag{4}$$

wo P_l die im wellenmechanischen Sinne gedeutete radiale Dichte des Bezugselektrons (im Quantenzustand n, l) bedeutet. Dieses Abstossungspotential unterscheidet sich von (1) durch das im Verhältnis zu D_l^2 kleine Glied $2D_l P_l$ in der Klammer, das zu einer kleinen Vergrösserung der Abstossung führt, die in den inneren Gebieten des Atoms unbedeutend ist, jedoch in den Randgebieten des Atoms die Abstossung merklich vergrössert. Hierdurch wird, wie aus unseren Resultaten hervorgeht, die in den vorangehenden Arbeiten gewonnene Elektronendichte im richtigen Sinne korrigiert, da die Elektronendichte durch diese Vergrösserung der Abstossung in den äusseren Gebieten des Atoms — wo sie sich in den früheren Arbeiten als zu klein ergab — angehoben wird.

Die Bestimmung der Dichteverteilung der Elektronen mit dem Besetzungsverbot-Potential (4) wurde für die Atome Ne und Ar und für das Rb^+-Ion durchgeführt und zwar auf ganz ähnliche Weise wie im Teil I dieser Arbeit.[5] Für die Exponenten $\varkappa_n$ im Ausdruck der Elektronendichten wurden

[3] Im Exponenten von r haben wir hier den in den vorangehenden Arbeiten in nicht konsequenter Weise mit n bezeichneten Variationsparameter mit $\varkappa_n$ bezeichnet.

[4] P. GOMBÁS, Zeitschrift für Physik, **172**, 293, 1963.

[5] P. GOMBÁS u. T. SZONDY, Acta Phys. Hung. **14**, 335, 1962.

die in I gewonnenen Werte (dort mit n bezeichnet) beibehalten. Die Parameter λ_n im Ausdruck der Elektronendichte haben wir geradeso wie in I wieder durch simultane Variation aus dem Energieminimum des Atoms bestimmt.

Die hier gewonnenen Werte der Variationsparameter sind in der Tabelle 1 angegeben. Zum Vergleich sind in dieser Tabelle auch diejenigen Werte dieser Parameter angeführt, die wir in I mit dem Ausdruck (1) des Besetzungsverbot-Potentials erhielten.

Tabelle 1

Werte der Variationsparameter $\varkappa_n$ und $2\lambda_n$, letzterer in $1/a_0$-Einheiten

	Resultate aus I mit dem Besetzungsverbot-Potential (1)		Hier gewonnene Resultate mit dem Besetzungsverbot-Potential (4)	
	$\varkappa_n$	$2\lambda_n$	$\varkappa_n$	$2\lambda_n$
Ne	1,0	19,65	1,0	19,74
	1,5	4,11	1,5	4,01
Ar	1,0	36,65	1,0	37,11
	2,0	13,36	2,0	12,87
	3,0	4,79	3,0	4,43
Rb^+	1,0	77,33	1,0	79,06
	2,0	32,11	2,0	30,99
	3,0	14,44	3,0	13,29
	4,0	6,18	4,0	5,41

Die Resultate für den Dichteverlauf der Elektronen sind in den Figuren 1, 2 und 3 angegeben und zwar ist in diesen die gesamte radiale Elektronendichte $D(r) = 4\pi r^2 \sum_n \varrho_n(r)$ als Funktion von r dargestellt. Zum Vergleich ist in den Figuren auch der mit dem in I benutzten Besetzungsverbot-Potential (1) gewonnene und der mit der Methode des self-consistent field ermittelte Hartreesche radiale Dichteverlauf[6] eingezeichnet. Wie aus den Figuren zu sehen ist, stimmt der hier gewonnene Dichteverlauf für das Ar-Atom und Rb^+-Ion mit dem Hartreeschen durchweg besser überein als der in I berechnete, besonders augenfällig ist dies in grösserer Entfernung vom Kern. Für das Ne-Atom ist die Übereinstimmung des hier gewonnenen Dichteverlaufs mit dem Hartreeschen für $r > 2a_0$ ebenfalls besser als die des in I gewonnenen Dichteverlaufes. Im Inneren des Ne-Atoms ist dies nicht der Fall. Es sei jedoch bemerkt, dass der hier dargestellte Hartreesche Dichteverlauf für das Innere

[6] Für Ne: F. W. BROWN, Phys. Rev. (2) **44**, 214, 1933. Für Ar: D. R. HARTREE u. W. HARTREE, Proc. Roy. Soc. Lond. (A) **166**, 450, 1938. Für Rb^+: D. R. HARTREE u. W. HARTREE, Proc Roy. Soc. Lond. (A) **151**, 96, 1935.

des Ne-Atoms keine ganz korrekte Vergleichsbasis gibt, da dieser Dichteverlauf mit Berücksichtigung der Austauschwechselwirkung zwischen den 1*s*- und 2*s*-Elektronen berechnet wurde, während in unserem Verfahren der Elektronenaustausch gänzlich unberücksichtigt bleibt.

Es erheben sich hier noch die folgenden Fragen: erstens in welchem Masse die von uns mit dem Variationsverfahren ermittelte Lösung die exakte

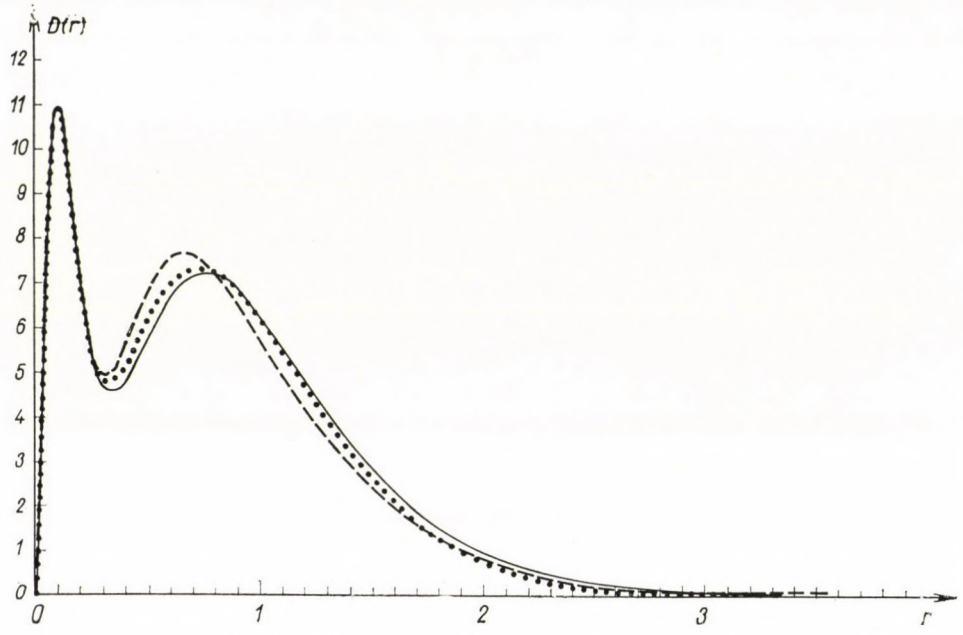


Fig. 1. Vergleich der radialen Elektronendichten D des Ne-Atoms, r in a_0- und D in $1/a_0$-Einheiten.
———— hier berechnet,
····· in I berechnet,
— — — nach HARTREE.

Lösung unseres Problems annähert und zweitens ob die exakte Lösung von der Hartreeschen nicht etwa stärker abweicht als die mit dem Variationsverfahren berechnete Näherungslösung? Zur Beantwortung dieser Fragen haben wir die Elektronendichte der äussersten Schale (N-Schale) des Rb^+-Ions auch exakt berechnet, indem wir die mit dem Variationsproblem equivalente erweiterte Schrödingersche Differentialgleichung numerisch exakt lösten. Der auf diese Weise nach unserem Modell exakt berechnete radiale Dichteverlauf der Elektronen in der N-Schale des Rb^+-Ions ist zusammen mit dem mit der Variationsmethode bestimmten und dem Hartreeschen Dichteverlauf in Fig. 4 dargestellt. Wie zu sehen ist, stimmt der auf numerischem Wege exakt berechnete Dichteverlauf mit dem Hartreeschen bedeutend besser überein als der mit dem Variationsverfahren bestimmte. Auf Grund

dieses Resultates und einiger früher durchgeführten ähnlichen Kontrollrechnungen kann man erwarten, dass dies allgemein der Fall ist. Die mit unserem Verfahren exakt berechneten Dichteverteilungen dürften also auch in allen anderen Fällen die Hartreeschen sehr gut annähern.

Einen Aufschluss über die Güte des Dichteverlaufes in den äusseren Gebieten der Atome erhält man durch Vergleich der diamagnetischen Suszepti-

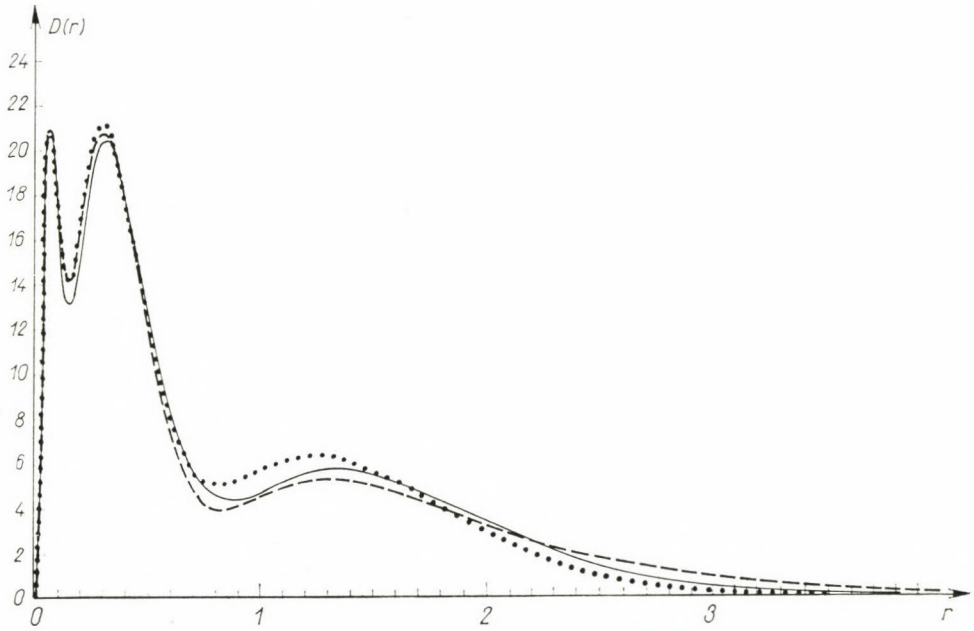


Fig. 2. Vergleich der radialen Elektronendichten D des Ar-Atoms, r in a_0- und D in $1/a_0$-Einheiten.
——— hier berechnet,
····· in I berechnet,
— — — nach HARTREE.

bilitäten mit der Erfahrung, da die diamagnetische Suszeptibilität zum Mittelwert von r^2 proportional ist, in welchem die äusseren Gebiete des Atoms stark betont sind. Die Werte der diamagnetischen Suszeptibilität, die wir mit dem hier gewonnenen Dichteverlauf erhielten, sind in der Tabelle 2 angegeben. Zum Vergleich sind auch diejenigen Werte der diamagnetischen Suszeptibilität angeführt, die man mit den Hartreeschen Verteilungen, sowie mit den in I erhaltenen Verteilungen erhält; ausser diesen enthält die Tabelle auch die empirischen Werte.[7] Wie aus der Tabelle ersichtlich ist, ergibt sich mit den Hartreeschen Werten in allen drei Fällen und mit den empirischen Werten für Ar

[7] LANDOLT-BÖRNSTEIN: Zahlenwerte und Funktionen, Atom und Molekularphysik, 1. Teil, Atome und Ionen. Herausgegeben von A. EUCKEN, Springer-Vlg., Berlin-Göttingen-Heidelberg, 1950, S. 394—395.

und Rb^+ eine bedeutend bessere Übereinstimmung, wenn man die Suszeptibilität mit den hier erhaltenen statt mit den in I erhaltenen Verteilungen berechnet. Der hier berechnete Wert der diamagnetischen Suszeptibilität für Ne stimmt mit dem empirischen etwas schlechter überein als der mit der in I gewonnenen Verteilung berechnete.

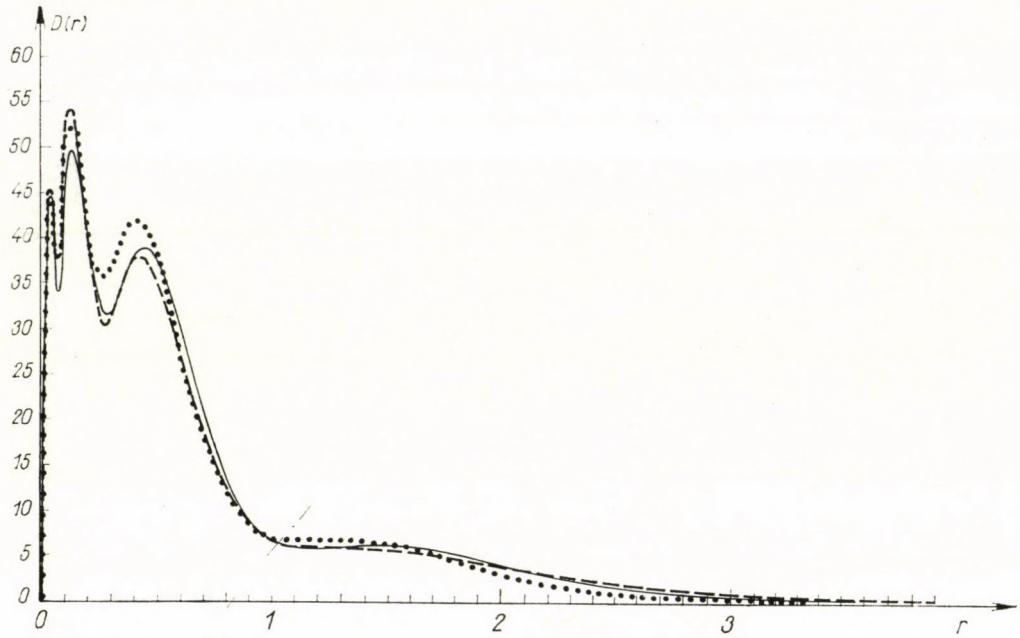


Fig. 3. Vergleich der radialen Elektronendichten D des Rb^+-Ions, r in a_0, und D in $1/a_0$ Einheiten.
———— hier berechnet,
····· in I berechnet,
— — — nach HARTREE.

Die Energie der Atome Ne und Ar und des Rb^+-Ions sind mit dem Minimum der Energie als Funktion der Variationsparameter γ_n und λ_n gleichzusetzen. Die hier gewonnenen Energiewerte sind zusammen mit den in I

Tabelle 2

Diamagnetische Suszeptibilität der Atome Ne und Ar und des Rb^+-Ions, in -10^{-6} cm³/mol-Einheiten

	Mit den Dichteverteilungen aus I berechnet	Mit den hier gewonnenen Verteilungen berechnet	Mit den Hartreeschen Verteilungen berechnet	Empirische Werte
Ne	7,6	7,9	8,0	7,2
Ar	16,5	19,2	24,8	19,4
Rb^+	18,9	24,2	29,5	22,5

berechneten und den halbempirischen Werten[8] in der Tabelle 3 angegeben. Zufolge der erhöhten Abstossung des Zusatzpotentials (4) im Verhältnis zu (1) liegen die hier mit dem Zusatzpotential (4) berechneten Energien etwas höher als die in I mit dem Zusatzpotential (1) berechneten. Diese geringe Er-

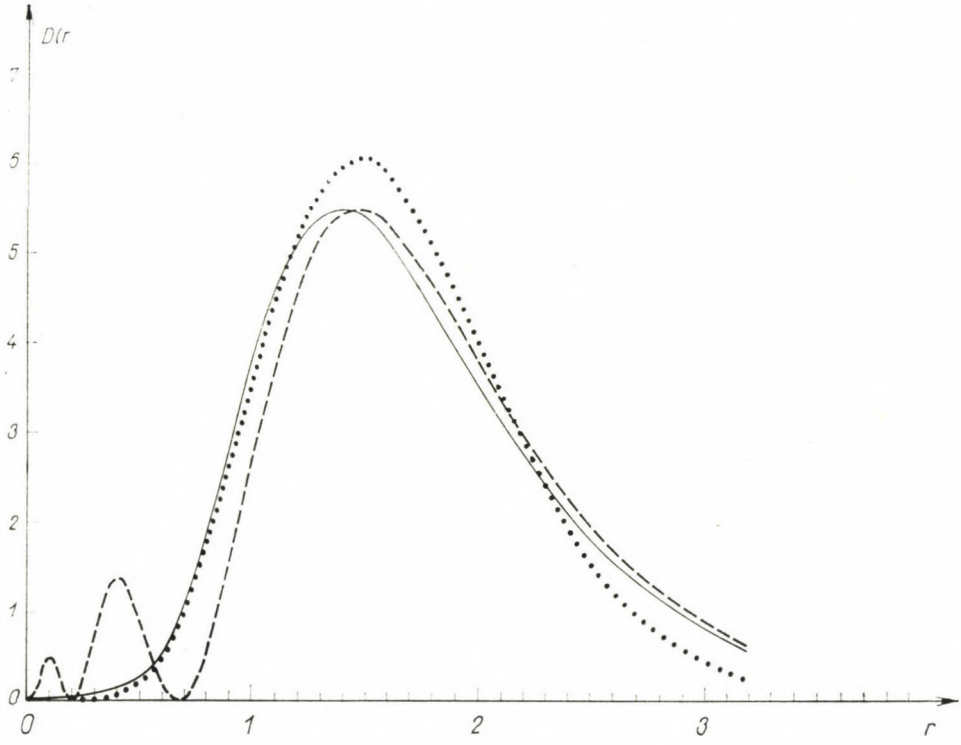


Fig. 4. Vergleich der radialen Elektronendichten D der N-Schale des Rb^+-Ions. r in a_0- und D in $1/a_0$-Einheiten.
····· hier mit dem Variationsverfahren berechnet,
———— hier auf numerischem Wege exakt berechnet,
— — — — nach HARTREE.

Tabelle 3

Energie der Atome Ne und Ar und des Ions Rb^+ in e^2/a-Einheiten

	Resultate aus I	Hier gewonnene Resultate	Halbempirische Werte
Ne	— 125,94	— 125,63	— 129,5
Ar	— 521,44	— 517,15	— 525,4
Rb^+	—2925,7	—2877,9	—2885,3

[8] J. C. SLATER, Phys. Rev. **36**, 57, 1930.

höhung der Energie ist sehr befriedigend, denn bei Hinzunahme der Austauschwechselwirkung und der Korrelation der Elektronen in den nächsten Näherungen wird die Energie etwas vertieft, so dass ein zu starkes Absinken der Energie in den weiteren Näherungen gerade durch die hier gewonnene Energieerhöhung voraussichtlich vermieden wird.

ВЫСШИЕ ПРИБЛИЖЕНИЯ СТАТИСТИЧЕСКОЙ МОДЕЛИ АТОМА, В КОТОРОЙ ЭЛЕКТРОНЫ СГРУППИРОВАНЫ ПО ГЛАВНЫМ КВАНТОВЫМ ЧИСЛАМ II

П. ГОМБАЩ и Т. СОНДИ

Р е з ю м е

Для статистической теории атома, сгруппирующей электроны по главным квантовым числам, разрабатывается одно высшее приближение, в котором применяется более точное выражение для потенциала запрета Паули, принимающего во внимание принцип запрета Паули для полностью заполненных электронами остова атома квантовых состояний. Вычисление распределения электронов для атомов Ne, Ar и для иона Rb^+ приводят к результатам, намного лучше согласующимся с результатами Хартри, чем это наблюдалось в предыдущих работах. Это особенно ярко выражается в случае диамагнитной восприимчивости. В энергии атомов Ne, Ar и иона Rb^+ наблюдается некоторое повышение

COMMUNICATIONES BREVES

ON THE INSTABILITY OF CLUSTERS OF GALAXIES

By

G. PAÁL

KONKOLY OBSERVATORY, BUDAPEST

(Received 20. III. 1963)

The probable value of the cosmological deceleration parameter (q_0) as determined from the red-shifts of extragalactic nebulae, is about 3 [1]. From this the probable order of magnitude of the cosmical constant, Λ, of the general theory of relativity may be estimated to be $\Lambda = -10^{-55}$ CGS. On the other hand, according to the results of SCHLÜTER [2], the law of gravitational attraction of a mass spherically distributed, including the influence of Λ, in "Newtonian approximation" can be written

$$F = -\frac{\gamma M}{r^2} + \frac{\Lambda c^2 r}{3}, \tag{1}$$

where F is the radial force per unit mass, r the radial distance, γ the gravitational constant, M the total mass within the radial distance r, c the velocity of light in vacuo. This means that the influence of Λ is equivalent to an additional gravitational attraction of a material having a density of $\varrho_\Lambda = -\frac{\Lambda c^2}{4\pi\gamma} \sim 10^{-28}$ CGS.

It is known that in spite of all morphological, statistical and astrophysical data indicating the stability of large globular clusters of galaxies, dynamical considerations make them appear to be extremely unstable [3]. According to computations, a density of ten to fifty times higher than that resulting from direct measurements would be required to ensure the stability of the clusters. Since the value of ϱ_Λ found above is of the same order of magnitude as the missing density, in our opinion the influence of the cosmical constant may make up for the attraction of the material missing in the large clusters of galaxies.

The condition of stability for a cluster, also considering the influence of ϱ_Λ, obviously may be written as

$$\frac{V_{\max}}{R} = \left\{\frac{4\pi\gamma}{3}(\varrho_\Lambda + \varrho)\right\}^{1/2}, \tag{2}$$

where $V_{\max}$ is the maximum specific velocity of all elements in the cluster of galaxies, ϱ the average density of the cluster, R the radius of the cluster. If, therefore, globular clusters of galaxies are really stable and their stability is a consequence of the relativistic effect described above, then for every cluster the empirical relation $\frac{V_{\max}}{R} = \text{constant}$ must hold true, as $(\varrho_\Lambda + \varrho)^{1/2}$ may be considered to be practically equal for the various clusters of galaxies. If we suppose that the velocity distribution of galaxies does not vary significantly from cluster to cluster, our condition may be replaced by the relation $\frac{\sigma_r}{R} = \text{const.}$ (σ_r is the line-of-sight velocity dispersion of the elements within the cluster.) Within the limits of accuracy of measurement it can be proved that all well-known clusters of galaxies of regular shape fulfil the above relation.

Globular clusters with measured velocities are the following [1, 4, 5, 6]:

	n	$\frac{\sigma_r}{R}$ km/sec/Mpc
Coma Cluster	46	130 h
Corona Borealis Cluster	8	135 h
Pegasus Cluster	5	131 h
Perseus Cluster	5	103 h
Fornax Cluster	5	160 h
Cancer Cluster	4	141 h

Here n is the number of galaxies with known velocity, $\mathrm{h} = \frac{H_0}{100}$, and H_0 is the Hubble parameter.

The mean value weighted by the number of measurements is $\frac{\overline{\sigma_r}}{R} = 133$ h km/sec/Mpc. Maximum relative deviations from the mean in case of incomplete sampling (of 4—5 elements with known velocity) is $\pm 22\%$. The average of the quotients obtained from incomplete measurements (based on 19 measurements) is $\frac{\overline{\sigma_r}}{R} = 134$ h, which shows a deviation of only 3% from the value determined for the Coma Cluster from 46 measurements.

It is seen that the relationship $\frac{\overline{\sigma_r}}{R} = \text{const.}$ holds true for all sufficiently known large globular clusters of galaxies, which empirically supports our conception as to the solution of the problem of instability of clusters.

Because of the uncertainty resulting from insufficient knowledge of the orientation of clusters in space for ellipsoidal clouds of galaxies of regular shape only a rough average value can be given. By using the averages of G.

DE VAUCOULEURS (Group A2) [3] we get $\frac{\overline{\sigma_r}}{R} \sim 160$ h km/sec/Mpc. For "association-like" groups including also blue giant galaxies (Group A1) we obtain $\frac{\overline{\sigma_r}}{R} \sim 530$ h km/sec/Mpc. As might be expected, these groups generally prove unstable, even if the A-force is being taken into account. Our considerations do not apply to such groups of galaxies of significantly higher densities and smaller dimensions, as, for instance, Stephan's Quintet.

We wish to point out that the above interpretation is not the only possible one. All that can be proved nowadays by dynamical considerations seems to be that the density of the material missing in the Metagalaxy on the one hand and in the clusters of galaxies on the other hand are approximately equal. Consequently our empirical data allow to suppose also a really existing but not observable substance (e.g. gas of neutrinos) of approximately constant density everywhere in the Metagalaxy.

Both interpretations require a certain revision of some ideas in current theories of cosmology, but we do not intend to enter here into the discussion of this question [7, 8].

REFERENCES

1. M. L. HUMASON, N. U. MAYALL and A. R. SANDAGE, ASTRON. J., **61,** 97, 1956;
   G. C. MCVITTIE, Fact and Theory in Cosmology, pp. 137, 144, 146. Eyre and Spottiswoode, London, 1961.
2. A. SCHLÜTER, Astron. J., **60,** 141, 1955.
3. G. DE VAUCOULEURS, Astron. J., **66,** 629, 1961.
4. F. ZWICKY, Handbuch der Physik, Vol. LIII, p. 398. Springer-Verlag, Berlin, 1959.
5. F. ZWICKY, Morphological Astronomy, p. 63, Table VIII. Springer-Verlag, Berlin, 1957,
6. F. ZWICKY, in Proc. Third Berkeley Symposium on Math. Stat. and Prob., Vol. III, p. 130, Table I, 1956.
7. O. HÉCKMANN, Astron. J., **66,** 599, 1961.
8. G. PAÁL, Astron. J., USSR, **39,** 911, 1962.

SCATTERING AMPLITUDE OF HIGH ENERGY KLEIN-GORDON AND DIRAC PARTICLES

By

T. TIETZ

DEPARTMENT OF THEORETICAL PHYSICS, UNIVERSITY OF ŁÓDŹ, ŁÓDŹ, POLAND

(Received 4. XII. 1963)

In this note we consider the scattering amplitude of high-energy KLEIN—GORDON and DIRAC particles. As known, the scattering amplitude of high-energy collisions gives information about the interactions of colliding particles, and about the reaction products they generate. We write the KLEIN—GORDON equation in the following form [1]:

$$\Delta\Psi(\vec{r}) + \left[k^2 - 2EV(\vec{r}) + V^2(\vec{r})\right]\Psi(\vec{r}) = 0. \tag{1}$$

Here E is the energy of the incident particle and $k = (E^2 - m^2)^{1/2}$ is the momentum. Eq. (1) has been written in relativistic units, $\hbar = c = 1$. In our treatment we shall assume that the particle is highly relativistic, $E \gg m$, and we shall neglect m as compared to E. In this case eq. (1) has the form

$$\Delta\Psi(\vec{r}) + \left[k^2 - 2kV(\vec{r}) + V^2(\vec{r})\right]\Psi(\vec{r}) = 0. \tag{2}$$

As the first step in the formulation of an integral equation we define the GREEN's function, $G(\vec{r}, \vec{r}')$, as a solution of the inhomogeneous wave equation $(\Delta + k^2)G(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}')$. We shall define $G(\vec{r}, \vec{r}')$ as the amplitude which corresponds to the radiation from a coherent source at $\vec{r}'$, so that our GREEN's function is

$$G(\vec{r}, \vec{r}') = -\frac{1}{4\pi}\,\frac{e^{ik|\vec{r}-\vec{r}'|}}{|\vec{r} - \vec{r}'|}\,. \tag{3}$$

The integral equation for $\psi(\vec{r})$ is given by

$$\Psi(\vec{r}) = e^{i\vec{k}\vec{r}} + \int G(\vec{r} - \vec{r}')\left[2kV\vec{r}') - V^2(\vec{r}')\right]\Psi(\vec{r}')\,d\vec{r}'\,. \tag{4}$$

In order to obtain the scattering amplitude $f(\vec{k};\vec{k})$ we consider the limiting case $r \to \infty$ and note that $|\vec{r} - \vec{r}'| \to r - \dfrac{\vec{r}'\,,\vec{r}}{r}$ for large $|\vec{r}| = r$. The scattering

amplitude is then given by

$$f(\vec{k}', \vec{k}) = -\frac{1}{4\pi}\int e^{-i\vec{k}'\vec{r}}\left[2k\,V(\vec{r}) - V^2(\vec{r})\right]\cdot\Psi(\vec{r})\,d\vec{r}\,. \tag{5}$$

If $\psi(\vec{r})$ is given one can calculate the scattering amplitude $f(\vec{k}', \vec{k})$. Unfortunately, the KLEIN—GORDON equation is exactly solvable only in a very limited number of cases, therefore a good approximation of $\Psi(\vec{r})$ is required for highly relativistic particles. To illustrate the method by which a good approximation may be obtained for $\psi(\vec{r})$ we write down the KLEIN—GORDON equation in one dimension:

$$\left(\frac{d^2}{dx^2} + k^2\right)\Psi(x) = \left(2k\,V(x) - V^2(x)\right)\Psi(x)\,. \tag{6}$$

The solution $\psi(x)$ we write in the following form: $\psi(x) = e^{ikx}\,\varphi(x)$, where $\varphi(x)$ is a function which varies slowly over a particle wave length. Substitution of $\psi(x)$ into eq. (6) gives us the differential equation for $\varphi(x)$

$$\left(2ik\frac{d}{dx} + \frac{d^2}{dx^2}\right)\varphi(x) = \left(2k\,V(x) - V^2(x)\right)\varphi(x)\,. \tag{7}$$

Since $\varphi(x)$ varies only slowly over a wave length we can drop $\dfrac{d^2}{dx^2}$. In this case we obtain for $\varphi(x)$ the approximation

$$\varphi(x) = e^{-i\int_{-\infty}^{x}[V(x') - V^2(x')/2k]\,dx'} \tag{8}$$

which fulfils the boundary condition $\varphi(-\infty) = 1$. Using the same method we can obtain an approximation for $\psi(x, y, z)$ in the case of three-dimensional problems, namely

$$\Psi(x, y, z) = e^{ikz - i\int_{-\infty}^{z}[V(x, y, z') - V^2(x, y, z')/2k]\,dz'}\,. \tag{9}$$

For a potential with azimuthal symmetry we obtain the following expression from the KLEIN—GORDON equation for the scattering amplitude $f(\vartheta)$ after simple calculations similar to those which have been made by GLAUBER [2] n the case of the Schrödinger equation:

$$f(\vartheta) = \frac{k}{2i}\int_0^{\infty} I_0\left(2k\,b\sin\frac{\vartheta}{2}\right)\left[e^{i\chi(\vec{b})} - 1\right] b\,d\,b\,, \tag{10}$$

where

$$\chi(\vec{b}) = -\int_{-\infty}^{+\infty}\left[V(\vec{b} + \hat{k}z) - V^2(\vec{b} + \hat{k}z)/2k\right]dz\,. \tag{11}$$

In eq. (10) I_0 denotes the zeroth order Bessel function, ϑ is the scattering angle and $b = |\vec{b}|$ is the impact parameter. In the last equation $\hat{k}$ is a unit vector, $|\hat{k}| = 1$. Any position vector $\vec{r}$ may be resolved into two components $\vec{r} = \vec{b} + \hat{k}z$ which means that $\vec{b}$ is a vector lying in a plane perpendicular to $\hat{k}$, and the positive z-axis lies in the direction of $\vec{k}$. The last two formulas are convenient for the calculation of the scattering amplitude of highly relativistic KLEIN—GORDON particles. The expression given by eq. (10) may be recognized as of the type of formula one would use to discuss the diffraction by a transparent obstacle.

