

Acta Physica Hungarica

VOLUME 62, NUMBER 1, 1987

EDITOR-IN-CHIEF

I. KOVÁCS

EDITORIAL BOARD

**Z. BAY, R. GÁSPÁR, I. GYARMATI, N. KÜRTI,
K. NAGY, L. PÁL, A. SZALAY, P. SZÉPFALUSY, I. TARJÁN,
B. TELEGDI, L. TISZA, E. WIGNER**



Akadémiai Kiadó, Budapest

ACTA PHYS HUNG APAHQ 62 (1) 1-100 (1987) HU ISSN 0231-4428

ACTA PHYSICA HUNGARICA

A JOURNAL OF THE HUNGARIAN ACADEMY
OF SCIENCES

EDITED BY
I. KOVÁCS

Acta Physica publishes original papers on subjects in physics. Papers are accepted in English, French, German and Russian.

Acta Physica is published in two yearly volumes (4 issues each) by

AKADÉMIAI KIADÓ
Publishing House of the Hungarian Academy of Sciences
H-1054 Budapest, Alkotmány u. 21

Subscription information

Orders should be addressed to

KULTURA Foreign Trading Company
1389 Budapest P.O. Box 149

or to its representatives abroad.

Acta Physica Hungarica is abstracted/indexed in Chemical Abstracts, Current Contents-Physical, Chemical and Earth Sciences, Mathematical Reviews, Science Abstracts, Physics Briefs, Risk Abstracts

© Akadémiai Kiadó, Budapest

ACTA PHYSICA HUNGARICA

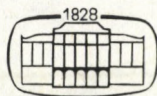
EDITORIAL BOARD

Z. BAY, R. GÁSPÁR, I. GYARMATI, N. KÜRTI, K. NAGY, L. PÁL,
A. SZALAY, P. SZÉPFALUSY, I. TARJÁN, B. TELEGDI, L. TISZA,
E. WIGNER

EDITOR-IN-CHIEF

I. KOVÁCS

VOLUME 62



AKADÉMIAI KIADÓ, BUDAPEST

1987

CONTENTS

Volume 62

GENERAL PHYSICS

Correlations in a nonequilibrium steady state: Exact results for a generalized kinetic Ising model. <i>Z. Rácz and M. Plischke</i>	203
A symbolic algorithm for finding exactly soluble statistical mechanical models. <i>P. Ruján</i>	287
Unusual maps and their use to approach usual ones. <i>Z. Kaufmann, P. Szépfalussy and T. Tél</i> ...	321

ELEMENTARY PARTICLES AND FIELDS

Infrared asymptotics of the quark propagator in a Bloch–Nordsieck-type model. <i>G. Pócsik and T. Torma</i>	101
Infrared asymptotics of the quark propagator in nonabelian gauge theories I. <i>G. Pócsik and T. Torma</i>	107
Remarks on the bilocal field theories. <i>Z. Fodor and E. Regős</i>	115
Energy-energy correlations for the $p\bar{p}$ collider. <i>F. Csikor and G. Pócsik</i>	155
Scalar Bethe–Salpeter equation and the relativistic bound state problem. <i>K. Ladányi</i>	161
On a problem of spontaneous compactification. <i>P. Forgács, Z. Horváth and L. Palla</i>	177
The nature of the deconfining phase transition in the SU(2) gauge theory in $1+\epsilon$ dimension. <i>A. Patkós</i>	185
Expansions at the parameter space boundary in the standard Higgs model. <i>I. Montvay</i>	191
Weak interactions of heavy quarks. <i>H. Pietschmann</i>	211
The possibility of avoiding the axion in the Peccei–Quinn mechanism. <i>T. Nagy</i>	237
Yang–Feldman formalism of dipole fields. <i>K. L. Nagy</i>	257
Canonical quantization of the relativistic theory of the Dirac monopole. <i>A. Frenkel</i>	307

NUCLEAR PHYSICS

Fragment yields in the microcanonical model of nuclear disassembly. <i>G. Fai</i>	277
Nucleon–nucleon interaction with ultra short range attraction. <i>I. Lovas, Erika Lovas and K. Sailer</i>	347
A quantitative analysis of the collective outward motion of nuclear matter in central Ca + Ca and Nb + Nb collisions. <i>L. P. Csernai and G. Fai</i>	355
A three-fluid model of heavy ion collisions. <i>J. Németh</i>	365
The effect of correlations on the entropy and hadrochemical composition in heavy ion reactions. <i>H-W. Barz, T. S. Biró, B. Lukács and J. Zimányi</i>	371

ATOMIC AND MOLECULAR PHYSICS

Spin orbitals and total energy calculated by the $X\alpha$ method including ab initio self-consistent exchange parameters α_{SCF} . <i>R. Gáspár and Á. Nagy</i>	131
---	-----

FLUIDS, PLASMAS AND ELECTRIC DISCHARGES

Hamilton principle for the vortex flow of an ideal fluid in special relativity. <i>I. Abonyi</i>	199
Inversion method to test radial structures in cylindrical plasmas. <i>K. G. Antal, J. F. Bitó and F. Halász</i>	225

CONDENSED MATTER

Preface to "CRYOSEM 3". <i>I. Kirschner</i>	5
Size- and high field effects in MOS-FET devices. <i>T. Porjesz, G. Zsolt, Gy. Kovács and T. Kármán</i>	7
Calculation of physical parameters of MOS structures at different temperatures. <i>G. Kovács, I. Kirschner, T. Porjesz, T. Kármán and G. Zsolt</i>	11
Resistivity measurements in silicon inversion layers. <i>G. Gombos, M. M. Mostafa and H. Al-Sharaby</i>	15
Avalanche current relaxation in p-Si MOSFET-s. <i>G. Zsolt, Gy. Kovács, T. Porjesz, T. Kármán and G. Gombos</i>	19
Irreversibility of magnetic flux penetration in superconductors. <i>I. Kirschner</i>	23
Microscopic calculations of vortex structure and magnetization curves for type II superconductors. <i>U. Klein, L. Kramer, W. Pesch, D. Rainer and J. Rammer</i>	27
Elementary quantum physical description of triplet superconductors. <i>K. Martinás and B. Lukács</i>	31
Simulation of I-V steps sharpening for Josephson junctions. <i>T. F. Refai and L. N. Shehata</i>	35
Computer simulation of the correlation length in type II superconductor in chaotic approach. <i>P. Samu</i>	39
Thermodynamic cross effects in type II superconductors. <i>G. Kiss</i>	43
Computer simulation of pinning in type-II superconductors and the theory of collective pinning (Abstract). <i>E. H. Brandt</i>	47
Analogies between granular superconductors and spin glasses (Abstract). <i>J. Rosenblatt, P. Peyral, A. Rabotou and C. Lebeau</i>	49
Anisotropy effects in superconductors down to millikelvin temperatures. <i>F. M. Sauerzopf and H. W. Weber</i>	51
On some developments of technology of multifilamentary wire V_3Ga . <i>J. Bánkúti, É. Hegyi Farkas and L. Lászlóffy</i>	55
Computer designed superconducting magnets, <i>T. Kármán, I. Kirschner and J. Bánkúti</i>	59
Biomagnetic measurements with a third order gradiometer. <i>K. Vad, S. Mészáros and G. Halász</i>	63
Magnetic properties of $(Fe_{1-x}M_x)_7Se_8$ (Abstract). <i>M. M. Abd-El Aal</i>	67
Magnetic properties and spontaneous distortion in phase laves intermetallic compounds. <i>M. M. Abd-El Aal, A. S. Ilyshin, V. I. Chechernikov and A. V. Pechenikov</i>	69
Low temperature magnetization of $Sm_{1-x}Er_xFe_2$ compounds. <i>M. M. Abd-El Aal, V. I. Chechernikov, L. I. Kazakova and A. V. Cheremushkina</i>	73
Influence of thickness on magnetic phase transitions in thin films of dysprosium and samarium. <i>J. Dudás, A. Fehér and S. János</i>	77
Residual resistivity investigations on high purity gallium. <i>D. Novák, S. Mészáros and K. Vad</i> ..	81
Study of non-isothermal phase transformation in $As_2Se_3-As_2Te_3$ glasses. <i>M. B. El-Den and M. K. El Mously</i>	85
Electron mobility in high density neon gas. <i>A. F. Borghesani, L. Bruschi, M. Santini and G. Torzo</i> ..	89
Are the stability and superconductivity of alloys connected? <i>A. Szász and J. Bánkúti</i>	93
Mechanical properties of teflons at low temperatures. <i>L. Lászlóffy and J. Bánkúti</i>	97
Zero modes and reparametrization invariance in the Ising spin glass. <i>I. Kondor and R. Németh</i> ..	219
The glass transition feedback: Review and comments. <i>T. Geszti</i>	267

ASTROPHYSICS

Entropy in the universe. <i>G. Marx</i>	139
Density peaks and large scale velocities. <i>Alexander S. Szalay and Lars Gerhard Jensen</i>	263

INTERDISCIPLINARY

Global-scale changes of the environment: Observations from space. <i>K. Ya. Kondratyev</i>	243
--	-----

C R Y O S E M ' 8 6

IIIrd International Seminar on
Low Temperature Physics

Visegrád, Hungary

May 26-31, 1986

Manuscript received by Akadémiai Kiadó: 19 June 1987
Manuscript received by the Printers:

Date of publication: 10 August 1987

PRINTED IN HUNGARY

Akadémiai Kiadó és Nyomda, Budapest

CONTENTS

PREFACE I. Kirschner	5
CONDENSED MATTER	
SIZE- AND HIGH FIELD EFFECTS IN MOS-FET DEVICES. T. Porjesz, G. Zsolt, Gy. Kovács and T. Kármán	7
CALCULATION OF PHYSICAL PARAMETERS OF MOS STRUCTURES AT DIFFERENT TEMPERATURES. G. Kovács, I. Kirschner, T. Porjesz, T. Kármán and G. Zsolt	11
RESISTIVITY MEASUREMENTS IN SILICON INVERSION LAYERS. G. Gombos, M. M. Mostafa and H. Al-Sharaby	15
AVALANCHE CURRENT RELAXATION IN p-Si MOSFET-S. G. Zsolt, Gy. Kovács, T. Porjesz, T. Kármán and G. Gombos	19
IRREVERSIBILITY OF MAGNETIC FLUX PENETRATION IN SUPER- CONDUCTORS. I. Kirschner	23
MICROSCOPIC CALCULATIONS OF VORTEX STRUCTURE AND MAGNET- IZATION CURVES FOR TYPE II SUPERCONDUCTORS. U. Klein, L. Kramer, W. Pesch, D. Rainer and J. Rammer	27
ELEMENTARY QUANTUM PHYSICAL DESCRIPTION OF TRIPLET SUPER- CONDUCTORS. K. Martinás and B. Lukács	31
SIMULATION OF I-V STEPS SHARPENING FOR JOSEPHSON JUNCTIONS. T. F. Refai and L. N. Shehata	35
COMPUTER SIMULATION OF THE CORRELATION LENGTH IN TYPE II SUPERCONDUCTOR IN CHAOTIC APPROACH. P. Samu	39
THERMODYNAMIC CROSS EFFECTS IN TYPE II SUPERCONDUCTORS. G. Kiss	43
COMPUTER SIMULATION OF PINNING IN TYPE-II SUPERCONDUCTORS AND THE THEORY OF COLLECTIVE PINNING /Abstract/. E.H. Brandt	47
ANALOGIES BETWEEN GRANULAR SUPERCONDUCTORS AND SPIN GLASSES /Abstract/. J. Rosenblatt, P. Peyral, A. Raboutou and C. Lebeau	49
ANISOTROPY EFFECTS IN SUPERCONDUCTORS DOWN TO MILLIKELVIN TEMPERATURES. F. M. Sauerzopf and H. W. Weber	51
ON SOME DEVELOPMENTS OF TECHNOLOGY OF MULTIFILAMENTARY WIRE V ₃ Ga. J. Bánkuti, É. Hegyi Farkas and L. Lászlóffy	55
COMPUTER DESIGNED SUPERCONDUCTING MAGNETS. T. Kármán, I. Kirschner and J. Bánkuti	59
BIOMAGNETIC MEASUREMENTS WITH A THIRD ORDER GRADIOMETER. K. Vad, S. Mészáros and G. Halász	63

MAGNETIC PROPERTIES OF $Fe_{1-x}^M/7Se_8$ /Abstract/.	
M. M. Abd-El Aal	67
MAGNETIC PROPERTIES AND SPONTANEOUS DISTORTION IN PHASE LAVES INTERMETALLIC COMPOUNDS.	
M. M. Abd-El Aal.	
A. S. Ilyishin, V. I. Chechernikov and A. V. Pechenikov	69
LOW TEMPERATURE MAGNETIZATION OF $Sm_{1-x}Er_xFe_2$ COMPOUNDS.	
M. M. Abd-El Aal, V. I. Chechernikov, L. I. Kazakova and A. V. Cheremushkina	73
INFLUENCE OF THICKNESS ON MAGNETIC PHASE TRANSITIONS IN THIN FILMS OF DYSPROSIUM AND SAMARIUM.	
J. Dudás, A. Feher and S. János	77
RESIDUAL RESISTIVITY INVESTIGATIONS ON HIGH PURITY GALLIUM.	
D. Novák, S. Mészáros and K. Vad	81
STUDY OF NON-ISOTHERMAL PHASE TRANSFORMATION IN As_2Se_3 - As_2Te_3 GLASSES.	
M. B. El-Den and M. K. El Mously	85
ELECTRON MOBILITY IN HIGH DENSITY NEON GAS.	
A. F. Borghesani, L. Bruschi, M. Santini and G. Torzo	89
ARE THE STABILITY AND THE SUPERCONDUCTIVITY OF ALLOYS CONNECTED?	
A. Szász and J. Bánkuti	93
MECHANICAL PROPERTIES OF TEFLONS AT LOW TEMPERATURES.	
L. Lászlóffy and J. Bánkuti	97

The final manuscript was received by the Editorial Office on 9 June 1987.

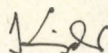
PREFACE

The CRYOSEM, our international seminar on low temperature physics and technology has been organized third time.

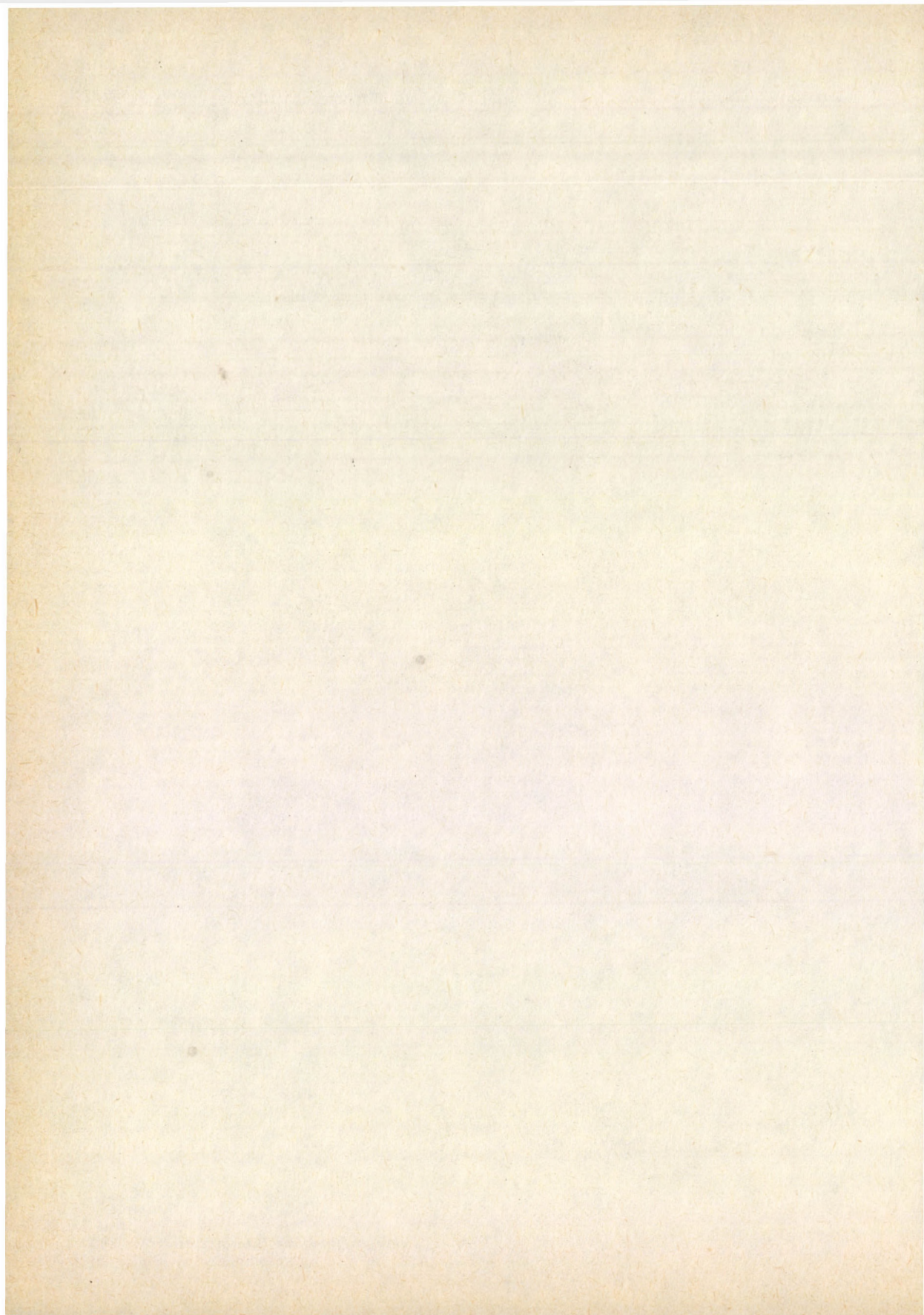
Participants from the different countries met between 26th and 31st May 1986 to discuss the research aims and achievements of their Institutes.

The CRYOSEM'86 was hosted by Visegrád, a small Hungarian holiday resort in the Danube bend.

The purpose of the meeting was to provide participants an opportunity for informal and free discussions, concerning any problem connected with low temperatures. The reports held at the CRYOSEM'86 will be published in this issue.



Prof. I. Kirschner



SIZE- AND HIGH FIELD EFFECTS IN MOS-FET DEVICES

T. PORJESZ, G.ZSOLT, GY. KOVÁCS, T. KÁRMÁN

Department for Low Temperature Physics, Roland Eötvös University
1088 Budapest, Hungary

One of the most prosperous chapters of the applied sciences is the solid state electronics. The ever decreasing sizes and the same time the increasing amount of semiconductor devices in one chip gave the importance to study the new physical phenomena of dense systems and small devices.

Introduction

The physical models describing the transport phenomena in solids are based mainly on the Boltzmann transport theory. This approximation is valid with strict assumptions only. The validity of these restrictions is doubtful in the case of the recent solid state devices [1].

Our investigations were carried out on semiconductor microstructures on dense systems of small devices. In the small devices the carriers are often in non-steady state. In a dense system of such devices the bulk-device picture loses its validity and surface effects and the interactions between devices will be dominant.

The understanding and the explanation of such phenomena and their correct physical interpretation have a crucial role both in solid state physics and the technical progress. Therefore the research of this field has a great importance.

Among our intentions the study of the role of hot carriers in an avalanche breakdown [2] and Van der Pauw measurements on MOS inversion layers [3] have been carried out so far.

Samples

The samples were designed and prepared according to our purposes. The samples are MOS FET systems of a rather simple type in order to avoid the disturbing effects due to a complicated structure.

The good availability and easy handling of the most widespread planar technology gave us the idea to apply it for sample construction.

Parallel p^+ doped regions were diffused into an n-type silicon layer with equivalent width and distances to each other giving a lattice-like structure. Deeper p^+ diffusion was used to get contact to the p-type substrate and a contact was made to the n-type layer as well. The last two made it possible to use a bias voltage in order to avoid the noise current. All this structure has been covered by a silicon-oxide insulated gate electrode. Five p^+ -type electrodes have been made in our samples (Fig. 1.).

The number of the electrodes can be increased up to 200 or more, but in this case individual connection to each one can not be made.

By the aid of this sample a tunable Kronig-Penney model has been realised completed with variable depth channels. For the investigation of the superlattice effects a higher number of electrodes is needed. For transport experiments in certain cases individual electrodes are more suitable.

The dimensions of the devices are dependent on the lithographical possibilities. Our first samples had the characteristic dimensions of $10 \mu\text{m}$. The second generation of our samples has the characteristic dimensions of $5 \mu\text{m}$. We should like to reach the $2 \mu\text{m}$ within a reasonable time.

Measuring system

A computer controlled measuring system has been designed and set up in order to study conduction phenomena and the transport properties of samples subjected to different effects. The system automatically permutes the current and voltage contacts in a given sequence. The block diagram is shown in [3].

Size effects

The conventional operational principles of MOS-FET work as far as the channel length is much higher than the mean free path and the transistors are far enough from each other to avoid the interaction between them.

The difficulties of the explanation can arise from the decreasing sizes or by lowering the temperature. Both can cause the mean free path to be comparable to the channel length.

To examine the size effects we use the low temperature measurement technique. In this way a transistor having a $10 \mu\text{m}$ channel can be treated as a small device at sufficiently low temperature, the mean free path being in the range of centimetres at helium temperature.

High field effects

The operating voltage for a transistor must not be less than a few volts. In our case, supposing 10 V for $10 \mu\text{m}$ channel we get 10^6 V/m field strength. In the smallest existing device - having 100 nm channel - for an operating voltage of 1 V 10^7 V/m field strength will be resulted. In the next future we have to consider even smaller devices and not lower operating voltages. There is no way to neglect the high field effects any more. We can say that below certain dimensions all transport phenomena should be regarded as hot effects. The energy of the electron will be much higher than that of the equilibrium one, exceeding the minimum energy that is enough for inelastic scattering, for avalanche effects, etc.

Conclusions

We found our measuring system to be applicable to collect automatically a large amount of data and the direct connection to the computer makes the evaluation easy at the same time. The detailed results are reported in this issue [2,3].

References

1. T. Porjesz, Acta Phys. Hung., 53, 411, 1982.
2. G. Zsolt, Gy. Kovács, T. Porjesz, T. Kármán, G. Gombos, Acta Phys. Hung., this issue, p.19.
3. G. Gombos, M.M. Mostafa, H. Al-Sharaby, Acta Phys. Hung., this issue, p.15.

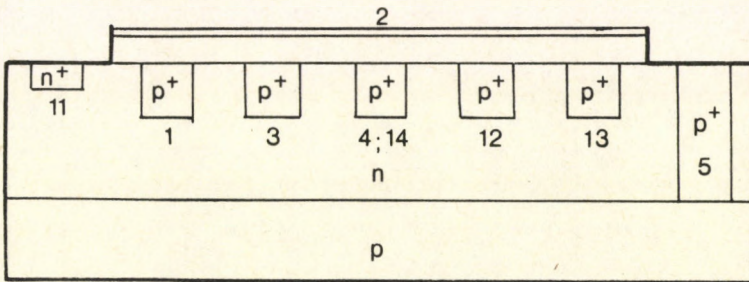
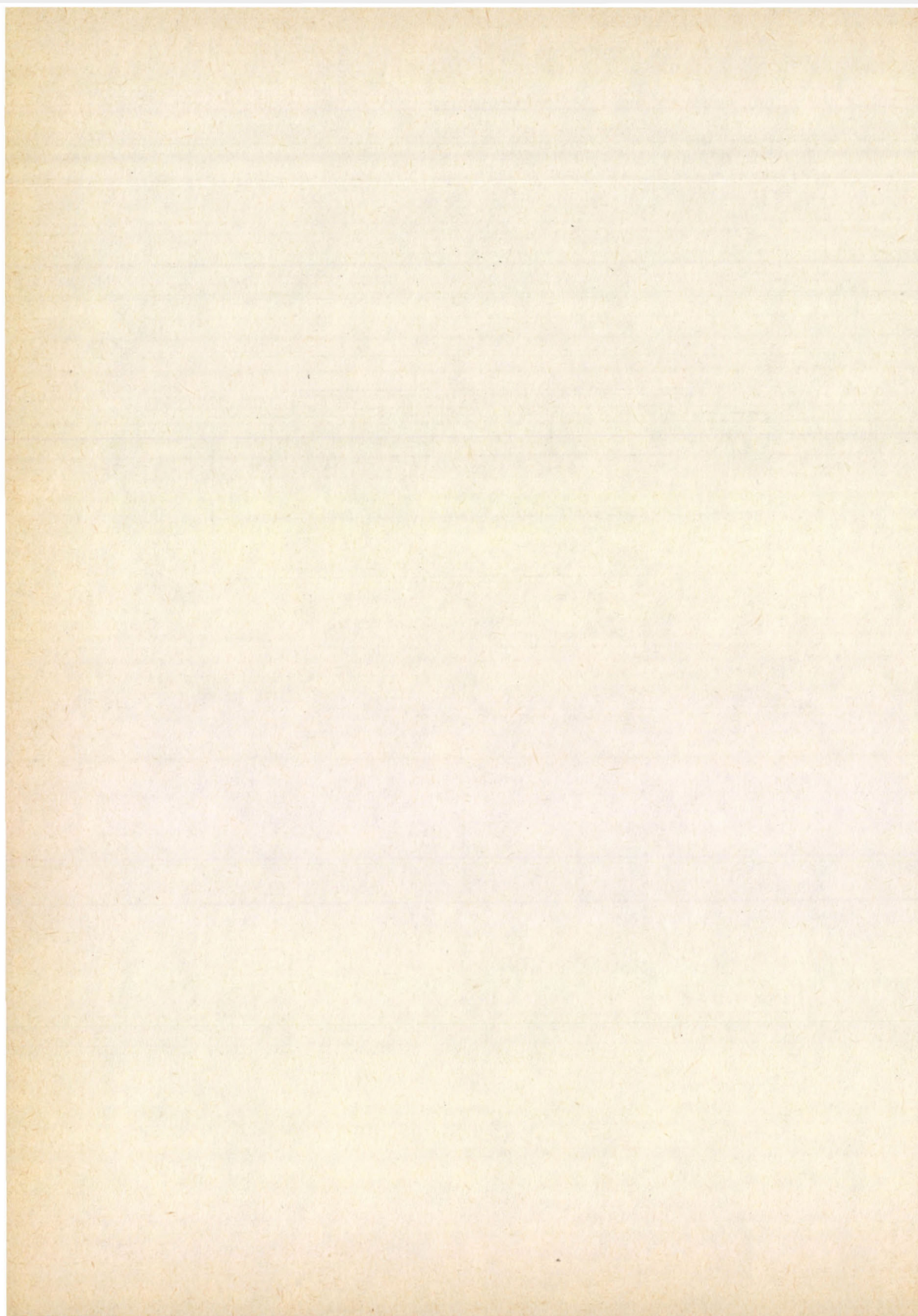


Fig. 1. The structure of the MOSFET device prepared for our experiments



CALCULATION OF PHYSICAL PARAMETERS OF MOS STRUCTURES AT DIFFERENT TEMPERATURES

G.KOVÁCS , I.KIRSCHNER , T.PORJESZ , T.KÁRMÁN , G.ZSOLT

Department for Low Temperature Physics Roland Eötvös University
1088 Budapest, Hungary

Exact theoretical equation will be given for the temperature dependence of the threshold voltage in MOS structures. Two different temperature dependences have been originated because of different assumptions for the bending of the impurity levels. The realistic case should be found between them.