PARZEN [3] has derived a formula which gives a relation between the scattering amplitude of KLEIN—GORDON $f(\vartheta)$ and that of DIRAC $f_D(\vartheta)$:

$$f_D = f(\vartheta) \cos^2\left(\frac{1}{2}\vartheta\right). \tag{12}$$

Since in our case $f(\vartheta)$ is given by equ. (10), one can calculate the scattering amplitude $f(\vartheta)$ in the DIRAC case for a potential with azimuthal symmetry. The method given in this note is better suited to practical calculations than the method given by the author previously [4].

REFERENCES

1. D. R. BATES, Quantum Theory, vol. 3, p. 44, Academic Press, New York and London, 1962.
2. R. J. GLAUBER, Lectures in Theoretical Physics, vol. I. p. 315, Interscience Publishers Inc. New York, 1959; for reference see also B. J. MALENKA, Phys. Rev., **95,** 522, 1956. and H. S. VOLK, Phys. Rev., **122,** 931, 1961.
3. G. PARZEN, Phys. Rev., **104,** 835, 1956.
4. T. TIETZ, Acta Phys. Hung., **11,** 85, 1960.

AVERAGE ENERGY, AVERAGE VELOCITY AND MEAN FREE PATH OF SLOW ELECTRONS IN HELIUM AND ARGON

By

O. J. ORIENT

DEPARTMENT OF ATOMIC PHYSICS, POLYTECHNICAL UNIVERSITY, BUDAPEST

(Received 28. I. 1964)

Introduction

Regarding the energy distribution of slow electrons first ALLEN [1] and then BARBIERE [2] did computations.

ALLEN's work relates to helium, neon and argon. In his computations ALLEN gave attention to the fact that, as resulting from RAMSAUER's measurements [3], the cross section of gases related to electrons is a function of energy. Using the energy distributions computed for the three rare gases mentioned above ALLEN determined drift velocity and temperature values for various electric field strengths. However, there are discrepancies between the results of his computations and those of measurements.

From BARBIERE's train of thought it follows that ALLEN has failed to notice some important points. For helium and argon BARBIERE has determined the distribution of electrons from $\frac{X}{p} x = 1$ to $\frac{X}{p} = 4 \frac{V}{\text{cm} \cdot \text{mm Hg}}$ (X is the electric field strength, p is the pressure). In this computation the following points were taken into consideration: *a)* as the mass of gas molecules is finite, the effect of the mass of the molecules cannot be neglected; *b)* inelastic collision taking place in the course of the excitation of the molecules should be taken into account; *c)* the scattering cross section (RAMSAUER) is a function of energy; *d)* the substitution of diffusion cross section (momentum-transfer cross section) for RAMSAUER's scattering cross section gives results of a better accuracy.

BARBIERE based his calculations on HOLSTEIN's [4] results who derived the differential equation defining the energy distribution in an electric field between plane-parallel electrodes, taking into account conditions *a)*, *b)*, *c)*, *d)*. From the differential equation BARBIERE determined the energy distribution function as regards inelastic collisions, taking only the first excitation energy into consideration. Electron drift velocity and electron temperatures computed by using the energy distribution were in fair agreement with the results of measurement (in the electric field strength range from $\frac{X}{p} = 1$ to $\frac{X}{p} = 4$ [2]).

ALLEN's results are in better agreement with the results of measurements in the case of helium, than in that of argon. In the case of argon the discrepancy

is large as here the first excitation energy is low ($E_1 = 11.57$ eV). For helium this value is higher ($E_1 = 19.7$ eV). In the case of argon the first excitation energy greatly influences the energy distribution of the electrons.

BARBIERE has computed the energy distribution function of electrons $F(E)dE$, (E is the energy) in the case of helium and argon, for four values of the electric field strength $\left(\frac{X}{p} = 1,\ 2,\ 3,\ 4\right)$ representing the distribution functions by graphical methods. The author of the paper presented here has made use of these curves to determine average energy $\bar{E}$, average velocity $\bar{v}$ and mean free path $\bar{\lambda}$ for helium and argon, for the four electric field strengths.

Results

If the energy distribution is known the average energy may be computed from the formula:

$$\bar{E} = \frac{\int_0^\infty E \cdot F(E)\, dE}{\int_0^\infty F(E)\, dE}$$

and with this the root-mean-square velocity may be computed from the formula:

$$\sqrt{\bar{v}^2} = \sqrt{\frac{2\bar{E}}{m}}$$

(m is the mass of the electron).

The average velocity $\bar{v}$ may be obtained from the energy distribution as:

$$\bar{v} = \frac{\int_0^\infty v^2 F\left(\frac{1}{2} m v^2\right) dv}{\int_0^\infty v F\left(\frac{1}{2} m v^2\right) dv}.$$

To compute the mean free path the diffusion (momentum-transfer) free path as a function of energy must be known. The relationship between diffusion cross section, Q_D, and diffusion free path, λ_D, is $\lambda_D = \frac{1}{Q_D}$. The dimension of

λ_D is cm, in case of 1 mm Hg pressure and 0° C temperature. The dimension of Q_D is cm^2/cm^3 in case of 1 mm Hg pressure and 0° C temperature. In his paper BARBIERE gives the diffusion cross section for helium and argon on the bases of RAMSAUER's measurements [3] for energies above 1 eV. From these data values of λ_D for energies higher than 1 eV have been obtained (Fig. 1).

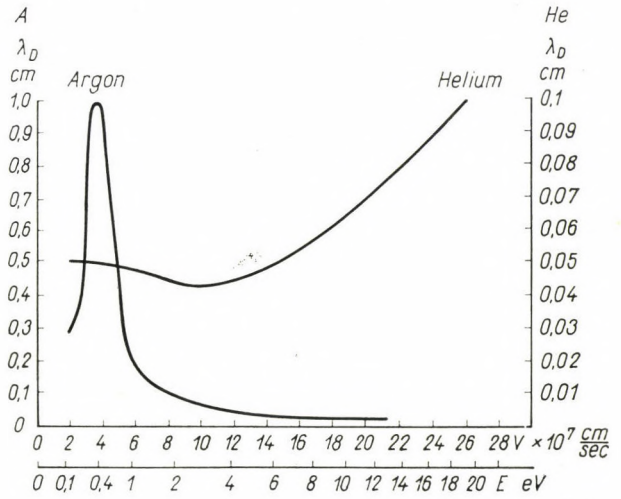


Fig. 1

For energies lower than 1 eV, λ_D has been determined as follows: In case of such low energies the difference between diffusion cross section and total cross section (RAMSAUER cross section, Q_R) may be neglected [5], i.e. $Q_D = Q_R$. Considering this $\frac{1}{Q_R}$ has been accepted as λ_D. The values of Q_R for helium and argon have been taken from results of measurements published in the literature [6]. Fig. 1 represents the dependence of λ_D on energy for helium and argon starting from 0,1 eV.

By using the diffusion free path, the mean free path is defined as

$$\lambda = \frac{\int_0^\infty \lambda_D(E)\, F(E)\, dE}{\int_0^\infty F(E)\, dE} .$$

The values of $\bar{E}$, $\sqrt{\bar{v}^2}$, $\bar{v}$ and $\bar{\lambda}$ computed for helium are listed in Table 1, for argon in Table 2.

Table 1

$\frac{X}{p}$ $\frac{V}{\text{cm} \cdot \text{mm Hg}}$	$\bar{E}$ eV	$\sqrt{\bar{v}^2}$ $\frac{\text{cm}}{\text{sec}}$	$\bar{v}$ $\frac{\text{cm}}{\text{sec}}$	$\bar{\lambda}$ cm
1	1,522	$7{,}319 \cdot 10^7$	$6{,}858 \cdot 10^7$	$4{,}553 \cdot 10^{-2}$
2	3,688	$11{,}377 \cdot 10^7$	$10{,}581 \cdot 10^7$	$4{,}657 \cdot 10^{-2}$
3	5,507	$13{,}921 \cdot 10^7$	$13{,}237 \cdot 10^7$	$5{,}143 \cdot 10^{-2}$
4	6,708	$15{,}347 \cdot 10^7$	$14{,}824 \cdot 10^7$	$5{,}633 \cdot 10^{-2}$

Table 2

$\frac{X}{p}$ $\frac{V}{\text{cm} \cdot \text{mm Hg}}$	$\bar{E}$ eV	$\sqrt{\bar{v}^2}$ $\frac{\text{cm}}{\text{sec}}$	$\bar{v}$ $\frac{\text{cm}}{\text{sec}}$	$\bar{\lambda}$ cm
1	3,871	$11{,}670 \cdot 10^7$	$10{,}597 \cdot 10^7$	$9{,}899 \cdot 10^{-2}$
2	4,864	$13{,}082 \cdot 10^7$	$12{,}435 \cdot 10^7$	$8{,}045 \cdot 10^{-2}$
3	5,052	$13{,}330 \cdot 10^7$	$12{,}789 \cdot 10^7$	$8{,}032 \cdot 10^{-2}$
4	5,115	$13{,}416 \cdot 10^7$	$13{,}068 \cdot 10^7$	$7{,}788 \cdot 10^{-2}$

Integrals occurring in the formulae have been determined by graphical methods. The error in the computations can safely be estimated as being lower than 3 percent in any of the data.

Conclusions

For values of $\bar{E}$ and $\bar{v}$ data may be found in the literature. TOWNSEND and BAILEY [7] have calculated these values from the drift velocity of electrons and from the measurement of the ratio of diffusion constant to drift velocity. They assumed Maxwell velocity distribution. These values are also found in LOEB's book [8]. To make a comparison possible, data found in the literature and those computed by us have been represented in the same figure.

Values of $\bar{E}$ for helium as functions of the field strength are shown in Fig. 2. Curve *a* represents data given by TOWNSEND and BAILEY, curve *b* shows our own results. Fig. 3 represents values of $\bar{\lambda}$: curve *a* those of TOWNSEND and BAILEY, curve *b* ours. Values $\bar{E}$ for argon are given in Fig. 4, curve *a* TOWNSEND's and BAILEY's data, curve *b* ours. Similarly, Fig. 5 shows the two sets of $\bar{\lambda}$ values.

In the case of helium no significant difference is found between the two values of the average energy. This may be explained by Q_D and λ_D not being strongly dependent on the energy, and the first excitation energy being relatively high (19,7 eV). Owing to this the distribution does not differ greatly

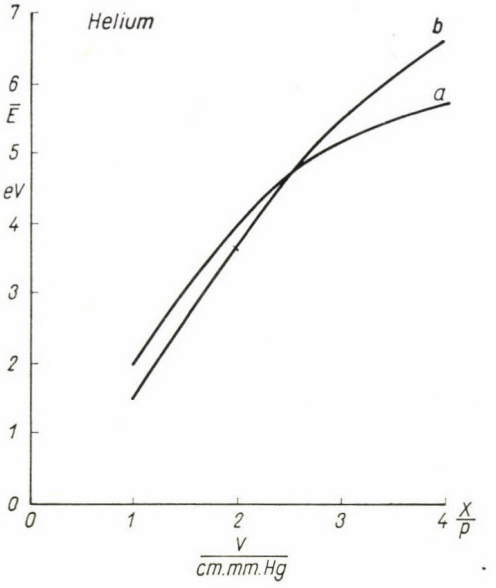


Fig. 2

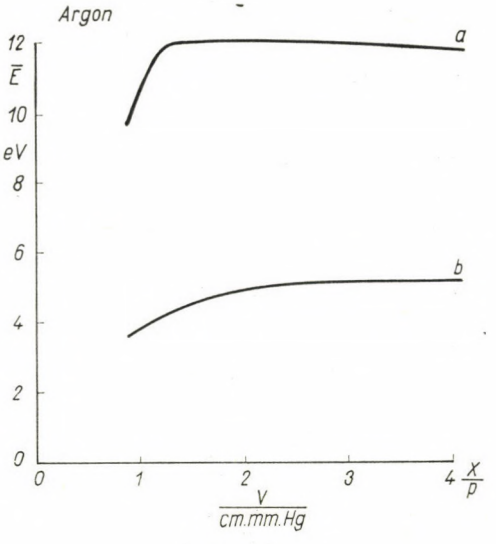


Fig. 3

from the Maxwell distribution. The largest discrepancy between the values of $\bar{\lambda}$ is 10 per cent. The values found in our work show that λ_D is not a linear function of the electric field strength. This is due to the dependence of λ_D on the energy.

In the case of argon average energies show very large deviations. This is due to Q_D (i.e. also λ_D) greatly changing with the energy of the electron, and the first excitation energy being relatively low (11,57 eV). These circumstances influence the distribution function to a great extent and make it differ considerably from the Maxwell distribution. According to TOWNSEND

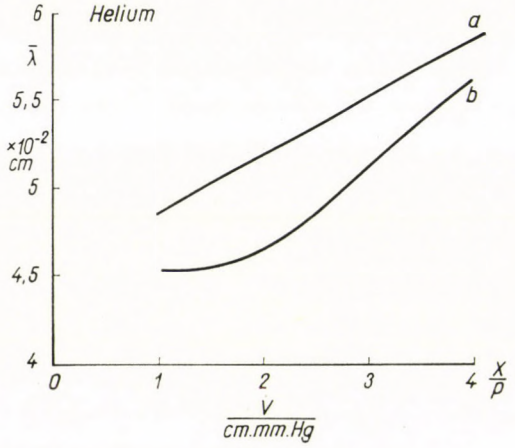


Fig. 4

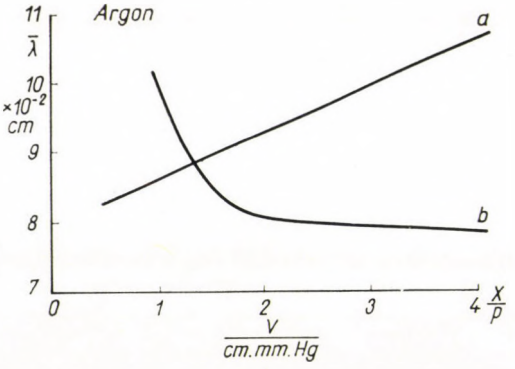


Fig. 5

and BAILEY above a field strength of $\frac{x}{p} = 1$ the average energy is higher than 11,57 eV i.e. higher than the first excitation energy. No explanation could be found for this. The deviation is high also for $\bar{\lambda}$. The increase of the values of $\bar{\lambda}$ cannot be explained by the increase of the field strength. λ_D decreases as the energy increases, owing to which the value found in our computation also decreases with increasing field strength.

Further references regarding the average energy of electrons in argon may be found in the literature [9], [10]. According to these, in case of $\frac{X}{p} = 0.95 \frac{V}{\text{cm} \cdot \text{mm Hg}}$ average energy, $\bar{E} = 8$ eV. Considering BARBIERE's starting assumptions and computation results, these values also are unjustifiably high.

REFERENCES

1. HARRIET W. ALLEN, Phys. Rev., **52,** 707, 1937.
2. D. BARBIERE, Phys. Rev., **84,** 653, 1951.
3. C. RAMSAUER and R. KOLLATH, Ann. d. Phys., **12,** 529, 1932.
4. T. HOLSTEIN, Phys. Rev., **70,** 367, 1946.
5. H. S. W. MASSEY and E. H. S. BURHOP, Electronic and Ionic Impact Phenomena, 1956, p. 15.
6. SANBORN C. BROWN, Basic Data of Plasma Physics, J. Wiley and Sons, Inc. New York, 1959, pp. 5 and 19.
7. J. TOWNSEND and V. A. BAILEY, Phil. Mag., **46,** 657, 1923.
8. L. B. LOEB, Basic Processes of Gaseous Electronics, University of California Press, Berkeley, 1955, p. 328.
9. J. K. THEOBALD, Journ. Appl. Phys., **24,** 123, 1953.
10. R. H. HEARLEY and J. W. REED, The Behaviour of Slow Electrons in Gases, 1941, p. 92.

THE "CLOCK-PARADOX" AND THE QUANTUM-MECHANICAL THEORY OF BIOLOGICAL AGEING

By

M. SÜVEGES

RESEARCH GROUP FOR THEORETICAL PHYSICS OF THE HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST

(Received 17. II. 1964)

There has been controversy among theoretical physicist whether the "clock paradox" [1] is really a consequence of relativity theory. More recently, however, CRAWFORD [2], FREMLIN [3] and others pointed out that experimental proof of this effect can be pieced together from experiments that had already been performed. FREMLIN, in particular, uses experiments involving mesons. He considers two systems of mesons each consisting initially of the same number of mesons. If now one of the systems is making a round-trip in an accelerator the different decay rates of the two systems will result in different numbers of final mesons. If one defines elapsed time by means of the number of mesons, then clearly an asymmetry in elapsed times is found. Obviously, in proving the clock paradox the essential point is the definition of a clock that measures elapsed time, as opposed to the rate of change of time. In the above experiment the number of mesons serves as a "clock" for its measurement and the purpose of the present note is to point out that a similar definition for the measurement of elapsed time seems to be possible also for biological systems.

Obviously, one might try to define a clock for a biological system in a way similar to the definition in the case of mesons. One could, for example, measure the C^{14} content of the system and define its elapsed time by means of the number of C^{14} atoms in the system. However, in connection with biological systems, the definition should fulfil the further requirement of indicating elapsed time also in a biological sense, i.e. the definition intrinsically should imply that if one of the systems had longer elapsed time it was nearer to biological death. Obviously, the definition by means of C^{14} would not, a priori, satisfy this requirement. Recently PER—OLOV LÖWDIN [4] putforward his hypothesis concerning the quantum-mechanical basis of biological ageing and we want to point out that it allows us to define a "clock" for biological systems in such a way that it satisfies also the above requirement.

It has been suggested (see e.g. YOCKEY [5]), that biological ageing is the result of accumulated errors in the genetic code in the transmission of genetic information and LÖWDIN's hypothesis is concerned with the origin of these errors.

According to LÖWDIN's quantum-mechanical analysis of the genetic code in the WATSON—CRICK model of DNA, pure genetic message does not exist, but there is always a finite probability error built into even the ground state of the DNA molecule. These errors will accumulate and, in the replication of DNA, biologically amplified. The reason for the errors is this. The hydrogen bonds between base-pairs play a fundamental role in the complementarity concept developed by WATSON and CRICK in the stereomodel of DNA. In DNA replication the sequence of bases in the new strand of the double helix is determined by the (complementary) configurations of those protons in the original chain which will then form the hydrogen bonds between base-pairs and the problem of the stability of the genetic code is hence concerned with the question of the motion and stability of protons in the hydrogen bonds. The problem of the motion of a proton in the hydrogen bond is, in a first rough approximation, a one-particle problem involving a fixed two-center potential. Each electron pair of the hydrogen bond attracts the proton and this attraction may be represented by a double-well potential. In this classical approximation one would consider only processes involving sufficient energy to take the proton above the potential barrier and this approximation would therefore correspond to the "pure", or stable, genetic code of the WATSON—CRICK model. However, the quantum-mechanical analysis of the wave packet associated with the proton shows that, even in the ground state of the molecule, the genetic message can never be entirely "pure", since the wave function is such that there is always a finite probability of proton tunnelling into the other well. As a consequence even in the ground state of DNA there is a small probability that some of the bases are in their unusual tautomeric forms. If the hydrogen bonds get released in this position, the rare tautomeric forms will lead to errors in the next replication, i.e. to somatic mutations. Since this error may cause a certain amount of loss of genetic information in each DNA replication, P. O. LÖWDIN puts forward the hypothesis that the phenomenon of biological ageing depends primarily on the accumulation of such loss of genetic information. In other words the cause of biological ageing might be the quantum-mechanical tunnel effect.

To return to the question of the definition of a biological clock it now seems that the LÖWDIN hypothesis allows of a definition which is completely analogous to that in the case of mesons and, further, which satisfies the above requirement concerning biological systems. One has only to define as "clock", measuring elapsed time, the number of somatic mutations in the system caused by the tunnel effect.

This definition has some advantages. First, it provides a direct relationship between the relativistic transformation properties of time and the mechanism of biological ageing. Since biological ageing is supposed to depend on the strictly time-dependent tunnel effect, it follows from the relativistic trans-

formation properties of time that the number of accumulated errors will be different for systems following different world lines. In particular, it will be smaller for the system making the round-trip and, consequently, it will be younger also in the biological sense of the word. This is in agreement with what would be expected from quite general relativistic considerations. Second, a direct experimental verification of the asymmetrical ageing of biological systems does not at present seem to be possible because of the small velocities, relative to that of light, to which biological systems can be accelerated. However, should the LÖWDIN hypothesis be confirmed on other grounds the asymmetric ageing effect of relativity could then be taken as experimentally proved also for biological systems, since "biological time" would then be identical with "physical time" and ageing would then depend on the tunnel effect.

In conclusion it must be said that the biological problem of ageing seems to be rather involved and we do not want to discuss this here but only to point out the relativistic consequences of the LÖWDIN hypothesis and its relevance to the definition of a biological "clock".

The author expresses his gratitude to Prof. G. MARX, Dr. J. LADIK and T. SZONDY for valuable discussions.

REFERENCES

1. C. MØLLER, The Theory of Relativity, p. 258, Oxford, University Press, London, 1952.
2. F. S. CRAWFORD Jr., Nature, **179,** 35, 1957.
3. J. H. FREMLIN, Proc. Phys. Soc., **80,** 1384, 1962.
4. P. O. LÖWDIN, Rev. Mod. Phys., **35,** 724, 1963.
5. H. P. YOCKEY, Information Theory in Biology, ed. by H. P. YOCKEY, pp. 50, 297, Pergamon Press, 1958.

A COMMENT ON: A NEW METHOD FOR FINDING THE PHASE SHIFTS FOR THE SCHRÖDINGER EQUATION

By

JOHN W. SHELDON

LEWIS RESEARCH CENTER
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
CLEVELAND, OHIO, USA

(Received 29. VI. 1964)

DR. TIETZ in his communication entitled "A New Method for Finding the Phase Shifts for the Schrödinger Equation"[1] has obtained phase shifts by writing the Schrödinger equation as two first order differential equations and obtaining asymptotic solutions for the two new dependent variables. His result is reported in terms of a difference between phase shifts:[2]

$$\eta_{l-1} - \eta_{l+1} = \frac{l+\frac{1}{2}}{k^2}\pi\int_0^\infty \frac{dU}{dr}\left[I_{l+\frac{1}{2}}(kr)\right]^2 dr. \qquad (1)^3$$

Integration of eq. (1) by parts yields

$$\eta_{l-1} - \eta_{l+1} = -\int_0^\infty \frac{2\pi\left(l+\frac{1}{2}\right)}{k} I_{l+\frac{1}{2}}(kr)\, I_{l-\frac{1}{2}}(kr)\, U(r)\, dr + \int_0^\infty \frac{2\pi\left(l+\frac{1}{2}\right)^2}{k^2 r}\left[I_{l+\frac{1}{2}}(kr)\right]^2 U(r)\, dr. \qquad (2)$$

Using the Bessel function identity

$$\left[I_{l+\frac{1}{2}}(kr)\right]^2 - \left[I_{l-\frac{1}{2}}(kr)\right]^2 = \frac{4\left(l+\frac{1}{2}\right)^2}{(kr)^2}\left[I_{l+\frac{1}{2}}(kr)\right]^2 - \frac{4\left(l+\frac{1}{2}\right)}{kr} I_{l+\frac{1}{2}}(kr)\, I_{l-\frac{1}{2}}(kr) \qquad (3)$$

[1] T. TIETZ, Acta Phys. Acad. Sci. Hung. 16, 289-292 (1963). A new method for finding the phase shifts for the Schrödinger equation.
[2] The nomenclature used here is that of DR. TIETZ.
[3] Eq. (16) of 1.

we get

$$\eta_{l-1} - \eta_{l+1} = -\frac{\pi}{2}\int_0^\infty \left[I_{l-\frac{1}{2}}(kr)\right]^2 U(r) r\, dr + \frac{\pi}{2}\int_0^\infty \left[I_{l+\frac{1}{2}}(kr)\right]^2 U(r) r\, dr \qquad (4)$$

which is recognized as the difference in the phases η_{l-1} and η_{l+2} in the Born approximation.[4]

[4] ALBERT MESSIAH, Quantum Mechanics, Vol. I, North Holland Publishing Company, Amsterdam, 1961.

RECENSIO

HERBERT E. SALZER, NORMAN LEVINE

Table of Sines and Cosines to Ten Decimal Places at Thousandths of a Degree

Pergamon Press, New York—Oxford—London—Paris 1962. XIV + 920 pp. $ 10.

As the title fully conveys the contents of the Tables, we may restrict ourselves to some remarks concerning its practical applicability.

The use of the Tables is considerably facilitated by the detailed discussion of the interpolation methods given on the first few pages of the book. This discussion involves detailed error estimates for direct and inverse interpolation showing that linear interpolation formulae are sufficient throughout the whole book, as more elaborate formulae lead to results, the errors of which are of the same order of magnitude as those of the linear ones.

The arrangement of the entries is highly practical. The values of sine and cosine belonging to the same argument are tabulated side by side, making it easy to compute values of tan x and cot x. Each entry gives all digits, saving the work of looking for the first few digits at other parts of the page.

T. SZONDY

Printed in Hungary

A iadásért felel az Akadémiai Kiadó igazgatója — Műszaki szerkesztő: Farkas Sándor

A kézirat nyomdába érkezett: 1964. VI. 11. — Terjedelem: 11,50 (A/5) ív, 46 ábra

64.59046 Akadémiai Nyomda, Budapest. — Felelős vezető: Bernát György

The *Acta Physica* publish papers on physics, in English, German, French and Russian.
The *Acta Physica* appear in parts of varying size, making up volumes.
Manuscripts should be addressed to:

Acta Physica, Budapest 502, Postafiók 24.

Correspondence with the editors and publishers should be sent to the same address.
The rate of subscription to the *Acta Physica* is 110 forints a volume. Orders may be placed with "Kultúra" Foreign Trade Company for Books and Newspapers (Budapest I., Fő u. 32. Account No. 43-790-057-181) or with representatives abroad.

Les *Acta Physica* paraissent en français, allemand, anglais et russe et publient des travaux du domaine de la physique.
Les *Acta Physica* sont publiés sous forme de fascicules qui seront réunis en volumes.
On est prié d'envoyer les manuscrits destinés à la rédaction à l'adresse suivante:

Acta Physica, Budapest 502, Postafiók 24.

Toute correspondance doit être envoyée à cette même andresse.
Le prix de l'abonnement est de 110 forints par volume.
On peut s'abonner à l'Entreprise du Commerce Extérieur de Livres et Journaux «Kultúra» (Budapest I., Fő u. 32. — Compte-couraut No. 43-790-057-181) ou à l'étranger chez tous les représentants ou dépositaires.

«*Acta Physica*» публикют трактаты из области физических наук на русском, немецком, английском и французском языках.
«*Acta Physica*» выходят отдельными выпусками разного объема. Несколько выпусков составляют один том.
Предназначенные для публикации рукописи следует направлять по адресу:

Acta Physica, Budapest 502, Postafiók 24.

По этому же адресу направлять всякую корреспонденцию для редакции и администрации.
Подписная цена «*Acta Physica*» — 110 форинтов за том. Заказы принимает предприятие по внешней торговле книг и газет «*Kultúra*» (Budapest I., Fő u. 32. Текущий счет: № 43-790-057-181) или его заграничные представительства и уполномоченные.

33,— Ft

INDEX

COMMUNICATIONES BREVES

RECENSIO

Acta Phys. Hung. Tom. XVII. Fasc. 3., Budapest, 20. IX. 1964

Index: 26.022

ACTA PHYSICA

ACADEMIAE SCIENTIARUM HUNGARICAE

TOMUS XVII FASCICULUS 4

AKADÉMIAI KIADÓ, BUDAPEST
1964

ACTA PHYS. HUNG.

ACTA PHYSICA

A MAGYAR TUDOMÁNYOS AKADÉMIA FIZIKAI KÖZLEMÉNYEI

SZERKESZTŐSÉG ÉS KIADÓHIVATAL: BUDAPEST V., ALKOTMÁNY UTCA 21.

Az *Acta Physica* német, angol, francia és orosz nyelven közöl értekezéseket a fizika tárgyköréből.

Az *Acta Physica* változó terjedelmű füzetekben jelenik meg: több füzet alkot egy kötetet.

A közlésre szánt kéziratok a következő címre küldendők:

Acta Physica, Budapest 502, Postafiók 24.

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Az *Acta Physica* előfizetési ára kötetenként belföldre 80 forint, külföldre 110 forint. Megrendelhető a belföld számára az Akadémiai Kiadónál (Budapest V., Alkotmány utca 21. Bankszámla 05-915-111-46), a külföld számára pedig a „Kultúra" Könyv- és Hírlap Külkereskedelmi Vállalatnál (Budapest I., Fő u. 32. Bankszámla 43-790-057-181 sz.), vagy annak külföldi képviseleteinél és bizományosainál.

Die *Acta Physica* veröffentlichen Abhandlungen aus dem Bereiche der Physik in deutscher, englischer, französischer und russischer Sprache.

Die *Acta Physica* erscheinen in Heften wechselnden Umfanges. Mehrere Hefte bilden einen Band.

Die zur Veröffentlichung bestimmten Manuskripte sind an folgende Adresse zu richten:

Acta Physica, Budapest 502, Postafiók 24.

An die gleiche Anschrift ist auch jede für die Redaktion und den Verlag bestimmte Korrespondenz zu senden.

Abonnementspreis pro Band: 110 Forint. Bestellbar bei dem Buch- und Zeitungs-Aussenhandels-Unternehmen »Kultúra« (Budapest I., Fő u. 32. Bankkonto Nr. 43-790-057-181) oder bei seinen Auslandsvertretungen und Kommissionären.