Introduction

One of the main trends in the technological progress is to develop different MOS structures.

Recently these devices are used mainly at room temperature, but in the big computers of high speed there is a tendency to apply them even at liquid He temperature. The higher speed of the devices is originated from the higher mobility of the charge carriers at low temperatures. Another advantage of the low operation temperature is the higher heat conductivity that is the higher dissipation capability.

The temperature dependence of the physical parameters of the MOS structures being not clear, therefore it is important to deal with them.

Calculations

Using the basic Eq./1 / to determine the Fermi energy with the following numerical values

$$\begin{aligned} N_c &= 5,53 \cdot T^{3/2} \cdot 10^{21} \text{ m}^{-3} \\ N_v &= 2,89 \cdot T^{3/2} \cdot 10^{21} \text{ m}^{-3} \\ E_c &= 1,12 \text{ eV} \\ E_v &= 0 \text{ eV} \\ E_d &= E_c - 0,05 \text{ eV} = 1,07 \text{ eV} \end{aligned}$$

for $T < 5 \text{ K}$ we got [1]:

$$F = \frac{E_c + E_d}{2} + \frac{kT}{2} \cdot \ln \frac{N_d}{2N_c} \quad /1/$$

After the electric field and the potential distribution in a silicon MOS capacitor is determined /Fig.1 /, we can calculate the electric charge distribution inside the semiconductor. In the presence of bias voltage the conduction, the valence bands and the Fermi energy are not bended. The electric field is given by the Poisson equation. The V_G' gate voltage:

$$V'_G = \frac{kT}{q} \left(U_g + X_O E_s \frac{K_s}{K_O} \right). \quad /2/$$

This V'_G is an ideal gate voltage, the real one is:

$$V_G = V'_G + \Phi_{MS} \quad /3/$$

because the metal-semiconductor work function, Φ_{MS} has to be taken into account. We have to point out that in the capacity measurements DC voltage and a small AC voltage are used. We consider the low frequency case.

Using the positive voltage, the capacity is approximately equal to the capacity of the oxide layer. Applying a small negative voltage, depletion will occur at a high negative voltage an inversion layer will be produced. In this case the total capacity will be the sum of the capacities of the oxide and of the semiconductor channel.

The gate voltage where the inversion begins is called threshold voltage at which the minority carrier concentration at the interface is equal to the majority carrier concentration in the bulk. It is a function of the drain voltage and the temperature. We have to point out that decreasing the temperature the threshold voltage is increasing and the dependence given in Eq. /4/ will be steeper as well.

$$V_T = \frac{kT}{q} \left(2 U_F - \frac{K_O}{E_O K_S} - 4 q N_d^+ K_S E_O \left(-2 \frac{kT}{q} U_F \right) \right). \quad /4/$$

Conclusions

In the Eq. /4/ the donor electron concentration is the most important quantity. It depends on the temperature and the bending. Neglecting the bending of the impurity level, we can find the temperature dependence of the threshold voltage shown in the Fig. 2 curve 'a'.

Assuming that the bending of the impurity level is equal to the bending at the interface all over the bulk, we can find the function, curve 'b', in Fig. 2. The realistic case should be between these two.

The interesting point of the results is that the significant difference between the two functions occurs below 100 K only.

Notations

- N_C effective density of state in the conduction band
- N_V effective density of state in the valence band
- E_C lower edge of the conduction band
- E_V upper edge of the valence band
- E_d donor energy level
- T absolute temperature
- F Fermi energy
- k Boltzmann constant

- V'_G ideal gate voltage
- q electron charge
- U_S reduced band width
- X_O oxide thickness
- E_S field strength at the interface
- K_S dielectric constant of the semiconductor
- K_O dielectric constant of the oxide
- V_G real gate voltage
- ϕ_{MS} metal-semiconductor work function
- V_T threshold voltage
- U_F reduced Fermi energy
- ϵ_0 dielectric constant of the vacuum
- N_d^+ ionized donor concentration

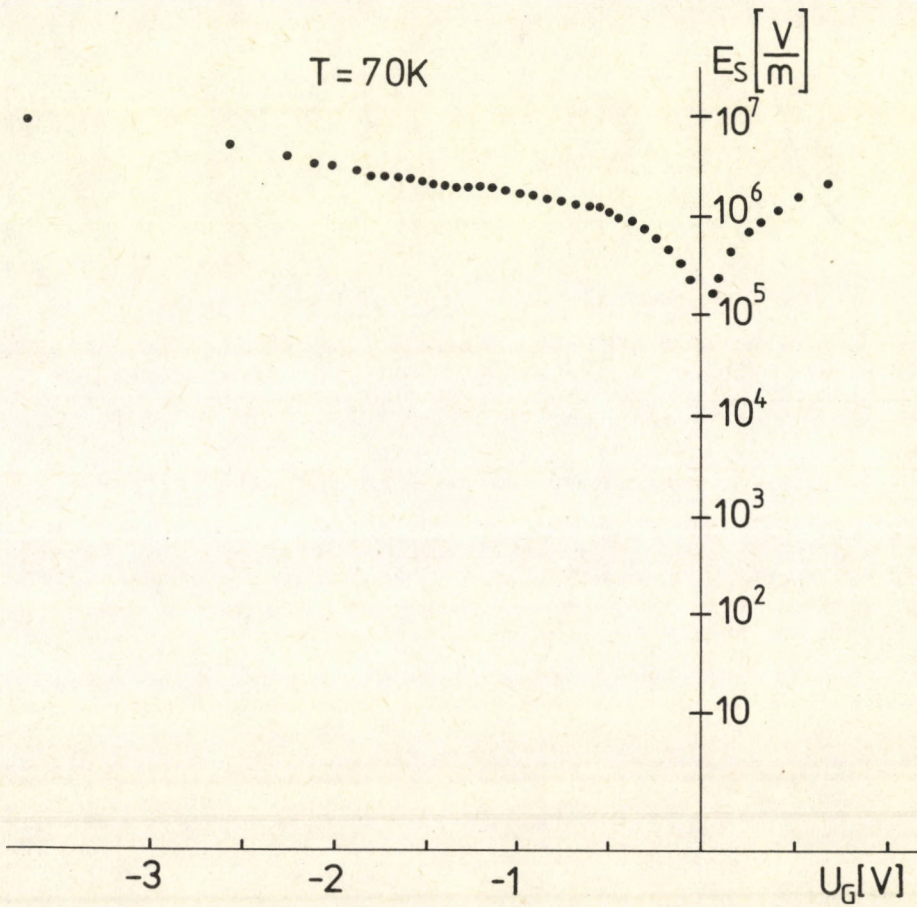


Fig.1. The electric field at the interface in a Si MOS capacitor as the function of the gate voltage

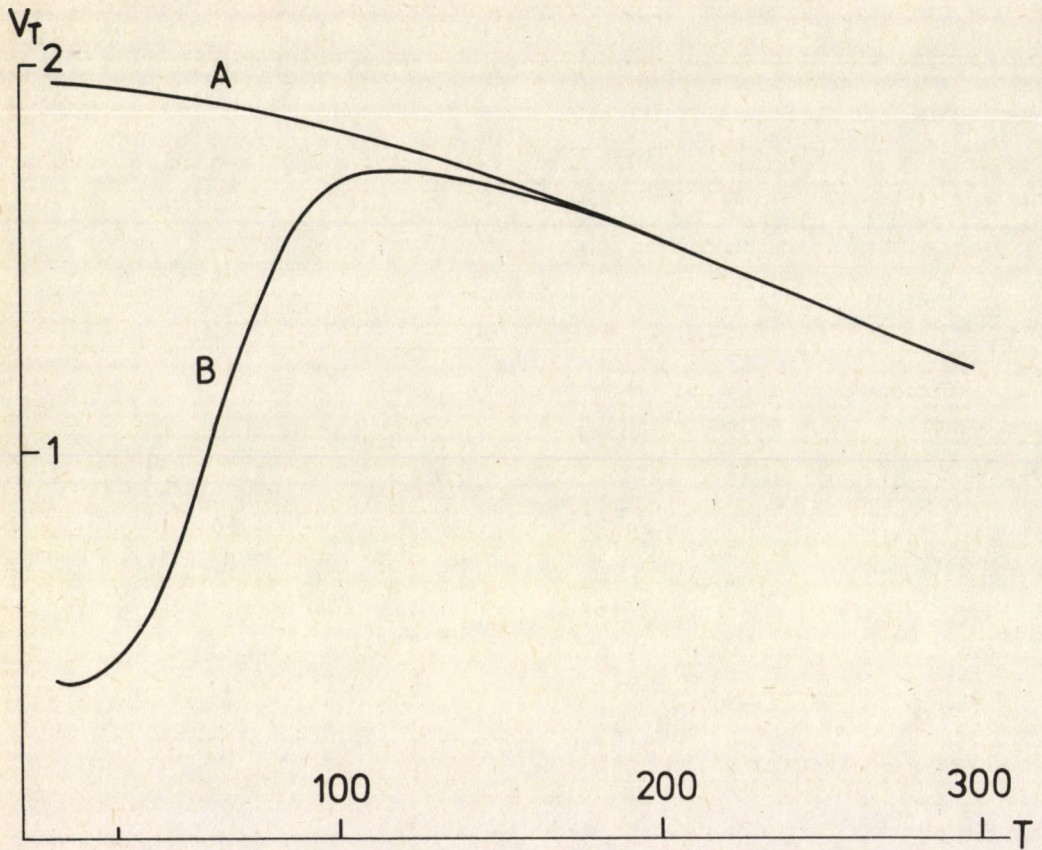


Fig.2. The temperature dependence of the treshhold voltage

References

1. P.Sz. Kirijev, *Félvezetők Fizikája*, Tankönyvkiadó, Budapest, 1974.

RESISTIVITY MEASUREMENTS IN SILICON INVERSION LAYERS

G. COMBOS¹, M.M. MOSTAFA², H. AL-SHARABY³

¹Department for Low Temperature Physics, Roland Eötvös University
Budapest, Hungary

²Faculty of Education, Alexandria University
Alexandria, Egypt

³Department of Physics, University of Baghdad
Baghdad, Iraq

In this paper a simple method is presented for resistivity measurements on silicon inversion layers. The classical Van der Pauw method is applicable only if the thickness of the sample is a given value. In MOSFET structures the layer depth is the function of the gate voltage. We evaluated the inversion layer thickness from the DC drain current-gate voltage characteristic by using the delta depletion theory. After this the Van der Pauw method can be used.

Four point method

We have shown the classical four point method for sample classification in Fig. 1 [1]. There are certain disadvantages of this method having some sources of errors. We have to make often examinations at low temperatures and different magnetic fields. A part of these measurements can be realized by the four point method only with rather great difficulties as it is shown in Fig. 1 but for another part of these measurements this method is not applicable at all. The advantage of the four point method is its quickness when industrial classification or comparison is needed without scientific exactness.

Van der Pauw method

Higher accuracy and more physical informations are needed for the researches where the Van der Pauw method [2] can be applied. This is a classical procedure to determine the resistivity of samples with given thickness. In MOS structures much more information is necessary because there is no direct measurement for the different layer thickness and charge carrier distribution depending on the control gate voltage.

The main advantages of the Van der Pauw method:

1. The measurements can be carried out on samples of arbitrary shape.
2. Evaluating the data from this type of measurements all of the parameters can be calculated which can be gained by using a conventional Hall probe.
3. Using samples cut in different orientations the complete resistivity tensor can be calculated.

After listing the advantageous properties of this measurement we are giving a brief description of realising this process. In Fig. 2 it is shown that we have four measuring contacts to the sample not considering the gate and substrate contacts for controlling the thickness of the inversion layer. The procedure differs from the four point method in the sense that the role of the contacts are changing during the measurements. Through two points for example 'A' and 'B' a given current is driven, while potential drop between 'D'

and 'C' is measured. This is the way to define a resistivity like quantity R_{ABCD} by dividing the voltage V_{DC} by the driving current through 'A' and 'B'. Permutating the contacts written in the subscript we can get different R_{WXYZ} . In this case the relationship given by Van der Pauw holds. From this equation the sheet resistance, the ratio of the specific resistivity to sample thickness can be calculated [3]. The computation can be carried out by dividing of intervals or by using an approximative formula.

The principle of the calculation is that the complex half-plane can be transformed by conformal mapping into a singly connected form of arbitrary shape.

The measurement can be carried out at different temperatures and magnetic fields as well [4]. By the aid of these measurements the variation of the resistance in the function of the temperature and the magnetic field and the measured sheet resistance give us the data to calculate the Hall mobility.

Results

Our investigations were carried out in MOSFET inversion layers. In this case the effective layer depth and the charge carrier distribution can be controlled by a voltage between the substrate and the gate electrode. In principle it is possible to determine experimentally the depth of the inversion layer from the gate voltage value giving zero channel current, although practically there is no zero current regime. This is the reason we worked out a method to determine the thickness.

We determined theoretically the drain current-gate voltage function. The so called delta depletion theory, where the depleted layer thickness is a Dirac delta, results in a parabolic behavior of the drain current-gate voltage relationship. According to this expectation a parabolic curve should be fitted to the measured data. Our results satisfy this relation in the range where the leakage current can be neglected (Fig. 3)

When MOSFET and Van der Pauw samples coexist on the same wafer they can be characterized by the same physical parameters. This is the way to use the value of the layer thickness determined in the MOSFET channel for the Van der Pauw samples.

Using contacts of arbitrary shape they can cause short-circuit for the covered area. To avoid this effect we designed the sample contacts relatively far from the sample as it is shown in Fig. 4. The driving current is led to the sample by conventional Reed relays while the voltages fed into a DVM by a Keithley Low Voltage Scanner. The arrangement of our experimental system is given in the Fig. 5. The computer controlled system permutes the current and voltage contacts in a given sequence.

Comparing the results of this type of measurements to the Hall measurements a good agreement was found. Therefore it can be expected that this method is well applicable both in the industry and research.

References

1. D. Perloff, F. Wahl and J. Conragan, *J. Electrochemical Soc.*, **124**, 582, 1977.
2. L. J. van der Pauw, *Philip's Technische Rundschau*, **20**, 230, 1958.
3. T. Porjesz, P. Debreceeni, I. Kirschner, M. F. Kotkata and Gy. Kovács, *Acta Phys.*, **53**, 361, 1982.
4. T. Porjesz, *Acta Phys.*, **53**, 411, 1982.

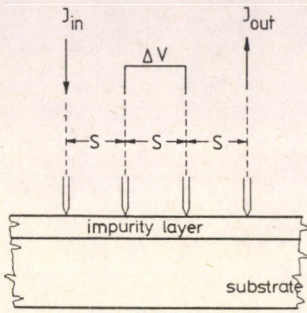


Fig.1. Contacts for the four point method

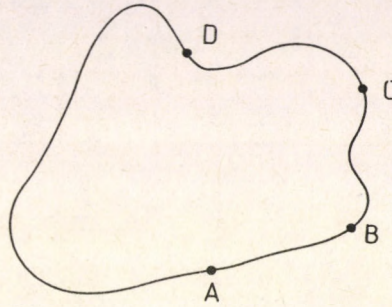


Fig.2. Contacts for the van der Pauw method

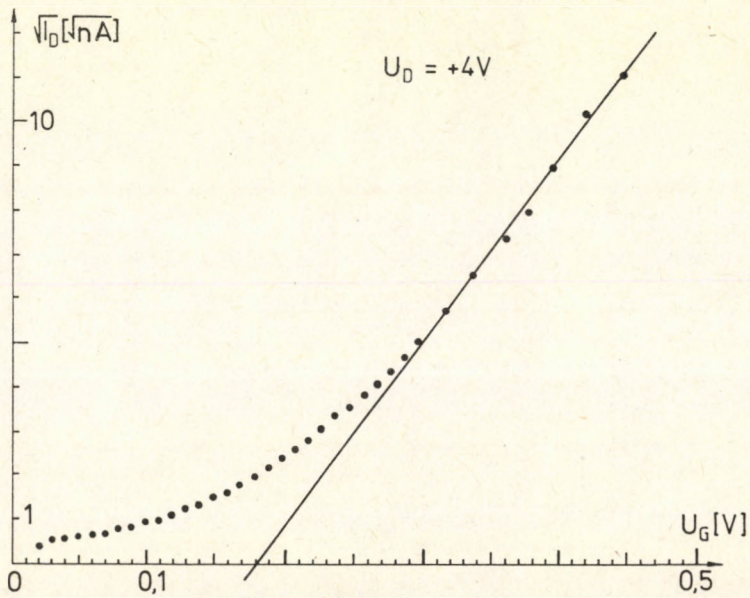


Fig.2. The theoretically expected I_D vs. U_G function and the experimental data

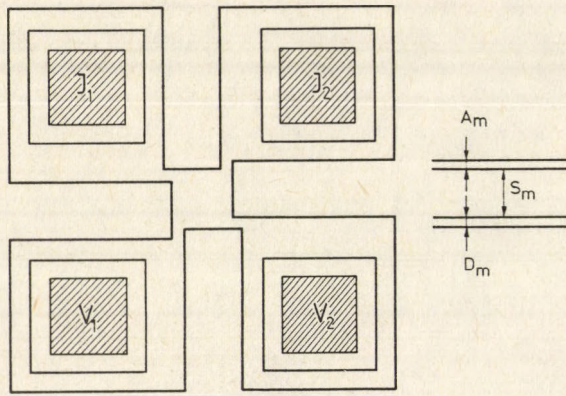


Fig.4. The shape of the sample designed for the Van der Pauw measurements

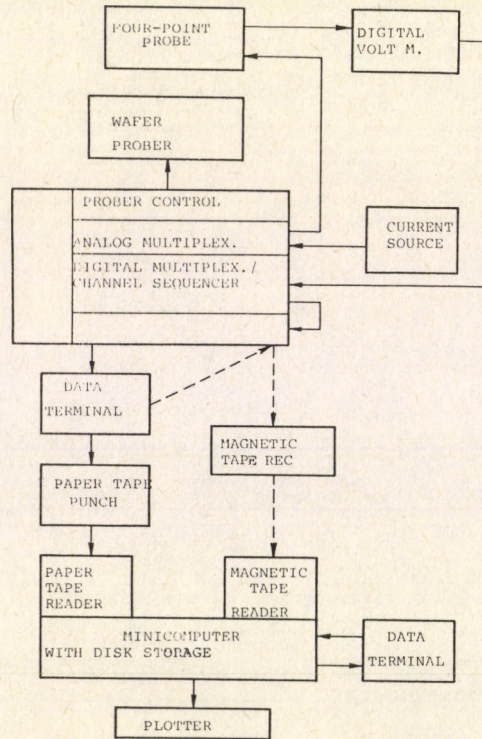


Fig.5. The block diagram of the experimental system

AVALANCHE CURRENT RELAXATION IN p-Si MOSFET-S

G. ZSOLT , GY. KOVÁCS , T.PORJESZ , T.KÁRMÁN , G.GOMBOS

Department for Low Temperature Physics, Roland Eötvös University
Budapest, Hungary

A characteristic time dependence of the avalanche current was found in p-Si MOSFET-s while the normal plateau current was constant. This can be explained by tunneling and trapping of hot carriers into the oxide layer.

Introduction

In the present state of the semiconductor microtechnology, the devices, materials and carriers are subjected to high electric fields because of the small dimensions. In this case we have to take into account the properties of the hot carriers [1,2].

In our paper we report some experimental results about hot carrier effects occurring in the avalanche breakdown of p-Si MOSFET-s.

Earlier experiments

In Fig.1 an experimental arrangement is shown for studying the hot electron emission from Si into SiO₂. The corresponding band structure diagram is shown as well. The electrons are optically generated. The carriers that overcome the potential barrier contribute to the gate current if they are not trapped in the SiO₂ layer [3].

In Fig.2 an experiment is shown to study the electron trapping in the SiO₂ film. Because of a partial trapping electric charge will be accumulated in the oxide, therefore the current changes in time Fig.3 [4].

Our experiments

We studied hot carriers generated in an avalanche breakdown in the surface region. A fraction of the carriers will have sufficient energy to surmount the energy barrier at the Si-SiO₂ interface. Some of them will get trapped in the oxide. In this case a time dependent drain current can be expected because of the influence of the accumulated charge. The rate of the charge trapping inside the oxide layer is related to the number of the hot carriers.

The experiments were carried out in samples detailed in [5].

Results

In the measurements carried out at room temperature we observed a U_o-I_o typical characteristic of a FET, with plateau and breakdown. At 77K we did not find breakdown within the same drain voltage range.

At room temperature we found a decreasing current vs. time function in

the avalanche region (Fig.4).

By a detailed analysis of such curves we can obtain informations about the energy distribution function of the hot carriers and the spatial distribution of the traps.

We carried out a reversed experiment too. We switched on a drain voltage causing a breakdown, waited until the drain current decreased then we switched it off. After that we periodically switched a constant drain voltage on for a very short time to measure the actual avalanche current. The resulting time dependence is shown in Fig.5. It can be seen that by detrapping the carriers the electric charge accumulated in the oxide gradually disappears and the conductivity of the channel will be regenerated.

References

1. I. Kirschner, T. Porjesz, Magyar Fiz. Folyóirat XVI. 111. 1968.
2. I. Kirschner, T. Porjesz, J. Bánkúti and P. Zental, Acta Physica 32, 289, 1972.
3. J. Y. Tang, and K. Hess, J. Appl. Phys., 54, 5145, 1983.
4. P. Olivo, B. Ricco and E. Sangiorgi, J. Appl Phys., 54, 5267, 1983.
5. T. Porjesz, G. Zsolt, Gy. Kovács and T. Kármán, Acta Physica this issue, page 7.

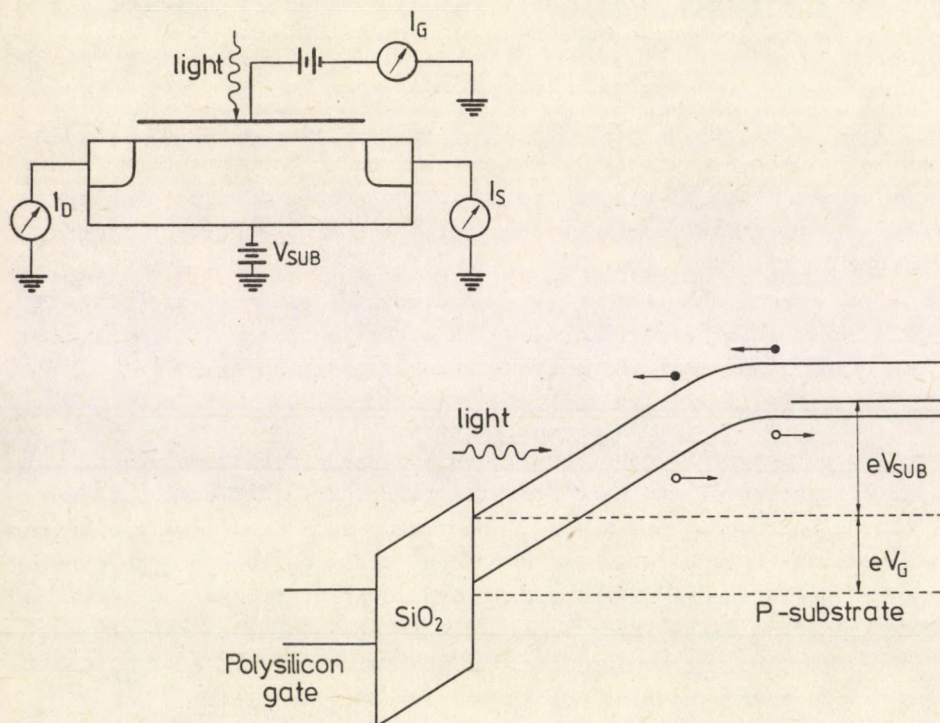


Fig.1. The experimental arrangement of [3] for studying the hot electron emission and the corresponding band structure diagram

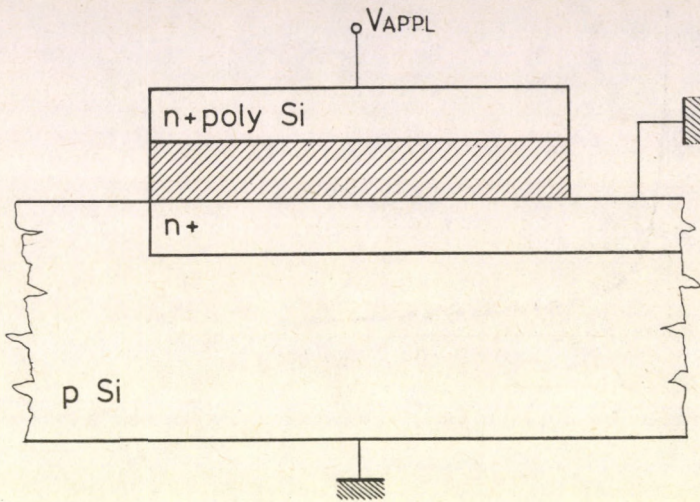


Fig.2. The experimental arrangement of [4] for studying the electron trapping

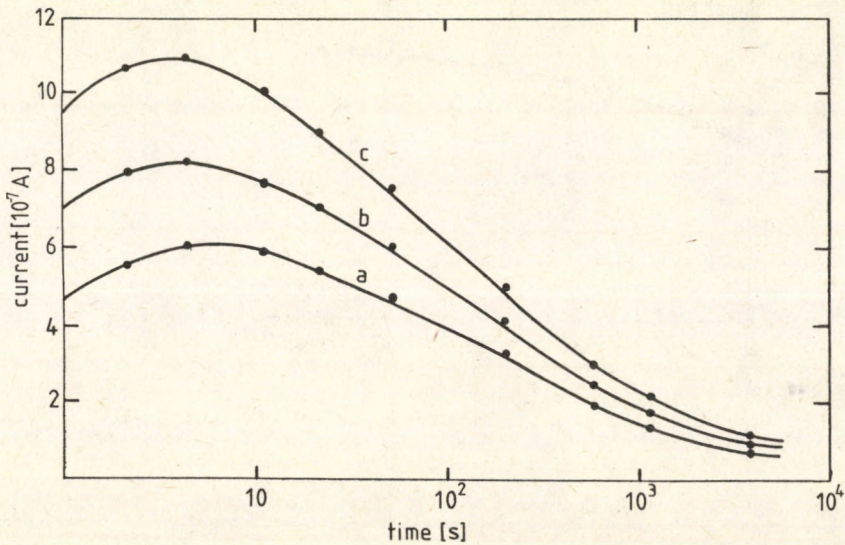


Fig.3. The current vs. time dependence in the experiment of [4]

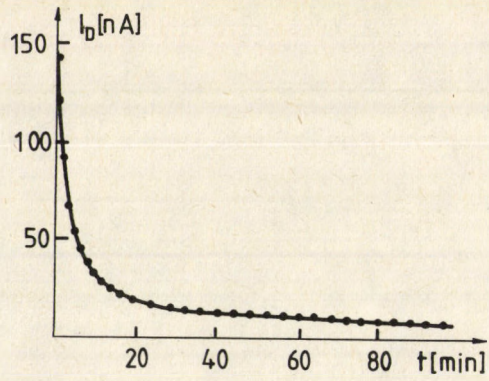


Fig.4. The avalanche current vs. time dependence in our experiment

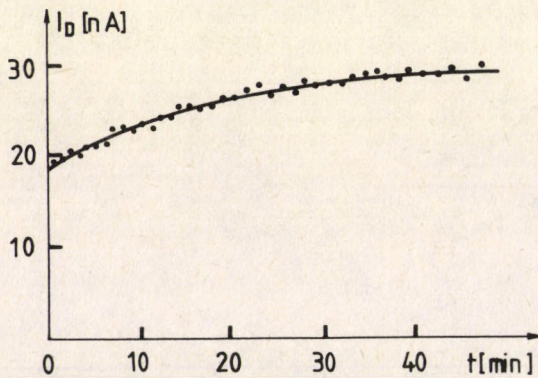


Fig.5. The dependence of the beginning value of the avalanche current impulse on the time

IRREVERSIBILITY OF MAGNETIC FLUX PENETRATION IN SUPERCONDUCTORS

I. KIRSCHNER

Department for Low Temperature Physics, Roland Eötvös University
1088 Budapest, Hungary

Thermodynamic behaviour of the fast, non-equilibrium, unstable, non-linear and irreversible magnetic flux penetration is investigated.

As a collective result of an external effect and an internal inhomogeneity such a phenomenon appears often in superconductors during which the non-equilibrium magnetic vortex distribution is followed by a sudden, avalanche-like equalization. It can lead to a local increase of the magnetic induction so that it is able to transform the sample into normal state. The abrupt change is accompanied by thermodynamic cross effects which result in a local increase in the energy density and temperature and help the transition into the normal state.

The phenomenon represents a degradation effect according to which the critical parameters of the short samples are much higher than those of the compact magnetic systems. At a given place of coils the value of the magnetic induction suddenly increases, the magnetic field penetration into the matter becomes jump-like, which goes with an increase of the local energy density and it results in reaching the upper critical magnetic field or the critical temperature at the place in question. In this way a phase transition leading into the normal state is induced and the magnetic /and thermal/ instability is able to propagate over the full sample.

This is a very fast, non-equilibrium, unstable, non-linear and irreversible dissipative process.

The equilibrium distribution of the magnetic vortex lines is a homogeneous and isotropic one which can be disturbed by the effect of pinning force F_p , vortex-vortex interaction force F_v and Lorentz-force F_L . F_p originates from the matter imperfections resulting in the flux pinning, F_v has mostly a repulsive character and F_L arises due to an electrical current introduced into the sample and plays role of an external generation. They are depending on the magnetic induction B and on the temperature T .

If a weak force is applied to the superconductors the flux lines will locally move a little. They leave their minimum energy sites ϵ_{\min} and move onto the neighboring minimum after an energy maximum ϵ_{\max} in between has been passed. The energy difference between the maximum and minimum sites can be dissipated[1].

The thermodynamic picture on the decay of the stability in the vortex line structure can be drawn up as follows.

A disturbance of the equilibrium means an addition of a flux increment to the interior of the specimen and the innermost flux lines move a small distance. Such a disturbance can be imagined as a density wave in the fluxoid structure which propagates much more faster than the vortex motion velocity.

In this way the disturbance changes alike all the forces F_L, F_p and F_v . Among them F_L has to be changed since for the same B the gradient differs from the earlier, F_p is also changed because of the local change in temperature in consequence of the energy dissipation, while F_v changes due to the change of actual local vortex density.

The non-equilibrium state succeeding an equilibrium can be described by the relation

$$\Delta F_L \stackrel{\text{sign}}{=} \Delta F_p + \Delta F_v, \quad /1/$$

where $\text{sign} <$ represents the stability, $\text{sign} >$ the instability and $\text{sign} =$ determines the stability limit.

The essence of the phenomenon is an abrupt change of the flux penetration, or in other words, a sudden change of the vortex current density.

The disturbance destroying stability may practically have three different forms:

1. a fast change in the vortex density,
2. a fast change in the vortex velocity and
3. a fast change in the energy density.

The thermodynamic cross effects of all the changes can naturally come into being as well.

The new non-equilibrium state is characterized by the existence of thermodynamic α -parameters α_i , general forces X_i and an irreversible change of the entropy α_s .

Taking into account the magnetic and energetic /thermal/ interactions, general forces and currents belonging to those are needed only for thermodynamic description.

In the biggest part of the thermodynamic changes of state in addition to the distribution of intensive parameters, external generations $G_i(t)$ and internal sources $Q_i(t)$ also participate in the development of the actual processes. Interactions occurring in this case may result in either permanent stationary non-equilibrium states, time-dependent forces and α -parameters, or fast changes. The motion equations of generated processes can be described in the form of [2]

$$\dot{X}(t) = -g LX(t) + gG(t) - gQ(t), \quad /2/$$

$$\dot{\alpha}(t) = -Lg\alpha(t) - G(t) + Q(t) \quad /3/$$

and their solutions with the condition of $G(t) = G_0 = \text{const}$ and $Q(t) = Q_0 = \text{const}$ are

$$X(t) = (-L^{-1}G_0 + L^{-1}Q_0 + X_{\text{init}}) e^{-gLt} + L^{-1}G_0 - L^{-1}Q_0, \quad /4/$$

$$\alpha(t) = (g^{-1}L^{-1}G_O - g^{-1}L^{-1}Q_O + \alpha_{init})e^{-Lgt} - g^{-1}G_O + g^{-1}L^{-1}Q_O, \quad /5/$$

which contain a component exponentially decaying in time if $g L > 0$ and $L g > 0$ and an other one being constant in time, where g is the entropy matrix and L is the conductivity matrix.

The thermodynamic force for the thermal interaction is

$$X_E = \frac{1}{T} \quad /6/$$

and the magnetic interaction force for a homogeneous temperature distribution consists of two parts

$$X_V = \frac{U}{T} = X_{VV} + X_{qV}, \quad /7/$$

where X_{VV} characterizes the state of the vortices and X_{qV} the current-vortex interaction [3].

$$X_{VV} = \frac{1}{8\pi t} \nabla \left\{ -\xi^2 h_c^2 (r) + \sum_k \Phi_O h_v (r_k - r_i) \right\}, \quad /8/$$

where the first term depends on the pinning character of the material and the second describes the direct vortex-vortex interaction.

$$X_{qV} = \frac{1}{4\pi t} \nabla (\Phi_O h_q (r - r_i)) = \frac{1}{cT} [j_q \times n_v \Phi_O], \quad /9/$$

which is the Lorentz-force arising from the effect of the electrical current j_q . Among the symbols ξ is the coherence length, h_c is the thermodynamic critical magnetic field, Φ_O is the magnetic flux quantum, h_v is the vortex-vortex interaction term, h_q is the magnetic field of transport current and n_v is the vortex density/.

On the base of these considerations it is possible to determine the conductivity coefficient of the vortex motion L_V and the value of external generation which is playing a role in equations /2/ and /3/

$$L_V = T \frac{n_v \Phi_O}{\eta_V} \quad /10/$$

and

$$G_V = \frac{1}{c} \frac{j_q n_v^2 \Phi_O^2}{\eta_V}, \quad /11/$$

where η_V is the viscosity coefficient of the vortex motion.

The magnetic vortex current density and the energy current density can be expressed by general thermodynamic motion equations of processes with an external generation and an internal source:

$$\dot{\alpha}_V(t) = -(L_{21}g_E + L_Vg_{12})\alpha_E(t) - L_{21}g_{12}\alpha_V(t) - G_V(t) + Q_V(t) \quad /12/$$

and

$$\dot{\alpha}_E(t) = -(L_Eg_E + L_{12}g_{21})\alpha_E(t) - (L_Eg_{12} + L_{12}g_V)\alpha_V(t). \quad /13/$$

These non-linear differential equations hint at the thermodynamic cross effects, reflect the non-linear character of the sudden magnetic penetration phenomenon, and provide a possibility for the chaotic behaviour, too [4].

The creation and propagation of this instability is a dissipative process and the entropy production σ or the dissipation D is the reason for irreversibility. The dissipation has three origins:

1. local movement of the individual vortices between the sites of ϵ_{\max} and ϵ_{\min} :

$$\sigma_{\text{loc}} = \frac{1}{T} n_v \frac{d(\epsilon_{\max} - \epsilon_{\min})}{dt}, \quad /14/$$

2. macroscopic flow of vortices which is dissipative in consequence of $\eta_v > 0$:

$$\sigma_{\text{flow}} = \frac{1}{T} \eta_v v_v^2 n_v \Phi_0, \quad /15/$$

3. electric current injected from outside into the sample (corresponding to the external generation):

$$\sigma_{\text{el}} = \frac{1}{T} \frac{d}{dt} \int_P \frac{q}{c} [v_v \times n_v \Phi_0] dP, \quad /16/$$

where q is the electric charge, v_v is the vortex velocity and P is the integration length.

The total dissipation can be determined by irreversible thermodynamic methods in the form of

$$D_{\text{tot}} = T\sigma_{\text{tot}} = -L_v (g_v \alpha_v X_v + g_{12} \alpha_v X_v) + Q_v X_v - L_E (g_E \alpha_E X_E + g_{12} \alpha_v X_E) - L_{12} \{ g_v \alpha_v X_E + g_{12} (\alpha_v X_v + \alpha_E X_E) + g_E \alpha_E X_v \}, \quad /17/$$

which contains the spontaneous terms, the external generation, the internal source and the cross effects as well.

References

1. S.L. Wipf, Phys. Rev., 161, 404, 1967.
2. I. Kirschner, Acta Phys. Hung, 30, 61, 1971.
3. I. Kirschner and K. Martínás, Journ. Low Temp. Phys., 14, 427, 1974.
4. I. Kirschner, Physica Scripta, 34, 842, 1986.

MICROSCOPIC CALCULATIONS OF VORTEX STRUCTURE AND MAGNETIZATION CURVES FOR TYPE II SUPERCONDUCTORS

U. KLEIN¹, L. KRAMER², W. PESCH², D. RAINER² and J. RAMMER²¹Institut für Theoretische Physik der Universität Linz
Linz, Austria²Physikalisches Institut der Universität Bayreuth
Bayreuth, F.R.G.

Up to now, important properties of low κ type II superconductors, as the amplitude and temperature-dependence of the jump of the induction B_0 at H_{C1} and the nature of the phase boundary between type II and type II/1 behaviour have not been explained quantitatively by theory. We calculated numerically the exact solution of the quasiclassical Eilenberger equations for a triangular fluxline lattice. We compare our results with previous theories and with recent experiments.

Introduction

The phase transition between the Meissner and mixed states becomes discontinuous for type II superconductors with sufficiently low values of the Ginzburg-Landau parameter. The GL theory gives some insight into the underlying physical mechanism:

The interaction energy of flux lines may be separated into a repulsive part /electromagnetic in origin/ and an attractive part /due to the order parameter overlap/. The latter favours the simultaneous entry of many flux lines at H_{C1} but always is smaller than the repulsive part for $\kappa > 1/\sqrt{2}$ near T_C [1]. The first order transition at H_{C1} /type II/1 behaviour/ indicates the importance of corrections to the GL theory such as nonlocality of the electromagnetic response or higher powers of the order parameter and its derivatives. All such effects are contained in the full set of microscopic equations of superconductivity. A detailed theoretical analysis of type II/1 superconductors requires a rigorous solution of these equations. Among previous theoretical studies we mention the asymptotic theory [2,3] and the theory of Brandt [4].

We outline the methods used in our calculation and report some preliminary results. We discuss the temperature-dependence of B_0 and compare our results with the asymptotic theory.

Methods

We start from Eilenberger's quasiclassical equations [7] in the weak-coupling, isotropic form, using the notation of [8]

$$[\omega_{\vec{k}} + \vec{k}(\vec{v} - i\vec{a})]f = \psi g + Fg - Gf, \quad (1)$$

$$[\omega_{\vec{k}} - \vec{k}(\vec{v} + i\vec{a})]f^+ = \psi^* g + F^+ g - Gf^+, \quad (2)$$

$$(2\vec{k}\vec{\nabla}g = \psi^*f - \psi f^* + F^*f - Ff^*), \quad (3)$$

$$g = (1 - ff^*)^{1/2}, \quad \text{Re}(g) > 0, \quad (4)$$

$$\psi \ln t = -2t \sum_{\ell=0}^{\infty} \left[\frac{\psi}{\omega_{\ell}} - \int \frac{d^2\vec{k}}{4\pi} f \right], \quad (5)$$

$$\vec{\nabla} \times \vec{\nabla} \times \vec{a} = -i(2t/\tilde{\kappa}^2) \sum_{\ell=0}^{\infty} \int \frac{d^2\vec{k}}{4\pi} \hat{k}g, \quad (6)$$

$$F^{(+)} = \int \frac{d^2\vec{q}}{4\pi} \left(\frac{1}{\tau} + \frac{3}{\tau_1} \vec{k}\vec{q} \right) f^{(+)}, \quad (7)$$

$$G = \int \frac{d^2\vec{q}}{4\pi} \left(\frac{1}{\tau} + \frac{3}{\tau_1} \vec{k}\vec{q} \right) g, \quad (8)$$

where f, f^*, g are the Green Functions; ψ the order parameter; \vec{a} the vectorpotential; F, F^*, G are "impurity potentials".

A self-consistent solution of (1)-(8) for the vortex lattice has been obtained by [8] using a one-dimensional circular-cell approximation. For clean superconductors the question of the type of phase transition at H_{c1} can be answered only by taking into account the actual structure of the vortex lattice.

We have solved Eqs. (1)-(4) for a triangular lattice by iterative methods starting with the potentials of Abrikosov's theory. We developed a complete theory of symmetry transformations for Green's functions on a triangular lattice which reduced the number of evaluations of Eqs (1)-(3) by a factor of 20. Symmetry properties were also used to determine the true boundary values of the Green's functions. We applied also an appropriate version of the "explosion method" [9]. Both methods are alternatively chosen according to their efficiency which depends on the characteristic length of variation of the Green's functions/i.e. depends on \vec{k} and ω .

Having obtained the Green's functions for a sufficiently large number of ω and \vec{k} , new potentials may be calculated from Eqs. (5)-(8). We found two iteration procedures of rapid convergence, the "subtraction of BCS-solutions" and the "shooting method".

From the final Green's functions and potentials the free energy of the flux-line lattice was calculated. Repeating the procedure for different flux-line distances we obtained the induction-dependence of the free energy. Its derivative gives the macroscopic induction \bar{B} vs. the external magnetic field.

Results

For reasons of symmetry we have calculated the order parameter and vector potential in one-twelfth of the Wigner-Weitz /WS/-cell (Fig.1). Fig. 2 shows the modulus of the order parameter ψ and the microscopic magnetic field b in the irreducible part of the cell along the line AB. In the GL regime b must be lower at point B /minimum/ than at point A /saddle-point/. According to [10] and [14] however, the positions of the minimum and saddle

-point fields are exchanged at $H = H_{C2}$ and $T = 0.62 T_C$. We verify these results, and find an additional structure at intermediate fields and temperatures. The magnetic field of a pure superconductor (Fig. 2. b.) is lower at point A than at point B ($b_B = 0.516$, $b_A = 0.498$) and has a maximum between them. With decreasing induction B the maximum comes closer to point A and the difference between b_A and b_B decreases. Below $\bar{B} \sim 0.3$ the usual situation described by the Abrikosov solution is recovered.

Increasing impurity content decreases the spatial variations of order parameter and magnetic fields. In addition, we find significant changes in the shape of both fields. Similar effects have been observed in the circular-cell approximation [8].

Magnetization curves for samples with different impurity content are shown in Fig. 3. The corresponding jumps of induction and the critical fields H_{C1} were obtained by means of a Maxwell construction. In the pure case our magnetization curves were similar in shape to those found by [4] and led to similar jumps. Fig. 4 shows the temperature-dependence of our B_0 and data of [6]. The agreement is satisfactory in particular with regard to the nearly temperature-independent behavior of B_0 for $T < 0.2 T_C$. Our reduced jump B_0/H_C , however, is higher than the experimental one by a factor of 1.4. This discrepancy between the experimental and theoretical values of B_0/H_C is related to similar discrepancies in κ_1 and κ_2 . These deviations are probably due to strong coupling corrections and anisotropy effects which are neglected in our calculation.

We compared our preliminary results on the phase boundary of the type II/1 region with the asymptotic theory and experimental data. Calculating magnetization curves for pure superconductors at $T = 0.5 T_C$ and various values of κ we found that B_0 vanishes more or less continuously at $\kappa \sim 1.1$. The prediction of the asymptotic theory [2,3] is 1.25.

In the impure case we performed calculations for two types of materials, the Nb samples of moderate impurity content, [6], and the impure Ta samples [5] (see Fig. 3). In both cases our results agree reasonably well with the predictions of the asymptotic theory. The possibility of an "attractive minimum" at finite flux-line distance has been discussed [2]. We found a continuous vanishing of B_0 at the phase boundary for both pure and impure superconductors.

Our theoretical II/1-II phase boundary falls below the experimental data on Ta-N [5]. This discrepancy may be due to our neglect of real metal effects and/or experimental inaccuracies. The Nb experiments agree reasonably well with both the asymptotic curve and our results.

Acknowledgements

We would like to thank H. W. Weber for stimulating discussions. This work has been supported by the Deutsche Forschungsgemeinschaft.

References

1. L. Kramer, Phys. Rev., **3B**, 3821, 1971.
2. L. Kramer, Z. Physik, **258**, 367, 1973.
3. M. C. Leung, J. Low Temp. Phys., **12**, 215, 1973.
4. E. H. Brandt, phys. stat. sol., **17**, 105, 1976.
5. J. Auer and H. Ullmaier, Phys. Rev., **B7**, 136, 1973.
6. M. Botlo, H. W. Weber and U. Klein, in Proc. of the 17th Int. Conf. on Low Temp. Physics, vol II, U. Eckern, A. Schmid, W. Weber and H. Wühl, eds., North-Holland, Amsterdam, 1984, p.1285.
7. G. Eilenberger, Z. Physik, **214**, 195, 1968.
8. W. Pesch and L. Kramer, J. Low Temp. Phys., **15**, 367, 1974.
9. E.V. Thuneberg, J. Kurkijärvi and D. Rainer, Phys. Rev., **B29**, 3913, 1984.
10. J. M. Delrieu, J. Low Temp., Phys., **6**, 197, 1972.
11. M. C. Leung and A.E. Jacobs, in Proc. of the 13th Int. Conf. on Low Temp. Physics, Vol. III, K. D. Timmerhaus, W. J. O'Sullivan and E. F. Hammer, eds., Plenum, New York, 1974, p.46.
12. A. E. Jacobs, Phys. Rev., **B4**, 3022, 1971.
13. L. Kramer, W. Pesch and R. J. Watts-Tobin, J. Low Temp. Phys., **14**, 29, 1974.
14. E. H. Brandt, Phys. Lett., **43A**, 539, 1973.

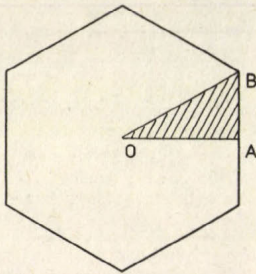


Fig.1. Hexagonal WS-cell with its irreducible part (hatched region)

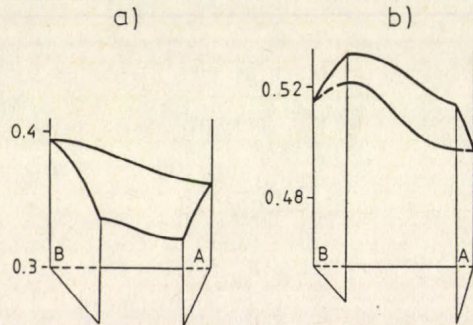


Fig.2.a. Order parameter, b. Magnetic field / Parameters: $T/T_c = 0.2$, $\kappa = \kappa_{\text{pure}} = 0.72$, $B = 0.6$

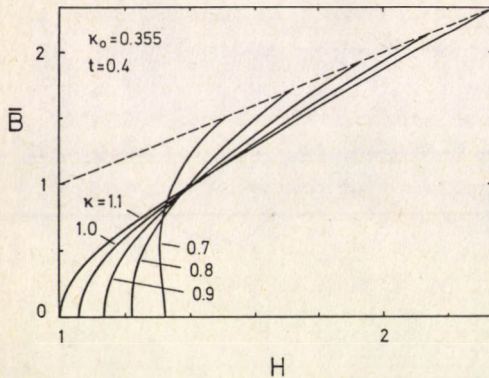


Fig.3. Magnetization curves with different impurity content. / $\kappa_{\text{pure}} = 0.355$, $\kappa = 0.7, 0.8, 0.9, 1.0, 1.1$ /

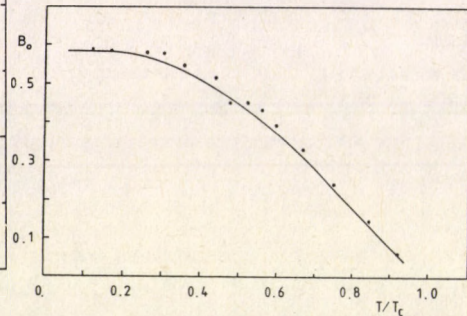


Fig.4. The induction jump vs. temperature and data of [6] / $\kappa = \kappa_{\text{pure}} = 0.72$ /

ELEMENTARY QUANTUM PHYSICAL DESCRIPTION OF TRIPLET SUPERCONDUCTORS

K. MARTINÁS¹ and B. LUKÁCS²¹Department for Low Temperature Physics, Roland Eötvös University
1088 Budapest, Hungary²Central Research Institute for Physics
1525 Budapest, Hungary

The elementary quantum physical description of superconductors using only three elementary facts of quantum physics, Bohr's quantum principle, the uncertainty relation and Pauli's exclusion principle, can reflect the basic bulk properties of superconductors, the effect of temperature and external magnetic fields. The same method for triplet superconductors gives that for every $H \neq 0$ magnetic field the perfect conductor state is thermodynamically more favourable, than the diamagnetic state; the perfect conductor state will cease at $H = H_{c2}$.

Introduction

There is an interesting discrepancy between the generality of superconductivity on one hand /more than 50% of metals is proven to possess a superconducting phase/ and the theoretical complexity of microscopic explanation. Weisskopf [1] already has demonstrated that a partial but quite detailed understanding can be achieved by using full quantum mechanics but not quantum field theory. Here we build up a simplified description of superconductivity; only elementary constants and the data of the lattice ions will be used, and, of course, the results are expected to be correct only up to number constants of order of unity.

EQP of superconductors

Consider an ideal metallic lattice with positive ions and a free electron gas. In first approximation these charges compensate each other. In second approximation, the moving electron disturbs the ion lattice, causing an effective positive charge near to its path, which acts on a second electron moving collinearly by a potential $U \sim -\sqrt{\frac{m}{M}} e^2 / r$, where e is the elementary charge, M is the ion mass and m is the electron mass. If this were a classical potential, there would be a bound state of the electrons with a characteristic energy 600 K. Nevertheless the uncertainty principle gives a simple, correct estimation. There are momentum and position uncertainties, so the ground state energy of a pair can be written as

$$\epsilon = \frac{3}{2} \frac{(\Delta p)^2}{m} - \frac{m}{M} \frac{e^2}{\Delta x}, \quad /1/$$

where

$$\Delta p \Delta x \sim \hbar/2.$$

Hence, looking for energy minimum, one obtains

$$\Delta E = -\frac{2}{3} \frac{1}{M} \left(\frac{me^2}{\hbar} \right)^2. \quad /2/$$

While the M dependence of this energy does not show the right isotope effect its numerical value is in the correct order of magnitude, 2 K for a metal of 50 atomic mass. So one can conclude that, via lattice oscillations, two electron states may appear with a binding energy.

Since the creation of such pairs is energetically favoured, one expects the sea of pairs in the $T=0$ ground state. Elementary symmetry and quantum considerations yield that the Cooper pair consists of two electrons being as collinear as possible, in order to maximize the attraction; but on the other hand it is a resonance with finite size $l \approx 2v_F \hbar / E_b$ and a minimal momentum uncertainty $l p_0 \sim E_b / v_F$, which forbids exactly zero total momentum. The optimal compromise is a state where the total momentum is p_0 , when it is greater, the binding is weaker, and it cannot be smaller, for details see Refs. [2] [3].

Consider now an external effect not disrupting but modifying the superconducting state. It can only change the total momentum of the pairs, as there are no other parameters to be modified. The change of the total pair momentum appears as an excess uncertainty.