ON THE RELATIONS BETWEEN SOME OF THE PARAMETERS OF THE DIRECT CURRENT MERCURY VAPOUR DISCHARGES

By

J. Bitó

INDUSTRIAL RESEARCH INSTITUTE FOR TELECOMMUNICATION TECHNIQUE, BUDAPEST

(Presented by G. Szigeti, — Received 30. V. 1964)

By the application of the probe measurement method of Langmuir the author was able to determine the dependence of the axial gradient and the electron temperature of direct current discharges on the discharge current, in the case of heated and unheated cathodes. With the help of the Nölle method of evaluation, the potential distribution of the discharge was determined through the measurement of the dependence of the cathode dark spaces on the current and through the measurement of the plasma potentials. Further, in the present paper the dependence of the cathode and anode fall on the current is discussed. The phenomenon of the axial cataphorasis is then examined and the connection between the process of cataphorasis and the change in time of the electron temperature is shown.

1. Introduction

Much research preceded the present achievements and methods of plasma physics. The subject of these investigations has been connected in the first place with the plasma of the simplest type of discharge, the positive column of the direct current discharge. Through the knowledge acquired by the investigation of direct current discharges it has become possible to apply these widely to-day both for scientific and technical purposes.

The physical processes going on in discharges of this type and their micro-characteristics are being further investigated also at the present time. Thus among others important researches are being carried on into the discharge losses, and the radiation and ionization effects of discharges. For these investigations several measuring methods may be considered. Of these the probe measurement methods worked out by Langmuir [1—3] are applied most extensively. The advantage of the Langmuir probe measurement method is that, if attention is paid to certain points, the measurement will disturb the discharge plasma to only a small degree and will not cause significant changes in it, making it thus possible to determine the fundamental plasma characteristics which otherwise cannot be measured externally without interference; such are the temperature T_e of the electrons, the electron concentration n_e, the plasma potential V_p of the positive column at the point in question, as well as the potential gradient E of the positive column.

If these are known it is possible to calculate further, rather important micro- and macro-parameters, thus the excitation probabilities of the individual spectral lines, the characteristics of the volumetric and wall recombination processes, diffusion and ionization processes and the conditions under which the phenomena around the electrodes take their course.

In the present article the relations between some of the parameters of the direct current mercury vapour discharges are discussed. Thus the dependence of the cathode fall and the anode fall on the current is determined. Further by extending the limits of the current range in which the micro-characteristics of the positive column plasma were formerly investigated it is shown that in this extended range some characteristics of the discharge take up extreme values. The part played by the different factors in the forming of a positive or negative characteristic of the discharge are discussed and in connection with the explanation of the cataphorasis appearing in the direct current discharge, the part played by the change of the electron temperature is pointed out.

2. Measuring method

In the present investigations — as was mentioned above — the probe measurement method [1—3] developed by LANGMUIR has been used. The essence of this method is that when a conductive electrode of small surface and of a potential that differs from that of one of the discharge electrodes is placed in a discharge, then the current flowing on this small-surface electrode — further called probe — will be characteristic of the investigated space of the discharge. If different voltages are applied to the probe, the current taken up from the discharge and flowing on the probe may be plotted as a function of the voltage [1—3]. This is called probe characteristics. Such a probe characteristic is shown in Fig. 1, where the probe current J_p was plotted as a function of the probe voltage V_p. LANGMUIR worked out not only the experimental method for the probe measurements, but jointly with MOTT—SMITH he investigated also theoretically [2] the physical process taking place.

According to his results the probe measurement curve shown in Fig. 1 may be interpreted as follows: In case of a negative probe as against the discharge space, the probe will take up the positive ions from the discharge space (Fig. 1, section $\overline{AB}$ "negative probe current"), then with the growing probe potential (passing through the current-free point E) more and more electrons will reach the probe (section $\overline{EC}$), until this will have reached the potential of the discharge space (C). Continued increase of the probe voltage will not change the electron current till the beginning of ionization before the probe (point F).

As will be explained later the plasma potential V_g can be determined from the position of point C of the probe characteristic described above, while the electron concentration and the electron temperature can be determined from the slope of the section $\overline{EC}$.

The following fundamental stipulations must be made for the method to be applicable [4]:

1. The surface of the probe should be so small as not to disturb significantly the equilibrium and the original condition of the discharge space to be investigated.

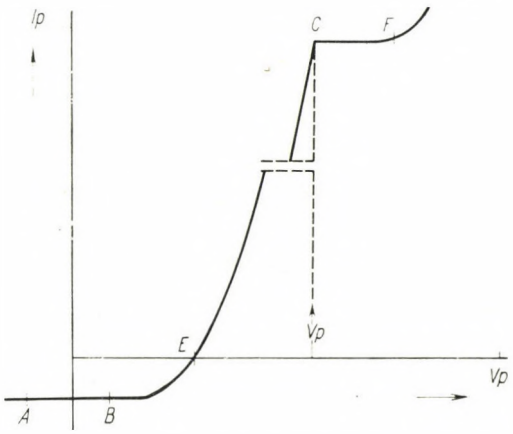


Fig. 1. Typical shape of probe characteristics in the LANGMUIR-probe measuring method

2. The velocity of the electrons and ions should follow the Maxwell—Boltzmann distribution.

3. In the discharge the number of carriers of negative and positive charges should be near equal (quasineutral plasma).

4. The negative charge carriers should be electrons only, other negative ions should not appear, at least not in a significant concentration.

5. The current carrying the positive charges should be made up of positive ions only, that is relative to the singly ionized atoms or molecules the number of the multiple ionized ones should not be significant.

6. During the measurement the work function of the probe should not change and its surface should not become contaminated.

The lack of any of the measuring conditions listed here or their not being satisfied to the required extent becomes also apparent from the shape of the probe characteristics which will then differ markedly from that of the ideal one [4—6].

It is usual to plot the probe characteristic in a semi-logarithmic scale and to base its evaluation on the Maxwell—Boltzmann distribution function taking the initial conditions into account [4].

The electric circuit employed in the course of the tests can be seen in Fig. 2, in the form of a block diagram. The discharge tube T has been fed by the stabilized direct current voltage supply source E with the symmetrically arranged ohmic resistances R_1, R_2 acting as restricting elements, the current

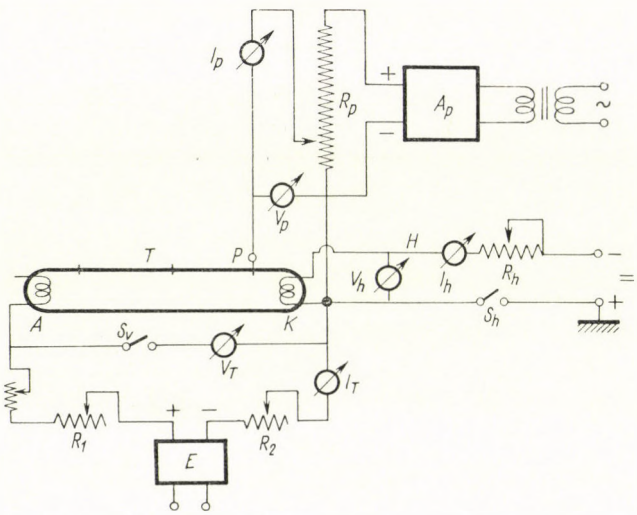


Fig. 2. Block diagram of the electrical arrangement applied in direct current probe measurements

and voltage of the discharge have been shown by the instruments I_T and V_T, respectively. In those experiments where the cathode K obtained separate external heating, the heating circuit H has been inserted, the heating current and heating voltage in which was shown by the instruments I_h and V_h, respectively; the heating current could be adjusted with the help of the resistance R_h.

In the investigations the probe circuit was developed with respect to the cathode. The voltage relative to the cathode on the probe P could be adjusted from the direct current source A through the variation of R_p. The current taken up by the probe from the discharge, as well as the voltage obtained by the probe could be read on the instruments I_p and V_p, respectively.

3. Test conditions

In the tests a cylindric discharge tube, length 1200 mm, external diameter 38 mm, made of glass of 1 mm wall thickness has been employed. At the ends of the tube two electrodes were placed which were identical both as

regards construction and arrangement. The electrodes consisted of a wolfram double-spiral provided with electron emission coating and of protecting auxiliary electrodes placed next to the electrodes, as well as of the corresponding electric terminal and support. During the test the auxiliary electrodes have been on a potential identical with that of the spiral.

Three probes were soldered in the wall of the discharge vessel; the two extreme ones were each at a distance of 200 mm from the ends of the tube, while the third was in the centre of the discharge, in the midpoint of the section between the two extreme probes.

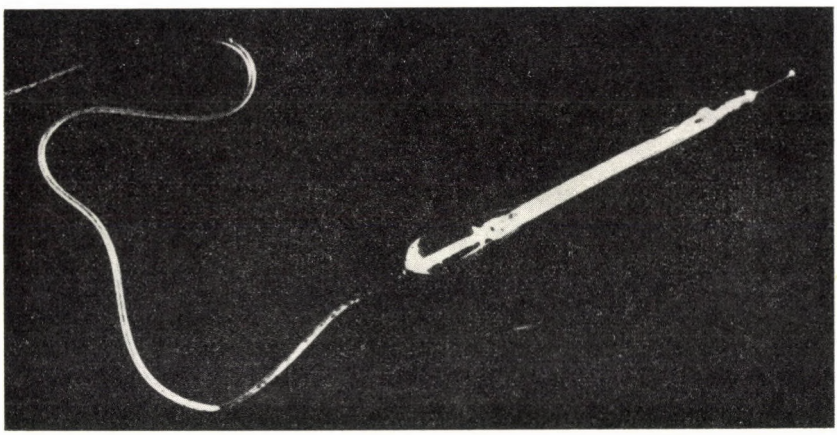


Fig. 3. Characteristic shape of the probe applied in the measurements

The probes reached radially up to the axis of the discharge. A nickel wire of 2 mm length, 0,95 mm diameter, formed their active part from the point of view of the discharge. In this way the surface of the casing of the probes has been larger by oner order of magnitude than the surface of the free base of the probe facing the discharge.

The employed probes were made with lead glass insulation, their shape is shown in the photograph in Fig. 3.

Following the vacuum technical treatment of the discharge tube developed in this way, the tube was filled with mercury of some 60 mg weight p. a. and 3 mmHg pressure, FeBa arc purified argon gas.

The measurements have been carried out at room temperature, between 22,5—24° C with the discharge tube placed in air surroundings. In the measurements of identical character (i.e. when the results were to be reproducible) the temperature of the ambient air differed at most by $\pm 0{,}5°$ C.

The measurements were always preceded by a surface cleaning of the probe through electron bombardment, in order to remove the contaminations and to keep the work function at a constant value.

According to experience [4] it appeared expedient to operate the discharge tube with symmetrically inserted limiting elements. It was possible thereby to reduce favourably the damaging oscillations appearing in the tube, and the noise level.

The heating cycle H developed at the cathode K has been fed by a direct current voltage, the positive pole of which has been on an earth potential the arc current finding its outlet at this point of the cathode. The measurement of the voltage V_T of the tube was effected relative to this point, and the probe cycle in the heated cathode measurements was developed also in respect of this point.

In those measurements where the cathode has not been separately heated the development of the probe cycle and also the probe measurement have been effected compared to the anode, for reasons of measuring technique. This did not mean any deviation from the previous series of measurements as regards the measuring results.

When measuring the probe characteristic the probe voltage could be adjusted to the desired value with the help of the resistance R_p. By exchanging the poles of the applied direct current voltage source A_p the probe could be put on lower and higher potentials as compared to the given electrode. For the determination of both the plasma potentials and the electron temperatures it is the electron current section of the probe characteristic that has to be investigated, consequently from the point of view of the tests the development of the saturation section of the ion current is indifferent. Accordingly, it was not necessary to employ a much more negative probe as compared to the discharge space.

In the course of the pre-test cleaning of the probe surfaces investigations have been carried out also concerning the cleaning effect of ion bombardment. This could be achieved through the adjustment of section $\overline{AB}$ of the probe characteristic shown in Fig. 1. A cleaning of this character proved to be, on the one hand, rather lengthy and on the other hand not so successful as the surface cleaning effected by electron current. This was shown by the flattening out of elbow C of the ideal probe characteristic represented in Fig. 1, and from the appearance of further elbows in the case of the recorded characteristics.

4. Results of measurement

In the first part of the investigations the object had been the determination of the characteristics of the heated cathode direct current discharges, by the probe measuring method briefly explained above under the described test conditions.

By measuring the plasma potentials on the individual probes, and with the known distances of the probes from the discharge, it was possible to calculate the potential gradient of the positive column of the individual discharge currents. When determining the potential gradient it is tacitly assumed that the potential is linear along the positive column of the plasma, consequently that there are no such regions in which the value of the electric field strength is higher or lower than that determined on the average. This is not so in reality, since on account of the stationary and moving striations [7] accompanying the discharges the parameters which could be considered constant otherwise, change along the axis of the positive column. In the case of field strengths, however, this change [7] can be neglected in order of magnitude as against the existing field strengths. The same may be said in good approximation of the electron temperatures and electron concentration as well, if the conditions along the axis of the discharge are considered. In radial direction, because of the wall diffusion current, the same cannot be said, here the change of electric field strength, electron temperature and electron concentration must be described by a function which is not of linear shape. The measurements showed that the electric field strength, the electron temperature and the electron concentration are indeed constant — as stated by LANGMUIR [1—3] — along the axis of the discharge in the positive column which may be considered as homogeneous and in a state of stationary equilibrium. This result may also have been due to the fact that it is not possible to show the change appearing in these parameters with the applied measuring method, the change being very fast in time yet small in percentage as regards its magnitude and falling outside the limits of accuracy of the method.

The measurements of the field strength have been carried through in the discharge range of 100—420 mA with direct current discharges and heated cathode. In the course of these tests the heating current of the cathode has been 500 mA, its heating voltage 5 V. The dependence on the current obtained is shown by the lower curve of Fig. 4. The values of the electric field strength vary in the range of 0,9—0,6 V/cm. It may be seen that the value of the electric field strength diminishes with increasing discharge current. VERWEIJ has carried out measurements of a similar character [6], however, under conditions not completely identical with those of the present tests. He determined the dependence of the electric field strength on the current for the very same current range and this dependence may be characterized similarly by a curve of falling character. The values of the gradient measured by him are as regards their order of magnitude corresponding to the values of the field strength shown in Fig. 4 between 1,0—0,7 V/cm. The dependence of the plasma potential on the current in this range as well as the corresponding tube voltages V_T which were obtained as the result of the present investigations, are shown in Table 1.

Table 1

I_T (mA)	V_T (Volt)	V_g (Volt)
100	135	64
200	119,5	57
300	112	51
420	102	49,5

The plasma potentials V_g shown in the Table 1 have been measured in a plasma that could be considered as homogeneous, in the middle of the discharge tube and in the axis of the discharge. It can be seen that the potential in an

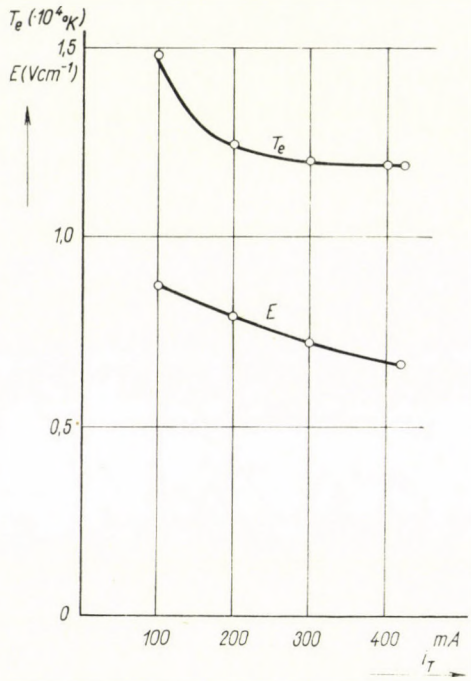


Fig. 4. The current dependence of the electron temperature T_e and electrical field strength E in the case of the direct current discharges without externally heated cathodes

arbitrary point of the plasma decreases with the increase of the discharge current, however, to a lesser extent than the voltage of the discharge tube.

The variation of the electron temperature of the positive column with the current of the discharge is similarly shown in Fig. 4. The electron temperature varies as $1/I_T$ with the discharge current I_T, at the beginning of the

investigated range it decreases more intensely and towards the end more slowly. A similar relationship has been shown by VERWEIJ [6] too, in the very same current range. The electron temperatures vary between 12 000° K and 15 000° K which — converted into electron energy — represent the electron energies between 1,4 and 2 eV.

Measurements of a similar character have been undertaken also in the case of direct current discharges, where the cathode was not heated externally, in

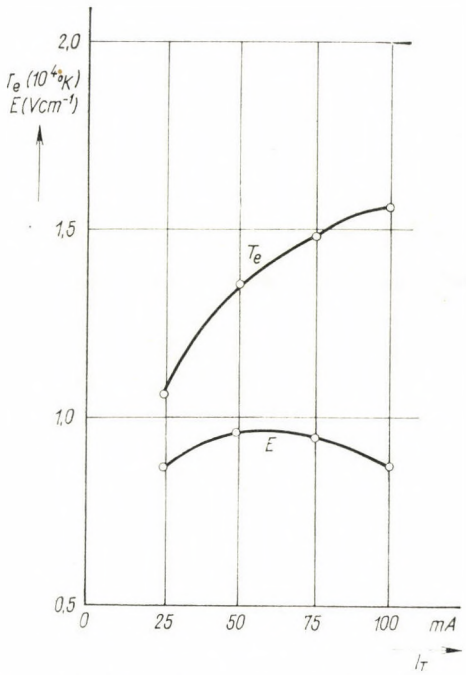


Fig. 5. The current dependence of the electron temperature T_e and electrical field strength E in case of the direct current discharges without externally heated cathodes

the current range between 25—100 mA. The results of the measurements may be seen in Fig. 5. These measurements with unheated cathode are very interesting in respect of the current range. Namely in the section of the discharge current which was investigated the burning voltage of the discharge shows a positive characteristic. The fact that contrary to expectation the character of the curves $T_e = T_e[I_T]$ and $E = E[I_T]$ shown in Fig. 5 does not agree with the character of the corresponding curves plotted in Fig. 4 may be connected with this.

In order to determine also the anode and cathode fall from the measured plasma potentials and the corresponding tube voltages, it was necessary to

determine the distance between the cathode and the positive column. Namely, in first approximation the length of the cathode fall agrees with that distance.

From the investigations described above we know the variation of the size of the cathode dark space (the distance from the cathode to the cathode-side meniscus of the positive column) and with this it is possible — by the help of NÖLLE's method [8] — to determine the value of the cathode fall belonging to the individual current values.

The dependence of this distance on the current as determined by supplementary tests at a constant heating current of 400 mA is shown in Fig. 6. The curve d_p plotted in the Figure shows the line of the cathode-side meniscus of the positive column at the individual discharge currents, while the curve d_g

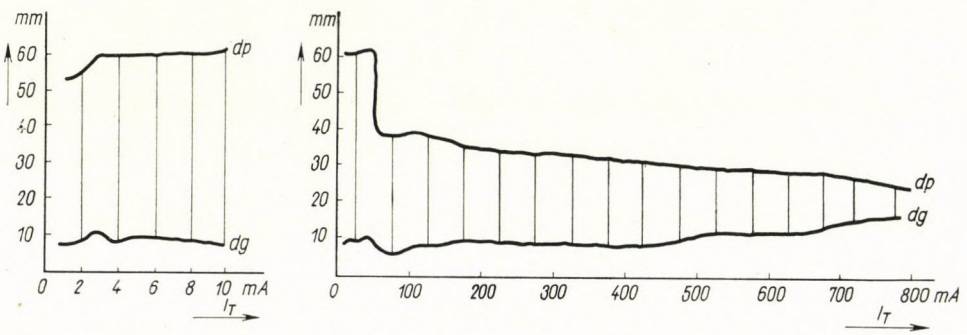


Fig. 6. The current dependence of the cathodic end d_p of the positive column and the end d_g of the glimm-space at $I_h = 400$ mA constant dc cathode heating. I_T = tube current

shows the dependence of the region of the glow light surrounding the cathode on the current. The cathode is placed at 0 mm. A separate diagram shows the curves at currents lower than 10 mA.

In the case of unheated cathodes the shifting of the spaces may be represented by similar curves. The difference was merely that here with the lower currents (under 20 mA) it was no longer possible to carry out measurements because of the diffusedness of the spaces, further also in the case of higher currents there is no significant relative dark-space shifting in the current range of 20—500 mA. These measurements were carried out in a dark place where the background provided a large contrast; a millimeter scale division was made use of.

The shaded area shown in Fig. 6 can in practice be distinguished as the dark space of the discharge. The length of the anode fall section is but a very small fraction of the overall length of the discharge space. As a rule this is

neglected [8] as it does not cause any considerable inaccuracy in the results of measurements of the kind carried out in the present investigations.

Under the above-mentioned conditions, with the help of the NÖLLE method [8] the value of the cathode and anode fall can be determined by extrapolation.

In the case of cathodes not heated from the outside the curve representing the potential of the discharge tube, under the investigated conditions and at a discharge current of 25 mA, is shown in Fig. 7.

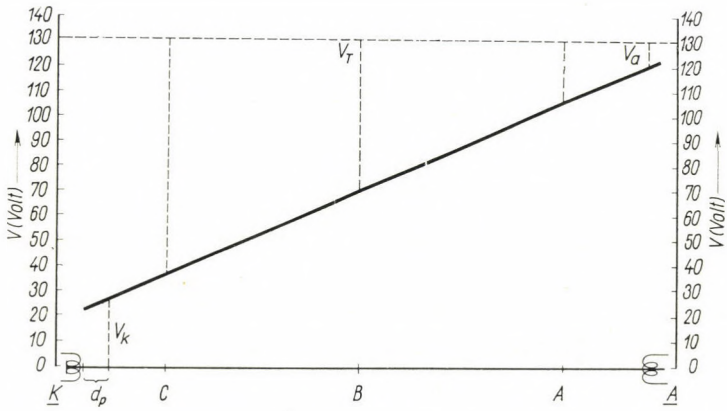


Fig. 7. The potential distribution of the direct current discharges with externally heated cathodes, at $I_T = 25$ mA tube current

The points marked A, B and C on the horizontal straight section correspond to the places where the probes were applied. This section is true to-scale between anode A and cathode K. The length D corresponding to this current and resulting from previous cathode space measurements, is shown at cathode K, while the voltage scale in volts is shown perpendicularly to the section $\overline{AK}$ corresponding to the axis of the discharge tube.

The voltage of the tube at the current in question was found to be 130 V. Since the probe measurement has been effected as compared to the anode, also the plasma potentials obtained at points A, B, C are shown accordingly in Fig. 7. Under the assumptions explained above the cathode fall resulting at the point of intersection is 28 V, the anode fall of the order of 8 V.

By extending this method of evaluation to the current range between 25—100 mA, it has been possible to determine the dependence of the anode fall V_a, the cathode fall V_k and the tube voltage V_T on the current in this range. The results obtained for the case of unheated cathodes are shown in Fig. 8.

5. Discussion

Figs. 4 and 5 show the dependence of the electron temperature T_e and that of the potential gradient E on the current for the case of heated and unheated cathode discharges.

As may be seen from the Figure the electron temperatures vary between 10—16 000° K, which corresponds to electron energies between 1—2 eV. With

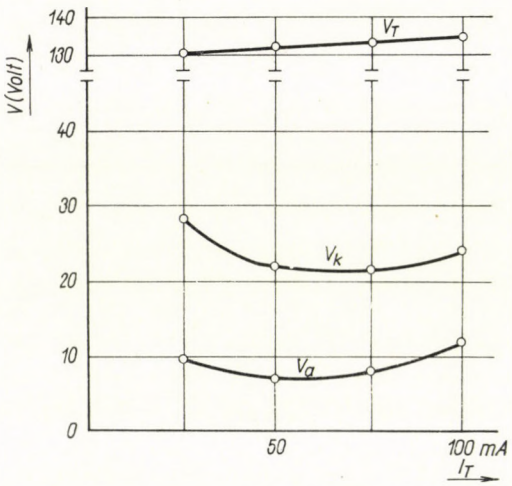


Fig. 8. The current dependence of the anode fall V_a, cathode fall V_k and tube voltage V_T in dc discharges without externally heated cathodes

these values it is possible to calculate the arithmetical mean $\bar{v}$ of the velocity of the electrons

$$\bar{v} = \left(\frac{8kT_e}{\pi \cdot m}\right)^{1/2}, \tag{1}$$

where k is the Boltzmann constant, T_e the temperature of the electrons, m the mass of the electrons.

From this relation the arithmetical mean velocity of the electrons in the above-mentioned electron energy range was found to be between $6.5 \cdot 10^8$—$8 \cdot 10^8$ cm/sec.

As now the mean velocities and the electron energies are known it is possible to determine the electron concentration from the probe characteristics on the basis of the following relation [1—3]:

$$i_e = \frac{1}{2} e \cdot n_e \cdot \bar{v} \cdot f \cdot \exp\left(-\frac{eV}{kT_e}\right), \tag{2}$$

where i_e is the electron current which flows through the probe,

e the charge of the electron,

n_e the electron concentration at the investigated point,

$\overline{V}_e$ the arithmetic mean velocity of the electrons,

f the current pickup surface of the probe,

V the potential of the probe as compared to the plasma,

k the Boltzmann constant,

T_e the electron temperature.

The values n_e of the electron concentration obtained from relation (2) were — under the given discharge conditions — of the order of 10^{11} electrons/cm^3.

When the electron temperature and the electron concentration are known it is possible to determine the DEBYE distance [9] for the case of electrons. As is known, the conditions prevailing in the plasma in many respects resemble those of strong electrolytes [9]. Certain relations of the DEBYE—HÜCKEL theory of strong electrolytes can therefore be well applied to describe the phenomena in the plasma. Thus, among others, the phenomenon of screening appearing in the case of strong electrolytes can be found in the case of ionized gases, i.e. in the case of charged particles, as well. As a result of the tendency at neutralization a screening cloud may develop around the individual charged particles so that beyond a certain distance — which after its discoverer is called DEBYE distance — the particle and the cloud around it will not exert any influence on the other particles, i.e. beyond this radius the particle will behave as a neutral atom, while upon charged particles inside this radius it will exert the well-known short-range force.

The value of this DEBYE radius l_e can be determined on the basis of the following relation [9]:

$$l_e = \left(\frac{k \cdot T_e}{4\pi \cdot n_e \cdot e^2} \right)^{1/2} , \tag{3}$$

where k is the Boltzmann constant, T_e the electron temperature, n_e the electron concentration, e the charge of the electron.

The DEBYE screening radius obtained from relation (3) was found to be of the order of 10^{-3} cm.

With the knowledge of the basic plasma and discharge characteristics it is possible to determine further important parameteres, such as the mobility [12] from the LANGEVIN equation, the ambipolar diffusion coefficients [13] from the SCHOTTKY relationship, the number of impacts, the number of ionized atoms and the ionization degree from the EGGERT—SAHA equation [13], space charging and vibration characteristics, the number of electrons per 1 cm of the positive column, the distribution functions of current density [13], wall charge density, etc.

As the electron energies are known as well as that the plasma is quasi-neutral it is possible to calculate also the velocity of the ions from the law of conservation of energy:

$$\frac{m_e \bar{v}_e^2}{2} = \frac{m_i \bar{v}_i^2}{2}, \tag{4}$$

where m_e and $\bar{v}_e$ are the mass and the velocity of the electron and m_i and $\bar{v}_i$ the corresponding quantities referring to the ion.

In this case such form of the law of conservation of energy is not applicable, because the ionisation degree of plasma is too low. In the investigated case the electron temperatures varied between 10 000—16 000° K, those of the ions amounted to 320—400° K in the axis of the discharge.

As now the ion temperatures and ion energies are known, it is possible — similar to the above calculations — to determine values of other parameters and processes characteristic of the ions.

Figs. 5 and 8 present the results of unheated cathode measurements in the 25—100 mA current range. Fig. 8 shows, for the sake of comparison, the dependence of the tube voltage on the current. It may be seen that here the discharge is characterized by a positive characteristic.

As can be seen from Fig. 8 the cathode and anode fall of the discharge has a minimum in the range between 50 and 75 mA. At the same time, however, according to Fig. 5 the potential gradient E has a maximum. This means that the voltage fall of the positive column is then the highest. This admits of the conclusion that in the processes around the cathode and the anode such changes have occurred at this current that have rendered superfluous the larger cathodic and anodic falls belonging to the previous currents.

The voltage fall of the positive column together with the cathode fall and the anode fall make up jointly the tube voltage of the discharge tube. Plotting this tube voltage (denoted by V_T in Fig. 8) as function of the discharge current the discharge characteristic is called positive or negative according to whether the direction tangent of the curve is positive or negative. From the comparison of Fig. 5 and 8 it becomes clear that in certain current ranges the positive or negative character of the discharge characteristic will be determined by the length of the positive column. It seems necessary to point this out because this question is not discussed sufficiently clearly even in the relevant technical literature [12—14].

The development of tube voltage, gradient, cathode fall and anode fall of the kind described above distinguishes in the present case the current range between 50—70 mA as against the currents (investigated up to 25 mA) of somewhat higher or lower values.

As is known a phenomenon characteristic for discharges fed off direct current voltage is the cataphorasis[10] which at times is used for gas separation. In our case in the mercury vapour discharges the ionized mercury atoms proceed towards the cathode on account of their positive charges and there become neutralized. A diffusion flow directed towards the anode and carrying mercury atoms attempts to reduce the change of concentration resulting in this way. The concentration of the neutral mercury and that of the argon atoms will, because of the cataphorasis, become permanent in the neighbourhood of the cathode after a certain time and increase up to a certain limit, their partial pressures will increase. This increase is considerable and of a magnitude which can be measured.

In the case of low pressure mercury vapour discharge tubes fed off the direct current voltage this phenomenon will appear also in a visible form [10]. A dark zone will be created in the neighbourhood of the anode after a certain time, which is due to the fact that the mercury resonance lines (1849 Å, 2537 Å) generated by the electrons will not make their appearance with the former intensity. With a 40 W fed output tube length 1200 mm diameter 38 mm this phenomenon may result in the complete darkening of the surroundings of the anode [10] following 18 hours of burning.

In the course of the present investigations measurements have been carried through also in connection with the phenomenon of cataphorasis, through comparison of the parameters which could be measured by probe measurement.

In the case of an unheated cathode the investigations have been carried out at a discharge current of 100 mA, in stationary discharge plasma, recording on the probe the probe characteristic and calculating from this the electron temperature following connection of the tube, upon the establishment of the state of equilibrium and after 14 hours of stable operation. The conditions of the measurements and the characteristic of the discharge tube agreed with that obtained before.