The disturbed quasiparticle possesses a greater size ξ_d ,

$$p_0^2 = (\Delta p_\alpha)^2 + (\hbar/\xi_d)^2. \quad /3/$$

The new binding energy is:

$$E_b = E_{b0} \sqrt{1 - \frac{(\Delta p_d)^2}{p_0^2}} \quad /4/$$

smaller than E_b , so the thermal excitation energy is within the energy uncertainty of the pairs, so it seems that the Fermi distribution of the electrons does not influence the possible excitations, i.e. A Boltzmann approximation can be used, so

$$\Delta p_t \sim k_B T / v_f. \quad /5/$$

The magnetic field can interact only with the individual electrons, as the usual Cooper pair, being a particle of 0 spin and 0 momentum, cannot feel the presence of the magnet field. The interaction via the momentum yields:

$$\Delta p_{HM} = p_0 \cdot \frac{\hbar}{\hbar c^2}, \quad /6/$$

where H_{c2} is the upper critical magnetic field, $H_{c2} = \frac{\Phi_0}{2\pi\xi^2}$.

In case of interaction via the spin, the field H can produce a change in the spin orientation, leading to a change in the potential energy

$$\Delta V = 2 \frac{eh}{mc} H, \quad /7/$$

while the kinetic energy of the Cooper pair changes by

$$\Delta E = (\Delta p)^2 / m. \quad /8/$$

They have to be in the same order of magnitude in dynamic equilibrium, so

$$(\Delta p)^2 \sim \rho_0^2 \frac{H}{H_{c2}}. \quad /9/$$

Combining the thermal and magnetic effects, the binding energy is as follows [2]:

$$E_b(T, H) = E_b(0, 0) \left\{ 1 - \left(\frac{T}{T_c} \right)^2 - \left(\frac{H}{H_{c2}} \right)^2 - \frac{H}{H_{c2}} \right\}^{1/2}. \quad /10/$$

For spin-1 Cooper pairs the magnetic field can interact with the pair as a whole, and the released interaction energy can be transferred into e.g. lattice vibrations, which is an external heat reservoir for the electron gas, therefore this interaction will not change the binding energy. Then, repeating, mutatis mutandis, the above steps, one gets eq. /10/ without its last term:

$$E_b(T, H) = E_b(0, 0) \left\{ 1 - \left(\frac{T}{T_c} \right)^2 - \left(\frac{H}{H_{c2}} \right)^2 \right\}^{1/2}. \quad /11/$$

Thermodynamics of the superconducting state

Since superconducting samples are handled at constant temperature and magnetic field, the actual state is selected by the minimum of the Gibbs potential G

$$G = E - TS - BH/4\pi. \quad /12/$$

The energy of the superconducting state can be approximated as [3]

$$E_s = E_n - \frac{1}{4} \xi(E_F) E_D^2, \quad /13/$$

where ξ is the state density, for a cold Fermi gas $\xi \sim E^{1/2}$ so, using eq. /11/ and the definitions

$$H_c = \Phi_0 / 2\pi\lambda\xi \quad \lambda^2 = \frac{1}{4\pi} \frac{mc^2}{ne} \quad /14/$$

one gets

$$E_s = E_n - \frac{1}{8\pi} H_c^2 \left(\sqrt{1 - \left(\frac{T}{T_c}\right)^2 - \left(\frac{H}{H_{c2}}\right)^2 - \frac{H}{H_{c2}}} \right)^2. \quad /15/$$

Again, this is true for spin-0 pairs; for spin-1 ones the last term is absent.

Now, let us indeed select the actual state by the minimum of G . For this one has to compare states of different structure. The list at least contains the following ones; normal $/E=E_n, B=H/$; Meissner $/E=E_s, B=0/$; mixed $/E=E_s(B), B < H/$ and a "perfect conductor" $/E=E_s(B), B=H/$.

For comparison the Gibbs potentials of the "usual" states /i.e. the first three ones/ can be found in Ref. 3. So we have to deal only with the "perfect superconductor" here. Its Gibbs potential can be directly obtained by using eqs. /12/, /15//with or without the last term according to the pair spin/ and the definition of that state.

After trivial calculations one gets for spin-0 Cooper pairs

$$G_P = G_M + \frac{1}{8\pi} \left\{ H_{c1} H + H^2 / \kappa^2 - H^2 \right\}, \quad /16/$$

where κ is the GL dimensionless parameter. The Meissner state is below the normal one until H_{c1} , and in this whole range the bracketed term is positive. Above H_{c1} one could easily show that $G_P > G_{\text{mixed}}$. This is just the standard result that the "perfect superconductor" state is not realised.

However, consider the case of spin-1 pairs. Here, according to eq. /11/ the term linear in H is absent, that is

$$G_P = G_M - \frac{H^2}{8\pi} \left\{ 1 - 1/\kappa^2 \right\}. \quad /17/$$

Now, obviously, this means that for the cases $\kappa > 1$ the "perfect superconductor" state is always preferred to the Meissner state. Thus, if this system has a superconducting state, then this state is a perfect superconductor until H_{c2} .

Conclusions

Here, we have demonstrated that fundamental quantum principles and thermodynamics do not rule out the possibility of a perfect conductor state, i.e. superconduction without diamagnetism. In fact, such states are rather predicted, but only when the Cooper pairs exist in spin-1 state. This is just the case of triplet superconductors [4], [5].

References

1. V. F. Weisskopf, *Contemp. Phys.*, **22**, 375, 1981.
2. K. Martinás and I. Kirschner, *Acta Phys. Hung.*, **40**, 297, 1976.
3. K. Martinás and B. Lukács, KFKI-1984-77.
4. K. Svozil, *Phys. Rev.*, **B33**, 602, 1986.
5. P. B. Allen and B. Mitrović, *Solid State Physics*, **37**, 1, 1983,

SIMULATION OF I-V STEPS SHARPENING FOR JOSEPHSON JUNCTIONS

T.F.REFAI¹, L.N.SHEHATA²¹Faculty of Engineering, Dept. of Physical and Mathematical Engineering
Ain Shams University, Cairo, Egypt²Atomic Energy Establishment, Dept. of Mathematics and Theoretical Physics
Cairo, Egypt

The effect of rf frequency on the induced I-V steps for Josephson junction is discussed. The resistively shunted junction is applied for the case of an applied rf rectified sinusoidal current in which the amplitude is modulated by a sawtoothed function. It is shown that the step rise sharpens on increasing the rf frequency. We also show that the first fundamental step is nearly absent for low rf frequency while on increasing the rf frequency, the step evolves and assumes its well-known configuration.

Introduction

There has been considerable interest to investigate the induced steps in the I-V characteristics of a Josephson junction driven by either ac- or dc-current sources or both. The effect of critical current amplitude modulation [1], radio and micro wave power, and thermal noise [2-5] on the induced steps has been investigated. In Fig.1 the effect of varying the rf frequency on the I-V characteristics is shown [6]. This Figure shows the evolution of the steps. Higher steps are better evolved than the first step, but still their step rise sharpens, and finally the overall curve rise increases /i.e. approaching the ohmic line faster/. To the best of our knowledge such phenomena have not been investigated and presented elsewhere.

Our purpose is the theoretical investigation of the effect of rf frequency on the features discussed above. We applied the resistively shunted junction model [7] /RSJ/, for a Josephson junction biased by an rf rectified sinusoidal current with sawtooth-modulated amplitude. The model works for three rf frequency /10, 50, 100 MHz/ at zero temperature and for an rf frequency /10 MHz/ at a normalized temperature /t=0.75/. The calculated I-V curves qualitatively agree very well with the I-V curves of Figure 1.

Theory

The response of a Josephson junction to an ac current source $I_{ac} \sin \omega t$, ω is the applied external frequency, while biased by a dc current source is given by:

$$C \frac{dV}{dt} + \sigma V + I_c \sin \phi = I_{dc} + I_{ac} \sin \omega t + I_N, \quad (1-a)$$

$$2eV = \hbar \frac{d\phi}{dt}, \quad (1-b)$$

where C is the capacitance, V is the voltage, σ is the conductance, I_C is the critical current, Φ is the phase difference across the junction, and I_N is the thermal current generated in the resistor $R=1/\sigma$ at a given temperature. We will only consider $C = 0$ and $I_{dc} = 0$. The rf current source will be replaced by:

$$I_0 \frac{t}{T} \left| \sin \omega t \right|, \quad (2)$$

where T is the modulation period of the sawtoothed function $\frac{T_0}{T} t$, and I_0 is the current amplitude at $t = T$. Next, by setting $u = \omega t$, Eq. /1-b/ can be expressed as

$$V = \frac{\hbar \omega}{2e} \frac{d\Phi}{du}. \quad (3)$$

From /1-a/ and /3/, we then have:

$$\frac{d\Phi}{du} = \frac{2eRI_C}{\hbar \omega} \left[\frac{i_0}{\omega T} u \sin u + i_N - \sin \Phi \right], \quad (4)$$

where i_0 and I_N are the rf and thermal currents normalized with respect to the critical current I_C . For the thermal current, the average distribution is given by:

$$\langle i_N \rangle_{\text{rms}} = \left(\frac{4k_B T \Delta f}{I_C^2 R} \right)^{1/2}, \quad (5-a)$$

$$\Delta f = \frac{2e R I_C}{\hbar} (2 \Delta u)^{-1}, \quad (5-b)$$

the Δu is the normalized time step involved in the numerical integration of Eq. /4/.

Eq. /4/ was solved numerically by the Runge-Kutta method. i_0 was taken to be 10^3 , $T=10^{-3}$, and $\Delta u = \pi/100$. The normalized voltage

$$\langle v \rangle = \left\langle \frac{V}{RI_C} \right\rangle = \frac{\hbar \omega}{2eRI_C} \left\langle \frac{d\Phi}{du} \right\rangle. \quad (6)$$

The averaging is taken over half rf cycle. Keeping T constant and changing the frequency has the same effect as if the opposite is done.

Discussion

Fig. 2 shows the I-V curves at zero temperature. In the regions where the normalized current is below 2, it is clearly shown how the first step evolves with increasing rf frequency. Fig.3 shows a general step shape induced by an applied rf current. On the path OA several rf cycles elapse before v reaches v_c . Along A_c the duration between transitions decreases. At C the phase transition occurs every cycle. Beyond C, the transition occurs every few cycles on both the positive and negative rf swings. At B the transition

rate is twice every cycle. At D a second step appears [5]. In Fig.3 we can compare the cases for 10 MHz and for 100 MHz rf frequency, and try to interpret how the first step is evolved. On increasing the current from 1 to $i=i_1$, the voltage on curve I is v_1 /point A/ which is less than v_2 /point B/ for curve II.

In the region where the current is above 2 /Fig.2/ subharmonic steps exist at 2, 2.5 and 3. This is due to the nature of the rf current considered in this work. Their general features are the same as discussed. Our work predicts that by changing the modulation frequency and keeping the junction at a constant rf frequency we would be able to observe such phenomenon. We have also solved the equations for a normalized temperature $t=0.74$ and an rf frequency 10 MHz, with no change in the modulation frequency. In Fig.4 the second subharmonic step is completely washed out, while the third one is less pronounced than that for $T=0$. The fundamental steps show a slight rounding on their knees. Hence the effect of thermal current is only very strong on subharmonic steps in agreement with [3].

Since curve II /Fig.3/ is at a higher frequency, there are more rf cycles involved than for curve I during the same range of current, hence a higher transition rate is expected for curve II which leads to $v_2 > v_1$. As the rf frequency is increased, the step rise /OB/ sharpens. For curve I, further increasing the current beyond C, a step is observed at $v \approx 1.5$, $i \approx 1.5$ which is a subharmonic step $/n=3/2/$. On curve II, increasing the current beyond B, a step width is observed. From B to C, the transition rate increases until at C we have a transition every complete rf cycle. If our rf current was not rectified, we would then have the step extended to $i \sim 2$ where two transitions occur every rf cycle. But since we have a rectified current the step width will terminate at D and a new riser starts again leading to the second step. Fig.3 and this discussion suggests that at low rf frequency there is a step /second subharmonic/, and as the rf frequency is increased the step is shifted downwards until it completely assumes the position of the first fundamental step with a width equal to half the critical current.

Conclusion

We have shown the effect of rf frequency on the evolution of the first step. We proved that such effect can also be observed by keeping the rf frequency constant, and varying the modulation frequency. Last, the rectified current considered in our work leads to the appearance of subharmonic steps. This last result suggests a new mean of experimentally detecting subharmonic steps.

References

1. C. Vanneste, A. Gilabert, P. Sibillot and D.B. Ostrowsley, *J. Low Temp. Phys.*, **45**, 517, 1981.
2. H. Seifert, *J. Low Temp. Phys.*, **37**, 595, 1979.
3. E. Ben Jacob and D.J. Bergman, *Phys. Rev.*, **A**, **29**, 2021, 1984.
4. T.M. Klapwijk, M. Sepers and J.E. Mooij, *J. Low Temp. Phys.*, **27**, 801, 1977.
5. L.D. Jackel and R.A. Buhrman, *J. Low Temp. Phys.*, **19**, 201, 1975.
6. Private communications with Prof. R. Peters, CUA, Washington D.C.
7. D.E. McCumber, *J. Appl. Phys.* **39**, 3113, 1968.

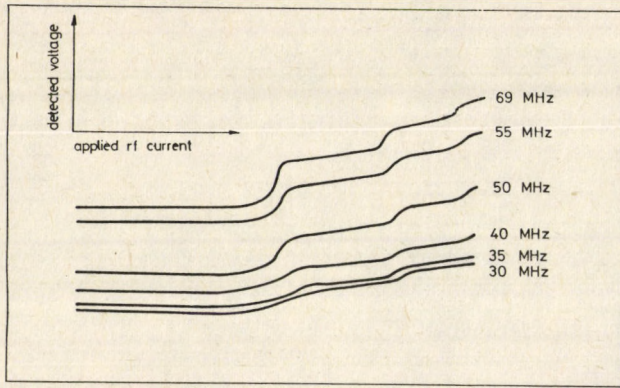


Fig. 1. Experimental I-Y curves for different rf frequencies

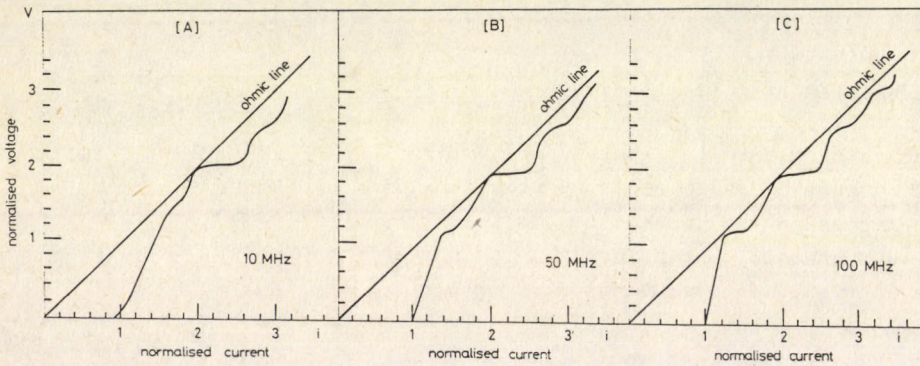


Fig. 2. Calculated I-V curves for different frequencies at zero temperature

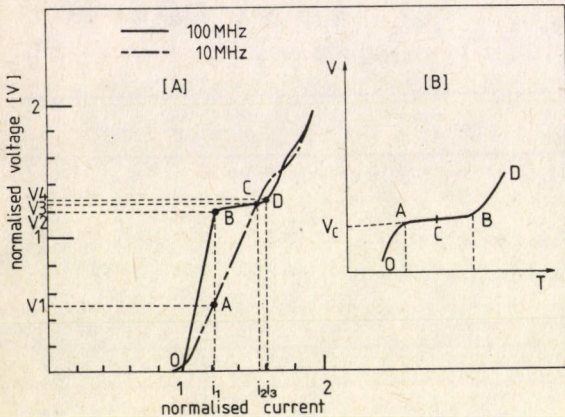


Fig. 3. The effect of the rf frequency on the sharpening of the first step

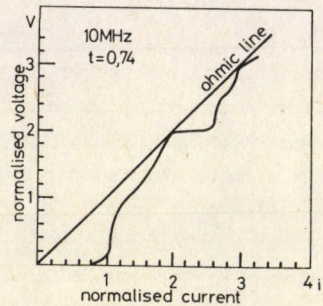


Fig. 4. Calculated I-V curve for 10 MHz and normalized temperature $t=0.74$

COMPUTER SIMULATION OF THE CORRELATION LENGTH IN TYPE II SUPERCONDUCTOR
IN CHAOTIC APPROACH

P. SAMU

Department for Low Temperature Physics, Roland Eötvös University
1088 Budapest, Hungary

It was shown that the pinning force in one dimensional case can be handled in a deterministic way, but the solution of the actual non-linear differential equations leads to a chaotic behaviour of the pinning-vortex interaction. This method provides the concept of the correlation length. The correlation length was determined with computer simulation and it is shown that there is sharp transition from linear to chaotic behaviour.

The large loss-free currents observed in type II superconductors are due to pinning of the vortices caused by inhomogeneities of the material. The Lorentz force exerted on the vortices by the current is balanced by pinning forces up to a critical current density j_c , where depinning occurs and the vortices start to drift and to dissipate energy. The sum of the elementary pinning forces to an average force density is the volume pinning force $j_c B$, where B is the magnetic induction. The pinning problem in type II superconductors is basically a 3-dimensional effect, nevertheless to investigate the fundamental features of it the 1-dimensional modelling also provides some useful informations.

The generally accepted assumptions about the flux line lattice and pinning centres are as follows [1]

- 1./ Non-interacting defects;
- 2./ One defect in the interaction range of the flux line;
- 3./ The defects are in randomly distributed positions;
- 4./ The vortex lattice can be treated as an elastic continuum;
- 5./ Thermal activation is negligible.

The pinning problem in type II superconductors averaging with randomly spaced weak pinnings generally is handled as a deterministic system in the sense that their differentia equation defining the position of the individual flux lines is as follows [2]

$$c \frac{r_{i+1} + r_{i-1} - 2r_i}{\frac{1}{2} (r_{i+1} - r_{i-1})^2} = f_p (r_i, R_k), \quad /1/$$

where c is the elastic constant, r_i is the position of the i th vortex, R_k is the position of the k th pinning centre and f_p is the pinning force acting on the i th flux line

$$f_p(r_i, R_k) = \begin{cases} 0 & \text{if } |r_i - R_k| > a_i/2, \\ f(x_i^k) & \text{if } |r_i - R_k| < a_i/2, \end{cases} \quad /2/$$

where $a_i = r_{i+1} - r_i$, is the lattice constant and

$$x_i^k = \frac{r_i - R_k + a_i/2}{a_i}. \quad /3/$$

By the help of this equation /1/ can be written in the form

$$\delta\left(\frac{1}{a_i^k}\right) = \frac{1}{c} f_p(x_i^k), \quad /4/$$

which shows that a change in the lattice constant occurs due to the existence of the pinning forces.

The other possible definition is as follows.

We should like to calculate the correlation length in case of 1-d collective pinning. There are at least two different methods yielding different correlation lengths.

One is the 1-d version of Larkin-Ovchinnikov approach [2], i.e. the regions of the vortex lattice in which relative shifts are less than the lattice constant will be called correlated regions. The linear dimension L_c of correlated regions is determined

$$\left\langle \left| U(L_c) - U(0) \right|^2 \right\rangle = a^2 \quad /5/$$

where a is the displacement vector, L_c is the correlation length.

The positional uncertainty of the i -th vortex

$$\Delta x_i = \Delta x_0 + \sum_{j \neq i} (i-j) \frac{1}{c} \frac{df_p}{dx} \Delta x_j, \quad /6/$$

where f_p is the individual pinning force length.

Calculating the average response due to the displacement of the first vortex:

$$(\Delta x_i)^2 = (\Delta x_0)^2 + \sum_{j \neq i} \frac{(i-j)^2}{c^2} \overline{\left(\frac{df_p}{dx}\right)^2} (\Delta x_j)^2. \quad /7/$$

From this we have [3]

$$L_c \sim L_p^{1/3} \cdot K^{-2/3}, \quad /8/$$

where

$$K \sim f_p/c \quad /9/$$

and L_p is the average distance between the pinning centres, measured in lattice constant dimension.

Equation /1/ can be transformed in the form being equivalent to the map of unit interval onto itself [4]. As the lattice constant changes are due to presence of pinnings, it is allowed to investigate only the flux lines possessing a pinning centre neighbourhood with distance less than the half of the lattice constant. The unit interval will be the interval of lattice constant length where the pinning centre is in the middle position. The position of the flux line relative to the nearest pinning centre is

$$x^{\alpha+1} = x^{\alpha} + N_{\alpha} a_{\alpha} - L_{\alpha}, \quad /10/$$

where L_{α} is the distance between the α th and $(\alpha+1)$ th pinning centres, measured in lattice constant and N_{α} is the number of flux lines between them and a_{α} is the lattice constant.

One vortex position is the function of the displacement of the first vortex which altered in every step a given Δx . To simulate the behaviour of the correlation length chaos parameter can be introduced [6]. The chaos parameter shows the linkage among one vortex sequence of three position.

The chaos parameter is

$$\text{chaos}_n = \left| y_{i-1}^1 - 2y_i^n + y_{i+1}^n \right| + \left| m_{i-1}^n - 2m_i^n + m_{i+1}^n \right|, \quad /11/$$

where $y_{i-1}^n, y_i^n, y_{i+1}^n$ are the positions of the n th pinned vortex and

$$m_i^n = \frac{a_i^n - a_0}{\sqrt{n} f_p}, \quad /12/$$

where f_p is the individual pinning force strength.

The results are in Fig.1-2 where L is the distance between the first vortex pinned by the n th pinning centre.

Figures 1-2 show that the transition is sharp from the linear to the chaotic behaviour.

In chaotic approach the correlation length yields [5]

$$L_c = \frac{a}{\Delta x_0} \left(\frac{f_p}{c} \right)^{-1} = \frac{\pi}{2} \frac{1}{\sqrt{n} f_p}. \quad /13/$$

The result of computer simulation is

$$L_c = g \frac{1}{\sqrt{n} f_p}, \quad /14/$$

where g is $g = 1,6 \pm 0,2$ numerical constant in accord with equation /13/.

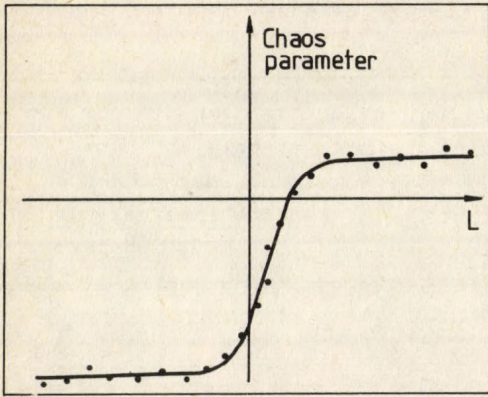


Fig.1 Chaos parameter dependence on the distance of the nth vortex in lattice constant measure

$$\Delta x = 0.001$$

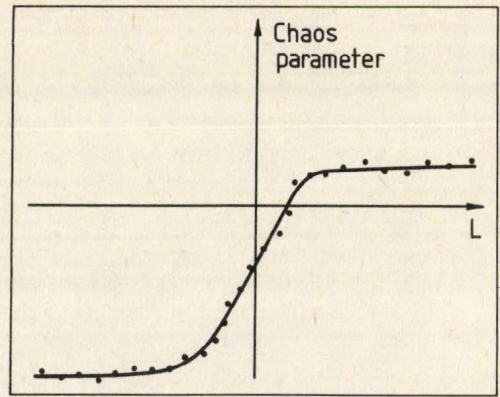


Fig.2 Chaos parameter dependence on the distance of the nth vortex in lattice constant measure

$$\Delta x = 0.01$$

References

1. A.M. Campbell and J.E. Evetts, *Adv. Phys.*, 21, 199, 1972.
2. I.Kirschner and K.Martinás, *Journ. Low Temp. Phys.*, 47, 105, 1982.
3. A.I.Larkin and Yu. N. Ovchinnikov: *Journ. Low Temp. Phys.*, 34, 409, 1979.
4. I.Kirschner and K. Martinás, *Proc. of the 17th Int. Conf. on Low Temp. Phys.* p.591, Karlsruhe, 1984.
5. K.Martinás, *ATOMKI Report E/7* p.467, 1985.
6. K. Martinás, to be published.

THERMODYNAMIC CROSS EFFECTS IN TYPE-II SUPERCONDUCTORS

G. KISS

Department for Low Temperature Physics, Roland Eötvös University
1088 Budapest, Hungary

Investigation of the galvanomagnetic and thermomagnetic effects in superconductors has a special role with respect to the case of normal conductors if these investigations are made in mixed state, because in this case only the normal electrons from the magnetic vortices take part in the effect. Their concentration is smaller than the concentration of the conduction electrons, so the effect taking place is bigger, and depends strongly on the pinning forces. This allows to get a lot of information from this measurement type.

Galvanomagnetic effects

a./ Transversal effects

- Isothermal Hall effect

The onset of the potential gradient in direction perpendicular to the external electric current $I_{ee} = I_x$ and magnetic field $B = B_z$, with the boundary conditions $J_{el,y} = 0$ and $\text{grad } T = 0$, in the case if the external magnetic field is greater than the lower critical magnetic field H_{c1} .

- Adiabatic Hall effect

Similar to the former one and the boundary conditions will be:

$$\frac{\partial T}{\partial x} = 0 \quad \text{and} \quad J_{sx} = 0$$

- Etingshausen effect

The onset of the temperature gradient perpendicular to the direction of external magnetic field B_z , and electric current J_x , with the boundary conditions:

$$J_{el,y} = 0, \quad J_{s,y} = 0 \quad \text{and} \quad \frac{\partial T}{\partial x} = 0$$

b./ Longitudinal effects

- Isothermal electric resistivity

- Adiabatic electric resistivity

In the case of superconductors the value of these effects is zero.

Thermomagnetic effectsa./ Transversal effects- Righi -Leduc effect

The onset of the temperature gradient perpendicular to the x direction temperature gradient and Z direction magnetic field, with the boundary conditions:

$$J_{el} = 0 \quad \text{and} \quad J_{s,y} = 0$$

/perhaps the magnetic field must be higher than H_{c1} /

- Isothermal Nernst effect

In the case of the applied magnetic field $B=B_z$ and heat current $J_s=J_{s,x}$ the onset of the potential gradient, with the boundary conditions:

$$J_{el} = 0 \quad \text{and} \quad \frac{\partial T}{\partial y} = 0$$

- Adiabatic Nernst effect

Similar to the former, but with the boundary condition $J_{s,y}=0$ instead of the 0 temperature gradient

b./ Longitudinal effects- Isothermal thermal conductivity- Adiabatic thermal conductivity

In the case of boundary conditions

$$J_{el}=0 \quad \text{and} \quad \frac{\partial T}{\partial y} = 0 \quad \text{or} \quad J_{s,y} = 0$$

The coefficient of thermal conductivity

$$\lambda = - \frac{T J_{s,x}}{\text{grad}_x T}$$

has a strong magnetic field dependence, because the Cooper-pairs do not take part in heat conduction, only the normal electrons appearing in magnetic vortices. At the same time the concentration of the vortices depends on the value of the magnetic field

- Isothermal Etingshausen-Nernst effect

Appearance of the potential gradient in the x direction due to the applied magnetic field $B=B_x$ and temperature gradient $\text{grad}_x T$, with boundary conditions

$$J_{el} = 0 \quad \text{and} \quad \frac{\partial T}{\partial y} = 0$$

- Adiabatic Etingshausen-Nernst effect

Similar to the former but in the boundary conditions we will have

$$J_{s,y} = 0$$

Because of the Onsager reciprocity relations there exists the Bridgman-type equation

$$T Q_i^t = \lambda_i \Phi^t$$

Q_i^t is the isothermal Nernst coefficient, λ_i the isothermal coefficient of

thermal conductivity and ϕ^t the Ettingshausen coefficient/.

Experimental results

We have made the experimental investigation of Nernst effect in superconducting samples of Pb-Sn with content of Sn 50 atomic percent. The effect takes place only if $H > H_{C1}$ because in the Meissner state the electric field cannot appear in superconducting sample. We have used cylindrical samples with a diameter of 4 mm and a length of 50 mm. They were prepared by quenching from liquid state to the nitrogen temperatures [2].

The system position of measurement is shown in the Fig. 1. The measurements allow a very precise determination of the lower critical magnetic field H_{C1} (see Fig. 2.) [3], because the onset of the potential gradient is very sharp when the magnetic field penetrates the sample. The results provided by this method are in good agreement with the values of H_{C1} determined by magnetisation measurements [4].

The appearance of the effect is similar to the appearance of Hall effect. If we have the magnetic vortices in our sample, they feel the effective thermodynamic force $-(\phi_0 S'/B) \text{ grad } T$ due to the heat current I_0 is the flux quantum, and S' is the entropy density/, and the vortices can move from the hotter place to the colder if this force is bigger than the pinning force. At the hot end of the sample the vortex density is higher, here the vortex lines permanently move into the sample, and at the cold end they permanently leave it. The Lorentz force acts on the normal electrons in moving vortices resulting in vortex density gradient in y direction, and E_y potential gradient. The effect disappears at H_{C2} , because with the magnetic vortices disappearing the moving electrons disappear too. The hysteresis we have found is the result of the pinning forces bounding a part from vortices. The value of the effect has strong pinning force dependence. If the pinning forces are strong enough the effect will disappear. In this case the thermodynamic force will be less than the pinning forces. If the vortices are bound the flow disappears. If we put the sample into the cryostat immediately after quenching it cannot recrystallise. In the other case strong pinning centres are formed and the effect disappears. The magnetisation curves show these phenomena very well.

References

1. S. Wisniewski, B. Staniszewski and R. Szymanik, Thermodynamics of Non-equilibrium Process, D. Reidel Publishing Company, Dordrecht-Holland Boston. U.S.A. 1976.
2. Дь. Ремени, А.А. Хариеди и Дь. Киши, Экспериментальное Термодинамических Эффектов в Сверхпроводящих Сплавах Pb-Sn Сообщ. 14. Междунар. Конф. Физ и Техн. Низк. Темп., стр., 54, Братислава, 1975.
3. Дь. Киши, Дь. Ремени и А.А. Хариеди, Термомангнитные Свойства Сверхпроводящих Сплавов Pb-Sn, Сообщ. 18. Междунар. Конф. Физ. и Техн. Низк. Темп., стр., 84, Дрезден, 1979.
4. Gy. Kiss, Gy. Reményi and A.A. Hariedy, Magnetic and Thermomagnetic Properties of Some Lead-Tin Alloys, Acta Phys. Hung., 50, 183, 1981.

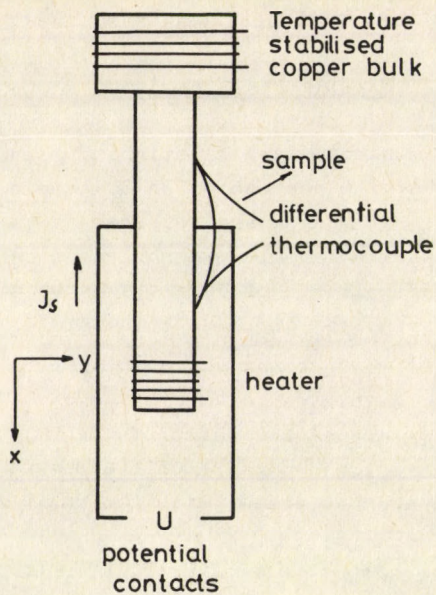


Fig.1 The system of measurement

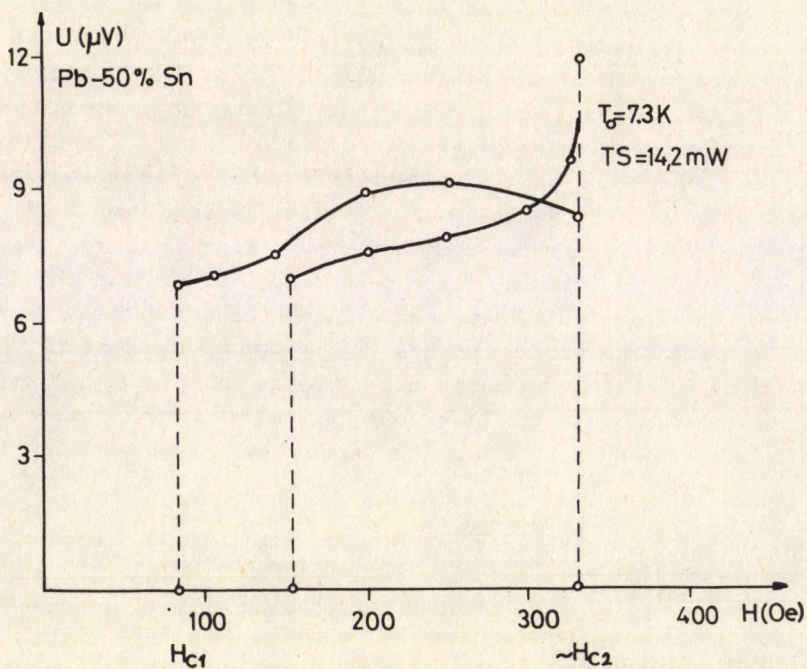


Fig.2. The Nernst effect

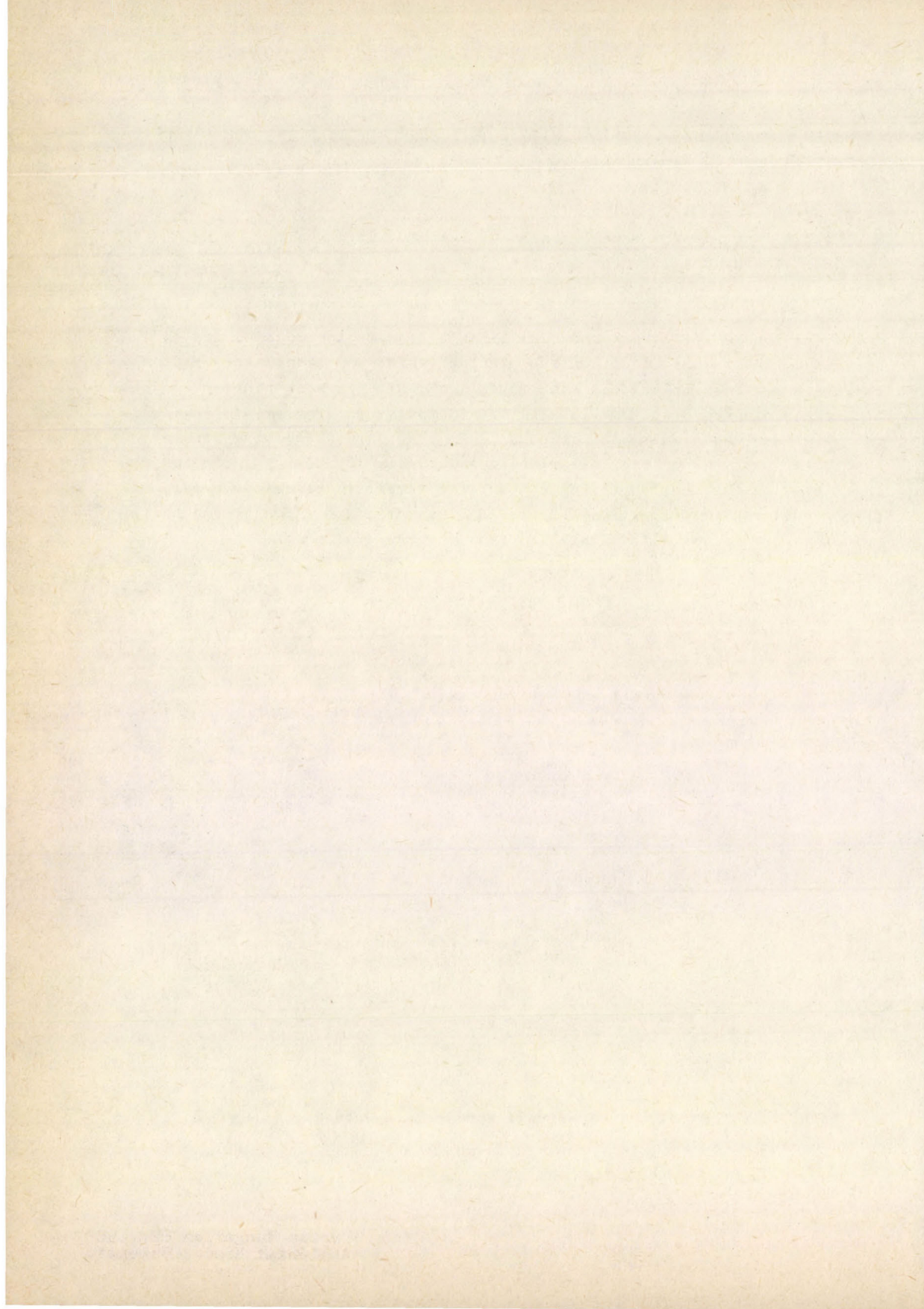
COMPUTER SIMULATION OF PINNING IN TYPE-II SUPERCONDUCTORS AND THE THEORY OF COLLECTIVE PINNING (Abstract) *

E. H. BRANDT

Max-Planck-Institut für Metallforschung, Institut für Physik,
Stuttgart, Germany

The problem of pinning of flux-lines by random inhomogeneities is discussed. The theory of collective pinning, which was established recently by Larkin and Ovchinnikov and has proven successful in weak-pinning amorphous alloys, is discussed in some detail. Then computer simulation of pinning is described and the results are discussed. They confirm collective pinning theory for weak two-dimensional pinning and modify it at strong pinning where the flux-line lattice becomes amorphous. The threshold-behaviour predicted by the "dilute limit theories" proves to be an artifact which is due to the non-existence of this limit in sufficiently large specimens.

* The full text of the paper will appear in a subsequent issue of Acta Physica Hungarica



ANALOGIES BETWEEN GRANULAR SUPERCONDUCTORS AND SPIN GLASSES (Abstract) *

J. ROSENBLATT, P. PEYRAL, A. RABOUTOU and C. LEBEAU
Laboratoire de Physique des Solides UA 786 au C.N.R.S.
I.N.S.A. 35043 RENNES Cedex France

The transition to coherence of a granular superconductor with grain size much larger than the characteristic superconducting lengths is described by an X-Y model with random couplings. This led us to measure the critical exponents (which of course need not coincide with those of mean field theory) for this transition.

To illustrate this work, we present a measurement of the correlation function exponent η and the "susceptibility" (order parameter fluctuation) exponent χ . Scaling laws allow then to obtain values of β and ν which are in good agreement with those resulting from previous measurements. The value obtained for $\eta = -0.1 \pm 0.2$, is in good agreement with calculations on the ordered X-Y model, but the exponents $\beta = 0.7 \pm 0.1$, $\chi = 2.8 \pm 0.3$, $\nu = 1.3 \pm 0.1$ are about twice the X-Y model values. Actually these exponents are rather close to those found experimentally in certain spin glasses.

We discuss a model which assumes that critical order parameter fluctuations are localised in the regions of highest coupling energy. This defines a temperature-dependent percolation problem, which in turn allows to relate the critical exponents of the coherence or spin glass transition to the percolation exponents. Fairly good agreement with experiment is obtained.

* The full text of the paper will appear in a subsequent issue of Acta Physica Hungarica

ANISOTROPY EFFECTS IN SUPERCONDUCTORS DOWN TO MILLIKELVIN TEMPERATURES

F. M. SAUERZOPF and H.W. WEBER

Atominstytut der Österreichischen Universitäten

A-1020 Wien, Austria

We report on two experimental techniques aimed at an analysis of anisotropy effects in superconductors. The first provides direct experimental evidence and is based on the measurement of the direction dependence of the upper critical magnetic field H_{c2} in single crystalline superconductors. Unfortunately, the theoretical situation is unsatisfactory at present and does not allow an evaluation of electron-phonon coupling anisotropies. The second technique is based on an analysis of the thermodynamics of type-I superconductors and its impurity dependence. In this case an excellent theoretical description in terms of Eliashberg theory is available, which allows a full evaluation of experimental results.

Introduction

A variety of experimental techniques have been devised to investigate anisotropies of the electron-phonon coupling strengths in superconductors (for a review of this subject cf., e.g., Ref.1.). Among these we have concentrated on anisotropy effects of the upper critical field H_{c2} in single-crystalline cubic type-II superconductors during the past few years [2-5] and obtained detailed and accurate information on the temperature - and impurity - dependence of H_{c2} -anisotropy, which was characterized by the coefficients A_l of a series expansion in terms of cubic harmonic functions [6]. The increasing precision of the experiments, which were carried out in the temperature range between 1.5 K and the transition temperature T_c , revealed the following features: 1) The temperature dependence of the expansion coefficients shows considerable discrepancies [4] with the predictions of the only presently available microscopic theory [7]; 2) at sufficiently low temperature ($t = T/T_c$, $t=0.3$) additional higher-order expansion coefficients ($l = 8, 10, 12$) are needed to describe the results on Nb and V to experimental accuracy [5]; and 3) even at 1.5 K, the coefficient A_6 was found to be negligibly small in Ta. (This could be caused, however, by the fact that a certain amount of impurities had to be introduced into Ta, in order to convert it into a type-II superconductor within the accessible temperature range [3,8]). Especially, in view of the last two observations, a new series of experiments was made on Nb-, V- and Ta-single crystals in a region down to milli-Kelvin temperatures. We are, however, aware that an evaluation in terms of electron-phonon coupling anisotropies is not at hand, due to the lack of an appropriate theoretical treatment.

On the other hand, the well-established technique [1] of studying the thermodynamics of superconductors with varying mean-free electron paths, in order to deduce mean-square values of the coupling anisotropy, has become attractive again, because a full theoretical treatment in terms of Eliashberg theory has become available [9,10]. From an experimental point of view, precise data on the T_c -depression as a function of impurity concentration, as well as accurate measurements of the deviation function, especially towards $T \rightarrow 0$, are essential.

H_{c2} -anisotropy in Nb, V and Ta

The experiments were made on 110 -oriented single crystals by measuring the a.c. susceptibility in transverse magnetic fields, either in a conventional bath cryostat or in a dilution refrigerator. We will restrict ourselves to results on four samples, namely Nb 8 (RRR = 2080), V 21 (RRR = 200) and Ta 4 and 5 (RRR = 34 and 30, respectively), which were doped with N_2 in order to provide type-II superconductivity in all crystal directions [8,11]. Typical results are shown in Fig.1. The analysis consists of a least-mean square fitting procedure of the experimental data to Eq. (1):

$$H_{c2}(\alpha, t; \vec{e}_H) = \sum_{l=0,4,6,8,10,12} A_l(\alpha, t) H_l(\vec{e}_H), \quad (1)$$

where $e_H = x, y, z$ denotes a unit vector in the direction of the magnetic field, $H_l(e_H)$ are the cubic harmonic functions defined in Ref. 6, and $A_l(\alpha, t)$ are the anisotropy coefficients, which depend on the impurity parameter $\alpha (= 0.882 \xi_0 / l)$ and the reduced temperature t only. $A_0 \equiv H_{c2}$ is conveniently used to normalize the higher-order anisotropy coefficients ($a_l = A_l / A_0$). From all the data on the anisotropy coefficients a_l , we wish to present an example result pertaining to the coefficients a_4 and a_6 in Fig. 2. In Nb the temperature dependence of a_4 follows a $(1-t)$ -law as predicted by theory [7] at low temperatures, whereas both in V and Ta, a_4 levels off towards $T = 0$. In the case of Ta, a_4 definitely remains finite at the conversion temperature T^* to type-I superconductivity. Concerning the coefficient a_6 the results on Nb again indicate agreement with the $(1-t)^2$ -law [7], even at the lowest temperatures, whereas saturation effects occur for V and Ta towards $T = 0$. The coefficient a_6 , identified for the first time in Ta, is much smaller than in the other two materials, which is partly related to the higher impurity parameters in the Ta samples, but definitely disappears at T^* , where type-II superconductivity disappears. None of the higher-order coefficients could be detected in Ta. a_6 in Nb shows a $(1-t)$ -dependence and in V again a plateau for $T \rightarrow 0$. a_{10} , which was identified in Nb only, is negative (as the leading term a_4) and displays again a $(1-t)$ -temperature dependence. This summary of experimental results shows clearly that through the uniform description of data in terms of cubic harmonic functions, a broad spectrum of effects is covered, whose variety becomes particularly pronounced at the lowest temperatures. It seems that the microscopic anisotropies of the specific material are responsible for the different phenomena observed in Nb, V and Ta. The disappearance of a_6 at T^* in Ta seems to be of special interest, because it is directly correlated to type-II superconductivity, as opposed to the behavior of the first coefficient a_4 .

Thermodynamics of anisotropic In

In order to study the thermodynamics of anisotropic superconductors, type-I materials are preferentially investigated, because the measurement of the thermodynamic critical field H_c can be done with much higher accuracy. The basic idea of the experiment and the subsequent theoretical analysis is as follows: Starting from the isotropic BCS-case, a separable pairing potential including a temperature independent anisotropy parameter a_k can be introduced, whose average over the entire Fermi surface is zero, but whose mean-square value $\langle a^2 \rangle$ denotes the average deviation of the coupling from the isotropic case. Consequently, in the clean limit the transition temperature T_c of the superconductor is enhanced compared to the isotropic case; but the addition of small amounts of nonmagnetic impurities provides new scattering centers and, hence, permissible k -states, which tend to smear out the original anisotropic distribution of states. This situation, i.e. the original anisotropic T_c and its subsequent reduction by anisotropy removal, can be treated accurately for any superconductor of arbitrary coupling strength in terms of the linearized anisotropic Eliashberg theory, if the electron-phonon spectral function $F(\omega)$ [12] is known for the clean material. With only one additional parameter, namely the lifetime of the electronic excited states due to normal scattering, a quantitative comparison between theory and the experimental T_c -depression as a function of impurity concentration yields the magnitude $\langle a^2 \rangle$ of the electron-phonon coupling anisotropy. Thus, having established all the parameters from this T_c -experiment, the full anisotropic Eliashberg theory can be used to calculate the other thermodynamic quantities, such as the temperature dependence of the thermodynamic critical field H_c and the electronic specific heat, without any free parameters. The corresponding experiments on $H_c(T)$ or the deviation function,

therefore, represent a very sensitive consistency check of the whole evaluation procedure and the concept of anisotropy and anisotropy removal by scattering centers. The experiments were made on high-purity indium and a series of InTl alloys whose superconductive properties have been investigated in much detail previously [13,14]. The results on the T_c -depression were analyzed following the procedure outlined above. The best agreement between theory and experiment is found for $0.035 \leq \langle a^2 \rangle \leq 0.040$, where the standard deviation of all the experimental data amounts to only 3 mK. In Fig. 3 we compare the experimental results on the deviation function $D(t) = \frac{H_c(T)H_c(0)}{H_c(0)^2} - [1-t^2]$ for pure indium and for the sample with the highest impurity content, with theory. In the clean limit we note significant discrepancies between experiment and the weak-coupling BCS-results, as well as with the isotropic Eliashberg calculation. On the other hand, satisfactory agreement is found for the anisotropy parameters $\langle a^2 \rangle$, as determined from the T_c -depression. Further evidence for the correctness of an anisotropic treatment of pure In is provided by the deviation function of $\text{In}_{90}\text{Tl}_{10}$, where a "smeared-out" isotropic treatment yields agreement with experiment. An additional quantitative comparison between theory and experiment can be made for the thermodynamic critical field H_c at $T = 0$. In this case, the experimental accuracy is much better than for $D(t)$, but the theoretical results require an additional quantity, which has to be taken from experiment and, therefore, contains experimental uncertainties, namely, the density of states at the Fermi level, which is calculated from the Sommerfeld constant [15]. The results demonstrate again, that best agreement is obtained for $\langle a^2 \rangle = 0.04$. In summary, all the evidence presented above and, in addition, a very careful and comprehensive analysis of the influence of experimental uncertainties and functional derivatives [10], show consistently that the concept of coupling anisotropy and its removal due to impurity scattering, represents the correct explanation of the experiments on In and provides a valuable tool for studying anisotropy effects in superconductors.

Conclusions

The experiments presented in this paper clearly emphasize the role of very low temperatures for an investigation of anisotropy effects in superconductors. Whereas in the case of H_{c2} -anisotropy, interesting features in the temperature dependence of the anisotropy coefficients were revealed only at temperatures below 1 K, millikelvin temperatures are a necessary condition for an accurate evaluation of the deviation function. Concerning the theoretical analysis in terms of electron-phonon coupling anisotropies, studies of the thermodynamics have to be preferred at present. However, because of the much higher precision of the H_{c2} -anisotropy data, future theoretical efforts should provide still more detailed insight into the nature of anisotropy effects in superconductors. We gratefully acknowledge the partial support of this work by Fonds zur Förderung der Wissenschaftlichen Forschung, Wien, under contract No. 3973.

References

1. H. W. Weber, Ed., Anisotropy Effects in Superconductors, Plenum Publishing Corporation, New York, 1977
2. E. Seidl, H.W. Weber and H. Teichler, J. Low Temp. Phys., **30**, 273, 1987.
3. J.F. Sporna, E. Seidl and H.W. Weber, J. Low Temp. Phys., **37**, 639, 1979.
4. E. Moser, E. Seidl and H.W. Weber, J. Low Temp. Phys., **49**, 585, 1982.
5. E. Moser, P. Hahn, E. Seidl, H.W. Weber and E. Schachinger, in Superconductivity in d- and f-Band Metals, W. Buckel, W. Weber, Eds., Kernforschungszentrum Karlsruhe p. 519, 1982.
6. F.C. von der Lage and H.A. Bethe, Phys. Rev., **71**, 612, 1947.
7. H. Teichler, in Ref. 1, p. 7.
8. H.W. Weber, J.F. Sporna and E. Seidl, Phys. Rev. Lett., **41**, 1502, 1978.
9. J.P. Carbotte, in Ref. 1, p. 183.
10. L. Niel, N. Giesinger, H.W. Weber and E. Schachinger, Phys. Rev., **B32**, 2976, 1985.
11. F.M. Sauerzopf, E. Moser, H.W. Weber and F.A. Schmidt, J. Low Temp. Phys., **66**, 191, 1987.
12. R.C. Dynes and J.M. Rowell, Tabulation of the Electron-Phonon Interaction, Part I, Bell Laboratories, Murray Hill (unpublished)
13. M.F. Merriam, J. Hagen and H.L. Luo, Phys. Rev., **154**, 424, 1967.
14. I. Kirschner and K. Sajo, J. Low Temp. Phys., **39**, 235, 1980.
15. H.R.O'Neal and N.E. Phillips, Phys. Rev., **137**, A748, 1965.

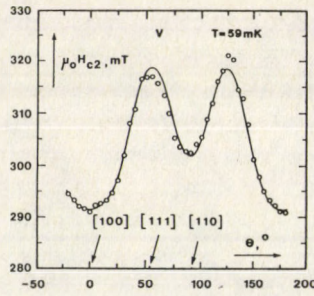


Fig. 1: Experimental results on the angular dependence of H_{c2} in a vanadium single crystal ($T = 59$ mK)

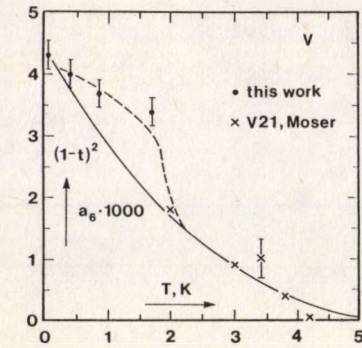
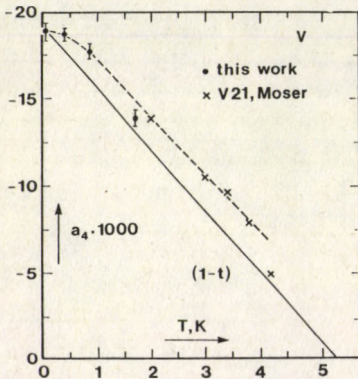


Fig. 2: Temperature dependence of a_4 and a_c in V ($RRR \sim 200$)

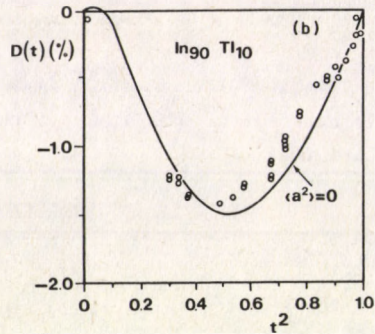
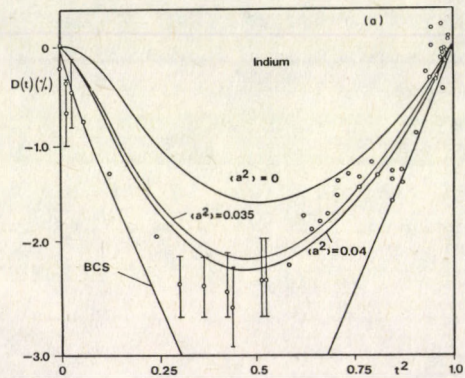


Fig. 3: Deviation function $D(t)$ for pure In (a) and $In_{90}Tl_{10}$ (b). The experimental data are compared with various solutions of Eliashberg theory and with the BCS-result.