According to the literature [10] in the direct current discharges when the tube is operated at a discharge current of 420 mA the reduction of the luminous flux mentioned earlier amounts at the anode to some 85% after 14 hours operation.

A considerable reduction of the luminous flux was obtained also in the course of the present tests, however, this amounted to only a fraction of the values given in the literature [10]. It seems to us that this difference may be ascribed to the deviation between the discharge currents. Namely it appears plausible that the difference in concentration resulting from the cataphorasis sets in faster in the case of a higher discharge current. With the increase of the current also the increase of the number of positive ions arriving at the cathode

in unit time may be expected and at the same time ions of higher velocity exert a higher resistance against the neutral atoms flowing back towards the cathode. Thus in the case of a higher current the flow of the ions can maintain the equilibrium against a higher concentration difference. In the case of mercury vapour discharges the neutralized particles are atoms. It is important to take this into account in the mass ratios of the collisions in any mathematical interpretation of the phenomenon.

In measurements carried out under identical conditions between which 14 hours elapsed it was found that the temperature of the electrons at the anode increased by 4270° K, while at the cathode it decreased by 150° K. The latter is not a significant value and is near to the limits of accuracy of the measurement.

In the qualitative interpretation of the results of the tests we have started from the observation by LAKATOS [11] according to which the change of electron temperatures of such a character is connected with the phenomenon of cataphorasis.

According to the above in the course of the cataphorasis the partial pressure of the neutral atoms diminishes in the surroundings of the cathode, and this results in the increase of the mean free path. Thus also the mean free path of the electrons will increase and the electrons will now — on identical mean free paths — collide less often with neutral atoms and thus also the energy transmitted by them (elastic, unelastic collisions) will be reduced, the energy coupling between the particles will become looser. Since in the course of the investigations the macro-parameters of the discharge did not change (discharge current, potential development), it may be assumed that also the space where the electrons were accelerated remained unaltered. This means that on the same mean free paths the electrons would gain the same amount of energy. However, with the inset of cataphorasis the mean free paths increase and the average kinetic energy increases correspondingly. Since there is a unique connection between the kinetic energy and the electron temperature, the above-mentioned experimentally observed increase of the electron temperature around the cathode may be interpreted in this way.

It follows that also the energy transmitted per unit volume to the mercury atoms becomes smaller. This, as already explained above, results in a reduced excitation of the mercury atoms, which is also visible to the naked eye. In the discharge — in a stationary state — mainly the mercury atoms will get excited and ionized because of their lower excitation and ionization potential as against the argon.

The excitation of the resonance lines of the mercury may be investigated if the excitation function which has a probability character, is known. According to the available literature [12] this excitation is a function of second order of

the electron energies and possesses a maximum. As the electron energies change because of the reduction in the number of collisions the excitation of resonance lines will be shifted from the optimal conditions of excitation.

In all probability the more detailed quantitative investigation of the cataphorasis can be carried out with the help of probe measurements by arranging the electrodes (probes) more closely to the anode and the cathode than in the present case making them reach to identical depths.

Further interesting results may be obtained by a refinement of the measurement which may make possible the interpetation of the process of cataphorasis by its dependence on the change of the electron temperature and the electron concentration in time in the neighbourhood of the anode and the cathode.

In addition to the process explained here which may be called "axial cataphorasis" simultaneously a "radial cataphorasis" arises as well [6], between the axis and the wall.

By taking into consideration the ambipolar radial diffusion, this radial cataphorasis may be treated theoretically as well [6]. VERWEIJ has shown under certain simplifying conditions, that the phenomenon may involve also a change in concentration of some 20%: the concentration of the mercury atoms will diminish to this extent along the axis of the discharge relative to the concentration of the mercury atoms to be found by the wall.

Through a discussion of the parameters described in the present article, it will be possible to give further relationships for the characteristic processes of the direct current discharges. Here also it will be expedient to rely on the values obtained by probe measurements which can be carried out under properly defined conditions.

6. Acknowledgement

The author wishes to thank Mr. G. LAKATOS for his valuable remarks and advice in connection with the investigations.

REFERENCES

1. I. LANGMUIR, J. Franklin Inst., **196,** 751, 1923.
2. I. LANGMUIR, H. MOTT-SMITH, Gen. El. Rev., **27,** 449, 538, 616, 762, 810, 1924.
3. I. LANGMUIR, K. T. COMPTON, Rev. Mod. Phys., **28,** 727, 1926.
4. J. BITÓ, Ph. D. thesis. Experimental Physical Institute, Szeged, 1960.
5. J. BITÓ, Hung. Phys. J., **X,** 411, 1962.
6. W. VERWEIJ, Ph. D. thesis, Utrecht, 1960.
7. J. BITÓ, Hung. Phys. J., **X,** 303, 1962.
8. E. NÖLLE, Ann. der Physik, **18,** 328, 1956.
9. J. L. DELCROIX, Introduction à la théorie des gaz ionisés, Dunod, Paris, 1959.
10. Vorschaltgeräte und Schaltungen für Leuchtstofflampen, B. B. C, 1954.
11. GY. LAKATOS, Verbal communication, 1960.

12. A. VON ENGEL, Ionised Gases, Clarendon Press, Oxford, 1955.
13. N. A. KAPZOW, Elektrische Vorgänge in Gasen und im Vakuum, Deutscher Verlag, Berlin, 1955.
14. K. G. EMELEUS, The Conduction of Electricity through Gases, Methuens, 1951.

О ЗАВИСИМОСТИ МЕЖДУ НЕКОТОРЫМИ ПАРАМЕТРАМИ ПАРО-РТУТНОГО РАЗРЯДА ПОСТОЯННОГО ТОКА

Й. БИТО

Р е з ю м е

Применяя метод измерения сондами Лэнгмюира, для случая нагретого и холодного катодов определяется зависимость аксиального градиента и электронной температуры разряда постоянного тока разряда. Применяя способ оценки по Нэлле, определением плазменного потенциала по зависимости катодных тёмных пространств от тока, автором даётся ход потенциала, характерного для разряда, излагается зависимость катодного и анодного падений от тока. Рассматриваются явления аксиального катодного форезиса, показывается зависимость между процессом катодного форезиса и временном изменении электронной температуры.

REFLECTIONS ON THE PROBLEM OF MEASURING THE VELOCITY OF LIGHT

By

L. JÁNOSSY

CENTRAL RESEARCH INSTITUTE OF PHYSICS, BUDAPEST

(Received 5. XI. 1963)

The phenomena connected with the propagation of light are analysed with the purpose of re-examining as to what conclusions can be drawn from the relevant experiments. It is found, that a very clear picture can be obtained supposing — contrary to the usually accepted view — that electromagnetic waves possess a carrier which may be taken as the ether. The philosophical questions concerning the re-introduction of the ether as the carrier of electromagnetic waves is treated by the author in other articles [13].

Introductory remarks

§ 1. In this article we attempt a renewed analysis of the methods for the determination of the velocity of light. On the one hand we start from MAXWELL's theory of the electromagnetic field, and on the other we shall make use of well-known results of optical and electromagnetic experiments. We shall endeavour to discuss the experiments in the form in which they were actually carried out, avoiding basing our conclusions on ideal experiments (Gedankenexperimente). We shall not make use of EINSTEIN's ideas on relativity of time and space — it is our aim to show that the phenomena discussed here, keeping strictly to the experimentally observed facts, may be theoretically explained entirely satisfactorily without the concept of relativity of time and space.

Naturally, every theory — and this is true also of the one proposed here — requires certain abstractions. In contrast to EINSTEIN's method we attempt, however, to limit ourselves to ideas of such a kind as correspond to the pre-EINSTEINian way of thinking.

§ 2. The numerical value c of the velocity of light appears as critical velocity in the MAXWELL equations, which can be determined numerically through the comparison of electrostatic and electromagnetic units. It is interesting to examine the role of the critical velocity in somewhat more detail.

a) The static measure of electric charge can be determined with the help of the Coulomb law. Consider a number of charges $e_1, e_2, \ldots, e_n$. The force (measured in dyne) with which two charges, say e_k and e_l at a distance r, act on each other is

$$F_{kl} = \frac{e_k e_l}{r^2}, \qquad k, l = 1,2,\ldots,n. \tag{1}$$

Considering three charges, say e_1, e_2, e_3, we may measure the forces F_{12}, F_{23}, F_{31} acting between the three pairs. With the help of (1) we find:

$$e_1 = r\sqrt{\frac{F_{12}F_{13}}{F_{23}}}, \qquad e_2 = r\sqrt{\frac{F_{23}F_{21}}{F_{31}}}, \qquad e_3 = r\sqrt{\frac{F_{31}F_{32}}{F_{12}}}. \tag{2}$$

Indeed, substituting (2) in (1) an identity results, provided the sign of the roots is chosen appropriately. If one of the charges is (arbitrarily) assigned the positive sign, the others are unambiguously defined and thus the measures of the charges are unambiguously determined.

b) The numerical values of the measures of magnetic pole strengths m_ν, $\nu = 1, 2, \ldots, n$ can be determined in a similar way with the help of the Coulomb law for magnetic poles

$$F_{kl} = \frac{m_k m_l}{r^2}. \tag{3}$$

A certain complication is due to the fact that in contrast to electrical charges single magnetic poles do not exist in nature. However, a thin bar magnet acts as if magnetic charges $\pm m_\nu$ of equal magnitude but opposite sign were placed at its ends, and by measuring the effect of bar magnets on each other — similar to the case of electric charges — the pole strength can be determined. Instead of (3) one may put

$$F_{kl} = \frac{m_k m_l}{r_{\text{eff}}^2},$$

where, if the two rods are forming a straight line:

$$\frac{1}{r_{\text{eff}}^2} = \frac{1}{r^2} + \frac{1}{(r+2l)^2} - \frac{2}{(r+l)^2}.$$

r means the distance between two poles facing each other and l the lengths of the rods.

c) Electrical charges act on magnetic poles only when they move relative to each other. The force with which a charge e_k moving with a velocity v acts on a magnetic pole m_l which is at rest, may be written

$$F_{kl} = \alpha \frac{e_k m_l}{r^2} \frac{|\mathbf{v} \times \mathbf{r}|}{r}, \tag{4}$$

where α is a constant. Comparing (4) with (1) or (3) we see that α must have the dimension of a reciprocal velocity, if e_k and m_l are expressed in electrostatic units. We may thus write

$$1/\alpha = c ,$$

where c is a velocity.

The value of c can be determined experimentally by measuring the force with which a moving charge acts on a magnet. Such measurements were carried out by ROWLAND and later by EICHENWALD [1]. In their investigations the effect of a rotating, electrically charged disc on a needle magnet was measured.

More exact experiments were carried out later, in which a condenser was periodically charged and discharged and c was determined with the help of the magnetic effect of the discharge current.

The detailed description of this experiment can be read in textbooks. The value of the critical velocity resulting from these measurements was found to be exactly equal to the value of the velocity of propagation of electromagnetic waves and to that of light.

§ 3. From MAXWELL's equations it follows that an electromagnetic field in vacuum is propagated with the critical velocity c. MAXWELL deduced from the numerical agreement of the value of the critical velocity resulting from electrical and magnetical measurements on the one hand with that measured for the velocity of light on the other that light consists of electromagnetic waves.

It must be emphasized, that it follows from MAXWELL's theory that the velocity of propagation of electromagnetic waves is independent of the state of motion of the source of the waves; this result of MAXWELL's theory is sometimes incorrectly regarded as a result (or postulate) of the theory of relativity.

To avoid misunderstanding we note that a radiating atom moving with a constant velocity $\mathbf{v}$ emits at every moment of its motion a new spherical wave. These successively emitted spherical waves spread each with the velocity c from the point from which they were sent out. Because of the motion of the emitting atom the generated waves are not concentric but have a distribution like that shown in Fig. 1.

The wave trains arriving successively in A will appear compressed because of the motion of the light source. A point B, however, from which the source moves away, will be reached by the individual wave fronts in longer time intervals than those between the actual emissions of the wave fronts.

The phenomenon is the DOPPLER-effect which indeed depends on the state of motion of the source. However, the individual wave fronts will pass A as well as B with the velocity c independent of the state of motion of the source.

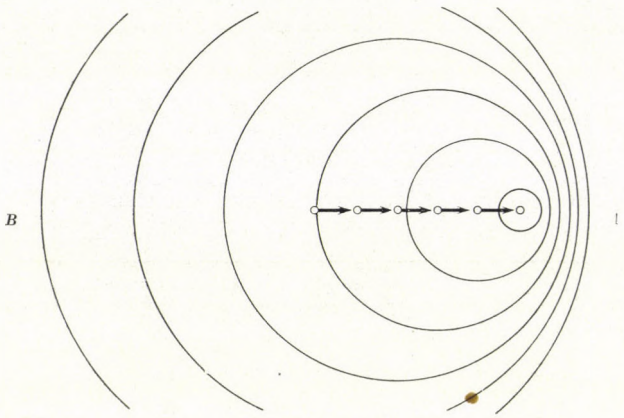


Fig. 1. The radiation field of a moving atom

Methods for the measurement of the velocity of propagation of light

§ 4. One might propose the following method for the measurement of the velocity of light. A short signal is being sent at $t = t_1$ from a source situated in the point A. The signal is observed at a point B which lies at a distance l from A. If the signal arrives at $t = t_2$ one has for the velocity of propagation

$$c = l/(t_2 - t_1) \,. \tag{5}$$

§ 5. The above method can, however, not be carried out directly, since in order to measure t_1 and t_2 one needs two synchronized clocks, U_A and U_B, one near A and the other near B. In principle, one could synchronize U_A and U_B by observing them from a point C, which is at the same distance from both. If the pictures of the clocks, U'_A and U'_B, as seen from C (e.g. in a telescope) do not show any phase shift one may suppose that the two clocks run synchronously and the measurement of the propagation of light signals can then be carried through according to equ. (5) with the help of the clocks thus synchronized. Using this method we must assume that light is propagated along all three paths, namely from A to B, from A to C and B to C with the same velocity.

How far such an assumption is justified, we shall discuss later on. The method just described has, however, not been put into practice.

In the following we discuss the experimental methods by which the velocity of propagation of light was measured.

In the experiments which have been carried out in practice the synchronization was attained in two ways.

Method of Römer

§ 6. One might synchronize the clocks U_A and U_B by placing them next to each other, synchronizing them, and then very cautiously transporting the one to A and the other to B, in the hope that the transport is not detrimental to the clock mechanism. We shall discuss later, how far this "hope" is realized.

After placing the clocks near the points A and B we may regard some given positions of the pointers of U_A as the signal sent out from A. Observing U_A from B with the help of a telescope a delayed picture U'_A of the clock U_A appears because of the finite velocity of propagation of light. Denoting the phase difference between the picture U'_A and the clock U_B by Δt we have

$$c = l/\Delta t \,.$$

U_A may be observed from points at various distances $l_1, l_2, \ldots$ from U_A. Always taking the clock U_B with us, we observe phase differences Δt_1, $\Delta t_2, \ldots$ between the pictures U'_A, $U''_A, \ldots$ and the clock U_B and the velocity of light can be determined from the change of the phase differences with the distance. We have $c = l_1/\Delta t_1 = l_2/\Delta t_2 = \ldots$ and thus

$$c = \frac{l_\mu - l_\nu}{\Delta t_\mu - \Delta t_\nu}, \quad \mu, \nu = 1, 2, \ldots \tag{6}$$

Equ. (6) has the advantage that it contains only changes of the distances, $l_\mu - l_\nu$, and changes of the phases, $\Delta t_\mu - \Delta t_\nu$.

§ 7. The method described above corresponds in principle to that used by the astronomer OLAF RÖMER for the first determination of the velocity of light. RÖMER by observing the motion of the satellites of Jupiter found apparent irregularities in their periods of revolution which obviously were connected with the variation of the distance between Earth and Jupiter.

So as to be able to apply our earlier considerations we regard a clock on the Earth as the clock U_B and Jupiter together with its satellites as the clock U_A. At the time $t = t_0$ suppose the Earth to be in opposition to Jupiter and their distance apart to be l_0. Observing now the revolutions of one of the satellites of Jupiter we obtain times $t_1, t_2, \ldots, t_n$ at which the successive revolutions of the satellite appear completed as seen from the Earth. The corresponding

distances be $l_1, l_2, \ldots, l_n$. The times t_ν, $\nu = 1, 2, 3, \ldots, n$ observed on the Earth differ from the times $t'_\nu = t_\nu - \Delta t_\nu$, $\nu = 0, 1, 2, \ldots, n$ at which the revolutions are actually completed. The delays Δt_ν are equal to the times for the light to travel from Jupiter to the Earth. Therefore

$$\Delta t_\nu = l_\nu / c \,, \qquad \nu = 0, 1, 2, \ldots . \tag{7}$$

Suppose $l_n = l_0$, i.e. at the time t_n the Earth be again in opposition to Jupiter. Observing the satellite at times t_0 and t_n from the same distance the corrections

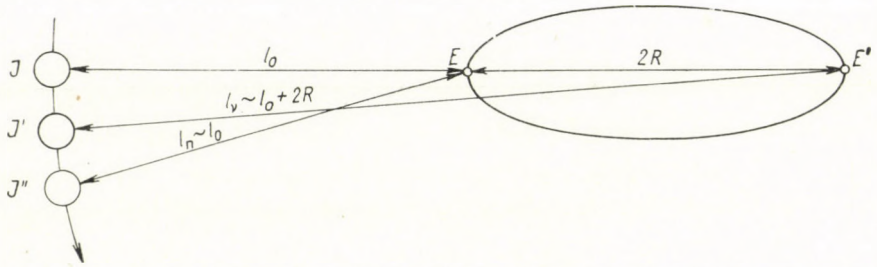


Fig. 2. Astronomical determination of the velocity of light. J, J', J'' subsequent positions of Jupiter, E, E' those of Earth

for the times of travel of the light are equal in both cases: $\Delta t_n = \Delta t_0$, and we have for the time the satellite takes in effect to complete n revolutions: $t_n - t_0 = t'_n - t'_0$. The time T for one revolution is thus

$$T = \frac{1}{n} (t_n - t_0) \tag{8}$$

and the phase shift the clock U_A on Jupiter shows after the ν-th revolution becomes

$$\Delta t_\nu = \Delta t_0 + (t_\nu - \nu T) . \tag{9}$$

With the help of (7), (8) and (9) the velocity of light is thus found to be

$$c = \frac{l_\nu - l_0}{\Delta t_\nu - \Delta t_0} = \frac{l_\nu - l_0}{t_\nu - \nu T} \,.$$

If ν is chosen so that l_ν corresponds to the largest distance (in the course of a year) between Earth and Jupiter, we find

$$l_\nu - l_0 = \text{diameter of the Earth's orbit} \approx 3 \cdot 10^8 \text{ km}.$$

As a result of observation one finds

$$t_\nu - \nu T \approx 1000 \text{ sec},$$

thus

$$c \approx 300\,000 \text{ km/sec}.$$

Methods of Fizeau and Foucault

§ 8. The need for the synchronization of clocks can be eliminated if instead of using two clocks one clock is compared with its image in a distant mirror.

Let a clock U_A be placed at the point A, at the point B at a distance l from A let there be a mirror normal to the line AB. With the help of a telescope near A we observe the image U'_A of U_A. Traveling from A to B and back again to A the light has gone along a path of length $2l$, we thus have

$$c = 2l/\Delta t,$$

where Δt denotes the phase shift observed between U_A and its image U'_A as seen through the telescope.

FIZEAU carried out this experiment using as clock a cogwheel which was rotating fast. With the help of a suitable optical arrangement the image of the cogteeth of the cogwheel could be observed through the teeth of the rotating cogwheel. When the velocity of rotation was such that during the time $\Delta t = 2l/c$ the wheel turned by a whole number of teeth the image of the cogteeth appeared lighter, whereas for intermediate velocities it was concealed by the teeth of the rotating cogwheel and thus appeared darker. Observing the darkening and lighting of the image of the cogteeth as a function of the velocity of rotation, one can determine the phase shift Δt and thus the velocity of propagation of light c.

FOUCAULT improved FIZEAU's arrangement by using as clock a rotating mirror. A beam of light falling on this rotating mirror is reflected by it and falls on a second fixed mirror placed at a distance l from the first, which again reflects the beam of light. This reflected beam returns to the rotating mirror after a time Δt. As, however, during the time Δt the mirror has turned through an angle $\alpha = \omega \Delta t$ (where ω is the angular velocity of the mirror), the beam reflected anew by the rotating mirror leaves the arrangement at an angle α to the direction of the incident beam. Measuring the angle α one obtains the value of c:

$$c = \frac{2\omega l}{\alpha}.$$

FOUCAULT's experiment furnished a very accurate determination of the light velocity. The best experimental value is according to BERGSTRAND [2] (see also COHEN et al. [3])

$$c = 299\,793{,}1 \pm 0{,}32 \text{ km sec}^{-1} \text{ (velocity of light).}$$

This value agrees — within the experimental error — with the best experimental value of the critical velocity determined by the comparison of electrostatic and

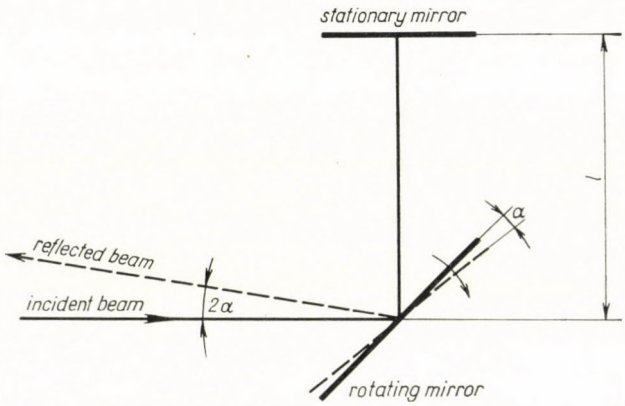


Fig. 3. Scheme of the FOUCAULT experiment

electromagnetic units (see § 2). This latter was found to be (according to ROSA—DORSAY [4])

$$c = 299\,784 \pm 15 \text{ km sec}^{-1} \quad \text{(critical velocity).}$$

The carrier of light

§ 9. When we state that light spreads with a velocity c we cannot avoid the question: In relation to what has this velocity the value c?

One might think that the light signal takes up the velocity of the light source by which it is emitted at the moment of emission, and then c would be the velocity of the signal relative to the light source. This so-called "ballistic theory" compares the light to a missile fired from a gun. The velocity of the missile will be $\mathbf{V} = \mathbf{v} + \mathbf{w}$, when the gun moves with a velocity $\mathbf{v}$ at the moment of firing and where $\mathbf{w}$ is the velocity relative to the gun the missile receives through the firing.

As we have already mentioned the ballistic theory does not apply to light. A light signal having left its source spreads with a velocity which is independent of how it originated. This phenomenon can be best understoop supposing, like Maxwell, that the electromagnetic waves are waves of the

ether and that the velocity c means the velocity these waves have relative to the ether.

The question of the ether has incorrectly been linked with the question of "absolute rest". It should be stressed here, that these two questions are not really connected with each other. We (following MAXWELL) suppose the ether to be the carrier of electromagnetic waves. It is of no consequence here whether the ether is at rest or whether it moves. In particular, statements according to which the ether has to represent "absolute rest" are purely metaphysical and irrelevant to the theory of light phenomena.

Further, there is no need to suppose the ether to be "rigid". One can well imagine that there are regions of the ether located in different parts of the universe which are in motion relative to one another, i.e. that the ether contains currents. Such currents would cause diffraction of light which might possibly be detected by astronomical observations. So as to avoid unnecessary speculation we note that for the interpretation of those experiments which have been carried through in practice it suffices to suppose that the ether flows nearly uniformly in the vicinity of the solar system.

§ 10. In §§ 4—8 we discussed the experiments by which the value of c has been determined. In those considerations we have implied that c is the velocity of light relative to the measuring apparatus. This assumption is justified only if the apparatus is at rest relative to the ether. This remark seems trivial; because of the confusion of concepts we should, however, like to go into some more detail.

Consider a system of coordinates K_0 which is at rest relative to the ether. As possibly different regions of the ether are moving relative to each other we suppose that the system of coordinates K_0 moves together with the ether in the vicinity of the Earth and we limit the system K_0 to such a region where the motion of different parts of the ether relative to each other can be neglected.

§ 11. Consider first two points A and B which move along the positive x-axis of K_0 with the same constant velocity $v > 0$. At the time $t = t_1 = 0$ let A be in the origin of the coordinate system K_0 emitting from there a light signal. We calculate the moment $t = t_2$ at which the signal reaches the point B. As the signal moves independently of the state of motion of the source A, at the time $t > 0$ it will have reached the surface of a sphere around the origin of the coordinate system of radius $r_s(t) = ct$. This sphere intersects the positive x-axis of K_0 at the time $t > 0$ in a point with x-coordinate

$$x_s(t) = ct. \tag{10}$$

At time t the x-coordinate of point B is

$$x_B(t) = l + vt, \tag{11}$$

where l means the (constant) distance between points A and B. At time $t_2^{(0)} = l/c$ the signal reaches a point B_0 the x-coordinate of which is l, i.e. it reaches the location at which point B was at time $t = 0$. However, at time t at which the signal has reached B_0, i.e. the original location of B, B itself is at a larger distance from the origin of the system than B_0. The signal reaches B at a time t_{12} at which the following relation is valid:

$$x_B(t_{12}) = x_s(t_{12}). \tag{12}$$

Substituting from (10) and (11) in (12) we obtain

$$l + vt_{12} = ct_{12} \tag{13}$$

and thus

$$t_{12} = \frac{l}{c - v}.$$

From this it can be seen that — provided $v > 0$, — $t_{12} > l/c$. Thus the signal reaches the point B, which is moving, later than the point B_0, which is at rest.

Introducing a coordinate system K, the points of which all move with the same constant velocity v parallel to the x-axis of K_0, A and B are two points at a distance l which are at rest relative to K. The velocity c_+ of the light signal relative to K can be determined from

$$l/t_{12} = c - v = c_+.$$

Considering further the signal which goes from B to A we find by a similar consideration for the time the signal takes for this distance

$$t_{21} = \frac{l}{c + v}.$$

The velocity c_- of the returning signal relative to K is thus

$$l/t_{21} = c + v = c_-.$$

We thus see, that relative to K there is $c_+ < c_-$, i.e. light does not spread isotropically relative to K.

§ 12. For the general case we consider two points A and B with coordinate vectors $\mathbf{r}_A$ and $\mathbf{r}_B$. So as to indicate that A and B move we write for the values of these vectors at time t

$$\mathbf{r}_A(t) \quad \text{and} \quad \mathbf{r}_B(t).$$

The light signal which was emitted at time $t = t_1$ from point A reaches at time $t > t_1$ the surface of a sphere, the points of which obey the following relation:

$$\left(\mathbf{r}_S(t) - \mathbf{r}_A(t_1)\right)^2 = c^2(t - t_1)^2. \tag{14}$$

The signal reaches point B at the time $t = t_2$ at which B lies on the surface of the sphere (14) so that

$$\left(\mathbf{r}_B(t_2) - \mathbf{r}_A(t_1)\right)^2 = c^2(t_2 - t_1)^2. \tag{15}$$

§ 13. We are in particular interested in the case in which A and B move with the same constant velocity $\mathbf{v}$. In this case we may write for the coordinate vectors of A and B at time t:

$$\mathbf{r}_A(t) = \mathbf{a} + \mathbf{v}t\,, \quad \mathbf{r}_B(t) = \mathbf{a} + \mathbf{l} + \mathbf{v}t\,,$$

where $\mathbf{a}$ and $\mathbf{l}$ mean constant vectors. $\mathbf{l}$ is the vector which starts in A and points to B and the absolute value of which is equal to the distance AB : $|\mathbf{l}| = l$. Writing $t_{12} = t_2 - t_1$ for the time the signal needs to go from A to B, we obtain substituting from (13) and (15)

$$(\mathbf{l} + \mathbf{v}t_{12})^2 = c^2\, t_{12}^2. \tag{16}$$

Putting $\mathbf{lv} = lv \cos a$ we find from (16)

$$(c^2 - v^2)t_{12}^2 - 2lv \cos a\, t_{12} - l^2 = 0. \tag{17}$$

a is the angle subtended by the vector $\mathbf{l}$ with the vector of velocity $\mathbf{v}$. The positive solution of (17) for t_{12} becomes:

$$t_{12} = l\,\frac{v \cos a + \sqrt{c^2 - v^2 \sin^2 a}}{c^2 - v^2}\,. \tag{18}$$

A simple rearrangement leads to

$$t_{12} = \frac{l}{\sqrt{c^2 - v^2 \sin^2 a} - v \cos a}\,. \tag{19}$$

It can be seen from (19) that light signals spread anisotropically relative to a coordinate system K, the points of which move with the constant velocity $\mathbf{v}$ relative to K_0. The velocity $c(a)$ in a direction which subtends an angle a with $\mathbf{v}$

is given by

$$c(\alpha) = l/t_{12} = \sqrt{c^2 - v^2 \sin^2 \alpha} - v \cos \alpha. \tag{20}$$

In particular, we have for $\alpha = 0, \frac{\pi}{2}$ and π, respectively,

$$c(0) = c - v, \quad c\left(\frac{\pi}{2}\right) = \sqrt{c^2 - v^2}, \quad c(\pi) = c + v. \tag{21}$$

Measurement of the velocity of light in moving systems

§ 14. The experiment of FIZEAU as well as that of FOUCAULT by which the velocity of light, c, was determined were carried through with the help of apparatuses fixed to the Earth. If we do not want to suppose that the ether moves together with the Earth, we have to assume, that the whole apparatus carries out a rotational as well as a translational motion relative to the ether. We thus have to revise the considerations of § 8, where it was supposed that the velocity of light is constant relative to the measuring arrangement.

The corrections which have to be expected are, however, small: In the experiments use is made of the time a signal needs to go to and fro between two points A and B. Denoting these times by t_{12} and t_{21}, respectively, both may be determined from (18) changing in the one case the role of A and B. This changing of roles corresponds to replacing α by $\pi - \alpha$, i.e. replacing $\cos \alpha$ by $-\cos \alpha$. Thus there is

$$t_{21} = l \frac{-v \cos \alpha + \sqrt{c^2 - v^2 \sin^2 \alpha}}{c^2 - v^2}. \tag{22}$$

With the help of (18) and (22) one thus finds for the time the signal needs to go from A to B and back again to A in a translational system:

$$t_{12} + t_{21} = \frac{2l}{c} \frac{\sqrt{1 - \frac{v^2}{c^2} \sin^2 \alpha}}{1 - \frac{v^2}{c^2}}.$$

Expanding in a series in powers of v^2/c^2 one obtains

$$t_{12} + t_{21} = \frac{2l}{c}\left[1 + \left(1 - \frac{1}{2} \sin^2 \alpha\right) \frac{v^2}{c^2}\right] + \text{terms of higher order.}$$

Putting the order of magnitude of the suspected translational motion of the Earth at

$$v \sim 30 \text{ km/sec},$$

i.e. the orbital velocity of the Earth around the sun, one obtains $v^2/c^2 \sim 10^{-8}$ and the relative error which arises further above when simply taking the running time of the signal to be $2l/c$ is smaller than the relative error of measurement of the most accurate determination of c (see § 8).