ON SOME DEVELOPMENTS OF TECHNOLOGY OF MULTIFILAMENTARY WIRE V_3Ga

J. BÁNKUTI, É. HEGYI FARKAS and L. LÁSZLÓFFY

Department for Low Temperature Physics, Roland Eötvös University
1088 Budapest, Hungary

The last twenty years have brought along a big evolution in the field of the application of the superconducting materials. It is a valuable task to improve the superconducting critical parameters, T_c , H_{c2} and I_c . As the experiments show some improvement can be achieved by adding other alloying elements to the superconductors in a low concentration. Taking into account of the advantageous properties of Al there are some experiments and considerations to make superconducting wire in Al matrix.

Introduction

The critical temperature T_c and the critical magnetic field H_{c2} depend on the microscopic physical properties of the basic material as it can be seen in Table I. However the critical current I_c is determined mainly by the technology, by the inhomogenities, the dislocations and the other crystal imperfections [1-2].

In mixed state of a superconductor the so-called pinning centres prevent the Lorentz motion of the flux line. The critical current density j_c is related to the mean pinning force F_p (B) by the critical state equation

$$\underline{j}_c \times B + \underline{F}_p (B) = 0,$$

where \underline{B} is the magnetic field. \underline{F}_p is determined by the interactions between the flux line and the crystal imperfections.

Recently some manufacturing technologies exist for BCC superconductors, first of all for Nb-Ti, Nb-Zr, because their critical parameters are not too sensitive for composition. This solid solution superconducting phase can be worked well mechanically too.

However from the point of view of the high critical parameters the compound of A15 structure is favourable. Among the A15 superconductors the Nb-Sn and the V-Ga systems have stable A15 phases in a wide range around the stoichiometric composition.

For multifilamentary superconductors V_3Ga there are two technologies:

1. Surface diffusion process (SDP);
 2. Composite diffusion process (CDP)
- By some reasonable modifications we can get higher values in the critical parameters extremely in the critical currents.

Some improvements in the critical parameters

In A15 superconductors the dominant pinning centres are the grain-boundaries [4]. When the grains are less we can get stronger pinning forces and the critical current is higher. The grain size can be limited by the temperature of heat treatment. However, at the heat treatment the thickness of the superconducting layer also decreases with the decreasing of the temperature.

Adding different alloying elements to the matrix or to the pure materials the diffusion constant can be increased due to the decrease of the activation energy [5-6] .

In the Fig. 1 the critical currents are plotted against the annealing time in some superconducting samples.

As it can be seen the critical currents increase with the annealing time. We mentioned that the changing of the critical current with the reaction time is influenced by two effects mainly. One of them is the growth of the superconducting layer and the other is the formation of the grain size. The critical current increases approximately monotonically with the thickness of the superconducting layer and so with the reaction time, but when the grain size will be larger j_c will decrease and so the strength of the pinning force.

From the results we can establish that the effect of the alloying of In is the highest . It surpasses the effects of Al and Zn. The element In owes the best effect to the largest superconducting layer due to the best diffusion of Ga. But after annealing of one hundred hours the grain size becomes too large and the critical current density will be lower.

In the Fig. 2 the critical temperature T_c is plotted against the annealing time in the same samples.

As we can see the solution of the additional elements in V_3Ga layer is a limited process because the critical temperatures change hardly.

Taking into account the advantages of element Al a new idea arises for the further modification of the V_3Ga technology. The element Al is very promising as a matrix material in the multifilamentary superconductive wires because its mass density is only one third of that of Cu, the electric and thermal conductivity is approximately that of Cu at low temperatures but it is much cheaper.

The first task in the CDP technique is to examine the Al-Ga system from the mechanical and electric points of view. In previous works we can find two different phase diagrams for Al-Ga. One of them is a simple eutectic system with limited solid solubility. The other one has three intermetallic compound phases.

In our early experiments the Al-Ga samples were produced by melting of the pure metals in a quartz tube under argon atmosphere. The compositions were next to the assumed intermetallic compounds. The content Ga prevented the development of the Al oxide layer and the samples were damaged in a relatively short time. The samples were porous, brittle and plastically undeformable. The unsuitable properties of the Al-Ga alloys demand the further modifications of CDP technique or other SDP techniques.

References

1. R. Flükiger, J.L. Jorda, Solid State Comm., 22, 109, 1977.
2. B.N. Das, J.E. Cox, Metallurgical Trans., 8A, 541, 1977.
3. K. Tachikawa, Cryogenics, 19, 309, 1979.
4. E. Nembach, K. Tachikawa, J. of Less-Common. Met., 19, 359, 1968.
5. K. Tachikawa, Proc. ICEC3, 1970. 389.
6. K. Fischer, S. Rohr, J. Mönch, W. Holzhäuser, Ch. Frenzel, Acta Phys. Hung., 53, 305, 1982.

Table I
Critical parameters of superconductive materials
of different crystal type

Crystal Type	Material	T_c (K)	$H_{c2}(4,2K)(T)$
BCC	Nb-Ti	9.8-10.2	12
	Nb-Zr	10.8	11
A15	Nb_3Ga	20.7	34.1
	Nb_3Al	18.6	25.2
	Nb_3Sn	18	23.5
	V_3Ga	15.9	22
	V_3Si	17	22.8
C15(Laves)	HfV_2	9.2	20
	ZrV_2	8.5	
	$Hf_{0.5}Zr_{0.5}V_2$	10.1	23
TMS(Chevrel)	$SnMo_6S_8$	11.7	
	$PbMo_6S_8$	14	

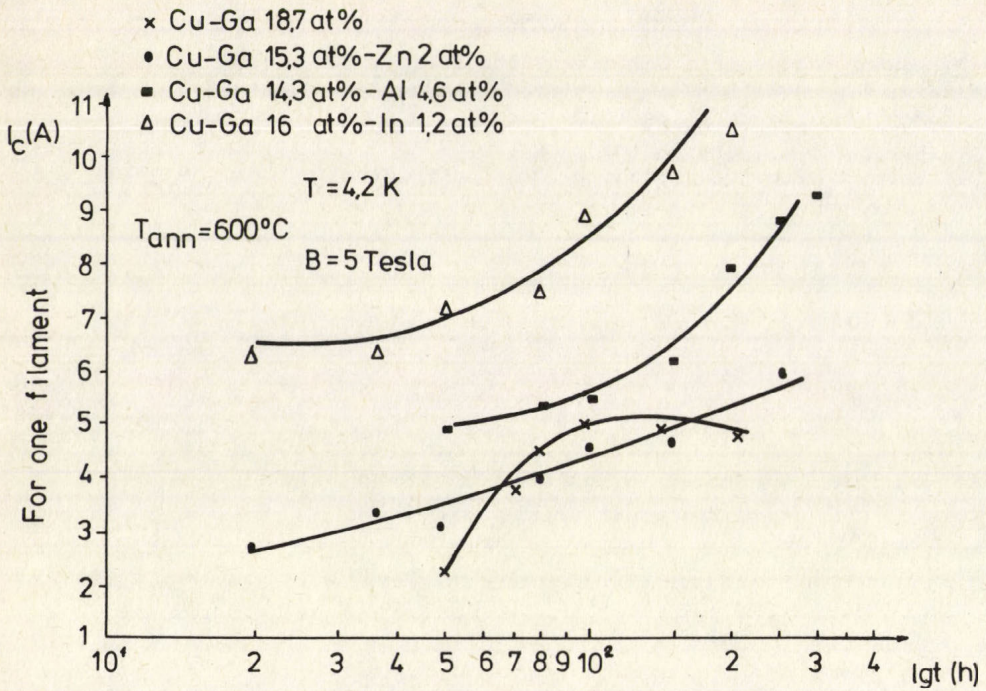


Fig.1. Critical current vs. reaction time of conductors with different compositions of the bronze

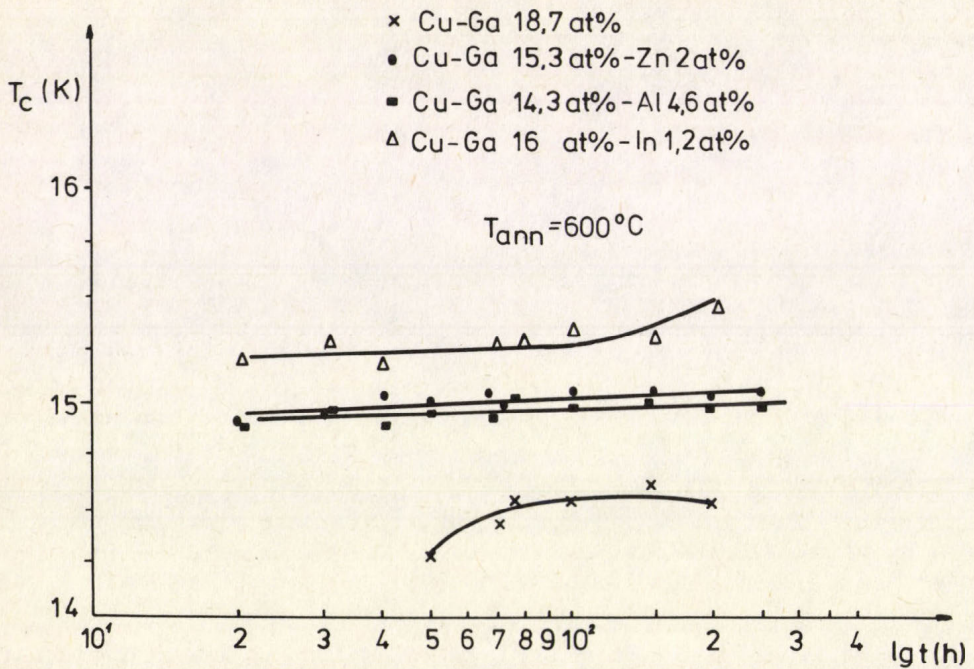


Fig.2. Critical temperature vs. reaction time of conductors with different compositions of the bronze

COMPUTER DESIGNED SUPERCONDUCTING MAGNETS

T. KÁRMÁN¹, I. KIRSCHNER¹ and J. BÁNKUTI¹Department for Low Temperature Physics, Roland Eötvös University
1088 Budapest, Hungary

The designing of specific magnetic field profiles is performed by a power series of the general field formula up to 11th term. This method instead of the usual integration increased the speed of the calculation by a factor of few hundreds. At given magnet parameters the program has a self-consistent procedure to gain the prescribed homogeneity taking into account the layer distribution. The operation can be extended for the planning of strongly inhomogeneous magnetic systems too. The elaborated way is applicable to the conventional, superconducting and multifilamentary magnets as well.

Introduction

Our aim was to improve a quick but versatile computer program to design different magnetic field profiles [1]. Solenoid type magnets of partially cylindrical symmetries were used to produce the magnetic field. The program was prepared on a personal computer to calculate the magnetic field profiles in the case of an arbitrary given solenoid arrangement. The only requirement for the calculation is the partially cylindrical symmetry.

The input data of the program are:

data of solenoids: length, inner diameter, maximum outer diameter

data of wires: diameter, density of turns, thickness of layers, intensity of current

minimum of magnetic field strength

data of homogeneity: volume of homogeneous magnetic field /length, diameter/, maximum inhomogeneity

other data: arrangement, profile

Using these data the program calculates the appropriate layer distribution by a self-consistent procedure.

To increase the speed of program running a power series was used up to 11th terms instead of the integrating Biot-Savart formula. The increasing of the calculation speed was by a factor of few hundred. The other advantage of power series calculation was the decreasing of the memory capacity needed for the stored data.

If the axial component at the axis of cylindrically symmetrical field, $H_{00}(z)$ is known then we can calculate the axial and radial components at any given r distance from the axis by the following expressions:

$$H_z / z, r / = \sum_{n=0}^{\infty} \frac{(-1)^n / n!}{|n|!^2} H_0 / 2n / / z / \frac{r}{2}^{2n}, \quad /1/$$

$$H_r/z, r/ = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{n! / (n+1)!} H_0 \frac{(2n+1)!}{(z/\frac{r}{2})^{2n+1}} \quad /2/$$

The value of $H_{00}(z)$ can be determined by this formula:

$$H_0/z/ = \frac{NI}{4b} \left(\frac{z+b}{(z+b)^2 + a^2} - \frac{z-b}{(z-b)^2 + a^2} \right) \quad /3/$$

where a , b , N and I are the half of diameter of solenoid, the half of length of solenoid, the number of turns and the intensity of driving current respectively and z is the axial co-ordinate beginning from the centre of solenoid.

The magnitude of the terms of the power series are decreasing very fast - about 1.5 order - both for the axial and radial components. The other cause of the high speed of the convergency is the sign-changing character of a given component.

The program consists of two parts. The first one calculates the main or basic magnetic field from layer to layer up to the prescribed value. The second part performs the appropriate outer correction layers - constructing it from turn to turn - to achieve the specified field profile or homogeneity.

By this procedure a very quick and versatile method has been developed for designing different normal and superconducting magnets but first of all our method is suitable for superconducting one where no Joule heating occurs.

Some superconducting magnets

To demonstrate the capability of our program we present a couple of examples of superconducting magnets and the magnetic field realized by them.

Simple outer corrected solenoid

At first a simple solenoid is presented for laboratory measurements. No rigorous requirements have been made. The length of solenoid is 10 cm, the inner diameter and the maximum of outer diameter are 4 cm and 20 cm, respectively. The wire has 0.1 cm diameter and the solenoid has to have a minimum magnetic field strength of 2 T, while the driving current is 100 A. The homogeneity requirement is 2 %/cm or less at least in cylindrical volume with 3 cm diameter and 4 cm length in the centre of solenoid.

By 24 layers the magnetic field strength reached the prescribed 2 T. After this the program constructed 36 layers to achieve the homogeneity. The length of the correction layers is decreasing from 2.9 dm to 1.7 cm in this way as it is shown in Fig.1.

At the end of procedure the homogeneity became better than 2 %/cm. During the correction procedure the magnetic field strength increases, in our case up to 2.8 T, but it is very simple to optimize the minimum layer number of the solenoid if it would be necessary.

Highly homogeneous solenoid

In a couple of measurements a highly homogeneous magnetic field is re-

quired in a given volume [2] [3] [4]. In our example the homogeneity had to be better than 10^{-6} /cm in a cylindrical volume with 1.5 cm diameter and 2 cm length in the centre of solenoid. The main data of solenoid are: length 15 cm, inner diameter 2.5 cm, maximum outer diameter 12.5 cm, diameter of wire 0.02 cm, intensity of current 50 A and we need a magnetic field strength of 10 T.

According to the calculation the solenoid needs 41 ground layers and 7 outer correction layers. The first one has 254 turns and the last layer has only 3 turns.

Linear varying magnetic field

Using two equivalent solenoids in front of each other we can obtain an interesting magnetic field profile between them. We have to point out that it differs from the Helmholtz solenoids being reversely coiled. We have examined the axial and radial components of the magnetic field in a volume having cylindrical symmetry of 2 cm length and 2 cm diameter in the centre of the solenoids. If we examine the radial component in the volume in question - see Fig.2 - then we can obtain two interesting facts. At first the radial component is constant in the function of the axial co-ordinate. Secondly the radial component is varying linearly in the radial direction. The axial component shows the opposite property. It is constant in the sense that it does not depend on the distance from axis and it has a linear variation in the axial direction.

In Figs.3 to 6 we have shown the validity of linearity and constancy. We have found that the deviations of above mentioned quantities have a deep minimum at a solenoid distance of about 3,2 cm.

We hope that these examples have shown the versatility of the use of our program. We are going to extend it to even more complicated works, too.

References

1. J. Bánkúti, I. Kirschner and G. Reményi, IEEE Trans. Magn., vol. Mag-17, no. 5, pp. 1999-2002, Sept., 1981,
2. G. Reményi, I. Kirschner and T. Porjesz, Cryogenics, no.17, pp. 565-568, Oct., 1977,
3. I. Kirschner, I. Kovács and L. Lászlóffy, Proc. ICEC 4., Eindhoven pp. 372-374, May 1972,
4. J. Bánkúti, I. Kirschner and L. Lászlóffy, Proc. MT-6, Bratislava, pp. 390-394, Sept., 1977,

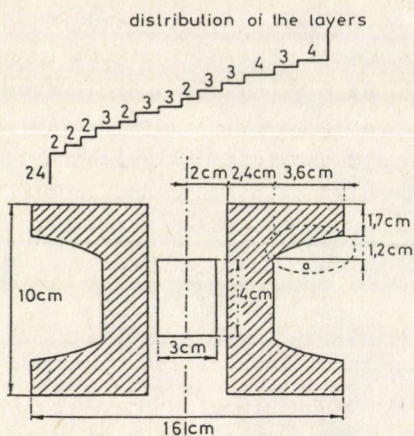


Fig. 1. Simple outer corrected solenoid

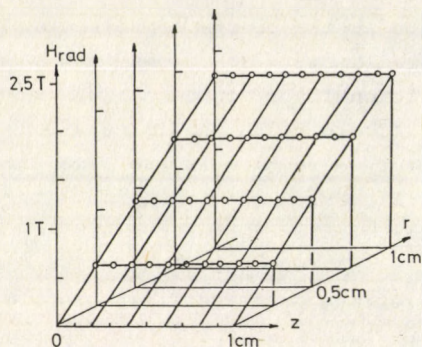


Fig. 2. Radial magnetic field components at a solenoid distance of 3.2 cm

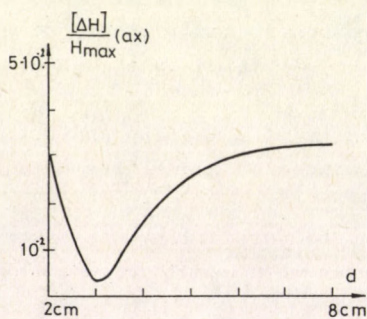


Fig. 3. Relative deviation of axial component from constancy

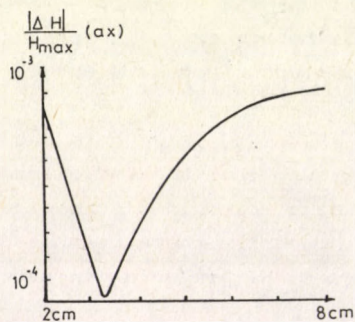


Fig. 4. Relative deviation of axial component from linear varying

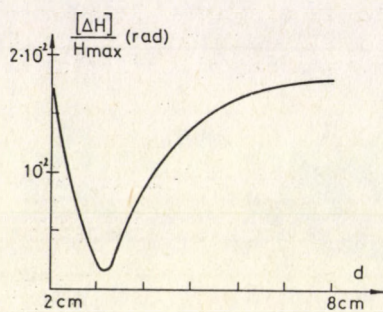


Fig. 5. Relative deviation of radial component from constancy

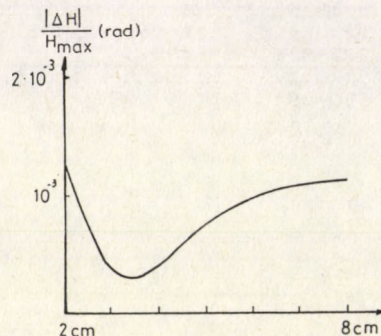


Fig. 6. Relative deviation of radial component from linear varying

BIOMAGNETIC MEASUREMENTS WITH A THIRD ORDER GRADIOMETER

K. VAD, S. MÉSZÁROS and G. HALÁSZ

Institute of Nuclear Research of the Hungarian Academy of Sciences
Debrecen, Hungary

A short introduction about biomagnetism and associated instrumentation is given in the first part of the paper. The application of SQUIDs and superconducting gradiometers in clinical biomagnetic instruments is reviewed in some detail. The construction and adjustment of a third order gradiometer is described later and a magnetocardiogram is displayed to demonstrate the capability of the instrument.

Introduction

It is well known that every electrophysiological phenomenon is accompanied not only by electric but magnetic signals as well. These biomagnetic signals are very weak, 4-7 orders of magnitude below the ambient magnetic noise. Fig. 1 shows some typical values of biomagnetic fields and the average amplitude of the ambient noise.

An apparatus designed to detect biomagnetic fields is to meet two requirements: it must have sufficient sensitivity and capability of producing adequate selection between the signal and the noise. In the majority of biomagnetic investigations SQUIDs are used as magnetic detectors [1]. So as to eliminate environmental magnetic noise, we should build a magnetically shielded room, or make a detection system that is appropriate for spatial discrimination, or apply filtering and averaging techniques.

Constructing shielded rooms is very expensive, so spatial discrimination with SQUIDs and superconducting gradiometers seems to be the best version for clinical applications.

Superconducting gradiometers

In Fig. 2. a schematic sketch of a flux transformer can be seen.

Its working principle is the flux quantization: if we change the magnetic flux in the pick up coil, it causes current change in the signal coil, which can be detected by a SQUID sensor. In biomagnetic measurements this configuration can only be used in a shielded room, because it cannot make the difference between the signal and background noise. Gradiometer-type flux transformers are more suitable for measurements under normal laboratory circumstances [2].

In Fig. 2b "a first order gradiometer" can be seen. This arrangement gives zero response to homogeneous magnetic fields and non-zero response to inhomogeneous ones. It is well known that the gradient of the magnetic field far from the source decreases faster in space than the field itself. So by measuring the gradient (either the first or the higher ones) spatial discrimination can be made between the signals generated by near and distant sources. Better discrimination can be achieved by higher order gradiometers.

In Fig. 3. some sorts of first and second order gradiometers can be seen.

A third order gradiometer

In our laboratory a third order gradiometer has been developed to take magnetocardiograms under laboratory circumstances. In other laboratories 2nd order gradiometers are more widely used, but nowadays some publications have come out where the advantages of the 3rd order types are carefully discussed [3][4]. The main dimensions of our gradiometer can be seen in Fig. 4. The inductance of the gradiometer is about 1.5 μ H, and in our measurements a 40 MHz SQUID electronics is used with a directional coupler [5]. The cryogenic part is located in a BMD-5 type cryostat [7].

The magnetic noise generated by distant sources can be approximated by the first two terms of the Taylor's series of the magnetic field:

$$B_{\nu}(\nu) = B_{\nu}(\nu_0) + \frac{\partial B_{\nu}(\nu_0)}{\partial \nu} (\nu - \nu_0) + \dots,$$

where $\nu = x, y, z$ and $\nu_0 = x_0, y_0, z_0$.

Usually a higher order gradiometer should be balanced in such a way that it gives zero response to homogeneous field and to first gradient of the field in axial direction. In mathematical form the following equations should be satisfied:

$$B_{\nu}(\nu_0) \sum_{i=1}^n N_i A_{i\nu} = 0, \quad (1)$$

$$\frac{\partial B_{\nu}(\nu_0)}{\partial \nu} \sum_{i=1}^n N_i A_{iz} Z_i = 0, \quad (2)$$

where N_i is the number and A_i is the effective area of the appropriate turn, Z_i is the distance between the i^{th} turn and the origin, n is the number of the coils.

In the equation (1), when for example $n=4$, A_1, A_2, A_3 are constant, the gradiometer against the homogeneous magnetic field can be balanced by changing A_4 . Using three superconducting tabs or rings in the three orthogonal directions, all the three conditions of equation (1) can be satisfied independently of one another. In the equation (2) the Z_i parameters can be used to balance the gradiometer against the first gradient part of the ambient magnetic noise. In practice this was realized in Ref. 6.

Because of mechanical simplicity in our gradiometer we used a non-orthogonal procedure: the effective areas in axial direction of two coils were changed. For this purpose a fourth tab was built into the interior of the gradiometer (Fig. 4.). At the best configuration the minimum noise level was about 290 fT/ $\sqrt{\text{Hz}}$ in a steel reinforced concrete laboratory building. Further reduction of noise level is in progress.

Application in magnetocardiography

A magnetocardiogram taken by this instrument can be seen in Fig. 5. A known magnetic field generated by a wire loop has been used to calibrate the instrument and to measure the sensitivity as the function of the source - gradiometer distance (Fig. 6.). It can be seen from this curve that the maximum sensitivity is at about 1 baseline distance and it drastically decreases as the distance increases.

References

1. O. V. Lounasmaa, Experimental Principles and Methods Below 1K, Academic Press, London, 1984.
2. Biomagnetism, Ed. by S. J. Williamson, G. L. Romani, L. Kaufman and I. Modena, Plenum Press, New York, 1983.
3. J. Vrba et al, Canadian Journal of Physics, 7, 1060, 1982.
4. A. C. Bruno et al, Cryogenics, 6, 324, 1983.
5. S. Mészáros, K. Vad, Application of a simple directional coupler for RF SQUIDS, 16. Int. Symp. Tieftemperaturphysik und Kryoelektronik, Bad Blankenburg, DDR, 3-7, Dezember 1984.
6. G. L. Romani, et al, Rev. Sci. Instr. 12, 1815, 1982.
7. S. H. E. Corporation, 4174 Sorrento Valley Blvd, San Diego, California, 92121

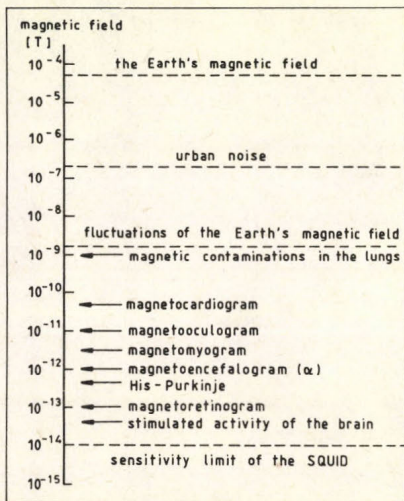


Fig. 1. The amplitudes of different biomagnetic fields originating from the human body

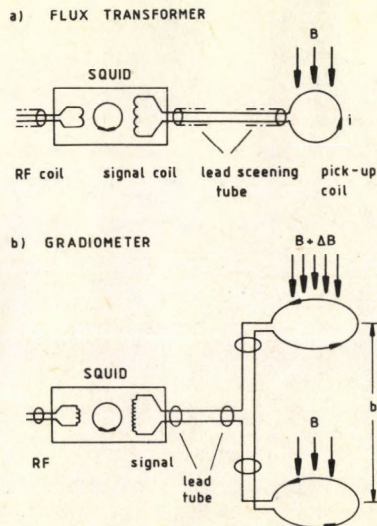


Fig. 2.a. Flux transformer b. First order gradiometer

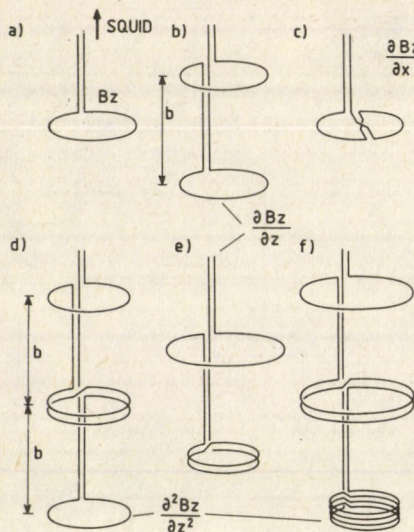


Fig. 3. First and second order gradiometers
The distance b is the baseline

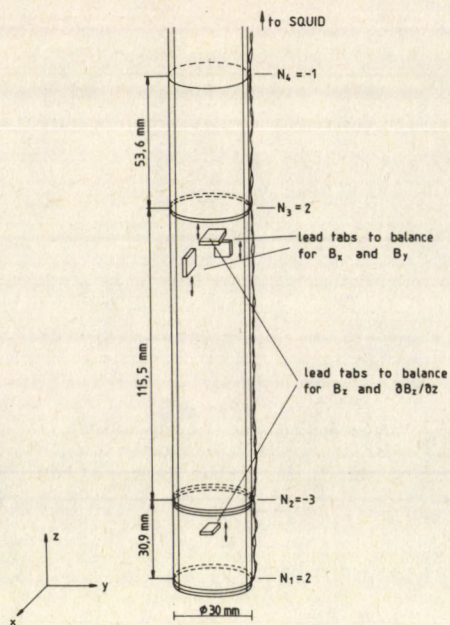


Fig. 4. Third order type gradiometer

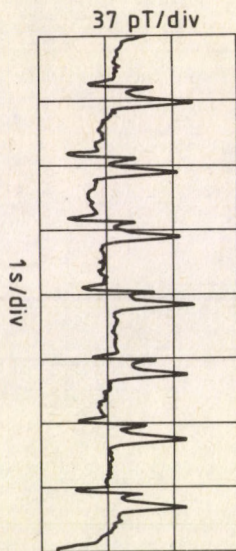


Fig. 5. Magnetocardiogram recorded in an unshielded laboratory. Bandwidth: 0,07-20Hz

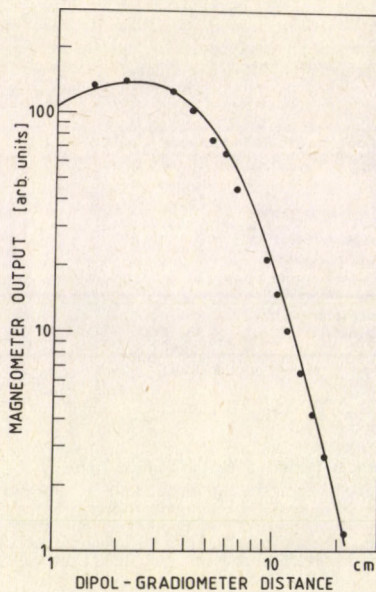


Fig. 6. Sensitivity calibration curve of the instrument

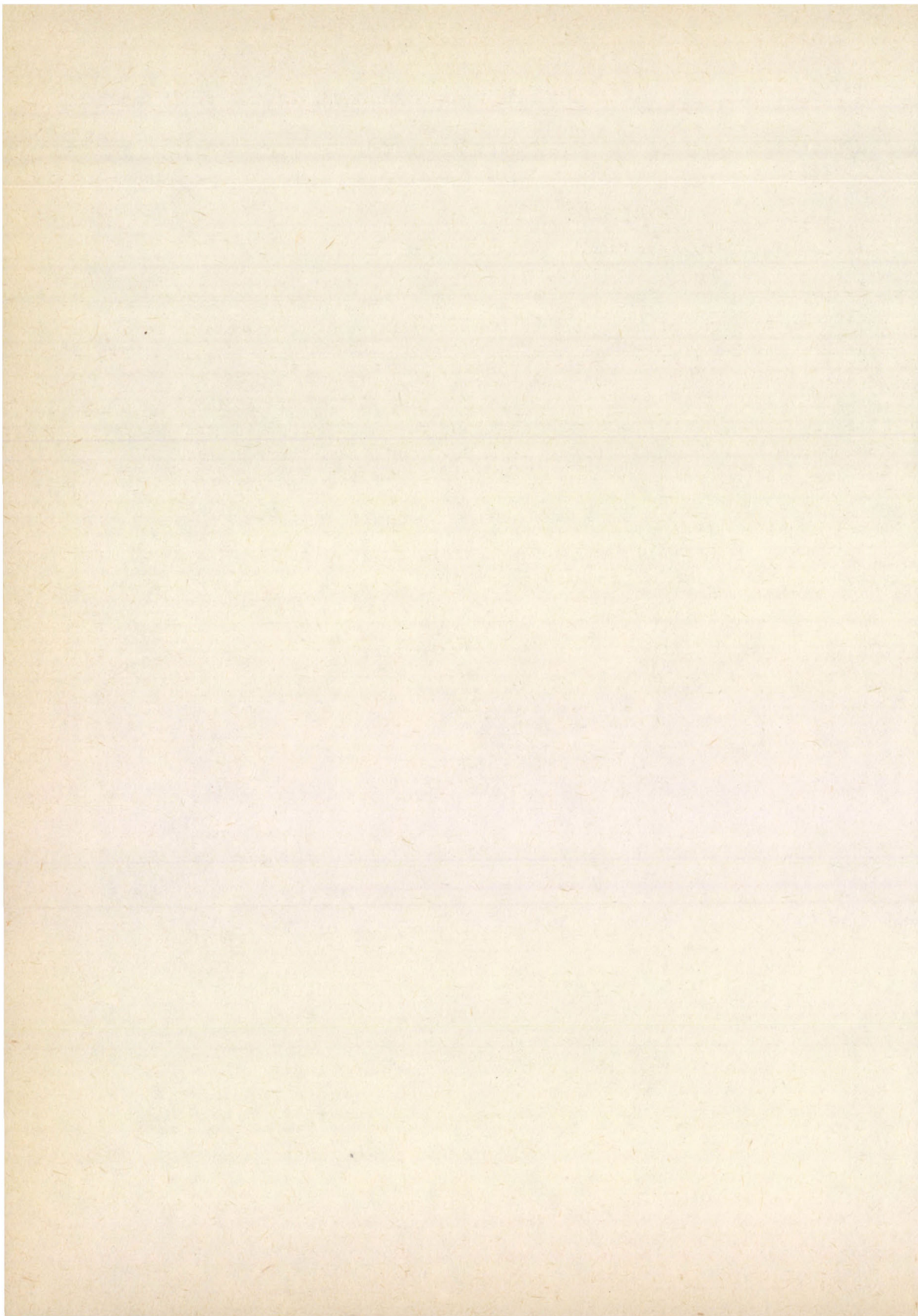
MAGNETIC PROPERTIES OF $(\text{Fe}_{1-x}\text{M}_x)_7\text{Se}_8$ (Abstract) *

M.M. ABD-EL AAL

Physics Department, Faculty of Women, Ain Shams University
Cairo, Egypt

The temperature dependence of the magnetization of the quenched and slowly cooled samples of Fe_7S_8 and $(\text{Fe}_{1-x}\text{M}_x)_7\text{Se}_8$ samples with $\text{M}=\text{Co}$ and Ni and $x = 0.02, 0.05$ and 0.08 were given. All the thermomagnetic curves obtained belong to the Weiss ferrimagnetic type. For some samples discontinuities indicating a magnetic transformation to antiferromagnetic order were obtained. The values of the magnetic moment μ_0 , μ_{78} were given. The reciprocal susceptibility-temperature dependence in the paramagnetic range were studied and the asymptotic Curie points were given. The values of the effective magnetic moment μ_{eff} and the number of unpaired electrons n were calculated. The thermal variation of the electrical conductivity of the host material Fe_7Se_8 in the paramagnetic region was studied.

* The full text of the paper will appear in a subsequent issue of
Acta Physica Hungarica



MAGNETIC PROPERTIES AND SPONTANEOUS DISTORTION IN PHASE LAVES INTERMETALLIC COMPOUNDS

M.M.ABD-EL AAL¹, A.S.ILYSHIN², V.I.CHECHERNIKOV² and A.V.PECHENIKOV²¹Ain Shams University, Faculty of Women
Cairo, Egypt²Moscow University, Faculty of Physics
Moscow, USSR

The magnetization of several polycrystalline compounds in the $Tb_{1-x}Ho_xCo_2$ series have been measured using magnetic fields up to 18 kOe. Linear variations of the magnetic moments at 4.2 K and the Curie temperatures T_c with x were found. X-ray diffracti-on analysis for the samples were carried out between 15 and 300 K. The cubic lattices of all the studied compounds distorted into rhombohedral lattices temperatures below T_c . The distortion parameter - temperature dependence were given. The spin orientation diagram of the system was determined.

Introduction

The compounds RB_2 , where R is a rare earth element and B is a transition metal have received a considerable attention recently [1,2]. These compounds are particularly interes-ting from the point of view of magnetism, while all possess the same $MgCu_2$ cubic Laves pha-se, they exhibit a wide variety of magnetic behaviour: In RFe_2 the magnitude of the magne-tic moment of the Fe atoms is slightly influenced by the different rare earths [3].

In the RNi_2 the magnetic properties depend exclusively on the different rare earths [4]. In RCo_2 the different rare earths can have a very pronounced influence on the magnetic be-haviour of cobalt. In the case of the heavy rare earth- Co_2 compounds, the rare earth moments align antiparallel to the cobalt moment [1,5].

Experimental

The compounds $Tb_{1-x}Ho_xCo_2$ with $x=0; 0.2; 0.4; 0.6; 0.8$ and 1.0 were prepared by arc melting in an argon atmosphere. The samples were annealed in evacuated quartz capsules at $900^\circ C$ for 200 h.

The lattice parameter was measured between 15 K and 300 K. The magnetic measurements were performed by means of a vibrating sample magnetometer in fields up to 16 kOe and between 4.2 K and 300 K. Ni was used for calibration of magnetization σ . The saturation magnetization moments were obtained by extrapolating the $\sigma - \frac{1}{H}$ curves to infinite field. The Curie temperatures were calculated from σ^2 -vs temperature plots.

Results and discussion

The X-ray diffraction pattern at room temperature can be completely indexed by the cubic $C15$ structure for all the prepared samples. The lattice parameters of $TbCo_2$ and $HoCo_2$ agree with the reported values [6,7]. With increasing x in the $Tb_{1-x}Ho_xCo_2$ system, the lattice parameter linearly decreased from 0,7203 nm for $TbCo_2$ to 0,7164 nm for $HoCo_2$.

The temperature dependences of the unit cell volume V are shown in Fig.1. For $TbCo_2$ an anomalous increase in V was found from 239 K to 148 K indicating a second order transiti-

on from cubic $Fd\bar{3}m$ phase to a rhombohedral $R\bar{3}m$ phase at 239 K. The cubic cell is distorted by elongation along the direction of the easy magnetization $[111]$ axis which becomes the rhombohedral three-fold axis. The distortion is measured by the shift in the angle $\epsilon = 90^\circ - \alpha_R$. The thermal variation of ϵ is shown in Fig.2. These results are in good agreement with [7] and [8].

Increasing x up to 0.4 no fundamental changes occurred. In samples with $x = 0.2$ and 0.4 the phase transition occurred at 220 and 150 K.

The phase transition for $x=0.6$ occurred at 142 K. But from 44 to 32 K, the character of the distortion could not be identified. Below 32 K, the $Tb_{0.4}Ho_{0.6}Co_2$ compound attained a monoclinic /rhombohedral/ phase with a parameter of distortion $\epsilon = \frac{c-a}{a}$. Thus, on decreasing the temperature below 44 K the easy axis of magnetization changes continuously from the $[111]$ to the $[110]$ direction.

The compound $Tb_{0.2}Ho_{0.8}Co_2$ is cubic until 98 K. From 98 to 44 K it has an intermediate phase. From 44 K to 15 K the compound attains the distorted monoclinic phase, /Fig.1 and 3/.

The $HoCo_2$ from 300 K to 79 K is cubic /C-15/. From 79 K to 14 K it is tetragonal with $\frac{c}{a} = 0.998$ where the easy axis of magnetization is $[100]$. The easy axis of magnetization is parallel to the $[110]$ axis. $HoCo_2$ has the smallest lattice parameter in the $Tb_{1-x}Ho_xCo_2$ system and the smallest absolute distortion parameter but the largest change in the spontaneous distortion. This large change in the spontaneous distortion as well as the high magnetic anisotropy may be attributed to the crystal field effects [6].

It is possible to determine the spin orientation diagram of the $Tb_{1-x}Ho_xCo_2$ system /Fig.3 /. Four zones can be identified with different stable phases.

The obtained spin orientation diagram is similar to that reported by [9-12].

In order to follow the spin orientations which take place in the studied compounds, the magnetization temperature dependences were studied from 15 K to T_C /Fig.4 /. T_C were obtained by plotting σ^2 versus T for several field strengths. The values of T_C with $H = 1$ kOe were the same as obtained by extrapolating to $H = 0$. The variation of T_C with composition shown in Fig.5 agrees well with that obtained from X-ray results. The relatively low T_C -s of the $Tb_{1-x}Ho_xCo_2$ system in comparison with that of the RFe_2 compounds could be understood by the single ion model. In the RFe_2 compounds the Fe-Fe exchange interaction is the dominant one, in RCo_2 compounds the R-Co interactions are dominant yielding much lower Curie temperatures and affecting only marginally the magnetic anisotropy characteristics.

In Fig.4, nonmonotonical changes in the magnetizations were obtained for $x=0;0.2;0.4$. The anomaly for $HoCo_2$ takes place at about 14 K analogous to the temperature of the spin orientation /Fig.1 and 3/. The anomaly for $x=0.8$ and $x=0.6$ takes place at $T=20$ and at 35 K. We can conclude that with increasing the Ho concentration the anisotropy energy decreases and the reorientation of the easy axis of magnetization occurs at lower temperature.

From the magnetisation - field strength dependences /at 4.2 K and up to 18 kOe/ the saturation magnetic moments were calculated. From the results obtained for the total magnetic moment of $TbCo_2$ and $HoCo_2$ and assuming a magnetic moment of cobalt $1\mu_B$ and opposite in direction to the rare earth ion moment, the rare earth moments are nearly equal to the corresponding values of the moments of the free ions. Similar results were obtained by [13] and [14]. The calculated magnetic moments of the ternary compounds increased linearly by increasing the Ho concentration. The experimental values of the magnetic moments for the pseudobinary compounds (μ_{comp}) show a good agreement with the values obtained by the addi-

tivity of the magnetic moments of the two binary compounds:

$$\mu_{\text{comp}} = (1-x) \mu_{\text{TbCo}_2} + x \mu_{\text{HoCo}_2}$$

According to the results, Co gives a constant magnetic contribution for the compounds of the $\text{Tb}_{1-x}\text{Ho}_x\text{Co}_2$ system. The Co moment (μ_{Co}) suggests an electronic configuration of $3d^9$ and zero charge on the Co atom, in agreement with [5] concept for the heavy rare earth-Co compounds, with antiparallel coupling between the spins of the rare earth and Co atoms.

References

1. K.N.R. Taylor, *Adv. Phys.*, **20**, 551, 1971.
2. W.E. Wallace, *Rare Earth Intermetallics*, Academic Press, 1973.
3. M. Cyrot and M. Lavagna, *J. Physique*, **40**, 763, 1979.
4. E. Gratz, H. Sassik and H. Novotny, *J. Phys. F: Metal Phys.*, **11**, 429, 1981.
5. D. Bloch, M. Edward, M. Shimizu and J. Voiron, *J. Phys. F: Metal Phys.*, **5**, 1217, 1975.
6. D. Gignoux, F. Givord and J. Schweizer, *J. Phys. F: Metal Phys.*, **7**, 1823, 1977.
7. N. Yoshimoto, J. Sakurai and Y. Komura, *J. Mag. and Magnetic Mat.*, **31-34**, 137, 1983.
8. D. Gignoux, F. Givord, R. Perrier de la Bathie and F. Sayetat, *J. Phys., F: Metal Phys.*, **9**, 763, 1979.
9. C.M. Williams, N.C. Koon, *Solid State Comm*, **27**, 81, 1978.
10. G. Dublom, U. Atzmony, *Phys. Rev. B*, **12**, 4628, 1975.
11. H. Klinker, M.P. Dariel and M.J. Rosen, *J. Phys. Chem. Solids*, **41**, 215, 1980.
12. U. Atzmony, M.P. Dariel, E.R. Bauminger, D. Lebenbaum, I. Nowik and S. Ofer, *Phys. Rev. B*, **37**, 4220, 1973.
13. J.W. Ross and J. Crangle, *Phys. Rev.*, **133**, A509, 1964.
14. R.M. Moon, W.C. Koehler and J. Farrell, *J. Appl. Phys.*, **36**, 978, 1965.

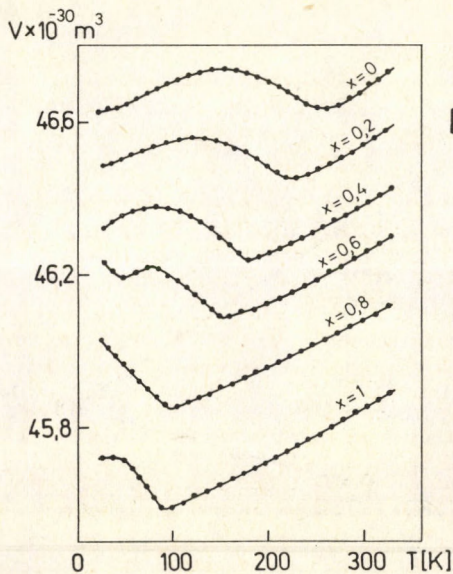


Fig.1. Temperature dependence of the unit cell volume for $\text{Tb}_{1-x}\text{Ho}_x\text{Co}_2$ system

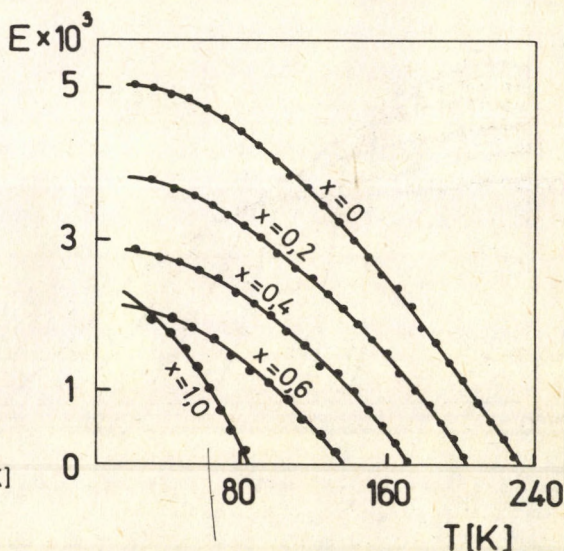


Fig.2. Distortion parameters-temperature dependence for $\text{Tb}_{1-x}\text{Ho}_x\text{Co}_2$ system.

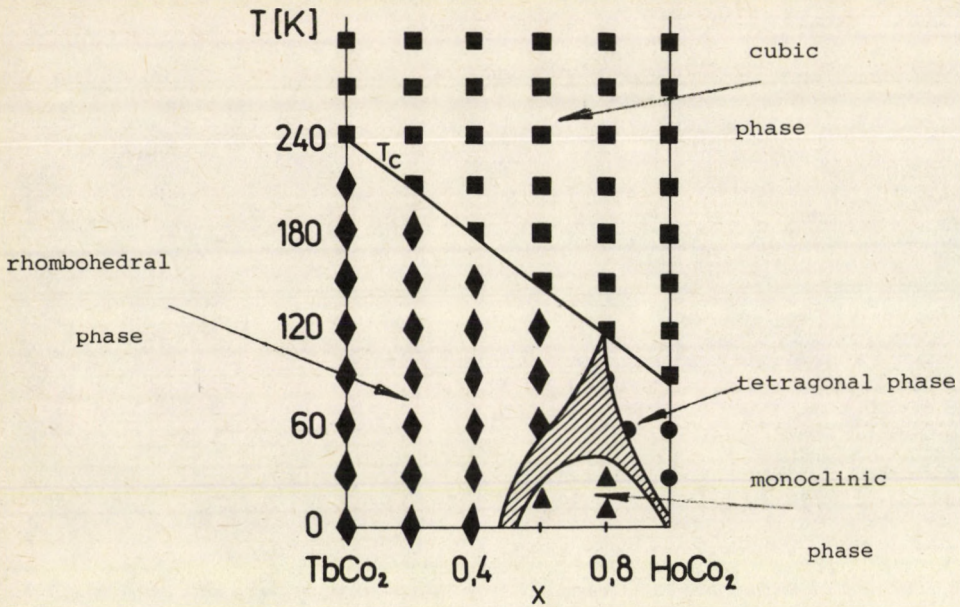


Fig. 3. Phase diagram of the intermetallic compounds $Tb_{1-x}Ho_xCo_2$

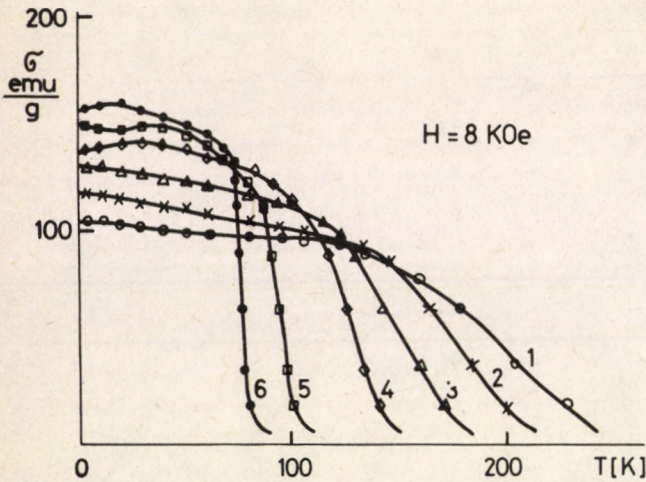


Fig. 4. Magnetization - temperature curves for $Tb_{1-x}Ho_xCo_2$.

1 - $TbCo_2$ 2 - $x = 0.2$ 3 - $x = 0.4$
 4 - $x = 0.6$ 5 - $x = 0.8$ 6 - $HoCo_2$

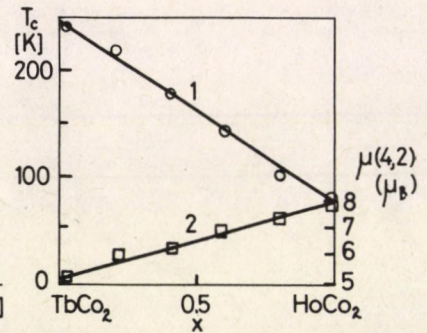


Fig. 5. 1. The Curie Temperature (T_c) K.
 2. The total magnetic moment at 4.2 K (μ_{comp}) (μ_B)

LOW TEMPERATURE MAGNETIZATION OF $\text{Sm}_{1-x}\text{Er}_x\text{Fe}_2$ COMPOUNDSM.M.ABD-EL AAL¹, V.I.CHECHERNIKOV², L.I.KAZAKOVA² and A.V.CHEREMUSHKINA²¹Ain Shams University, Cairo, Egypt²Moscow University, Moscow

USSR

The magnetization of ten compounds in the $\text{Sm}_{1-x}\text{Er}_x\text{Fe}_2$ series has been measured from 4.2K to Curie temperatures in magnetic fields up to 110 kOe. At the composition $x=0.2$ the minimum value of the saturation magnetic moment at 4.2 and 77 K is obtained. In the small erbium concentration region a decrease of the magnetic moment of the compounds in comparison with SmFe_2 is found. This may be related to the increase of the magnetic moment of samarium which is caused by the changes of the energy of anisotropy, the crystalline and exchange field in the compounds. The results are interpreted using the single-ion model and also in the sight of Néel's theory of ferrimagnetism.

Introduction

In the present work the magnetic properties of the pseudobinary system $\text{Sm}_{1-x}\text{Er}_x\text{Fe}_2$ have been investigated. The intermetallic compounds SmFe_2 and ErFe_2 exhibit interesting magnetic characteristics. They possess high Curie temperatures, huge magnitudes of anisotropy energy and negative magnetostriction [1]. ErFe_2 is ferrimagnetic with high magnetic moment of the rare earth sublattice at low temperatures. In contrast to ErFe_2 , the magnetic structure of SmFe_2 had not been established yet because the Sm^{3+} ion magnetic moment is very small and the excited states $^6\text{H}_{9/2}$ and $^6\text{H}_{7/2}$ are close to the ground $^6\text{H}_{5/2}$ state of Sm^{3+} ion. The admixture of these excited states leads to the sign change of conduction electron polarisation relative to the iron magnetic moment causing ferrimagnetic coupling in the SmFe_2 compound [2,3]. For this reason it is interesting to investigate the effect of samarium replacement by erbium on the magnetic properties of $\text{Sm}_{1-x}\text{Er}_x\text{Fe}_2$ compounds.

Experimental results

The specimens of $\text{Sm}_{1-x}\text{Er}_x\text{Fe}_2$ compound with $x=0, 0.02, 0.03, 0.04, 0.05, 0.10, 0.20, 0.4, 0.6, 1.0$ have been prepared by arc melting stoichiometric amounts of the constituents in a purified argon atmosphere. The purity of the starting materials was 99.99% for the rare earth and 99.98% for iron. The samples were annealed in high vacuum at 600°C for 70-120 h and then homogenized in evacuated ampules for 10 days. X-ray, microstructural and spectral analysis revealed that all the specimens are single phase having MgCu_2 cubic structure. The magnetization measurements were recorded as a function of the applied field in "Solenoid" apparatus [4], at 4.2 and 77K in magnetic fields up to 110 kOe. The magnetization measurements were also carried out from 4.2.K to the Curie temperatures in magnetic fields up to 10 kOe.