The velocity of the solar system in the galaxy is estimated at 300 km/sec. Substituting for v this value, one still obtains very small corrections.

The rotation of the system together with the Earth does not lead to corrections as will be shown later.

Determination of motion relative to the ether of an experimental arrangement

§ 15. As we have thus seen, the measurements which were carried out to determine the velocity of light were — from the very beginning — much too inaccurate to show any possible effect which might arise from the motion relative to the ether of the measuring arrangement. We have, however, to put the question, whether this motion might not be determined with the help of other more suitable experiments.

If we want to discuss this question in more detail we have to distinguish between rotation and translation. To begin with we discuss the properties of a system rotating with respect to the ether.

§ 16. We thus consider a system K which rotates relatively to K_0. The rotation may be described by a vector ω, the direction of which is parallel to the axis of rotation and the absolute value of which gives the angular velocity of the rotation. We assume the origin of K to be the same as that of K_0 and the axis of rotation to pass through the common origin. The coordinate vector $\mathbf{r}_B$ of a point B moving together with K is a variable vector with respect to K_0. One finds

$$\frac{d\mathbf{r}_B}{dt} = \omega \times \mathbf{r}_B. \tag{23}$$

A signal leaving A at the time $t = t_1$ reaches the point B at the time t_2, where according to § 12 equ. (15) the following relation holds:

$$\big(\mathbf{r}_A(t_1) - \mathbf{r}_B(t_2)\big)^2 = c^2(t_2 - t_1)^2. \tag{24}$$

Consider now the case in which A is located on the axis of rotation. In that case $\mathbf{r}_A = \text{const.}$, and thus in particular $\mathbf{r}_A(t_1) = \mathbf{r}_A(t_2)$

$$\big(\mathbf{r}_A(t_1) - \mathbf{r}_B(t_2)\big)^2 = \big(\mathbf{r}_A(t_2) - \mathbf{r}_B(t_2)\big)^2 = l^2.$$

From (24) one finds thus

$$t_{12} = t_2 - t_1 = l/c\ ,$$

i.e. the time of travel t_{12} from A to B in this case is not influenced by the rotation.

The result is similar in the case when the receiver B lies on the axis of rotation, when thus $\mathbf{r}_B(t) = \text{const}$. In this case:

$$t_{12} = t_{21} = l/c.$$

Thus, if A or B lie on the axis of rotation the times of travel to go from A to B and from B to A are equal and their value is that which one obtains in the case of a system at rest.

§ 17. In the general case when neither A nor B lie on the axis of rotation we can solve (24) by putting

$$\mathbf{r}_B(t_2) - \mathbf{r}_A(t_1) = \mathbf{l}(t_1) + \left(\mathbf{r}_B(t_2) - \mathbf{r}_B(t_1)\right), \tag{25}$$

where in general

$$\mathbf{l}(t) = \mathbf{r}_B(t) - \mathbf{r}_A(t)$$

means the vector going from A to B at the time t. The distance AB is constant, we have thus at any time t

$$\mathbf{l}^2(t) = l^2 = \text{const}. \tag{26}$$

Substituting (25) in (24) we find with the help of (26)

$$l^2 + 2\mathbf{l}(t_1)\left(\mathbf{r}_B(t_2) - \mathbf{r}_B(t_1)\right) + \left(\mathbf{r}_B(t_2) - \mathbf{r}_B(t_1)\right)^2 = c^2(t_2 - t_1)^2. \tag{27}$$

Assuming $t_2 - t_1 = t_{12}$ to be a short time and neglecting terms of second and higher order in $\omega\, t_{12}$ we have according to (23)

$$\mathbf{r}_B(t_2) - \mathbf{r}_B(t_1) \approx t_{12}\frac{d\mathbf{r}_B(t_1)}{dt_1} = t_{12}\,\boldsymbol{\omega} \times \mathbf{r}_B(t_1).$$

We can now write instead of (27)

$$l^2 + 2t_{12}\,\mathbf{l}(t_1)\left(\boldsymbol{\omega} \times \mathbf{r}_B(t_1)\right) \approx c^2 t_{12}^2. \tag{28}$$

Similarly, we find for the time t_{21} of the travel back

$$l^2 - 2t_{21}\,\mathbf{l}(t_1)\left(\boldsymbol{\omega} \times \mathbf{r}_B(t_1)\right) \approx c^2 t_{21}^2. \tag{29}$$

We note that according to the rules for the multiplication of vectors

$$\mathbf{l}(t_1)\left(\boldsymbol{\omega}\times\mathbf{r}_B(t_1)\right) = \boldsymbol{\omega}\left(\mathbf{r}_B(t_1)\times\mathbf{l}(t_1)\right)$$

and further

$$\left|\,\mathbf{r}_B(t_1)\times\mathbf{l}(t_1)\,\right| = S_{OAB}\,,$$

where S_{OAB} means the area of the triangle OAB, which is formed by the points A, B and the origin O of the coordinate system. Further we may write

$$\boldsymbol{\omega}\left(\mathbf{r}_B(t_1)\times\mathbf{l}(t_1)\right) = 2\omega S'_{OAB}\,,$$

where S'_{OAB} means the area of the projection of the triangle OAB on to a plane which is normal to $\boldsymbol{\omega}$. The sign of S'_{OAB} has to be chosen to be positive or negative according to whether the cycle $O \to A \to B \to O$ corresponds to a rotation with positive or negative angular velocity.

Similarly we find

$$\mathbf{l}(t_1)\left(\boldsymbol{\omega}\times\mathbf{r}_A(t_1)\right) = 2\omega\, S'_{OAB}$$

Subtracting (28) from (29) and dividing by $t_{12} + t_{21}$ we find

$$t_{12} - t_{21} = 4\omega\, S'_{OAB}/c^2. \tag{30}$$

§ 18. The relation (30) has the following physical meaning. We denote the times of travel from O to A and back by t_{01} and t_{10}, respectively, and similarly the times of travel from O to B and back by t_{02} and t_{20}. A light signal starting from O and reflected by suitable mirrors in A and B needs the time $t^+ = t_{01} + t_{12} + t_{20}$ to travel around OAB if the direction of travel is $O \to A \to B \to O$.

For the time of travel in the opposite direction one can write

$$t^- = t_{02} + t_{21} + t_{10}.$$

Considering that according to § 16 $t_{01} = t_{10}$, $t_{20} = t_{02}$ we find

$$t^+ - t^- = t_{12} - t_{21} = 4\,\omega\, S'_{OAB}/c^2.$$

The difference in the times of travel around OAB in the positive and negative direction is thus proportional to the projection S'_{OAB} of the triangle OAB on a plane normal to $\boldsymbol{\omega}$ and proportional to ω.

This result can easily be generalized. Consider a polygon formed by the points $O, A_1, A_2, \ldots, A_k$ which are at rest with respect to K. Denoting the times of travel from A_l to A_k, $l, k = 1, 2, \ldots, n$ by t_{lk} and the times of travel from O to A_k and from A_k to O by t_{0k} and t_{k0}, $k = 1, 2, \ldots, n$, respectively, one obtains for the times of travel around the polygon in opposite directions

$$t^+ = t_{01} + t_{12} + t_{23} + \ldots + t_{n-1\,n} + t_{n0}$$

and

$$t^- = t_{01} + t_{n\,n-1} + \ldots + t_{32} + t_{21} + t_{10}.$$

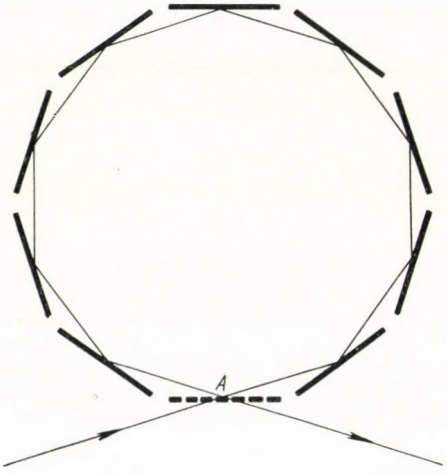


Fig. 4. Beam of light guided by mirrors around a disk

The difference in the times of travel is thus, using relation (30) for the individual times of travel,

$$t^+ - t^- = \frac{4\omega}{c^2}\,(S'_{012} + S'_{023} + \ldots + S'_{0n-1n} + S'_{0n1}).$$

Taking into account that the projections S'_{0kk+1} have to be taken as positive and negaive, respectively, according to whether the cycle $O - A_k - A_{k+1} - O$ represents a rotation in the positive or negative direction we conclude from a simple geometrical consideration that the sum in the bracket on the right-hand side is equal to the area S'_p, i.e. equal to the projection of the area of the polygon on a plane which is normal to ω. We have thus generally

$$t^+ - t^- = \frac{4\omega}{c^2}\,S'_p. \tag{31}$$

§ 19. The relation (31) can be understood in the following simple way. Consider a disk which rotates around its own axis with an angular velocity ω and at the periphery of which mirrors are located tangentially. (Fig. 4.) By the help of these mirrors a light signal starting in a point A on the periphery is reflected around the disk to reach A again. It travels along a polygon which in good approximation, however, may be regarded as a circle.

The length of the circular path is $2\pi R$ and the time of travel $T = 2\pi R/c$, where R means the radius of the disk. If the disk does not rotate the signal

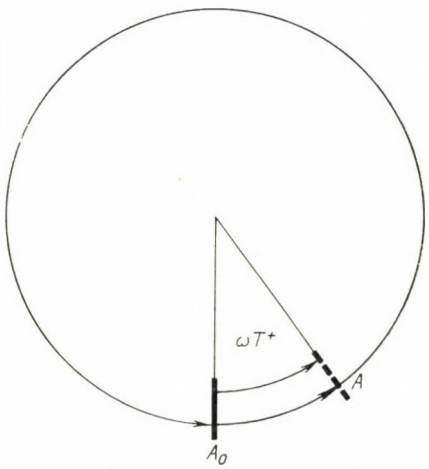


Fig. 5. Time T in which a beam starting in A_0 again reaches that point after circling round the disk. Time T^+ in which a beam starting in the moving point A circles round the disk and catches up with A. $T = 2\pi R/c$, $T^+ = (2\pi + \omega T^+)R/c$

emitted from A will come back to A after the time T. If, however, the disk rotates, the point A moves with it and at the time T the signal reaches only the original location A_0 of A, which now no longer coincides with A. In case the direction of travel is the same as that of rotation, A will have hurried in front of the light signal and the time t^+ in which the signal reaches actually the point A is longer than T.

Analogous to the considerations of § 11, where the translational motion was dealt with, we find here, that the light signal reaches the point A at that time at which it has made up for the length of path $2\pi R$. We thus have

$$2\pi R + \omega R t^+ = c t^+$$

or

$$t^+ = \frac{2\pi R}{c - \omega R} . \tag{32a}$$

In case the signal travels around the disk in a direction opposite to the rotation, the point A runs towards it and the signal meets the point A at a time

$t^- < T$. We have

$$2\pi R - \omega R t^- = c t^- ,$$

thus

$$t^- = \frac{2\pi R}{c + \omega R} . \tag{32b}$$

From (32a) and (32b) we have a

$$t^+ - t^- = \frac{4\pi R^2 \omega}{c^2 - \omega^2 R^2} . \tag{33}$$

Neglecting $(\omega R)^2$ in comparison with c^2 we find

$$t^+ - t^- \sim 4\omega S/c^2 , \tag{34}$$

where $S = R^2\pi$ is the area around which the light travelled. Equ. (34) is a particular case of the general relation (33).

The Sagnac—Michelson experiment

§ 20. The considerations of §§ 16—19 may be used to determine the rotational motion of some arrangement relative to the ether. Experiments based on this principle were actually carried out by SAGNAC [5], POGÁNY [6] and MICHELSON and GALE [7] and the scheme of one such experimental arrangement is shown in Fig. 6.

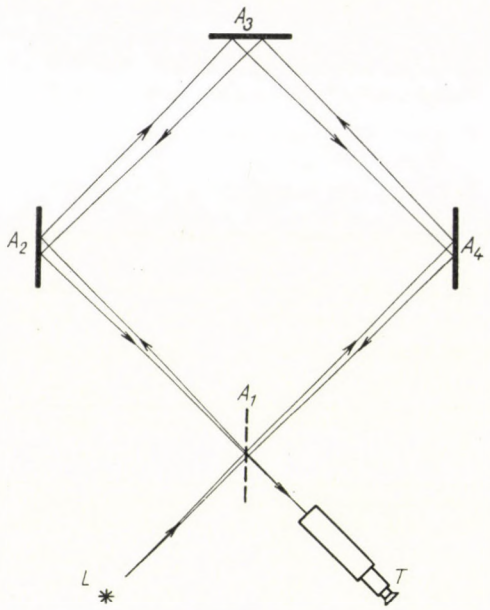


Fig. 6. Scheme of the SAGNAC experiment

A light signal falls on a semitransparent mirror A_1 at an angle of 45° and here is split into two components: the one passes through the mirror, the other is reflected. The component of the beam which is transmitted is reflected back to A_1 with help of further (not transparent) mirrors A_4, A_3 and A_2 which together with A_1 form the corners of a square. The returning beam is split again by A_1, the transmitted component falling in the telescope F. The reflected component of the beam which originally fell on A_1 travels around the square formed by the mirrors in the opposite direction to that of the transmitted component of the original beam. With help of interferometric methods the mirrors are adjusted in such a way that the geometric paths of the light beams travelling in opposite directions are of equal length. If the arrangement does not rotate, one thus expects that the times of travel of the two light beams are equal and that the two beams enter the telescope with equal phase. In case the beam incident on A_1 is narrow, we observe interference fringes in A_1.

Making the arrangement rotate (together with light source and telescope), the light beams travelling in opposite directions show a phase shift

$$\Delta\lambda = c(t^+ - t^-)$$

in the rotating system and the interference fringes shift in the telescope F.

The experiment described here was actually carried out and the phase shift was observed.

It occurred to MICHELSON [8] to demonstrate the rotation of the Earth relative to the ether with help of the described arrangement. He succeeded (see [9] and [7]), using an arrangement of the kind shown in Fig. 6 which was fixed to the Earth, in showing that there is a phase shift $\Delta\lambda$ which corresponds to the angular velocity of the Earth:

$$\omega_E = 2\pi/1 \text{ day} \approx 10^{-4} \text{ sec}^{-1}.$$

We may remark that in the original form of the SAGNAC experiment the arrangement at first is at rest relative to the Earth and is then rotated with a great angular velocity, the phase shift caused by the rotation being observed directly. In the MICHELSON—SAGNAC experiment, however, such a comparison cannot be made, as the angular rotation of the Earth is constant.

This difficulty was overcome by MICHELSON by adjusting interferometrically the individual paths A_1A_2, A_2A_3, ... These adjustments are not being influenced by the rotation of the Earth, as only such beams are being used, which travel to and fro on the same path, and which thus do not enclose any area. This corresponds thus to the case $S'_p = 0$ and therefore according to equ. (31) as a result of the rotation no phase shift occurs between the beams which were used for the individual interferometric measurements. Thus if the

distances $A_1A_2, A_2A_3 \ldots$ are determined and adjusted interferometrically the rotation of the Earth is of no account. It is thus possible accurately to adjust the position of the mirrors A_1, A_2, A_3 and A_4 and then — with the help of the arrangement so adjusted — to determine the effect of rotation. The detailed description of the difficult experiment can be found e.g. in the book by LAUE [10].

Experiments for the determination of the translational motion

§ 21. It follows from the SAGNAC experiment that the ether near the Earth does not take part in the rotation of the Earth. One may thus assume that the Earth does not take the ether with her and that she is in translational motion relative to the ether.

It seems obvious to ask whether it is possible to determine the translational motion of some apparatus relative to the ether with the help of an experiment of the type of the SAGNAC—MICHELSON experiment.

Equ. (30) makes it plain that this is not the case. A translational motion may be regarded as the limiting case of rotation with disappearing angular velocity and thus one concludes from (30) that $t^+ - t^-$ disappears in the case of translation.

The conclusion, however, is not quite exact as (30) was derived by neglecting certain quantities. However, $t^+ - t^-$ may be determined exactly from the strict relations (18) and (22). Subtracting these equations one finds after a simple transformation

$$t_{12} - t_{21} = 2\mathbf{l}\mathbf{v}/(c^2 - v^2). \tag{35}$$

Consider thus the polygon $A_1, A_2, \ldots, A_n$ introduced in § 18. Assuming that the polygon moves with the constant velocity $\mathbf{v}$ relative to K_0 the difference in the times of travel is found to be

$$t^+ - t^- = t_{12} + t_{23} + \ldots + t_{n-1n} + t_{n1} - (t_{21} + t_{32} + \ldots + t_{n\,n-1} + t_{1n}). \tag{36}$$

Denoting the vector pointing from A_k to A_{k+1} by $\mathbf{l}_{k,k+1}$ it follows from (35) that

$$t_{k,k+1} - t_{k+1,k} = 2\mathbf{l}_{k,k+1}\,\mathbf{v}/(c^2 - v^2),$$

and thus according to (36)

$$t^+ - t^- = 2(\mathbf{l}_{12} + \mathbf{l}_{23} + \ldots + \mathbf{l}_{n-1n} + \mathbf{l}_{n1})\,\mathbf{v}/(c^2 - v^2).$$

As, however, the sum of the vectors inside the bracket is zero for a closed polygon, we find that in case of translation the relation

$$t^+ - t^- = 0$$

is valid exactly.

Thus, by comparison of times of travel no information can be obtained as to the translational motion of the apparatus.

MAXWELL [11] proposed another method, later carried through by MICHELSON [12] by which it was hoped to be able to determine the motion of the measuring arrangement relative to the ether. The experiment, however, did not come up to expectation and led to a negative result. In the following we analyse the experiment.

§ 22. Consider three points A, B and C which are at rest relative to K, i.e. which move with a constant velocity $\mathbf{v}$ relative to K_0. The line AB be parallel and AC normal to $\mathbf{v}$. Denoting the distances $AB = AC$ by l the times a signal needs starting from A to reach B and C, respectively, and to return to A are given by the following expressions:

$$\begin{aligned} T_1 &= t_{12} + t_{21} = \frac{2lc}{c^2 - v^2}, \\ T_2 &= t_{13} + t_{31} = \frac{2l}{\sqrt{c^2 - v^2}} \end{aligned} \tag{37}$$

(compare § 13 equ. (21)), thus

$$T_2 - T_1 = \frac{-2l}{c^2 - v^2}\left(c - \sqrt{c^2 - v^2}\right) = \frac{-2lv^2}{(c^2 - v^2)\left(c + \sqrt{c^2 - v^2}\right)}.$$

Neglecting higher powers of v^2/c^2 we find approximately

$$T_2 - T_1 \approx \frac{-lv^2}{c^3}. \tag{38}$$

The difference $T_2 - T_1$ of the times of travel can be measured with the help of interference. Place at A a semi-transparent mirror which subtends angles of 45° with the line AB and AC (see Fig. 7) and in B and C be placed mirrors which are normal to the lines AB and AC. A light beam incident on A and parallel to the direction AC is split into two components. The transmitted component falls on the mirror in C and is reflected. This reflected beam is again reflected in A and falls in the telescope F. The original beam reflected in A falls on B and is again reflected to fall back on A and is here transmitted

to the telescope in F. The beams which reached the telescope F on two different paths give an interference pattern.

Provided the system moves with the velocity v parallel to AB we expect according to equ. (38) a phase shift

$$\Delta\lambda = c(T_2 - T_1) = -lv^2/c^2$$

between the two beams falling into F.

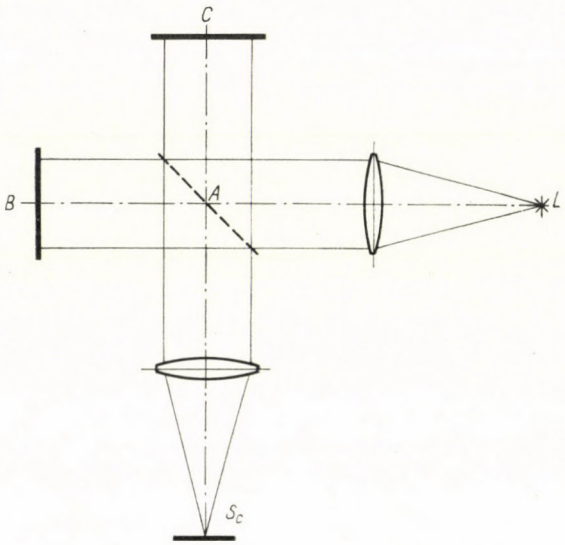


Fig. 7. Scheme of the MICHELSON interferometer. L light source, Sc screen

§ 23. The experiment dealt with above cannot, however, be carried out in practice. So as to draw some conclusion from the observed phase shift as to v we have to adjust the mirrors in such a way that the distances AB and AC are exactly equal. However, the only available method to adjust distances to the required accuracy is just MICHELSON's interferometric method which itself is influenced by the motion of the system. Thus, so as to be able to make a statement concerning the value of v one would have to make the distance AB and AC equal by some method which is not based on the propagation of light.

It is interesting here to compare the problem of adjusting the mirrors, on the one hand in the experiment by SAGNAC and in the experiment described above — on the other. In the SAGNAC experiment it is possible to adjust the mirrors exactly with the help of light which is reflected there and back, i.e. with the help of interference, as this kind of adjustment is not disturbed by the

rotational motion of the arrangement. Such adjustment of the arrangement of Fig. 7 is, however, not possible, as here it is changed on account of the translation by just that amount which we wish to measure.

§ 24. The difficulty described above may apparently be overcome in the following way. To begin with mirrors B and C are adjusted so that $\Delta\lambda = 0$. In case AB is parallel to v, there is $\Delta\lambda = 0$ for such distances $AB = l_1$ and $AC = l_2$ for which $T_1 = T_2$, i.e. according to equ. (37) if

$$l_1 = l_2 \sqrt{1 - v^2/c^2} \,. \tag{39}$$

Adjusting the arrangement so that $\Delta\lambda = 0$ we know that the lengths of the arms l_1 and l_2 are related to each other according to (39). Turning now through 90° the apparatus which was adjusted to show no phase shift, i.e. for which $\Delta\lambda = 0$, so that now AB becomes normal, AC, however, parallel to v, the times of travel there and back should now be given by

$$T_1' = \frac{2l_2\, c}{c^2 - v^2} \,,$$

$$T_2' = \frac{2l_1}{\sqrt{c^2 - v^2}} = 2l_2/c.$$

The difference between the times of travel after the rotation becomes thus

$$T_2' - T_1' = -\frac{2l_2\, v^2}{c(c^2 - v^2)} \,.$$

We thus expect that as a result of the rotation the phase shift becomes now

$$\Delta\lambda' = \frac{-2l_2\, v^2}{c^2 - v^2} \,.$$

Turning the apparatus through some arbitrary angle α the phase shift similarly turns out to be

$$\Delta\lambda'(\alpha) = \frac{-2l_2\, v^2}{c^2 - v^2} \sin^2\alpha. \tag{40}$$

Thus slowly turning the apparatus one should expect the interference pattern seen in the telescope to wander periodically sideways.

The experiment by Michelson and Morley

§ 25. The experiment actually carried through by MICHELSON and MORLEY as well as the very exact later experiments showed that contrary to expectation

$$\Delta\lambda'(a) = \text{const.}$$

The accuracy of the measurements was such, that fringe shifts (40) corresponding to a velocity v of the order of 1 km/sec should still have been observed. Physicists were greatly surprised at the negative outcome of MICHELSON's experiment. Had not MICHELSON during the many years in which he prepared his experiment been regarded as an oddity, who tried to prove the self-evident, namely the fringe shift?

The MICHELSON experiment, had it been carried out once only, could have been interpreted by the assumption, that by chance the Earth was at rest relative to the ether at the very time of the experiment. It would then have been possible to assume that the solar system moved with such a velocity (30 km/sec) relative to the ether that at the time of the experiment the motion of the Earth round the sun by chance compensated the translational motion of the solar system.

This assumption is, however, not correct, which follows from the fact, that the MICHELSON experiment, repeated at various times of the year, always had a negative result. Had by chance the Earth been at rest relative to the ether at the time of the first experiment, her velocity after half a year — when namely the direction of her motion was the opposite — should have been 60 km/sec relative to the ether.

The explanation given by LORENTZ and FITZGERALD of the negative outcome of the MICHELSON experiment is the most obvious, it was, however, rejected by EINSTEIN — for philosophical reasons. That the difference between the two concepts is indeed only a philosophical one was shown elswhere [13].

Interpretation of the Michelson experiment

§ 26. We endeavour in the following to formulate the idea of LORENTZ and FITZGERALD — somewhat more rigorously than this was done originally.

If there is no fringe shift when rotating the interferometer and the circumstances are such, that from simple geometrical considerations one would have to expect such a shift, this must be taken as a proof, that the lengths of the arms of the interferometer changed during the rotation, thus compensating

the expected phase shifts. In this way *the Michelson experiment may be regardep as an interferometric measurement of the contraction and dilatation, respectively, of the arms of the interferometer.*

The measurement of the deformations of the interferometer by help of the MICHELSON experiment has certain flaws and because of these, generally EINSTEIN's interpretation of the effect has been preferred. We are of the opinion that, although these defects may be deplorable, they do not affect the conclusions.

§ 27. Suppose, that a rod which is in translation (e.g. one arm of an interferometer) changes its length when rotated. Denote its length by $l(\alpha)$ and α be the angle which the rod subtends with the direction $\mathbf{v}$ of its translational motion. The result of the MICHELSON experiment may be understood if one supposes that

$$l(\alpha)/c(\alpha) = \text{const.}\,, \tag{41}$$

where $c(\alpha)$ is given by the expression in equ. (20) § 13. However, (41) cannot be correct, as it leads to the expectation $l(\alpha) \neq l(\pi - \alpha)$. This would mean that a rod after rotation through 180°, i.e. by a simple reversion, has to change its length.

A reasonable assumption removing the above difficulty and suitable to explain the outcome of the MICHELSON experiment is the following:

$$l(\alpha)\left(\frac{1}{c(\alpha)} + \frac{1}{c(\pi - \alpha)}\right) = \text{const.} \tag{42}$$

It follows from (42) with the help of equ. (20) § 13 choosing the constant suitably

$$l(\alpha) = l\sqrt{1 - \frac{v^2}{c^2}}\Big/\sqrt{1 - \frac{v^2}{c^2}\sin^2\alpha}\,. \tag{43}$$

Instead of (43) we may also write

$$l^2 = \left(l(\alpha)\cos\alpha\Big/\sqrt{1 - \frac{v^2}{c^2}}\right)^2 + (l(\alpha)\sin\alpha)^2\,. \tag{44}$$

It is obvious from (43) that $l = l\left(\frac{\pi}{2}\right)$ is the biggest length a rod can take up.

The geometrical meaning of (43) — as this can be seen from equ. (44) — is the following. If the length of the rod AB is l when the rod is normal

to the direction of velocity v and the rod is turned into a direction which subtends an angle α with v the component which is normal to v remains unaltered. The component, however, which is parallel to v contracts by a factor $\sqrt{1 - v^2/c^2}$.

Equ. (43) and (44) express the hypothesis of contraction of LORENTZ and FITZGERALD.

Wishing merely to explain the MICHELSON experiment, instead of (42) one might start from the more general relation

$$l(\alpha)\left(\frac{1}{c(\alpha)} + \frac{1}{c(\pi - \alpha)}\right) = l\left(\alpha + \frac{\pi}{2}\right)\left(\frac{1}{c\left(\frac{\pi}{2} - \alpha\right)} + \frac{1}{c\left(\frac{\pi}{2} + \alpha\right)}\right). \quad (45)$$

Considering, however, in addition the results of other experiments one may exclude those generalizations of (42) which are contained in (45).

§ 28. A blemish of the above considerations, which has been overemphasized, is that we cannot determine the value of v; in an actual case we thus cannot calculate by how much the arm of an interferometer changes during rotation,

This blemish, however deplorable it may be, does not alter the consequences. Even if we do not know v, i.e. the velocity of our apparatus relative to the carrier of the electromagnetic waves, and this means relative to the ether, we know for certain, that this velocity has changed in the course of a year. Since v is changing continuously its value must differ from zero for most of the year. At all times for which $v \neq 0$ the interferometer is being deformed when rotated, as otherwise — this follows from the elementary considerations which were discussed at length above — the rotation should have caused fringe shifts.

The result of the experiments thus proves, that deformations of e.g. the form (43) arise, although there are no ways or means to measure the actual deformation produced in a concrete case.

The philosophical aspect of this question has been discussed in detail elsewhere [13], in any case, it seems to us misleading to deny the existence of deformations, only because it is not possible in a concrete case to determine unambiguously which among the different possibilities is actually realized.

Measurement of the critical velocity in the system K.

§ 29. We have shown at the beginning of this article, that the critical velocity can be determined by measuring the forces between electrical charges and magnetic poles. The question may be asked, whether the described experi-

ments when carried out in a system K which moves with a velocity $\mathbf{v}$ relative to the ether, lead to the same result as when carried out in K_0 which is at rest with respect ot the ether.

We discuss first a specific problem, namely we calculate the Coulomb effect of a charge $-e$ on a charge $+e$, when both are at rest in K, thus when both move with a constant velocity $\mathbf{v}$ with respect to K_0.

The effect of a moving charge — to be exact — *would* have to be calculated with the help of the retarded potentials. In the following for the sake of simplicity we neglect the effect of retardation, since this, as is shown by a more exact calculation, influences only slightly the effect under consideration.

Denote by $\mathbf{r}$ the vector which points from $-e$ to $+e$. The Coulomb force which acts on $+e$ is then

$$\mathbf{F}_c = -\frac{e^2\,\mathbf{r}}{r^3}\,.$$

Besides of the Coulomb force the magnetic field of the moving charge $-e$ also acts on $+e$. Denoting by H the magnetic field strength the magnetic force is

$$\mathbf{F}_M = \frac{e}{c}\,(\mathbf{v}\times\mathbf{H}).$$

On the other hand H is given by

$$\mathbf{H} = -\frac{e}{c}\,(\mathbf{r}\times\mathbf{v})/r^3,$$

thus the magnetic force becomes

$$\mathbf{F}_M = -\frac{e^2}{r^3\,c^2}\,(\mathbf{v}\times(\mathbf{r}\times\mathbf{v})) = -\frac{e^2}{c^2}\left(\frac{v^2\,\mathbf{r}}{r^3} - \frac{\mathbf{v}(\mathbf{vr})}{r^3}\right).$$

The second term in the bracket in the last expression gives a component of force parallel to $\mathbf{v}$ which in general does not point in the direction from the one charge to the other; this component results in a torque $\mathbf{M}$. Writing for the total force $\mathbf{F} = \mathbf{F}_C + \mathbf{F}_M$ there is

$$\mathbf{M} = \mathbf{r}\times\mathbf{F} = +\frac{e^2}{c^2}\;\frac{(\mathbf{r}\times\mathbf{v})\,(\mathbf{rv})}{r^3}\,.$$

Denoting further by a the angle subtended by $\mathbf{r}$ with $\mathbf{v}$ we may write for the absolute value of the torque

$$M = \frac{1}{2}\;\frac{e^2\,v^2}{rc^2}\sin 2a. \tag{46}$$

The experiment by Trouton and Noble

§ 30. Equ. (46) seems to provide a means by which the velocity v of a system relative to the ether may be determined. The following experiment had this aim and was carried out by TROUTON and NOBLE.

A charged condenser was hung on an elastic thread in such a way that the condenser plates subtended an angle of 45° with the direction of the motion of the Earth. After adjusting the position of the condenser it was arrested with respect to the frame and together with the frame was turned through 90° around a vertical axis. In the new position the arrestation of the condenser was then released and it was investigated whether or not the condenser remained in the position into which it was turned.

The theory of the effect is as follows: if the torque (equ. 46) acts upon the condenser in its original position then it will be necessary to give an opposite torque to the thread upon which it is suspended, so as to achieve equilibrium. Turning round the arrangement the electric torque should change its sign and therefore in the new position, if the elastic torque has remained the same, the elastic and electric torques should no longer compensate each other and the condenser after release should turn through a certain angle.

The result of the experiments was, however, negative. When the whole frame was rotated the condenser remained in the same position relative to the frame that it had in the original arrangement.

§ 31. One may be tempted to conclude from the negative outcome of the TROUTON—NOBLE experiment that the moment (46) does not exist at all. Such a conclusion, however, does not appear to be justified. If the moment (46) would not exist, this would be in contradiction to MAXWELL's equations which have been verified to a high degree of accuracy.

The correct explanation of the negative result seems to be, that a moment **M** and simultaneously also an opposite moment $-\mathbf{M}$ arises in the arrangement, the latter exactly compensating the effect of the electromagnetic moment.

At closer inspection this is not so very surprising. The effect of the force acting between the charges $+e$ and $-e$ has to be exactly compensated by mechanical forces to keep the plates of the condenser at a constant distance. So as to understand the behaviour of the system when rotated we have to consider the effect of translation not only on the electrical forces but also on the mechanical ones compensating the former. It seems obvious to suppose that the mechanical forces which exactly compensate the electromagnetic forces in one arrangement change — when the system is rotated — in such a manner as to compensate these exactly also in the new arrangement so that no observable effect remains.

§ 32. From the generalization of the negative results of the MICHELSON, the TROUTON—NOBLE and a series of similar experiments, which aimed at determining the velocity $\mathbf{v}$ of the apparatus relative to the ether, one suspects that here we do not have to do with a series of chance events, but that the negative results are based on a general law of nature.

This law of nature on which these phenomena are based has elsewhere been formulated by us as the LORENTZ principle.

A detailed analysis of the LORENTZ principle shows that the measurement of the critical velocity carried through with the help of an arrangement in translation relative to the ether leads to the same result as that obtained with help of an arrangement that rests relative to K_0.

A remark upon Römer's experiment

§ 33. In § 6 we have described the method of RÖMER which was used to determine the velocity of light by observing the satellites of Jupiter.

In this experiment the velocity of the light coming from Jupiter and reaching the Earth is measured relative to the solar system.

If the solar system moves with a velocity $\mathbf{w}_0$ relative to the ether, one will have to expect that the result of the experiment depends on the angle subtended by the direction Jupiter—Earth with the direction of the velocity $\mathbf{w}_0$ at the time of observation. Because of the orbital motion of Jupiter this angle changes and thus one might expect that the values of the velocity of light repeatedly determined by the method of RÖMER during one revolution of Jupiter would show variations. If, for instance, the solar system moves with a velocity $\mathbf{w}_0$ relative to the ether the time by which the satellites of Jupiter are seemingly late during half a year is given by

$$T = \frac{2R}{c - \mathbf{w}_0 \mathbf{e}} .$$

Here $\mathbf{e}$ means the unity vector which points from the Earth to Jupiter and R is the mean radius of the orbit of the Earth. (The motion of Jupiter during one revolution of the Earth is neglected here.)

Denote by $\mathbf{w}'$ the projection of $\mathbf{w}_0$ on the ecliptic. At a time at which $\mathbf{w}'$ points in the direction Earth—Jupiter we hawe $\mathbf{w}_0 \mathbf{e} = \mathbf{w}'$ and the observed delay has a minimum value:

$$T_{\min} = \frac{2R}{c + w'} . \tag{47}$$

After half a revolution of Jupiter w′ points in the opposite direction and we obtain for the delay T the maximum value

$$T_{\max} = \frac{2R}{c - w'} . \tag{48}$$

The value of the difference

$$\Delta T = T_{\max} - T_{\min} \approx \frac{4Rw'}{c^2 - w'^2} \approx \frac{4Rw'}{c^2} \tag{49}$$

should give information on w' and from the configurations Earth—Jupiter at which the difference has its maximum and minimum value, respectively, one should be able to determine the direction of w'.

According to astronomical estimates the velocity with which the solar system moves in the Galaxy is about 300 km/sec, and its projection on the plane of the ecliptic is about half this value. Identifying this velocity with w_0 we may assume $w' \approx 150$ km/sec and with this value one obtains from (49) $\Delta T \approx 1$ sec. The accuracy with which the motion of the satellites of Jupiter can be observed is at present not sufficient to determine whether this effect exists.

§ 34. In this analysis we have aimed at discussing the result of such experiments which have actually been carried out. The experiment described above to determine the translational motion w_0 of the solar system — it is true — has not been carried out and seems to be hopeless with present-day techniques. As the experiment, however, takes up a problem which is of principal importance we analyse it here — basing ourselves on real experiments — and make a guess as to its outcome. Our suggestion is that the velocity of light as measured by RÖMER's experiment would always have the same value independent of the constellation of Jupiter and that thus the experiments are not suitable for a determination of the value of w_0.

§ 35. Our suggestion is based on the assumption that the LORENTZ principle is valid also for the mechanism of the solar system. With help of the LORENTZ principle it can be shown exactly that the various effects resulting from the translation of the solar system relative to the ether fully compensate each other, and further that the experiment of RÖMER would be likely to yield the value c for the velocity of light, independent of the translational motion of the solar system.

We do not reproduce here the exact calculation, but instead content ourselves with an approximative consideration, which at the same time has the advantage that it makes much clearer the physical mechanism by which the various effects compensate each other.

§ 36. We use the rotating Earth as clock in astronomical observations.* If, however, the orbital velocity of the Earth changes we have to expect that the angular velocity of the rotation also changes.

Denoting by $m_E(w)$ the mass of the Earth we expect

$$m_E(w) = \frac{m_E(0)}{\sqrt{1 - w^2/c^2}}$$

at a time at which the velocity of the Earth relative to the ether is $\mathbf{w}$. As the mass depends on the velocity also the moment of inertia and thus according to the principle of conservation of angular momentum also the angular velocity ω depend on the velocity. We expect

$$\omega(w) = \omega(0)\sqrt{1 - w^2/c^2}. \tag{50}$$

By using the rotation of the Earth for the measurement of time the variations of the angular velocity which are caused by the variations of $\mathbf{w}$ are disregarded. We use as measure of the time, e.g.

$$t' = \frac{1}{\omega(0)} \int_0^t \omega(w)\, dt\,; \tag{51}$$

t' is thus the measure of the time elapsed between an arbitrary moment $t = 0$ and a moment t.

The connection between the measure t' and t becomes clear when substituting (50) in (51) and neglecting terms of higher order. We find

$$t' = t - \frac{1}{2c^2} \int_0^t w^2\, dt.$$

§ 37. If the Sun moves with a constant velocity $\mathbf{w}_0$, we may put

$$\mathbf{w} = \mathbf{w}_0 + \mathbf{v}\,,$$

where $\mathbf{v}$ means the velocity of the Earth in her orbit around the Sun.

Taking the orbit of the Earth to be circular, we may put $\mathbf{v}^2 = v^2 = \text{const.}$ We then have

$$\frac{1}{2c^2} \int_0^t w^2\, dt = \frac{t}{2c^2}(v^2 + w_0^2) + \frac{1}{c^2}(\mathbf{w}_0\, \mathbf{r}),$$

* Thus we obtain the measure of stellar time which deviates to some extent from the measure of ephemeris time. We shall not discuss here the problem implied.

where $\mathbf{r} = \int_0^t \mathbf{v}\, dt$ is the vector which describes the displacement of the Earth relative to the Sun during the time interval $0 \to t$. We thus find

$$t' = t\left(1 - \frac{1}{2c^2}(v^2 + \mathbf{w}_0^2)\right) - \frac{1}{c^2}(\mathbf{w}_0\,\mathbf{r}). \tag{52}$$

The measures $T'_{\min}$ and $T'_{\max}$ of the times $T_{\min}$ and $T_{\max}$ as given in (47) and (48) are obtained — putting in (52) $t = T_{\max}$, $\mathbf{r} = 2R\mathbf{e}$ in the first and $t = T_{\min}$, $\mathbf{r} = -2R\mathbf{e}$ in the second case — as

$$T'_{\max} = \frac{2R}{c - w'}\left(1 - \frac{1}{2c^2}(v^2 + w_0^2)\right) - \frac{2Rw'}{c^2},$$

$$T'_{\min} = \frac{2R}{c + w'}\left(1 - \frac{1}{2c^2}(v^2 + w_0^2)\right) + \frac{2Rw'}{c^2}.$$

Expanding these expressions in series of powers of $1/c^2$ we find, neglecting terms of fourth order

$$T'_{\max} - T'_{\min} = 0.$$

The above analysis thus shows, that the effect which is to be expected according to § 37 does not make itself felt in the observed measures, because the translation of the solar system disturbs our measurement of time to such a degree that the errors of the disturbed clocks exactly compensate the expected effect.

Concluding remarks

§ 38. We see from the considerations presented here that the experimental results relating to the velocity of propagation of light can be easily understood assuming that the propagation takes place isotropically with the critical velocity c. The propagation is isotropic relative to the carrier of electromagnetic waves, thus, as we are not afraid of the old-fashioned expression, relative to the ether.

By the aid of measurements of the propagation of light of the kind of SAGNAC's experiment one can determine whether the apparatus rotates relative to the ether. However, one cannot determine whether the apparatus is in translational motion relative to the ether or whether it is not.

Translation although influencing the apparatus to a noticeable degree does not modify the final results of measurement as the various effects brought about by it compensate each other.

The mechanism by which such a compensation takes place may be understood with the help of the LORENTZ principle, which has not been discussed here.

§ 39. The conclusion which EINSTEIN has drawn from the facts is that the effects, which compensate each other exactly and thus cannot be observed, do not exist. Thus it is generally concluded according to EINSTEIN that light is propagated isotropically relative to every system of reference K which does not show rotation.

EINSTEIN's conclusion, however, appears to us wrong philosophically. Above all the statement that light is propagated isotropically relative to every coordinate system K which is in translational motion, is such as to be refuted by primitive kinematic considerations (see §§ 9—13).

We, of course, do not want to suggest, that EINSTEIN was not aware of these kinematic considerations. EINSTEIN's argument amounts to saying — although EINSTEIN has used a quite different terminology — that it is possible to introduce in a system K, which is in translation relative to K_0, such *measures* for coordinates and times that one obtains with these the same *measures* for the velocity of light in the various directions.

Thus introducing in the various systems K different conventions for the measurement we obtain, choosing these conventions suitably, apparent isotropy in the various systems K; we then find that

$$c'(\mathbf{e}) = \text{independent of } \mathbf{e}\,,$$

where $c'(\mathbf{e})$ means the velocity of propagation of light in the direction $\mathbf{e}$, measured in the unit of the corresponding conventions.

These conventional units are of great practical value, as they well adapt themselves to certain dynamic properties of matter and of the electromagnetic field. However, we must not deceive ourselves as regards the following: If the propagation of light is found to be isotropic in two systems K' and K'' which are in motion relative to each other, so that

$$c'(\mathbf{e}) = \text{const.} \quad \text{and} \quad c''(\mathbf{e}) = \text{const.} \tag{53}$$

this can solely be ascribed to the fact that the measures in K' and K'' had been chosen in such a way as to conceal the difference in the propagations of light relative to K' and K'', respectively.

§ 40. Consider again MICHELSON's interferometer experiment. The result, that the propagation of light is isotropic can be deduced from the interferometer experiment on the *assumption* that during rotation of the apparatus the

arms do not change their lengths. If thus according to this assumption the lengths of the arms of the interferometer are taken to be the *definition* of the unit of length in the various directions, we can conclude that using units thus defined propagation of light is found to be isotropic in the sense of equ. (45).

Repeating now the experiment at another time of year, at which the translation of the Earth differs from that in the first experiment, we may again *define* the unit of length by the length of the arms of the interferometer and we again find isotropy in the new units.

The situation when comparing the two results of measurement can be expressed in the way of equ. (53). As, however, the two results are expressed in different units one cannot conclude from them that nothing has changed. No longer can we "define" that the two definitions of units are identical. It is no longer a question of definitions but of facts, whether e. g. a rod after acceleration has or has not changed its inner physical state and with this also its length.

It is just the negative outcoume of the MICHELSON experiment obtained at different times of the year which proves clearly that the definitions introduced above are not identical, and that the dimensions of the interferometer change when its state of translation changes.

Apart from the result of the MICHELSON experiment it follows from purely dynamic considerations, that a closed physical system suffers deformations when its state of translation is being altered. These considerations were presented in another paper. We thus have good reasons to suspect that the change of the orbital velocity of the Earth is not without effect on the interferometer — and in this way we can understand the apparently paradoxical result of the MICHELSON experiment.

§ 41. The mistake in the EINSTEINian argumentation lies in that the difference between measure and real length is being denied. In our interpretation it is of utmost importance to distinguish clearly between measure and the object of measurement. The distance between two towns is something which is objective, but we may express this distance in cm or km, and we obtain accordingly different *measures*.

We here quote a typical statement taken from a popular article on relativity theory:

"If namely the older physicist, and thus also LORENTZ talked about the length of some rod he was convinced that he talks about an inherent property of the rod which unambiguously sticks to it and which exists without any measurement. Nothing is thus more natural than that LORENTZ could imagine the shortening of the arm of the interferometer only in the way that in the rod an objective physical change of state occurs. Today, according to EINSTEIN, we already regard this in a completely different way. Bodies have no inherent

properties whatsoever which without measurement could even be given meaning."

This quotation demonstrates the confusion between measure and a quantity which is objectively existing.

Naturally — as this has been correctly seen by LORENTZ and many others — a rod has a length independent of the fact whether or not we have given a definition of length — such quantities have even existed long before there were any human beings and professors who thought it necessary to give definitions.

If we want to characterize objectively existing quantities by measures, we have to make it quite clear by which method we intend to obtain the measures. Definition is given a certain latitude by the circumstance that the method of measurement is not unambiguously determined. We have, however, to be wary of definitions which do not correspond to the inner structure of the object to be measured — such definitions lead to contradictions, as is the case e.g. when choosing definitions of time and space coordinates which are not appropriate.

REFERENCES

1. A. EICHENWALD, Ann. d. Phys. **11**, 1, 1903.
2. E. BERGSTRAND, Arkiv Fysik, **2**, 119, 1950; **3**, 479, 1951.
3. R. W. COHEN, K. M. CROWE and S. W. M. DUMOND, Fundamental Constants of Physics, Interscience Publ., N.Y. 1957, p. 111.
4. E. B. ROSA and N. E. DORSAY, Bull. Bureau Stand., **3**, 605, 1907.
5. G. SAGNAC, Comptes Rendus, **15**, 708 and 1410, 1913.
6. B. POGÁNY, Ann. d. Phys., **8**., 217, 1926.
7. A. A. MICHELSON and H. G. GALE, Astrophys. J., **61**, 140, 1925.
8. A. A. MICHELSON, Phil. Mag., **8**, 716, 1904.
9. A. A. MICHELSON, Astrophys. J., **61**, 137, 1925.
10. M. v. LAUE, Die Relativitätstheorie, Fr. Vieweg and Sohn, Braunschweig, 1955, Vol. I, p. 15 ff.
11. Ibid. p. 22.
12. Ibid. p. 197; footnote 29.
13. L. JÁNOSSY, Philosophical Analysis of the Special Theory of Relativity. In Russian: Вопросы Философии **8**, 101; **9**, 89, 1961; in Hungarian: Filozófiai Szemle, **6**, 153, 1962; preprint in English.

РАЗМЫШЛЕНИЯ О ПРОБЛЕМЕ ИЗМЕРЕНИЯ СКОРОСТИ СВЕТА

Л. ЯНОШИ

Резюме

Анализируются явления, связанные с распространением света чтобы заново рассмотреть какие выводы могут быть сделаны из самих опытов. Мы приходим к выводу, что получается вполне ясная картина если, в противоположность обычно принятой точке зрения, предполагать, что электромагнитные волны имеют носитель, который можно принять за эфир. Философские вопросы, касающиеся введения вновь эфира как носителя электромагнитных волн рассмотрены автором в других статьях [6].

A CALORIC METHOD FOR THE DETERMINATION OF THE ANODE FALL OF DISCHARGES

By

G. Szigeti, G. Lakatos and J. Bitó

INDUSTRIAL RESEARCH INSTITUTE FOR TELECOMMUNICATION TECHNIQUE, BUDAPEST

(Received 17. XII. 1963)

A discharge without a positive column is produced by the authors and its cathode fall is determined by means of a caloric method. The surface current loading and the surface watt loading of the anode and cathode resp. are adjusted to a previously determined value. The experimental conditions are described and the theoretical and experimental principles of the testing method are discussed in detail. The authors examine the conditions under which the method can be applied and report their results obtained in the case of mercury-argon-neon-filled discharge tubes, which agree well with those of their previous examinations.

1. Introduction

For the determination of the anode fall of certain discharge forms various electronoptical-, probe-, measurement- and other experimental methods are known [1]. Generally, the value of the anode fall is only a fraction of that of the cathode fall and at the same time the region of the anode fall is confined to a relatively small part of the field. Because of this often difficulties arise in the course of investigations which aim at a more exact knowledge of the basic processes taking place at the anode of the discharge tubes and the energy transfer- and oscillation-phenomena connected with them.

The methods most generally used for the examination of the anode fall are the probe measuring procedures [1]. With their help the value of the anode fall can be given to different accuracies and the dimensions of the anode dark spaces constituting the anode-fall region can be determined by optical observations. From the dimensions of the anodic dark spaces and the extent of the anode fall the electric field strength within this domain can be determined on the basis of the following relation given by von Engel [2] which represents a parabola:

$$E_x = \frac{3}{2} \frac{V_a}{d} \left[1 - \left(\frac{2x}{d} - 1 \right)^2 \right], \tag{1}$$

where E_x means the electric field strength in the direction of the discharge axis X (Volt $\cdot$ cm^{-1}), V_a the value of the anode fall (Volt), d the length of the anodic dark space (cm), x the distance from the anode (cm). The function E_x

has a maximum in the point $x = d/2$ and its value is zero in the points $x = 0$ (anode) and $x = d$ (limit of the positive column at the anodic side).

The examinations performed until now indicated [1] that discharges of positive, negative and zero anode fall are also possible; the character of the discharge processes around the anode varied accordingly.

In order to examine these processes more fully the exact knowledge of the anode fall itself is indispensable. In the course of this article a thermal method is described which was used for its determination. In these investigations the energy taken up by the ions arising incidentally within the anodic field was neglected. In certain cases it may be imagined that e.g. the electrons coming from the positive column and entering the anodic field, there obtain an energy such that during their way to the anode they are able to ionize once again. This occurs especially in case of discharges where the anode fall is high and also the region of anode fall is extensive enough in the direction of the drift current. Then the electrons can obtain adequate energy for the ionization and further it becomes very probable that they will collide with neutral atoms when traversing the extensive anode fall section.

In our case the value of the anode fall was about 3 V, i.e. it was small. The dimension of the anodic dark space was of the order of some mm and this is also a relatively small distance. Therefore the loss of energy of the electrons could indeed be neglected and the method led to good results.

The method described here can also be applied to general cases. However, when also ionization processes take place within the anodic field the results will be of only limited accuracy.

2. Experimental conditions

The examinations were carried out in ambient air of a temperature of $25 \pm 1°$ C, with discharge tubes of 38 mm outer diameter and 1 mm wall thickness. The length of the tubes was chosen in such a way that within the examined discharge there were only cathode-fall and anode-fall regions. The positive column was missing. The electrodes protruded to a distance of 100 mm into the discharge tube. The cathode consisted of a tungsten coiled coil provided with an electron-emitting coating and on both sides of the cathode there were nickel auxiliary electrodes placed parallel to each other. The auxiliary electrodes had the same potential as the cathode coil.

The anode of the tube was an iron cylinder of high thermal inertness, 7 mm outer and 1 mm inner diameter jacket height, in the interior of which an insulated tungsten coil was placed which was provided with the necessary terminals.

Also the anode reached into the discharge tube, 100 mm from its end. The tungsten coil within the iron cylinder provided with an insulating cement coating was a cathode used for radio valves of the type Tungsram EF 89 and its terminals were electrically insulated by means of quartz tubes.

The heat capacity of the iron cylinder was found to be 0,36 cal/degree.

During the usual vacuum technical treatment 60 mg mercury and a basic gas of 2 mmHg pressure were put into the discharge tube the pressure of which was set to an accuracy of $\pm 0{,}05$ mmHg. As a standard gas a gas mixture was used containing 80% neon and 20% helium.

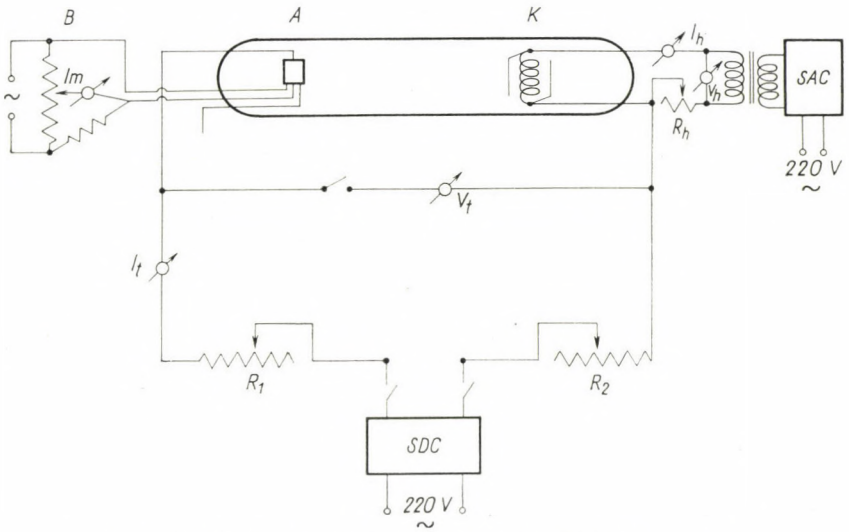


Fig. 1. Circuit diagram applied in the course of the experiments

The cathode was heated by a constant a.c. heating, the voltage of which was supplied by a stabilizer. The tube was started by a high-frequency keep-alive voltage, ensuring adequate cathode heating.

Fig. 1 shows the electric circuit diagram of the experimental arrangement. The constant heating of the cathode K was assured by the stabilized a.c. voltage source SAC. The heating current and the heating voltage could be read on the instruments I_h, resp. V_h. The ohmic resistance R_h served for the exact adjustment of the heating current to the desired value.

The discharge was fed by the d.c. source SDC and its current was limited by the symmetrically placed ohmic resistances R_1 and R_2. During the whole series of measurements the discharge current had a value of 760 mA throughout. The discharge current could be read on the instrument I_t and the burning voltage of the discharge on the instrument V_t.

The heating current of the cathode was chosen in such a way that the total specific surface watt loading arising from the discharge and the heating of the cathode reached a value determined in advance. Accordingly, the total power loss on the cathode was 6,2 W.

The dimensions of the anode cylinder were chosen so that the specific surface current loading of the anode had a value determined in advance, i.e. 2,6 mA/mm^2. By choosing these dimensions the anode fall governed by the specific surface current loading arising on the spiral electrode in the anodic half period could be adjusted and determined in the case of a given type of a.c. supplied fluorescent tube.

The length of the discharge space which — as was described above — consisted of only the anode-fall and cathode-fall sections, was 23 mm, under the given discharge conditions.

The insulated coil within the anode cylinder was chosen in such a manner that its heat capacity was by some orders of magnitude lower than that of the anode cylinder in order that it should be able to follow satisfactorily the changes of temperature of the anode cylinder.

In the present case the ratio of the two heat capacities was 1 : 540. Consequently, the coil could take up the temperature of the iron cylinder very quickly and to a good approximation the temperature of the coil agreed with that of the iron cylinder.

As the cold resistance of the coil (the resistance measured at 20° C) was known, the temperature of the coil could be determined by measuring the heat resistances arising at the various temperatures and in this way the determination of the temperature of the iron cylinder could be reduced to measuring the changing coil resistance. This was measured in the bridge circuit *B* shown in Fig. 1, in case of an a.c. supply voltage. The measurement was preceded by test measurements where the change of the resistance was determined at 1 mA current intensity by the aid of an instrument which was also supplied by alternating current.

All these measurements were preceded by control measurements which indicated that at alternating currents of the intensities in question the inductance of the coil was still negligible beside the ohmic resistance. At a current of 1 mA no warming up could be observed, therefore also in this way the cold resistance of the coil could be determined.

The resistance measurements executed with the help of the bridge circuit B and represented in Fig. 1 were reproducible to a satisfactory accuracy.

Before the beginning of the measurements the discharge operated for 30 minutes under the previously described discharge circumstances and the conditions which were to be investigated.

3. Measuring method. Results

The anode fall measuring method was based on the fact that from the discharge current and from the heat energy transported by the electrons to the anode, which can be measured, the potential difference at which the electrons accelerate to the anode can be calculated. The potential difference corresponds to the anode fall.

In the course of the execution of the measurements the change of the temperature of the anode had to be determined. This happened with the help of the above-mentioned resistance measuring method by determining the cooling of the anode cylinder which set in when the discharge was extinguished.

The course of the measurements and calculations was the following.

When the cathodic watt loading (with a.c.) and anodic current loading (with d.c. discharge) were adjusted to the values given above the discharge equilibrium practically set in after 30 minutes under the conditions described. Accordingly, the parameters to be examined, the anode temperature, the gas temperature and the anode fall, represented then stationary quantities. When the discharge was extinguished the cooling curve of the anode cylinder was recorded. The quantity of heat Q stored by the iron cylinder relative to the ambient gas can be calculated from the following expression:

$$Q = c \cdot M(T - T_g)\,, \tag{2}$$

where

c is the specific heat of iron (cal gramm^{-1} degree^{-1})

M the mass of the iron anode cylinder (g)

T the temperature of the anode cylinder (degree)

T_g the temperature of the gas surrounding the anode cylinder in case of a stationary discharge (degree).

One part of the heat quantity taken up from the discharge and stored by the anode after the stopping of the discharge is given off by heat transfer through the surface of the anode. This quantity Q_1 (cal) may be calculated from the relation

$$dQ_1 = F \cdot \alpha(T - T_g)dt\,, \tag{3}$$

where

F is the surface of the anode cylinder (mm^2), the surface heat transfer coefficient of the anode cylinder (cal. mm^{-2} sec^{-1} degree^{-1}) and

dt the corresponding time (sec).

The heat quantity dissipated by radiation is neglected in the calculations. During the cooling the heat quantity dQ_2 lost by the anode cylinder in the time dt is given by

$$dQ_2 = cMdT\,, \tag{4}$$

where dT is the temperature by which the anode cylinder has cooled down. As there must be

$$dQ_1 = dQ_2 \tag{5}$$

the relation

$$c \cdot M \cdot dT = F \cdot a(T - T_g) \cdot dt \tag{6}$$

must also hold. Transforming equ. (6)

$$\frac{dT}{T - T_g} = \frac{F \cdot a}{cM} dt \,. \tag{7}$$

Integration equ. (7) results in

$$T = \exp\left(\frac{F \cdot a}{cM} t + B\right) + T_g, \tag{8}$$

where B is the integration constant. If we denote

$$\frac{Fa}{cM} = A, \tag{9}$$

relation (8) can be written in the following form

$$T = \exp(At + B) + T_g. \tag{10}$$

If we take into account that at the moment $t = 0$ the temperature T_0 of the cylinder cannot be regarded as equal to the temperature T_{0g} of the gas, the value of B can be given from this initial condition as

$$T_0 = \exp B + T_{g0}, \tag{11}$$

$$T_0 - T_{g0} = \exp B. \tag{12}$$

From this the equation of the cooling curve of the anode cylinder is found to be:

$$T = (T_0 - T_{g0}) \exp At + Tg, \tag{13}$$

where generally

$$T_{g0} \neq T_0. \tag{14}$$

The tangent of the curve is given by

$$\frac{dT}{dt} = (T_0 - T_{g0}) A \cdot \exp(At). \tag{15}$$

As the aim of the present investigations is the determination of the heat energy taken up by the anode from the discharge, the values of the cooling

curve taken up immediately after the cutoff are important, i.e.

$$\lim_{t\to 0}\frac{dT}{dt}=\lim_{t\to 0}\left[(T_0-T_{g0})\,A\exp AT\right], \tag{16}$$

$$\lim_{t\to 0}\frac{dT}{dt}=(T_0-T_{g0})\,A. \tag{17}$$

Determining the limiting values in different points of the cooling curve and plotting the various discharge curves, the value of the three unknowns in the equation A, T_g and T_0 can be given. From the relation

$$dQ=F\cdot\alpha(T_0-T_{g0})dT \tag{18}$$

follows

$$\frac{dQ}{dt}=F\cdot\alpha\,(T_0-T_{g0}) \tag{19}$$

in units of cal $\sec^{-1}$. Converting this into watts, the value of the anode fall can be given as the discharge current is known.

In those cases when before the discharge has started the temperature of the gas agrees with that of the iron cylinder, the calculation becomes even simpler.

The cooling curve of the anode cylinder measured under the discharge conditions described earlier is to be seen in Fig. 2. The scale to the left of the vertical axis gives the heat resistance R_m of the measuring coil in ohms, while that to the right gives the ratio of the heat resistance R_m and the cold resistance R_h as well as the temperature of the cylinder T calculated from it in °K. On the horizontal axis the cooling time t is marked in sec. The cooling curves were generally registered during a cooling time of 5 minutes.

As only the initial section of the curves, i.e. the environment of the point corresponding to $t=0$ is important from the point of view of the calculations, in Fig. 2 only that part of the measured cooling curves, i.e. the curves for the time immediately after the stopping of the discharge is represented. The initial values belonging to the curves shown in Fig. 2 and the corresponding results are summarized in Table 1, with the notations introduced previously.

Table 1

Discharge current i_t	Tube voltage V_t	Heating voltage V_h	Heating current I_h	Temperature of the cylinder T_c	Temperature of the gas T_g	Anode performance W_a	Anode fall V_a
760 mA	16,0 V	8,5 V	730 mA	317° C	71° C	2,66 W	3,5 V

As can be seen from the Table the value of the anode fall obtained by means of the described method was $V_a = 3{,}5$ V.

The value of 3,5 V for the anode fall agrees well with the results of investigations of a different character performed previously by the authors. The value of 12,5 for the cathode fall calculated in the present case from the difference between the burning voltage and the anode fall is in good agreement with the results of the probe measurements carried out previously under similar circumstances.

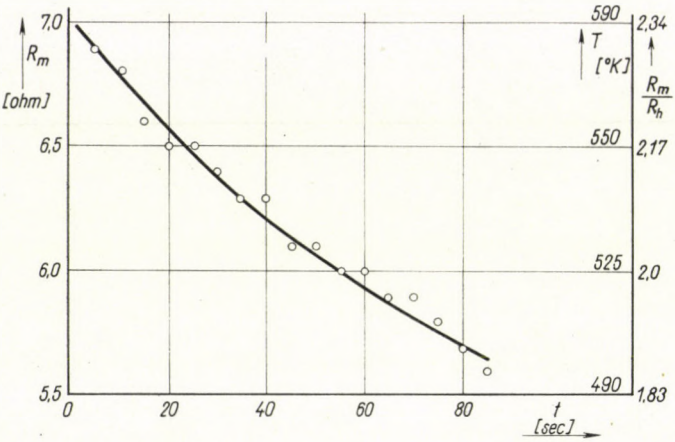


Fig. 2. Cooling curve of the iron anode cylinder. R_m the heat resistance in ohms, R_m/R_h the ratio of the heat and cold resistances, T the corresponding temperature in °K, t the cooling time in sec

In the course of the experiments also such measurements were carried out in which instead of the a.c. cathode heating d.c. cathode heating of the same intensity was applied. The aim of these additional experiments was to establish, whether the character of the cathode heating influences the value of the anode fall. In agreement with the results of investigations of a similar character performed, however, under different circumstances by SAGGAU [3] it was found here also that the value of the anode fall in this case is not affected by the character of the cathode heating.

REFERENCES

1. J. BITÓ, Magyar Fizikai Folyóirat, **10,** 411, 1962.
2. A. VON ENGEL, Phil. Mag., **32,** 417, 1941.
3. B. SAGGAU, Proc. 4th Int. Conf. on Ioniz. Phen. in Gases, **II,** 280, 1960.

КАЛОРИЧЕСКИЙ МЕТОД ДЛЯ ОПРЕДЕЛЕНИЯ АНОДНОГО ПАДЕНИЯ ГАЗОВЫХ РАЗРЯДОВ

Г. СИГЕТИ, Г. ЛАКАТОШ и Й. БИТО

Р е з ю м е

Авторами создается разряд без положительного столба, в котором определяется анодное падение калорическим методом. Значение поверхностного тока и ваттовая нагрузка анода и катода определяются заранее. Описываются экспериментальные условия, подробно анализируются теоретические и экспериментальные основы исследования. Рассматриваются условия применимости. Приводятся результаты, полученные в случае ртутно-аргонно-неонной разрядной трубки. Эти результаты хорошо согласуются с предыдущими исследованиями авторов.

PIEZOGALVANOMAGNETIC EFFECT IN *n*-TYPE GERMANIUM

By

A. ZAWADOWSKI*

RESEARCH INSTITUTE FOR TECHNICAL PHYSICS OF THE HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST

(Presented by G. Szigeti. — Received: 14. II. 1964)

The elastogalvanomagnetic coefficient of *n*-type germanium has been calculated in the HERRING—VOGT relaxation time tensor approximation choosing the direction of the sample and of the magnetic field to be [1, 0, 0] and [0, 1, 0], respectively. If the ratio of the relaxation time tensor components is energy-independent, the elastogalvanomagnetic constant is exactly one. This assumption is satisfied for scattering purely on acoustic phonons or purely on impurities. A discrepancy can be expected for intervalley scattering or for scattering on phonons and impurities. The latter effect decreases the elastogalvanomagnetic coefficient by not more than a few percent.

§ 1. Introduction

Transport phenomena due to electrons and holes in semiconductors can be approximately described by the BOLTZMANN-type transport equation. In general the exact solution of the BOLTZMANN equation cannot be given, but by introducing the concept of relaxation time a good approximate solution of the equation can be obtained for several, physically interesting cases. During the last decade experimental data supplied an increasing amount of information on relaxation time. One group of the measurements was devoted to measuring the galvanomagnetic effect (measurement of the variation of resistivity, Hall constant, magnetoresistance as a function of temperature in magnetic and electric fields of various intensities). The other major group of measurements was concerned with the effects due to the mechanical deformation of matter (the variation of piezoelectric resistance, acoustoelectric effect).

R. W. KEYES [1] was the first to investigate the combined effect of mechanical stress and magnetic field on the conductivity of *n*-type Ge. The effect of purely mechanical deformation on the conductivity of *n*-type germanium was first observed by C. S. SMITH [2]. The theory of the variation of piezoresistance for many-valley semiconductors was developed by C. HERRING [3], [4]. In HERRING's theory the energy of the bottom of each valley is shifted owing to mechanical deformation, which has two important consequences:

* Present address: Central Research Institute of Physics, Budapest.

1. the probability of the scattering of the electrons from one valley to the other changes, and thus the relaxation time describing the so-called intervalley scattering also depends on stress.

2. In the valleys of lower energy more electrons are located in accordance with thermal distribution. The electrons of each valley contribute to conducti vity in a different way corresponding to the directions of external fields, thus conductivity is a function of the number of electrons in each valley.

At high temperature the first factor prevails and at low temperature the second. The temperature-dependence predicted by the theory was also established experimentally by F. S. MORIN, T. H. GABALLE and C. HERRING [5]. However, it is difficult to estimate the coupling constant of intervalley scattering from these measurements. R. W. KEYES [1] stretched an n-type germanium sample in the direction of its axis, and applying a transversal magnetic field perpendicular to this direction measured the resulting change of resistivity (elastogalvanomagnetic effect). The direction of the sample was [1, 0, 0] and that of the magnetic field [0, 1, 0]. KEYES developed the theory of the effect under the assumption that the collisions can be described by a scalar relaxation time. He found that the piezogalvanomagnetic coefficient G introduced by him was exactly unity, when the coupling constant W_I of the intervalley scattering was zero and that in first approximation the difference from unity was proportional to the coupling constant W_I. In his calculations R. W. KEYES expanded in series in powers of the magnetic field H and determined the terms proportional to H^2. Thus in his measurements the intensity of the magnetic field had to be restricted. From his experimental results he estimated the coupling constant W_I. Later G. WEINREICH and his collaborators determined the value W_I more accurately from the results of the acoustoelectric effect [7], ($W_I = 10^{11 \pm 0{,}3}\ \mathrm{sec}^{-1}$).

C. HERRING and E. VOGT [4], and later SAMOILOVICH and his collaborators showed that a better approximation of the solution of the BOLTZMANN equation can be given if a relaxation time tensor is introduced. The approximation for phonon scattering was investigated by all these authors and their collaborators [4] [6] and [8]. C. HERRING and his collaborators found that the ratio of the two relaxation times introduced is $\frac{\tau_l}{\tau_t} = 1{,}18$. For scattering on an ionized impurity SAMOILOVICH and his collaborators [8], [9] obtained for the value of $\frac{\tau_l}{\tau_t}$ a result between 9—11, which is in good agreement with HAM's [10] previous result. If the ratio $\frac{\tau_l}{\tau_t}$ is energy-independent, by modifying the values of the effective masses transport phenomena can be treated with the aid of the scalar relaxation time.

The above results show that for scattering purely on an ionized impurity or purely on a phonon the deviation of the elastogalvanomagnetic constant G from unity is still proportional to the coupling constant of intervalley scattering. Here we have assumed that for scattering on ionized impurities the ratio $\frac{\tau_l}{\tau_t}$ is energy-independent. However, this is only valid to an accuracy of approx. 10%. Should the two types of scattering mechanisms be mixed, the constant $\frac{\tau_l}{\tau_t}$ sensitively depends on the energy.

In the present paper assuming a relaxation-time tensor we have determined the value of G for the directions of the axis on the sample and the magnetic field as chosen by H. W. KEYES. Based on these calculations the deviation of the piezogalvanomagnetic constant from unity is composed of two terms:

1. the term proportional to the coupling constant W_I,

2. the term due to the energy-dependence of $\frac{\tau_l}{\tau_t}$.

The results obtained will be applied to the case when the two types of scattering mechanisms (scattering on impurity and on phonon) are equally significant. Our calculations are valid up to magnetic field intensities at which quantum phenomena can still be neglected. If only the energy dependence of $\frac{\tau_l}{\tau_t}$ is taken into account, the deviation from unity of the constant G amounts to a few percent only, but it is still possible to observe it experimentally by increasing the intensity of the magnetic field. On the basis of the theory given here it can be expected that the elastogalvanomagnetic constant has a local minimum at the temperature where the two scattering mechanisms are mixed; or, at a given temperature, at that impurity concentration where the two kinds of scattering mechanisms are mixed to the greatest extent.

§ 2. The Boltzmann equation and its solution

For the band structure of germanium the many-valley model is assumed. HERRING and VOGT[4] described the mobility of electrons and holes, respectively, by a transport equation for the general case when the semiconductor contains N_v energy valleys.

In the conduction band of semiconductors the equienergetic surfaces near the energy minimum are generally ellipsoids; the effective mass tensor $\mathbf{M}^{(i)}$, the components of which are energy-independent, can be interpreted in this way for the valley (i). The energy measured from the bottom of the energy valley is denoted by $\Delta\varepsilon^{(i)}$ and $\boldsymbol{J}^{(i)}(\Delta\varepsilon)\,d(\Delta\varepsilon)$ is the current density due to the

electrons in the valley (i) the energy of which is in the energy interval $(\Delta\varepsilon, \Delta\varepsilon + d(\Delta\varepsilon))$; further $n^{(i)}$ is the total number of electrons in the valley (i). $\boldsymbol{\tau}^{(i)}(\Delta\varepsilon)$ denotes the relaxation time tensor of the valley (i), the principal axes of which agree with the principal axes of the effective mass tensor. According to HERRING and VOGT [4] in the relaxation time tensor approximation the Boltzmann equation is of the following form:

$$\frac{\partial \boldsymbol{j}^{(i)}(\Delta\varepsilon)}{\partial t} = \frac{2\Delta\varepsilon^{(i)}}{3kT} e^2 n^{(i)}(\Delta\varepsilon)\,\mathbf{M}^{(i)-1}\boldsymbol{E} - \\ - \frac{e}{c}\mathbf{M}^{(i)-1} d\boldsymbol{j}^{(i)}(\Delta\varepsilon) \times \boldsymbol{H} - \boldsymbol{\tau}^{-1}(\Delta\varepsilon)\, d\boldsymbol{j}^{(i)}(\Delta\varepsilon)\,, \tag{1}$$

where k is the BOLTZMANN constant, T is the temperature, e is the absolute value of the charge of the electron, c is the velocity of light, $\boldsymbol{E}$ and $\boldsymbol{H}$ are the intensities of the electric and the magnetic field, respectively. $n^{(i)}(\Delta\varepsilon)$ is the energy distribution of the electrons in the valley (i), which in the given approximation agrees with the Boltzmann distribution. Let us introduce the following simplifying notations:

$$\boldsymbol{\varkappa}^{(i)}(\Delta\varepsilon) = \boldsymbol{\tau}^{(i)}(\Delta\varepsilon)\,\mathbf{M}^{(i)-1} \tag{2}$$

as well as

$$\frac{2e^2}{3kT} = \lambda \qquad \text{and} \qquad \frac{e}{c}\boldsymbol{H} = \boldsymbol{h}$$

and let us define for the valley (i) and the energy $\Delta\varepsilon$ the conductivity $\boldsymbol{\sigma}^{(i)}(\Delta\varepsilon)$ by the following equation:

$$\boldsymbol{j}^{(i)}(\Delta\varepsilon) = \boldsymbol{\sigma}^{(i)}(\Delta\varepsilon)\,\boldsymbol{E} \tag{3}$$

with which the partial conductivity of the valley (i) becomes

$$\boldsymbol{\sigma}^{(i)} = \int_0^\infty \boldsymbol{\sigma}^{(i)}(\Delta\varepsilon)\, d(\Delta\varepsilon). \tag{4}$$

The conductivity of the substance is the sum of the partial conductivities, i.e.

$$\boldsymbol{\sigma} = \sum_{i=1}^{N_v} \boldsymbol{\sigma}^{(i)}\,. \tag{5}$$

The solution of the Boltzmann equation of the form (1) for small electric field intensities, and for arbitrary magnetic field intensities becomes then:

$$\boldsymbol{\sigma}^{(i)}(\Delta\varepsilon) = \lambda\Delta\varepsilon^{(i)} n^{(i)}(\Delta\varepsilon)\, \frac{\boldsymbol{\varkappa}^{(i)} + \Delta(\boldsymbol{\varkappa}^{-1}\boldsymbol{h}) \times + \Delta\boldsymbol{h} \circ \boldsymbol{h}}{1 + \Delta(\boldsymbol{h}\boldsymbol{\varkappa}^{-1}\boldsymbol{h})}\,,^1 \tag{6}$$

[1] Here $\times$ is the vector and $\circ$ means the diadic product.

where

$$\Delta(\Delta\varepsilon) = \det \boldsymbol{\varkappa}^{(i)}|_{\Delta\varepsilon}. \tag{7}$$

The resistivity ϱ can be determined from the reciprocal of the conductivity tensor. Measuring the resistivity in the direction of the unity vector $\boldsymbol{e}$ we have

$$\varrho_e = \boldsymbol{e}\,\sigma^{-1}\,\boldsymbol{e}. \tag{8}$$

For n-type germanium the well-known model of four energy minima is used. In this model the energy surfaces near the minima are ellipsoids of revolution

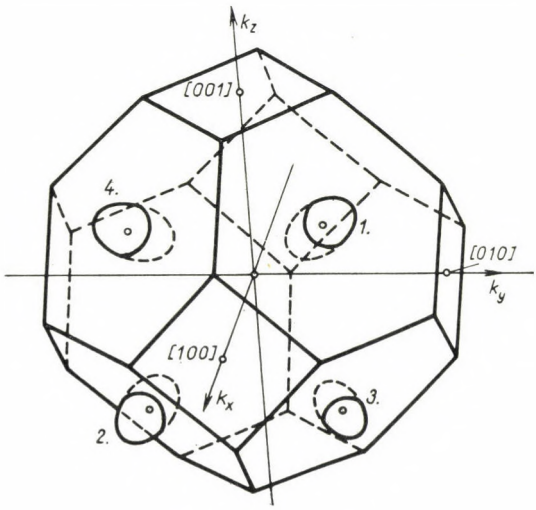


Fig. 1

and thus two effective masses are required to characterize them. The effective mass tensor is transformed to the principal axis in the form

$$\mathbf{M} = \begin{pmatrix} m_l & o & o \\ o & m_l & o \\ o & o & m_l \end{pmatrix}. \tag{9}$$

The energy minima are numbered as shown in Fig. 1.

The form of the mass tensor for the various valleys is given in the Appendix in a coordinate system determined by the crystal axes. From symmetry considerations it follows that the relaxation time tensor is of a similar form as the mass tensor and thus the tensor $\boldsymbol{\varkappa}$ derived from the above two tensors is also of the structure shown in the Appendix.

Let the relaxation time tensor transformed to the principal axis be the following:

$$\tau(\Delta\varepsilon) = \begin{pmatrix} \tau_t(\Delta\varepsilon) & 0 & 0 \\ 0 & \tau_t(\Delta\varepsilon) & 0 \\ 0 & 0 & \tau_l(\Delta\varepsilon) \end{pmatrix} = \tau(\Delta\varepsilon) \begin{pmatrix} a_t(\Delta\varepsilon) & 0 & 0 \\ 0 & a_t(\Delta\varepsilon) & 0 \\ 0 & 0 & a_l(\Delta\varepsilon) \end{pmatrix}. \quad (10)$$

If the ratio $\frac{\tau_t(\Delta\varepsilon)}{\tau_l(\Delta\varepsilon)}$ is independent of $\Delta\varepsilon$, a_t and a_l can be chosen to be energy-independent. From the expression (6) it can be seen that the resisitivity only depends on the tensor $\boldsymbol{\varkappa}$ and does not depend separately either on $\boldsymbol{\tau}$ or on $\mathbf{M}$. Thus, as far as $\frac{\tau_t(\Delta\varepsilon)}{\tau_l(\Delta\varepsilon)}$ is energy independent, the value of conductivity equals the value calculated with the scalar relaxation time, if the following substitutions are carried out: 1. the relaxation time is replaced by $\tau(\Delta\varepsilon)$; 2. the values of the effective mass are modified to be

$$m_t \to \frac{m_t}{a_t}, \qquad m_l \to \frac{m_l}{a_l}.$$

KEYES [1] based the theory of the elastogalvanomagnetic effect on the assumption of scalar relaxation time. Thus it is obvious from what we have said above that his results can easily be generalized for $\frac{\tau_t(\Delta\varepsilon)}{\tau_l(\Delta\varepsilon)} = \text{const.}$ by carrying out the appropriate substitutions. Thus, *if the ratio of the components of the relaxation-time tensor is energy-independent and the coupling constant W_I of intervalley scattering is zero, the value of the elastogalvanomagnetic constant G is one.*

In the following we shall be concerned with the case when the ratio of the components of the relaxation-time tensor is energy-dependent. First we calculate the value of the variation of magnetoresistance for the case considered.

§ 3. Determination of the variation of magnetoresistance if the magnetic field is of direction [0,0,1]

As regards symmetry transformations the energy valleys taken into account in undeformed semiconducturs play identical roles and thus the number of electrons is the same in each valley. Accordingly, in the undeformed case $n^{(i)}(\Delta\varepsilon)$ is independent of (i) and the index can be omitted.

The conductivity can be given with the aid of (4), (5) and (6). The result is, in general, rather intricate. Considerable simplification can be obtained if

the value of the denominators is an identical scalar for each valley. Then the final result is the integral of an expression similar to (6), where the tensors $\boldsymbol{\varkappa}$ are replaced by mean values which may obtained by the aid of (A. 3). The denominators can be made identical by choosing the direction of the magnetic field to be [0, 0, 1], as in this case the relative position of the individual valleys with respect to the magnetic field is identical.

Let us determine first the value of $\sigma^{(i)}$. The integration involved in it corresponds to a BOLTZMANN averaging as the integrand includes the electron density $n^{(i)}(\Delta\varepsilon)$ which follows the BOLTZMANN distribution. Denoting the BOLTZMANN average by $< >$ we obtain for (4) from (6) the following:

$$\sigma^{(i)} = \lambda_n^{(i)} < \frac{\Delta\varepsilon}{1 + \Delta(\boldsymbol{h}\,\boldsymbol{\varkappa}^{(i)-1}\,\boldsymbol{h})} (\boldsymbol{\varkappa}^{(i)} + \Delta(\boldsymbol{\varkappa}^{(i)-1}\,\boldsymbol{h}) \times + \Delta \boldsymbol{h} \circ \boldsymbol{h}) >, \tag{11}$$

where $n^{(i)}$ denotes the total number of electrons in the valley (i), which, expressed in terms of the density of conduction electrons in the case without deformation is $\frac{n}{N_v}$. The expression (5) of conductivity by using (6) and (11) simply becomes

$$\sigma = A + B\boldsymbol{h}\boldsymbol{x} + C\boldsymbol{h} \circ \boldsymbol{h}, \tag{12}$$

where A, B and C are the following scalar expressions:

$$\begin{aligned} A &= \frac{\lambda n}{N_v} \sum_{i=1}^{N_v} \langle g\boldsymbol{\varkappa}^{(i)} \rangle, \\ B &= \frac{\lambda n}{N_v} \sum_{i=1}^{N_v} \langle g\Delta\boldsymbol{\varkappa}^{(i)-1} \rangle, \\ C &= \lambda n \langle g\Delta \rangle. \end{aligned} \tag{13}$$

Here the function $g(\Delta\varepsilon)$ is

$$g(\Delta\varepsilon) = \frac{\Delta\varepsilon}{1 + \Delta(\boldsymbol{h}\boldsymbol{\varkappa}^{-1}\,\boldsymbol{h})}. \tag{14}$$

For the resistivity tensor we obtain from (12)

$$\boldsymbol{\sigma}^{-1} = \frac{A - B\boldsymbol{h} \times + \dfrac{B^2 - AC}{A + C|\boldsymbol{h}|^2} \boldsymbol{h} \circ \boldsymbol{h}}{A^2 + B^2|\boldsymbol{h}|^2}. \tag{15}$$

§ 4. Theory of the elastogalvanomagnetic effect

In calculating the elastogalvanomagnetic effect the geometrical arrangement is given in a coordinate system, the axes of which are the crystal axes. The magnetic field intensity in the arrangement chosen is [0, 0, 1]. The axis of the sample is, as shown in Fig. 2, in the plane [0, 0, 1] and subtends an angle φ with the direction [1, 0, 0]. The sample is pulled in the direction of its axis.

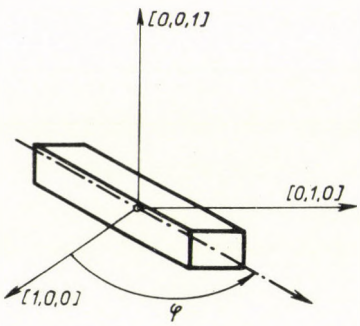


Fig. 2

Owing to deformation the energy of the bottom of the energy valleys is shifted. For symmetry reasons the shift is identical for the valleys 1 and 3, as well as for 2 and 4, respectively. As an effect is only brought about by relative displacement, we can write for the shift $\delta\varepsilon^{(i)}$ of the individual valleys

$$\delta\varepsilon^{(i)} = \begin{cases} -\delta\varepsilon & \text{if } i = 1, 3, \\ +\delta\varepsilon & i = 2, 4. \end{cases} \tag{16}$$

In HERRING's [3] [4] theory the change of resistance due to deformation, as was mentioned in the Introduction, is composed of two parts. The first part results from the change of $n^{(i)}$. As the $n^{(i)}$ follow the BOLTZMANN distribution the multiplying factors corresponding to the total number of particles will be different. It can readily be seen that if the energy of the bottom of the valley (j) is shifted by $\delta\varepsilon^{(j)}$, the variation of $n^{(i)}$ is

$$\frac{\partial n^{(i)}}{\partial \delta\varepsilon^{(j)}} = -\frac{n}{N_v kT}\left(\delta_{ij} - \frac{1}{N_v}\right),$$

whence for the special case given by (16)

$$\delta n^{(i)} = (-1)^i \delta n, \tag{17a}$$

where

$$\delta n = -\frac{\delta\varepsilon}{kT}\frac{n}{N_v}. \tag{17b}$$

The second part of the change of resistivity is due to the variation of the transition probability of intervalley scattering. The latter can formally be taken into account in the given relaxation time approximation by means of the change of the relaxation time tensor. Let there be $\boldsymbol{\tau}^{(i)}_{\text{int r}}$ the relaxation time tensor of the valley (i) corresponding to the electron transitions between the valleys, i.e. to intervalley scattering and $\boldsymbol{\tau}^{(i)}_{\text{intra}}$ the relaxation time tensor of an electron transition in the valley, i.e. of intervalley scattering. It is well known that $\boldsymbol{\tau}^{(i)}$ is a resultant relaxation time tensor, which describes the effect of both types of scattering

$$\boldsymbol{\tau}^{(i)-1} = \boldsymbol{\tau}^{(i)-1}_{\text{inter}} + \boldsymbol{\tau}^{(i)-1}_{\text{intra}}. \tag{18}$$

The energy shift of the bottom of the valleys results in altering the probability of intervalley scattering. The change of the relaxation time tensor of the valley (i) can be given with the aid of a β function, which can easily be obtained from the transition probability of intervalley scattering [1]. Thus

$$\left.\frac{\partial\boldsymbol{\tau}^{(i)-1}}{\partial\delta\varepsilon^{(k)}}\right|_{\delta\varepsilon^{(k)}=0} = \frac{\partial\boldsymbol{\tau}^{(i)-1}_{\text{inter}}}{\partial\delta\varepsilon^{(k)}} = \left(\delta_{ik} - \frac{1}{N_v}\right)\beta\,\mathbf{I}, \tag{19}$$

where we have used the fact that intervalley scattering can be described by a scalar relaxation time tensor. The function β is an intricate function of the energy $\Delta\varepsilon^{(i)}$ proportional to the coupling constant W_I of intervalley scattering. The right-hand side of expression (19) can be written in such simple form, only if the relative position of any two valleys is the same. For germanium this can directly be seen in Fig. 1. At the same time for silicon the relative positions of the valleys are of two types, as the axes of the ellipsoids of revolution corresponding to the valleys are in the direction [1, 0, 0] and the centres of the valleys are in the interior of the Brillouin zone. For the latter material the right-hand side of expression (20) is more intricate and involves two β functions. The change of the function $g(\Delta\varepsilon)$ given by the expression (14) which depends on $\boldsymbol{\tau}$ will also be required. By using the definition (14)

$$\frac{\partial g^{(i)}(\Delta\varepsilon)}{\partial\delta\varepsilon^{(j)}} = -f(\Delta\varepsilon)\,\beta(\Delta\varepsilon)\left(\delta_{ij} - \frac{1}{N_v}\right), \tag{20a}$$

where

$$f(\Delta\varepsilon) = \frac{\Delta\varepsilon}{(1 + \Delta(\boldsymbol{h}\boldsymbol{\varkappa}^{-1}\boldsymbol{h}))^2}\,\Delta(\boldsymbol{h}\mathbf{M}(\mathbf{I} - \boldsymbol{\tau}^{-1}\,\text{Spur}\,(\boldsymbol{\tau}))\boldsymbol{h}). \tag{20b}$$

The two effects dealt with in detail above equally result in the change of σ. From the form given by expressions (12) and (11) of the conductivity tensor σ

it can readily be seen that the change $\delta\sigma$ of the conductivity tensor due to deformation can be written in the form

$$\delta\sigma = \mathbf{a} + (\mathbf{bh}) \times + \mathbf{ch} \circ \boldsymbol{h}, \tag{21}$$

whence the change of the tensor of resisitivity ϱ in linear approximation is

$$\delta\varrho = \delta\sigma^{-1} = -\sigma^{-1}\,\delta\sigma\,\sigma^{-1}. \tag{22}$$

The linear approximation is justified for stresses usually applied, as $\delta\sigma$ is sufficiently small. According to expression (8) the change $\delta\varrho$ of the specific resistivity can be expressed, by using (5), (11), (12), (13) and (21) in terms of the tensors **a**, **b** and **c**. Using the fact that $(\boldsymbol{eh})$ is zero in the arrangement given in Fig. 2 we obtain

$$\delta\varrho = (\boldsymbol{e}\delta\sigma^{-1}\boldsymbol{e}) =$$

$$= \frac{1}{(A^2 + B^2\,|\,\boldsymbol{h}\,|^2)}\Big(e\Big\{\big[A^2\,a + 2AB(\boldsymbol{hbh})\big] + \tag{23}$$

$$+ \big[-AB\mathbf{a} - B^2(\boldsymbol{hbh})\,\boldsymbol{h}\times + \boldsymbol{h}\times AB\mathbf{a} + A^2(\mathbf{bh}) \times + B^2\,\boldsymbol{h}\times \mathbf{ah}\times\Big\}\boldsymbol{e}\Big),$$

where $\boldsymbol{e}$ is the unity vector pointing in the direction of the axis of the sample.

It can be seen that by means of the tensors **a** and **b** in the expression (21) of the tensor $\delta\sigma$ the change of resisitivity due to deformation can be given. The tensor **c** does not appear in the expression (23).

In the following $\delta\sigma$ will be given which consists of two parts: $\delta\sigma_{W_I}$ due to the change of $\boldsymbol{\tau}^{(i)}$ and $\delta\sigma_n$ due to the change of $n^{(i)}$. From (11), using the expressions (17), (19) and (20) it can readily be seen that

$$\delta\sigma = \delta\sigma_{W_I} + \delta\sigma_n\,, \tag{24}$$

$$\delta\sigma_n = \Big\langle \sum_i \lambda\delta n^{(i)}\big[\boldsymbol{\varkappa}^{(i)} + \Delta(\boldsymbol{\varkappa}^{(i)-1}\,\boldsymbol{h}) \times + \Delta\boldsymbol{h} \circ \boldsymbol{h}\big]\,g\Big\rangle =$$

$$= -\sum_i \lambda\,\big\langle g\big[\boldsymbol{\varkappa}^{(i)} + \Delta(\boldsymbol{\varkappa}^{(i)-1}\,\boldsymbol{h}) \times + \Delta\boldsymbol{h} \circ \boldsymbol{h}\big]\big\rangle\,\frac{n\delta\varepsilon}{N_v kT} + \tag{25a}$$

$$+ 2\sum_{i=1,3} \lambda\,\frac{n\delta\varepsilon}{kTN_v}\,\big\langle g(\boldsymbol{\varkappa}^{(i)} + \Delta(\boldsymbol{\varkappa}^{(i)-1}\,\boldsymbol{h}) \times + \Delta\boldsymbol{h} \circ \boldsymbol{h})\big\rangle\,,$$

$$\delta\sigma_{W_I} = \Big\langle \sum_{ij} \lambda\boldsymbol{n}^{(i)}\,\frac{\partial}{\partial\delta\varepsilon^{(j)}}\Big\{\big[\boldsymbol{\varkappa}^{(i)} + \Delta(\boldsymbol{\varkappa}^{(i)-1}\,\boldsymbol{h}) \times + \Delta\boldsymbol{h} \circ \boldsymbol{h}\big]\,g\Big\}\Big\rangle\,\delta\varepsilon^{(j)} =$$

$$= \Big\langle \sum_{ij} \lambda\boldsymbol{n}^{(i)}\,\frac{\partial}{\partial\delta\varepsilon^{(j)}}\Big\{\big[\boldsymbol{\varkappa}^{(i)} + \Delta(\boldsymbol{\varkappa}^{(i)-1}\,\boldsymbol{h}) \times + \Delta\boldsymbol{h} \circ \boldsymbol{h}\big]\,g\Big\}\Big\rangle\,\delta\varepsilon - \tag{25b}$$

$$- 2\Big\langle \sum_{j=1,3} \lambda\boldsymbol{n}^{(i)} \sum_i \frac{\partial}{\partial\delta\varepsilon^{(j)}}\Big\{\big[\boldsymbol{\varkappa}^{(i)} + \Delta(\boldsymbol{\varkappa}^{(i)-1}\boldsymbol{h}) \times + \Delta\boldsymbol{h} \circ \boldsymbol{h}\big]\,g\Big\}\Big\rangle\,\delta\varepsilon\,,$$

where the first term of the expression (25) is zero, because the identical shift of the bottom of the valleys cannot result in the change of σ, as then τ_{inter} describing intervalley scattering does not change, as the scattering probability does not change either. Thus, by comparing the expression (21) with (24) and (25) respectively, we obtain

$$\frac{\mathbf{a}}{\delta\varepsilon} = -\sum_{i=1}^{4} \frac{\lambda n}{N_n} \left\{ \langle g\boldsymbol{\varkappa}^{(i)} \rangle \frac{1}{kT} + \langle \boldsymbol{\tau}^{(i)} \boldsymbol{\varkappa}^{(i)} g\beta \rangle + \langle \boldsymbol{\varkappa}^{(i)} f\beta \rangle \right\} + \\ + 2 \sum_{i=1,3} \frac{\lambda n}{N_v} \left\{ \frac{1}{kT} \langle g\boldsymbol{\varkappa}^{(i)} \rangle + \langle \boldsymbol{\tau}^{(i)} \boldsymbol{\varkappa}^{(i)} g\beta \rangle + \langle \boldsymbol{\varkappa}^{(i)} f\beta \rangle \right\}, \tag{26}$$

where the following relation derived from (20a) as well as (2) and (19) was used:

$$\frac{\partial \boldsymbol{\varkappa}^{(i)}}{\partial \delta\varepsilon^{(j)}} = -\boldsymbol{\varkappa}^{(i)} \boldsymbol{\tau}^{(i)} \beta \left(\delta_{ij} - \frac{1}{N_v} \right), \; N_v = 4 \; .$$

To calculate $\delta\sigma_{W_I}$ we shall require the following two expressions of similar character

$$\frac{\partial \Delta}{\partial \delta\varepsilon^{(j)}} = -\Delta (\text{Spur}\, \boldsymbol{\tau})\, \beta \left(\delta_{ij} - \frac{1}{N_v} \right)$$

and

$$\frac{\partial \Delta \boldsymbol{\varkappa}^{-1}}{\partial \delta\varepsilon^{(j)}} = \Delta M (I - \boldsymbol{\tau}^{-1} \text{Spur}\, \boldsymbol{\tau})\, \beta \left(\delta_{ij} - \frac{1}{N_v} \right).$$

Thus for the expression (25) of $d\sigma_\tau$ the following is readily obtained

$$\frac{\mathbf{b}}{\delta\varepsilon} = -\sum_{i=1}^{4} \frac{\lambda n}{N_v} \left\{ \frac{1}{kT} \langle g\Delta^{(i)} \boldsymbol{\varkappa}^{-1} \rangle + \right. \\ + \langle g\beta\Delta \mathbf{M}^{(i)} (\mathbf{I} - \boldsymbol{\tau}^{(i)-1} \text{Spur}\, \boldsymbol{\tau}^{(i)}) \rangle + \langle \Delta\beta \boldsymbol{\varkappa}^{(i)-1} f \rangle \Big\} + \tag{27}$$
$$+ 2 \sum_{1,3} \frac{\lambda n}{N_v} \left\{ \frac{1}{kT} \langle g\Delta\boldsymbol{\varkappa}^{(i)-1} \rangle + \langle g\beta\Delta \mathbf{M}^{(i)} (I - \boldsymbol{\tau}^{(i)-1} \text{Spur}\, \boldsymbol{\tau}^{(i)}) \rangle + \right. \\ \left. + \langle \beta\Delta f \boldsymbol{\varkappa}^{(i)-1} \rangle \right\}.$$

In the expression of **b** given here sums of tensors are contained. The concrete form of the tensors is given in the Appendix. From the tensors in (2) it is easy to construct the expression (27) which is characterized by the fact that when a vector **h** of the direction of the z axis is applied to it it becomes zero. Thus in our further expressions, of the tensors **a**, **b** and **c** only **a** occurs in the expression (23) of the change $\delta\varrho$ of resisitivity. The remaining vector and tensor ex-

pressions, respectively, in (23) can be simplified still further by taking into account that in the geometrical arrangement chosen $(\boldsymbol{e} \cdot \boldsymbol{h}) = 0$. By introducing the notation

$$\boldsymbol{h}^* = \boldsymbol{h} \times \boldsymbol{e},$$

$$(\boldsymbol{e}, \boldsymbol{h} \times \mathbf{a}\boldsymbol{e}) = -(\boldsymbol{e}\mathbf{a}\boldsymbol{h}^*)$$

and

$$(\boldsymbol{e}\boldsymbol{h} \times \mathbf{a}(\boldsymbol{h} \times \boldsymbol{e})) = -\boldsymbol{h}^* \mathbf{a} \boldsymbol{h}^*$$

therefore expression (23) of $\delta\varrho$ is simplified further

$$\delta\varrho = \frac{1}{A^2 + B^2 |\boldsymbol{h}|^2} \left\{ A^2 (\boldsymbol{e}\mathbf{a}\boldsymbol{h}^*) - B^2 (\boldsymbol{h}^* \mathbf{a} \boldsymbol{h}^*) \right\}. \tag{28}$$

In the geometrical arrangement shown in Fig. 2

$$\begin{aligned} \boldsymbol{h} &= (0, 0, 1)\, \boldsymbol{h} \\ \boldsymbol{e} &= (\cos\varphi, \sin\varphi, 0), \\ \boldsymbol{h}^* &= (-\sin\varphi, \cos\varphi, 0)\, |\boldsymbol{h}|. \end{aligned} \tag{29}$$

In the following $\delta\varrho_{ph}$ and $\delta\varrho_p$ mean the change of resistivity due to deformation with and without a magnetic field. $\delta\varrho_{ph}$ can be determined by the relations (28), (29), (26) and (A. 2) of the Appendix. In this way:

$$\begin{aligned} \delta\varrho_{ph} = \frac{1}{A^2 + B^2 |\boldsymbol{h}|^2} \frac{\lambda n}{3} \left\{ \left\langle \left(\frac{g}{kT} + \beta f \right) \left(-\frac{\boldsymbol{\tau}_t}{\boldsymbol{m}_t} + \frac{\boldsymbol{\tau}_l}{\boldsymbol{m}_l} \right) \right\rangle + \right. \\ \left. + \left\langle g\beta \left(-\frac{\boldsymbol{\tau}_t^2}{\boldsymbol{m}_t} + \frac{\boldsymbol{\tau}_l^2}{\boldsymbol{m}_l} \right) \right\rangle \right\} \delta\varepsilon \sin 2\varphi. \end{aligned} \tag{30}$$

In the similar expression of $\delta\varrho_p$ it should be taken into account that from (14) and (20), respectively, $f = 0$ and $g = \Delta\varepsilon$. If we introduce the notation $A|_{h=0} = A_0$, $\delta\varrho_p$ is obtained as a special case of $\delta\varrho_{ph}$ i.e.

$$\begin{aligned} \delta\varrho_p = A_0^{-2} \frac{\lambda n}{3} \left\{ \left\langle \frac{\Delta\varepsilon}{kT} \left(-\frac{\tau_t}{\tau} + \frac{\tau_l}{m_l} \right) \right\rangle + \right. \\ \left. + \left\langle \Delta\varepsilon\beta \left(-\frac{\tau_t^2}{m_t} + \frac{\tau_l^2}{m_l} \right) \right\rangle \right\} \delta\varepsilon \sin 2\varphi. \end{aligned} \tag{31}$$

The resistivity ϱ_0 of the sample is obtained from (2):

$$\varrho_0 = \frac{1}{A_0} \tag{32}$$

and its magnetoresistance ϱ_h from (15):

$$\varrho_h = \frac{A}{A^2 + B^2 |\boldsymbol{h}|^2}$$

and therefore the change of magnetoresistance is

$$\delta\varrho_h = \varrho_h - \varrho_0 = \frac{A}{A^2 + B^2 |\boldsymbol{h}|^2} - \frac{1}{A_0}. \tag{33}$$

KEYES [1] introduced the elastogalvanomagnetic constant G by using the following definition:

$$G = \frac{\delta\varrho_{ph} - \delta\varrho_p}{\delta\varrho_p} \left(\frac{\delta\varrho_h}{\varrho_0}\right)^{-1}. \tag{34}$$

By substituting the corresponding expressions (30), (31), (32) and (33) we obtain

$$G = 1 + \Delta G_\tau + \Delta G_{W_I}, \tag{35}$$

where

$$\Delta G_\tau = \frac{\lambda n}{3} \frac{A_0}{kT} K \cdot \left\{ \frac{\left\langle g\left(-\frac{\tau_t}{m_t} + \frac{\tau_l}{m_l}\right)\right\rangle \left\langle \Delta\varepsilon\left(2\frac{\tau_t}{m_t} + \frac{\tau_l}{m_l}\right)\right\rangle}{\left\langle \Delta\varepsilon\left(-\frac{\tau_t}{m_t} + \frac{\tau_l}{m_l}\right)\right\rangle} - \right. \tag{36}$$

$$\left. - \frac{\left\langle \Delta\varepsilon\left(-\frac{\tau_t}{m_t} + \frac{\tau_e}{m_l}\right)\right\rangle \left\langle g\left(2\frac{\tau_t}{m_t} + \frac{\tau_l}{m_l}\right)\right\rangle}{\left\langle \Delta\varepsilon\left(-\frac{\tau_t}{m_t} + \frac{\tau_l}{m_l}\right)\right\rangle} \right\}$$

and

$$\Delta G_{W_I} = K A_0^2 \frac{\left\langle \beta\left\{ f\left(-\frac{\tau_t}{m_t} + \frac{\tau_l}{m_l}\right) + \left(g - \Delta\varepsilon\frac{A}{A_0}\right)\left(-\frac{\tau_t^2}{m_t} + \frac{\tau_l^2}{m_l}\right)\right\}\right\rangle}{\left\langle \Delta\varepsilon\left(-\frac{\tau_t}{m_t} + \frac{\tau_l}{m_l}\right)\right\rangle + \left\langle \Delta\varepsilon\beta\left(-\frac{\tau_t^2}{m_t} + \frac{\tau_l^2}{m_l}\right)\right\rangle} \tag{37}$$

with

$$K = \frac{1}{AA_0 - (A^2 + B^2 |\boldsymbol{h}|^2)}.$$

In expression (35) ΔG_{W_I} is different from zero only if the coupling constant W_I of intervalley scattering is not equal to zero. Thus it is justified to call

ΔG_{W_I} the term due to intervalley scattering. This term was also determined by KEYES [1]; his results are contained in the relations derived by us as a special case. The term ΔG_τ is obviously equal to zero, if $\frac{\tau_l}{\tau_t}$ is an energy-independent constant; thus ΔG_τ is characteristic of the energy-dependence of $\frac{\tau_l}{\tau_t}$. HERRING and VOGT [4] as well as SAMOILOVICH and his collaborators [8] have shown that this condition is equally satisfied for phonon scattering and in good approximation for scattering on impuirities, i.e. $\frac{\tau_l}{\tau_t}$ is energy-independent. Nevertheless, if the two scattering mechanisms play equally important roles, i.e. both scattering mechanisms are effective at the same time, the ratio $\frac{\tau_l}{\tau_t}$ may sensitively depend on the energy. This means that *in relaxation time approximation only if the two scattering mechanisms are mixed can $\Delta G\tau$ be regarded as a new correction to Keyes' results.* Naturally, if $\frac{\tau_l}{\tau_t}$ is energy-independent, the value of the effective mass in KEYES' results has to be modified in the way described in § 2 in order to arrive at the results obtained by us.

KEYES simplified his results by expanding in series in powers of the magnetic field and then neglecting the terms of higher order than $\mathbf{H}^2$. It can be seen that our results obtained so far are quite independent of the intensity of the magnetic field. To determine each term numerically we have also carried out the appropriate expansion in series.

§ 5. The part of the elastogalvanomagnetic coefficient G proportional to the square of the intensity

We now expand the expression for the elastogalvanomagnetic constant G in series in powers of the magnetic field intensity and calculate the terms proportional to its square. The value obtained for G in this approximation will be denoted by $G^{(2)}$. By expanding in series the relations (36) and (37) the calculations give the following results:

$$G^{(2)} = 1 + \Delta G_\tau^{(2)} + \Delta G_{W_I}^{(2)}, \tag{38}$$

where

$$\Delta G_\tau^{(2)} = \frac{X}{(Z+V)\,W} \qquad \text{and} \qquad \Delta G_{W_I} = \frac{Y}{(Z+V)\,W}. \tag{39}$$

Here the meaning of X, Y, V, W and Z is the following:

$$Z = \left\langle \frac{\Delta\varepsilon}{kT}\left(-\frac{\tau_t}{m_t}+\frac{\tau_l}{m_l}\right)\right\rangle, \quad V = \left\langle \Delta\varepsilon\beta\left(-\frac{\tau^2}{m_t}+\frac{\tau_t^2}{m_l}\right)\right\rangle, \tag{40a}$$

$$W = \left(\left\langle \Delta\varepsilon\left(2\frac{\tau_t}{m_t}+\frac{\tau_l}{m_l}\right)\right\rangle\left\langle \Delta\varepsilon\left(2\frac{\tau_t}{m_t}+\frac{\tau_l}{m_l}\right)\left(2\frac{\tau_l}{m_l}+\frac{\tau_t}{m_t}\right)\frac{\tau_t}{m_t}\right\rangle - \left\langle \Delta\varepsilon\frac{\tau_t}{m_t}\left(2\frac{\tau_l}{m_l}+\frac{\tau_t}{m_t}\right)\right\rangle^2, \tag{40b}$$

$$X = -\left\langle \Delta\varepsilon\left(2\frac{\tau_t}{m_t}+\frac{\tau_l}{m_l}\right)\right\rangle \cdot \left\{\left\langle \Delta\varepsilon\left(2\frac{\tau_t}{m_t}+\frac{\tau_l}{m_l}\right)\right\rangle\left\langle \Delta\varepsilon\beta\frac{\tau_t}{m_t}\left[\left(-\frac{\tau_t}{m_t}+\frac{\tau_l}{m_l}\right)\left(\frac{\tau_t^2}{m_t}+\frac{\tau_t\tau_l}{m_l}+\frac{\tau_l^2}{m_l}\right)+\left(-\frac{\tau_t^2}{m_t}+\frac{\tau_l^2}{m_l}\right)\left(2\frac{\tau_l}{m_l}+\frac{\tau_t}{m_t}\right)\right]\right\rangle - \left\langle \Delta\varepsilon\beta\left(-\frac{\tau_t^2}{m_t}+\frac{\tau_l^2}{m_l}\right)\right\rangle\left\langle \Delta\varepsilon\left(2\frac{\tau_t}{m_t}+\frac{\tau_l}{m_l}\right)\frac{\tau_t}{m_t}\left(2\frac{\tau_t}{m_t}+\frac{\tau_l}{m_l}\right)\right\rangle\right\}. \tag{40c}$$

After the subsitution ΔG_{W_I} the expression $\tau_t = \tau_l$ gives KEYES' results [1]. In what follows we shall estimate the correction ΔG_τ. In the denominator of ΔG_τ the terms due to intervalley scattering can be neglected, as at sufficiently low temperatures the intervalley scattering is insignificant. Introducing the abbreviation $\left\langle \Delta\varepsilon\left(\frac{\tau_t}{m_t}\right)^h\left(\frac{\tau_l}{m_l}\right)^m\right\rangle = \langle n, m\rangle$ we have according to (39)

$$\Delta G\tau|_{W_I=0} = \tag{41}$$

$$= 3\,\frac{1+2\frac{\langle 1,0\rangle}{\langle 0,1\rangle}\langle 0,1\rangle\left[\langle 3,0\rangle+2\langle 2,1\rangle\right]-\langle 1,0\rangle\left[\langle 2,1\rangle+2\langle 1,2\rangle\right]}{1-\frac{\langle 1,0\rangle}{\langle 0,1\rangle}\left(2\langle 1,0\rangle+\langle 0,1\rangle\right)\left(2\langle 3,0\rangle+5\langle 2,1\rangle+2\langle 1,2\rangle\right)-3\left(2\langle 2,0\rangle+\langle 1,1\rangle\right)^2}.$$

In the following we shall take the joint effect of the phonon and impurity scattering into account. The resultant relaxation time tensor is given by the appropriate tensors of the two scatterings:

$$\frac{1}{\tau_\alpha} = \frac{1}{\tau_{i\alpha}} + \frac{1}{\tau_{ph\alpha}} \qquad \alpha = t, l\,, \tag{42}$$

where τ_i and τ_{ph} can be calculated from impurity and lattice scattering, respectively. HERRING and his collaborators [6] obtained for phonon scattering

$$\left(\frac{\tau_l}{\tau_t}\right)_{ph} = 1{,}18 \tag{43}$$

while SAMOILOVICH and his collaborators obtained [8], [9]

$$\left(\frac{\tau_l}{\tau_t}\right)_{i} = 10\,. \tag{44}$$

The well-known energy-dependence of the relaxation time tensor components is expressed by

$$\tau_{i\alpha} = \nu_{i\alpha}\,\Delta\varepsilon^{3/2} \quad \text{and} \quad \tau_{ph\alpha} = \nu_{ph\alpha}\,\Delta\varepsilon^{-1/2} \quad \alpha = t, l\,. \tag{45}$$

With the notations

$$\left(\frac{\nu_i}{\nu_{ph}}\right)_t = \alpha\,, \qquad \left(\frac{\nu_i}{\nu_{ph}}\right)_l = \beta\,, \tag{46a}$$

and from (43) and (44)

$$\frac{(\nu_{ph})_l}{(\nu_{ph})_t} = 1{,}18\,, \qquad \frac{(\nu_i)_l}{(\nu_i)_t} = 10\,, \qquad \lambda = \frac{\beta}{\alpha} = \frac{10}{1{,}18} \approx 8{,}5 \tag{46b}$$

we obtain from (42), (45) and (46) the components of the relaxation time tensor:

$$\tau_l = (\nu_i)_l \frac{\Delta\varepsilon^{3/2}}{1 + \lambda\alpha(\Delta\varepsilon)^2}\,, \tag{47a}$$

$$\tau_t = (\nu_i)_t \frac{\Delta\varepsilon^{3/2}}{1 + \alpha(\Delta\varepsilon)^2}\,. \tag{47b}$$

It can be seen that the ratio $\frac{\tau_l}{\tau_t}$ may depend on the energy. To characterize this energy-dependence let us introduce the following expression:

$$\frac{\left(\frac{\tau_l}{\tau_t}\right)_{\Delta\varepsilon = E}}{\left(\frac{\tau_l}{\tau_t}\right)_{\Delta\varepsilon = 2E}}\,. \tag{48}$$

This ratio attains it maximum value when $E^2_{\max} = \dfrac{1}{2\alpha\sqrt{\lambda}}$. For the values of E given here the value of the expression (48) is 1,93, i.e. at the end of the energy interval $(E, 2E)$ the value of $\dfrac{\tau_l}{\tau_t}$ decreases to almost one half of the value assumed at the beginning of this interval. Naturally, the value of $E^2_{\max}$ sensitively depends on the temperature, as the constant α in (47a) and (47b) greatly varies with temperature and impurity concentration. In our formulae the energy-dependence of $\dfrac{\tau_l}{\tau_t}$ attains significance when the maximum of the energy distribution of electrons is between E and $2E$, i.e. the sensitive energy-dependence of $\dfrac{\tau_l}{\tau_t}$ characterized by the expression (48) is in such a region of the energy spectrum where according to the Boltzmann distribution many electrons are located. Let us choose the temperature or the impurity concentration so that $E = kT$, where T is the absolute temperature and treat this case numerically. Let there be in general

$$E_{\max} = kT\gamma,$$

and

$$\Delta\varepsilon = E_{\max}\, x = kT\gamma x\,. \tag{49}$$

In the special case chosen by us $\gamma = 1$.

The mean values to be calculated can be written in the following form:

$$\langle k, l\rangle = C\int_0^\infty \Delta\varepsilon^{3/2}\left(\frac{\tau_t}{m_t}\right)^n\left(\frac{\tau_l}{m_l}\right)^l e^{-\frac{\Delta\varepsilon}{kT}}\, d(\Delta\varepsilon)\,, \tag{50}$$

where C is the normalizing constant chosen to be unity.

The concrete form of the integral can be given with the aid of (45)—(49)

$$\langle k, l\rangle_\gamma = E^{3/2(k+l+1)}\left(\frac{v_{it}}{m_t}\right)^k\left(\frac{v_{il}}{m_l}\right)^l\int_0^\infty x^{3/2}\left(\frac{x^{3/2}}{1+\frac{x^2}{2\sqrt{\lambda}}}\right)^k \cdot \left(\frac{x^{3/2}}{1+\frac{\sqrt{\lambda}x^2}{2}}\right)^l e^{-x\gamma}\, dx\,. \tag{51}$$

Let there further be

$$\langle k, l\rangle_\gamma = E^{3/2(k+l+1)}\left(\frac{v_{it}}{m_t}\right)^{k+l}\langle k, l\rangle^*_\gamma\,. \tag{52}$$

It can readily be seen that in expression (41) of ΔG_τ which is to be calculated the multiplying factor of $\langle k, l \rangle_\gamma^*$ in (52) is cancelled. Therefore in the case $\gamma = 1$ only the calculation of the integrals $\langle k, l \rangle_\gamma^*$ is required. The values of the effective mass are well-known from the literature [11]:

$$m_t = 0{,}0815 \quad \text{and} \quad m_l = 1{,}588 .$$

The numerical values of the integrals are shown in the following Table

Integral	
$\langle 1, 0 \rangle_{\gamma=1}^*$	2,044
$\langle 0, 1 \rangle_{\gamma=1}^*$	0,249
$\langle 2, 0 \rangle_{\gamma=1}^*$	3,618
$\langle 3, 0 \rangle_{\gamma=1}^*$	6,76
$\langle 1, 1 \rangle_{\gamma=1}^*$	0,390
$\langle 2, 1 \rangle_{\gamma=1}^*$	0,681
$\langle 1, 2 \rangle_{\gamma=1}^*$	0,0757

Substituting these values in expression (41) of ΔG_τ we obtain

$$\Delta G_\tau |_{W_I = 0} = -0{,}023 .$$

This value corresponds to the particular temperature or impurity concentration in the case $\gamma = 1$. At other values of γ the value of this expression may exceed the one given above, it is, however, probable that it will not be larger than 3—4%, as we have attempted to select the value of γ so that the average values very sensitively react to the energy-dependence of the ratio $\frac{\tau_l}{\tau_t}$.

§ 6. Summary

The deviation of the elastogalvanomagnetic coefficient G from unity can be attributed to two phenomena: the intervalley scattering and the energy-dependence of $\frac{\tau_l}{\tau_t}$. The former becomes extremely strong only at higher temperatures; at low temperature and appropriate impurity concentration the two effects are approximately equally strong. As the term due to the energy-dependence of $\frac{\tau_l}{\tau_t}$ can amount to only a few percent, high precision measure-

ments are required for the experimental check. The accuracy of the measurement can be improved by increasing the intensity of the magnetic field. KEYES carried out his measurements at low temperature with a magnetic field of an intensity of 3000 Gauss and the error of measurement was approximately 3—4 percent. The minimum characterizing the mixing of the two scattering mechanisms (e.g. at a given temperature, varying the impurity concentration) amounts to a few percent, and so by tripling the magnetic field intensity the value of the quantities to be measured increase by a factor of 8 to 9, and the error decreases accordingly. In addition to the mixing of the two scattering mechanisms the relaxation time approximation can also be checked by measuring the elastogalvanomagnetic coefficient in the temperature range of impurity scattering. At sufficiently low temperatures the latter type of conduction leads to an approximately energy-independent ratio $\frac{\tau_t}{\tau_l}$ so G must be unity. For a possible deviation from unity the relaxation time tensor approximation is responsible.

My thanks are due to Prof. ZALÁN BODÓ, Doctor of Physical Sciences, for numerous discussions on transport phenomena and valuable advice.

Appendix

The tensors diagonalized in the coordinate system of the relevant valley have the following form

$$\mathbf{T} = \begin{pmatrix} t_t & o & o \\ o & t_t & o \\ o & o & t_l \end{pmatrix}.$$

These will now be transformed to the coordinate system determined by the crystal axes. (The number of the valleys is illustrated by Fig. 1).

Let there be $\mathbf{T}^{(i)}$ the appropriate tensor of the valley (i), which can be written in the form

$$\mathbf{T}^{(i)} = \frac{2t_t + t_l}{3}\mathbf{I} + \frac{t_l - t_l}{3}\mathbf{T}'^{(i)}, \tag{A,1}$$

where $\mathbf{I}$ is the unity tensor and

$$\mathbf{T}'^{(1)} = \begin{pmatrix} o & -1 & -1 \\ -1 & o & -1 \\ -1 & -1 & o \end{pmatrix}, \qquad \mathbf{T}'^{(2)} = \begin{pmatrix} o & 1 & 1 \\ 1 & o & -1 \\ 1 & -1 & o \end{pmatrix},$$

$$\mathbf{T}'^{(3)} = \begin{pmatrix} o & -1 & 1 \\ -1 & o & 1 \\ 1 & 1 & o \end{pmatrix}, \qquad \mathbf{T}'^{(4)} = \begin{pmatrix} o & 1 & -1 \\ 1 & o & 1 \\ -1 & 1 & o \end{pmatrix}. \tag{A,2}$$

Obviously

$$\Sigma \mathbf{T}'^{(i)} \equiv 0$$

and hence

$$\frac{\Sigma \mathbf{T}^{(i)}}{4} = \frac{2t_l + t_l}{3}\mathbf{I} = \frac{1}{3}(\mathrm{Spur}\,\mathbf{T})\,\mathbf{I}\,,$$

which means that the tensor **T** averaged over the four energy minima is a scalar.

REFERENCES

1. R. W. Keyes, Phys. Rev., **103,** 1240, 1956.
2. C. S. Smith, Phys. Rev., **94,** 42, 1954.
3. C. Herring, Bell. System Techn. J., **34,** 237, 1955.
4. C. Herring and E. Vogt, Phys. Rev., **101,** 944, 1956.
5. F. S. Morin, T. H. Gaballe and C. Herring, Phys. Rev., **105,** 525, 1957.
9. C. H. Herring, T. H. Gaballe and J. E. Kunzler, Bell. System, Techn. J., **38,** 657, 1959.
7. G. Weinreich, T. M. Sanders and H. G. White, Phys. Rev., **114,** 33, 1959.
8. *А. Т. Самойлович, И. М. Коренблит, И. В.* Даховский и В. Д. Искра, Ф. Т. Т. 3285, 1961.
9. *А. Т. Самойлович,* И. М. Коренблит и И. В. Даховский, ДАН СССР, 139, 355, 1961.
10. F. Ham, Phys. Rev., **100,** 1251, 1955.
11. B. W. Levinger and D. R. Frankl, J. Phys. Chem. Solids, **20,** 281, 1961.

ПЬЕЗОГАЛЬВАНОМАГНИТНОЕ ЯВЛЕНИЕ ГЕРМАНИЯ ТИПА *n*

А. ЗАВАДОВСКИ

Р е з ю м е

Расчитан эластогальваномагнитный коэффициент германия типа *n*, при направлении магнитного поля и оси образца [1,0,0] и [0,1,0] соответственно с помощью приближения тензора времени релаксации Герринга—Фогта. Если отношение компонент тензора времени релаксации не зависит от энергии, то эластогальваномагнитный коэффициент равен точно единице. Это условие выполняется в случае чисто акустического рассеяния и при рассеянии только на примесях. Отклонения от этого ожидается в случае рассеяния между различными минимумами энергии или в случае смешивания рассеяния на фононах и на примесях. Последний эффект приводит в крайнем к несколько процентному уменьшению эластогальваномагнитного коэффициента.

Printed in Hungary

A kiadásért felel az Akadémiai Kiadó igazgatója Műszaki szerkesztő: Farkas Sándor

A kézirat nyomdába érkezett: 1964. VII. 20. — Terjedelem: 7,25 (A5) ív, 19 ábra

64.59218 — Akadémiai Nyomda, Budapest — Felelős vezető: Bernát György

The *Acta Physica* publish papers on physics, in English, German, French and Russian.

The *Acta Physica* appear in parts of varying size, making up volumes.

Manuscripts should be addressed to:

Acta Physica, Budapest 502, Postafiók 24.

Correspondence with the editors and publishers should be sent to the same address.

The rate of subscription to the *Acta Physica* is 110 forints a volume. Orders may be placed with "Kultúra" Foreign Trade Company for Books and Newspapers (Budapest I., Fő u. 32. Account No. 43-790-057-181) or with representatives abroad.

Les *Acta Physica* paraissent en français, allemand, anglais et russe et publient des travaux du domaine de la physique.

Les *Acta Physica* sont publiés sous forme de fascicules qui seront réunis en volumes.

On est prié d'envoyer les manuscrits destinés à la rédaction à l'adresse suivante:

Acta Physica, Budapest 502, Postafiók 24.

Toute correspondance doit être envoyée à cette même andresse.

Le prix de l'abonnement est de 110 forints par volume.

On peut s'abonner à l'Entreprise du Commerce Extérieur de Livres et Journaux «Kultúra» (Budapest I., Fő u. 32. — Compte-couraut No. 43-790-057-181) ou à l'étranger chez tous les représentants ou dépositaires.

«*Acta Physica*» публикют трактаты из области физических наук на русском, немецком, английском и французском языках.

«*Acta Physica*» выходят отдельными выпусками разного объема. Несколько выпусков составляют один том.

Предназначенные для публикации рукописи следует направлять по адресу:

Acta Physica, Budapest 502, Postafiók 24.

По этому же адресу направлять всякую корреспонденцию для редакции и администрации.

Подписная цена «*Acta Physica*» — 110 форинтов за том. Заказы принимает предприятие по внешней торговле книг и газет «*Kultúra*» (Budapest I., Fő u. 32. Текущий счет: № 43-790-057-181) или его заграничные представительства и уполномоченные.

24,– Ft

INDEX

Acta Phys. Hung. Tom. XVII. Fasc. 4, Budapest 30. XI. 1964.

Index: 26.022