Magnetization-magnetic field ($\sigma - H$) curves for all the compositions of the compound $\text{Sm}_{1-x}\text{Er}_x\text{Fe}_2$ at 4.2 and 77 K are shown in Fig.1. The saturation magnetic moments M_s calculated from σ_s which were obtained by extrapolating the magnetizations to infinite field ($\frac{1}{H} - \sigma$) are drawn in Fig.2. It is seen from this Figure that on replacing samarium by erbium in a small concentration the magnetic moment M_s noticeably decreases to a minimum value for compound with $x = 0.2$. This can be explained by the fact that in the binary compound SmFe_2 the main contribution to the magnetization of the compound is given by the iron atoms and the samarium and iron magnetic moments are coupled antiparallel. Thus with increasing the erbium concentration in the $\text{Sm}_{1-x}\text{Er}_x\text{Fe}_2$ the rare earth magnetic moment increases decreasing the total magnetic moment of the compound. For the compound with $x = 0.2$ a compensation of the magnetic moments by the rare earth sublattice takes place. By further increasing the erbium concentration the main contribution to the magnetization is given by the rare earth atoms leading to an increase of the total magnetic moment M_s /Fig.2/.

If we assume that in SmFe_2 compound the iron electrons configuration remains the same as in other rare earth - Fe_2 compounds with iron magnetic moment $1.5 - 2\mu_B$ [1] it is possible to apply Néel's theory of ferrimagnetism. The total saturation magnetic moment of the RFe_2 compound is given by

$$M = 2M_{\text{Fe}} - M_{\text{R}}. \quad (1)$$

The samarium magnetic moment in the compounds $\text{Sm}_{1-x}\text{Er}_x\text{Fe}_2$ at 4.2 K can be calculated according to the relation

$$M = 2M_{\text{Fe}} - (nM_{\text{Sm}} + mM_{\text{Er}}), \quad (2)$$

where M_{Fe} , M_{Sm} and M_{Er} are the magnetic moments of iron, samarium and erbium, n and m are the molar concentration of samarium and erbium. The calculated values of M_{Sm} are given in Fig.2. It is seen from this Figure that increasing the erbium concentration x up to 0.04 the samarium magnetic moment increases up to a value comparable to that of the free ion moment for samarium.

The increase of Sm magnetic moments with increasing the erbium content in the small erbium concentration range may be caused by changing the crystal and the exchange field effects and also due to the changes of Sm^{3+} ion anisotropy energy.

As a confirmation of the previous assumption the temperature dependence of the magnetization of the given compounds in 8 kOe applied magnetic field has been studied /Fig.3/.

It is clear from this Figure that there are nonmonotonical changes in the total magnetization in the low temperature range 140 - 240 K in SmFe_2 . It is recorded that a spin reorientation of the direction of the easy axis of magnetization in the 140 - 240 K temperature range [5] takes place. The easy axis of magnetization of SmFe_2 rotates from the direction [011] to [111] with temperature increasing. We can attribute the occurrence of different di-

rection of magnetization to single ion crystal field anisotropy at the Sm site [6].

Investigating the pseudobinary compounds $\text{Sm}_{1-x}\text{Er}_x\text{Fe}_2$ in the region of small erbium concentration till $x = 0.04$ it is found that the nonmonotonical changes in magnetization temperature curves shift to a lower temperature region in comparison with the SmFe_2 compound. For the compounds with $x = 0.05 - 0.20$ the nonmonotonically changing regions are noted at higher temperatures. Analogous to the binary compound SmFe_2 we can propose that these regions of the nonmonotony in the magnetization of pseudobinary compounds are characteristic of the phenomena of the reorientation of the easy axis of magnetization /the direction for which the free energy $/F = -K_B T \ln Z/$ reaches its lowest value/. Thus, with increasing the erbium concentration to 0.04 the anisotropy energy /at low temperatures/ decreases.

Analysing the region of the spin reorientation as a function of composition and temperature in the sight of the single ion theory we can consider that the anisotropy energy K of the SmFe_2 compound at 4.2 K is nearly 10^6 J/m^3 .

In summary, in the $\text{Sm}_{1-x}\text{Er}_x\text{Fe}_2$ compounds the noticeable decrease of the magnetic moment in the small concentration erbium region may be attributed to the increase of the samarium magnetic moment which approaches the magnetic moment of the free samarium ion Sm^{3+} and which is directed oppositely to the iron moment. This may be related to the decrease of the compound anisotropy energy and also to the changing of the crystal and exchange field effects.

References

1. A.E. Clark, *Ferromagnetic Materials*.
2. K.H.J. Buschow, A.W. Van Diepen and H.W. de Wijn, *Phys. Rev. B.*, **8**, 5134 1973.
3. G. Dublon, M.P. Dariel and V. Atzmony, *Phys. Lett.*, **51A**, 262, 1975.
4. V.G. Veselago, L.P. Maksimov and A.M. Prokhorov, *Proc. ICM 73*, Vol. 1, p.250, 1973.
5. M. Rosen, H. Klimker, V. Atzmony and M.P. Dariel, *Phys. Rev. B.*, **9**, 254, 1974.
6. A.M. Van Diepen, H.W. de Wijn and K.H.J. Buschow, *Proc. ICM-73*, Vol.1, p.227, 1973.

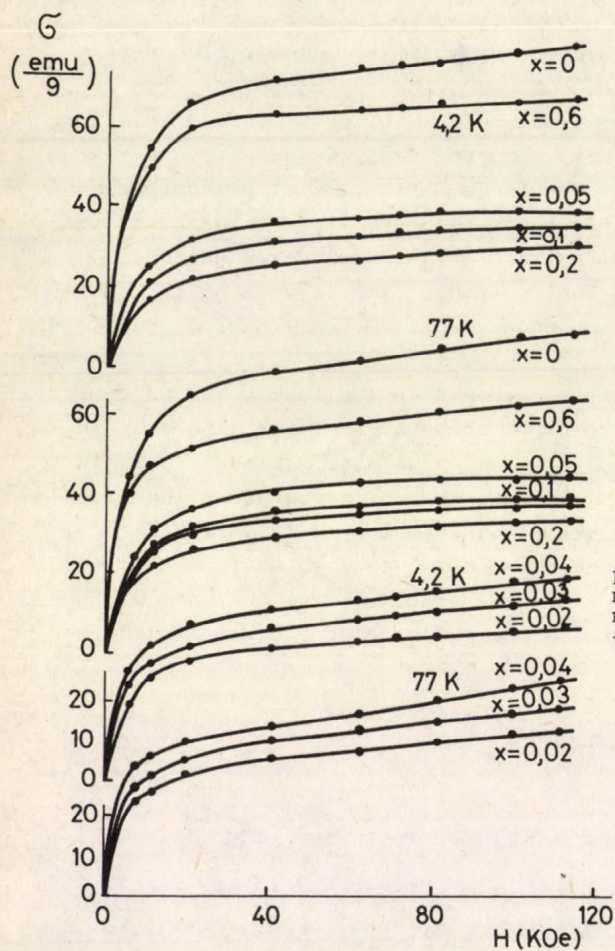


Fig. 1. Isothermal magnetization of $\text{Sm}_{1-x}\text{Er}_x\text{Fe}_2$ system

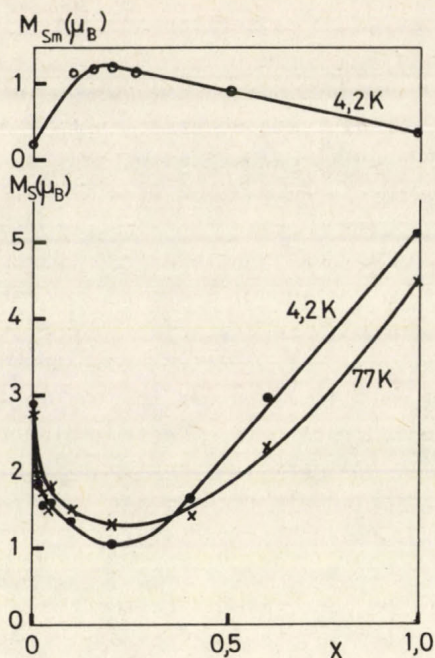


Fig. 2. The saturation magnetic moment M_{S_m} and the calculated magnetic moment of Sm (M_{S_m}) vs. the concentration of Er.

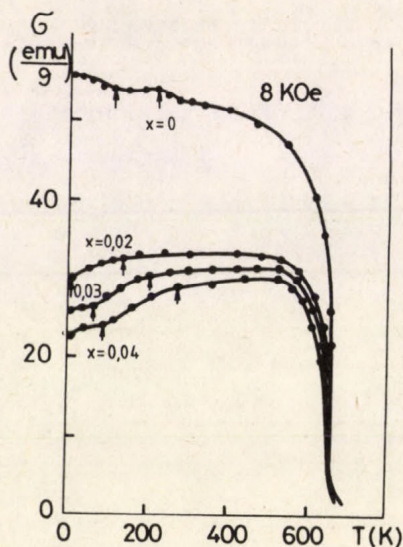


Fig. 3. Thermal variation of the magnetization (σ) in the $\text{Sm}_{1-x}\text{Er}_x\text{Fe}_2$ compounds

INFLUENCE OF THICKNESS ON MAGNETIC PHASE TRANSITIONS IN THIN FILMS OF
DYSPROSIUM AND SAMARIUM

J. DUDÁS¹, A. FEHER², S. JÁNOS³

¹Department of Theoretical Electrotechnics and Circuit Techniques, Technical
University
Kosice, Czechoslovakia

²Department of Experimental Physics, Safarik University
Kosice, Czechoslovakia

³Department of Solid State Physics, Komensky University
Bratislava, Czechoslovakia

The electrical resistivity temperature dependences of bulk and thin film samples of dysprosium, samarium and thulium have been investigated in the temperature interval from 4.2 K to 300 K in the thickness range from 25 nm to 350 nm. A decrease of the Néel temperatures with decreasing film thickness has been observed for Dy and Sm thin films. The observed anomalies in resistivity vs. temperature dependences of Tm thin films have been caused by magnetic as well as by structural transitions due to the presence of hydrogen.

Introduction

The behaviour of rare earth transport properties is dominated by contributions which have their origin in the various spin configurations of ferro- and antiferromagnetic states, that occur mainly at low temperatures.

The transport properties of rare earth metal thin films, mainly the electrical resistivity size effect at room temperature, have created considerable interest [1]. Only few papers have been devoted to the low temperature study of RE thin films.

The aim of this paper is to study the influence of thickness on magnetic phase transitions in thin films of dysprosium and samarium.

Experimental procedure

Thin Dy and Sm films were prepared by evaporation in the vacuum 10^{-4} Pa onto pre cleaned glass substrates. A conventional four probe d.c. arrangement was used to measure the temperature dependences of electrical resistivity in the temperature range from 4.2 K to 300 K. The temperature of thin films was measured using calibrated Ge and Pt thermometers with an accuracy ± 0.1 K. The crystal structure of films was determined using X-ray diffraction techniques.

Our attention has been concentrated to the two RE metals- Dy and Sm, the magnetic structures of that represent their wealth in RE metal group [2].

Results and discussion

Prior to the thin film study we have measured resistance R vs. temperature T dependence of Dy and Sm bulk samples, that have been used as reference samples.

We have observed two anomalies on R vs. T curve of 99.98 % pure Dy with residual resistance ratio $RRR=52$. Using the method described in [3] we obtained dR/dT vs. T dependence, the local extremes of that yielded Néel temperature $T_N = 180.5$ K and Curie temperature $T_C = 92$ K.

Thin Dy films were prepared from this bulk in a thickness range from 26 nm to 350 nm. Their resistance was measured from 4.2 K to 300 K. T_N values, obtained by the above method, are illustrated in Fig.1. We can see in this figure that T_N value of all Dy films is lower than that of bulk and decreases with decreasing film thickness. RRR value of Dy films, illustrated also in Fig.1., decreases with decreasing film thickness.

X-ray analysis has shown h.c.p. phase of Dy and a small amount of dysprosium dihydride.

The R vs. T curve of 99.9 % pure Sm with $RRR = 12$ exhibited two anomalies in the temperature range from 4.2 K to 300 K. Using the method mentioned we obtained phase transition temperatures $T_N^h = 106.8$ K and $T_N^c = 13.4$ K, that correspond to the magnetic moment arrangement of ions in hexagonal respectively cubic sites.

Thin Sm films were prepared from this bulk in the thickness range from 37 nm to 115 nm. Both anomalies are clearly seen on their R vs. T curves. Values of T_N^h , T_N^c , and RRR are illustrated in Fig.2. and are lower than that of bulk sample and decrease with decreasing film thickness.

We assume, that the observed T_N vs. d dependence of Dy and Sm films is caused by one of the following reasons or by their combination:

- 1/ the decrease of T_N value is caused by the increasing relative contamination with decreasing film thickness.
- 2/ the decrease of T_N value is caused by increasing internal stresses and
- 3/ the T_N decrease as predicted by theories.

The RRR vs. d dependence suggests the mechanism 1/ to be prevailing.

References

1. M. Gasgnier, Phys. Stat. Sol, a 57, 11, 1980.
2. R.S. Elliott, Magnetic Properties of Rare Earth Metals, Plenum Press, London, 1972.
3. K.V. Rao, Ö. Rapp and D.J.W. Geldart, J. Phys., C6, L231, 1973.

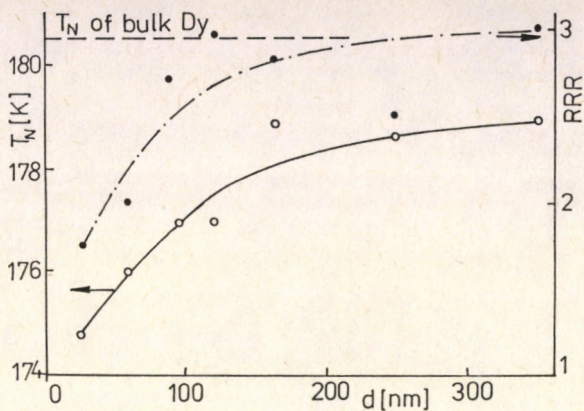


Fig. 1. T_N vs. thickness and RRR vs. d dependences of Dy thin films

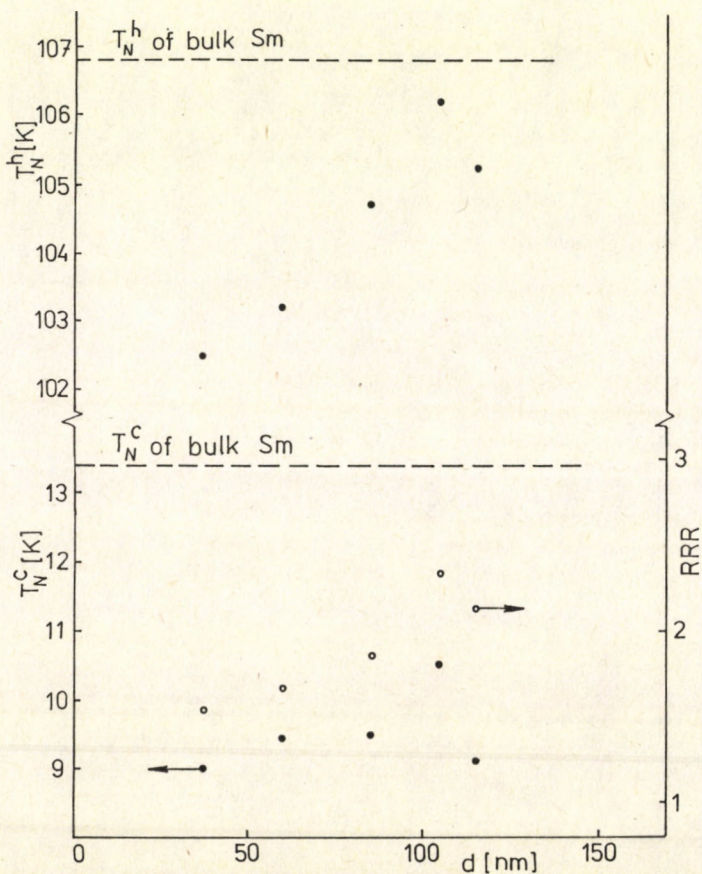
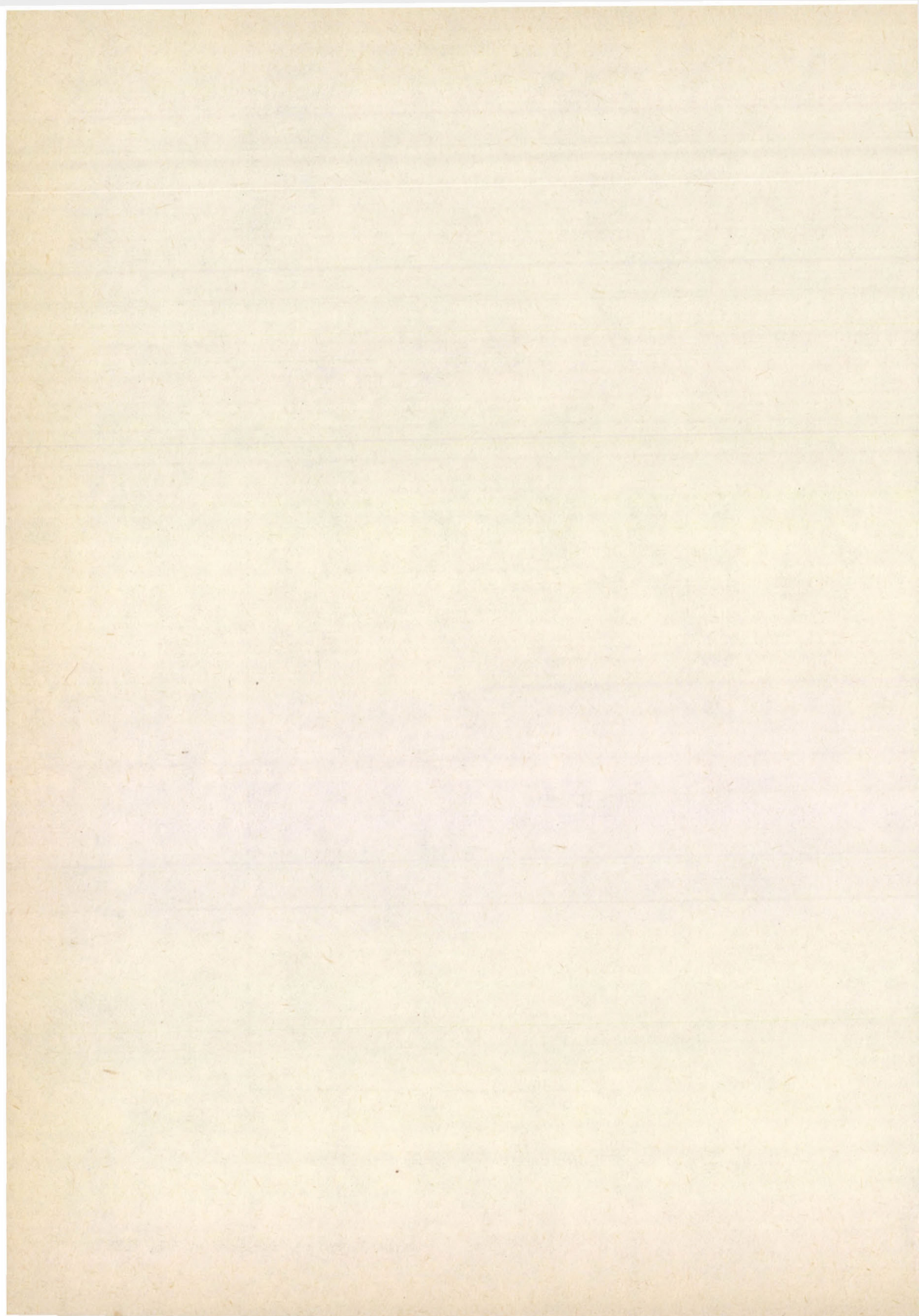


Fig. 2. T_N^h vs. d , T_N^c vs. d and RRR vs. d dependences of Sm thin films



RESIDUAL RESISTIVITY INVESTIGATIONS ON HIGH PURITY GALLIUM

D. NOVÁK, S. MÉSZÁROS, K. VAD

Institute of Nuclear Research of the Hungarian Academy of Sciences
4001 Debrecen, Hungary

Different factors have been studied that influence the residual resistivity analysis of gallium. A proper method to separate the impurity contribution from the measured resistivity has been established and the suitable experimental apparatuses and procedures have been developed. The empirical correlation between the impurity contribution of resistivity and the integral impurity concentration of gallium metal has been studied on the basis of data obtained from measurements on a large number of gallium samples of various purity grades.

Introduction

Investigation of low temperature resistivity of metals is a useful tool to study conduction electron scattering processes. On the other hand, the residual resistivity is widely used for practical purposes to characterize the impurity content /or purity grade/ of high purity metals.

The work reported here is of practical nature. For this purpose instead of resistivity the residual resistivity ratio is used to characterize the samples:

$$R^x = R(T)/R(295 \text{ K}),$$

where $R(T)$ and $R(295 \text{ K})$ are the measured resistances at "low" and "room" temperatures respectively. Methods have been elaborated to measure R^x for gallium samples. In order to set proper experimental conditions and find empirical correlation between R^x and chemical purity a systematic study has been performed on different factors that influence the residual resistivity analysis of gallium. Investigations have been performed on gallium materials of different origin though most of the samples has been provided by Hungalu Company, Budapest.

Experimental method

According to Matthiessen's rule the resistivity comes from several additive contributions. The same is valid for R^x :

$$R^x = R_{ph}^x + R_d^x + R_s^x + R_i^x, \quad //1/$$

where:

- R^x is the experimentally measured value of the resistivity ratio;

- R_{ph}^x comes from scattering of electrons on phonons
- R_d^x, R_s^x, R_i^x are contributions from electron scattering on lattice defects, on the surface of the specimen and on impurity atoms, respectively.

In finding correlation between impurity content and residual resistivity ratio the elimination of the first three terms in /1/ is necessary. R_i^x is not expected to be a single-valued function of the integral impurity concentration because different impurity elements have different scattering cross sections. Instead of it a wider correlation band can be determined experimentally as an empirical relation to be used for the estimation of the purity grade in practical measurements.

To eliminate R_{ph}^x the measurements of R^x have to be performed as a function of temperature. The R^x extrapolated to $T=0$ K is free from contribution of R_{ph}^x . However for practical purposes R^x data measured at $T=4.2$ K can be accepted. /Fig. 1 and 2 /. To get rid of R_d^x one has to use defect free annealed single crystals as specimens. In case of gallium the use of single crystals with definite orientation is preferable because of its resistivity anisotropy, though R^x is not too sensitive to the orientation [1]. Elimination of R_s^x can be performed by extrapolating to infinite specimen size which needs the measurement of the size dependence of R^x [1].

Our measurements have been carried out on cylindrical Ga specimens in a diameter range of 0.6 to 6.3 mm. The single crystal samples have been grown in teflon moulds or polyethylene tubes by seeding. The crystallographic orientation of each specimen has been determined from the room temperature resistivity.

For resistance measurements four-contact d.c. method has been used in two arrangements. One is a conventional system with a commercial nanovoltmeter, the other one uses a SQUID picovoltmeter with a sensitivity of 10^{-13} V [2]. Measuring currents of 0.5-2.5 A were used for the conventional arrangement and 0 - 100 mA for the picovoltmeter with current reversal in both cases.

Results

Our aim was to find an empirical correlation between R_i^x and integral impurity concentration $C = \sum C_i$, where the C_i are the concentrations of different impurity elements. From the R^x values measured at finite size and temperature the contributions of $R_{ph}^x + R_s^x$ have been subtracted /here we assumed that R_d^x is negligible due to the proper sample preparation and handling/. The integral impurity concentrations have been determined by optical spectroscopy and/or mass spectrometry, C is taken as the sum of the concentrations of the detectable impurities. The obtained correlation diagrams /Fig.3. and 4 / show that on the basis of residual resistivity ratio measurements [3] the integral impurity concentrations in gallium can be estimated with an uncertainty of an order of magnitude.

References

1. D. Novák, and K. Botos, ATOMKI Report C/4 49, 1986.
2. S. Mészáros, K. Vad, and D. Novák, Pribori i Technika Eksperimenta No.6 167, 1985.
3. D. Novák, S. Mészáros, K. Vad, K. Botos and M. Mészáros, ATOMKI Report C/1, 121, 1984.

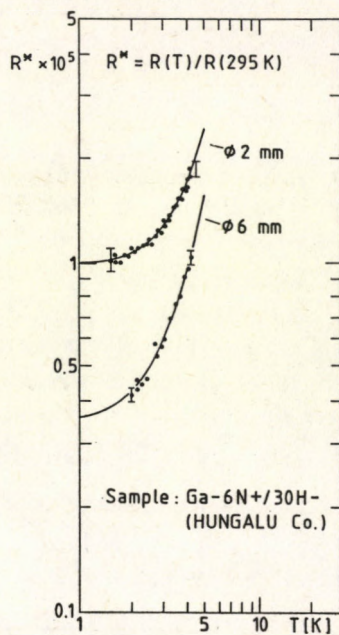
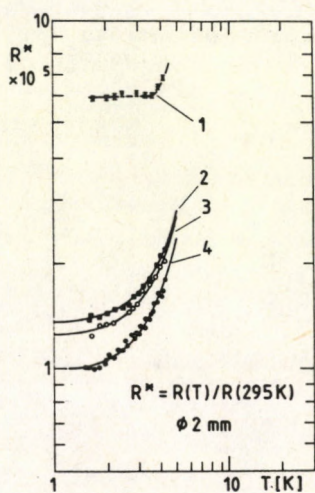


Fig.1. Temperature dependence of R^x measured on gallium samples of various purity grades; 1: technical /4N/; 2: 5N; 3: 6N; 4: 6N+

Fig.2. Temperature dependence of R^x of Ga - 6N⁺/30 H sample /Hungalu Comp./ measured on ϕ 2 mm and ϕ 6 mm specimens

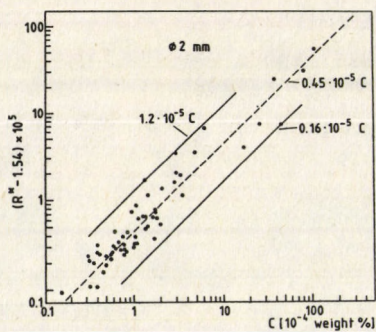


Fig. 3. Correlation diagram for Ga samples of various purity grades. R^x data obtained from ϕ 2 mm specimens [3]

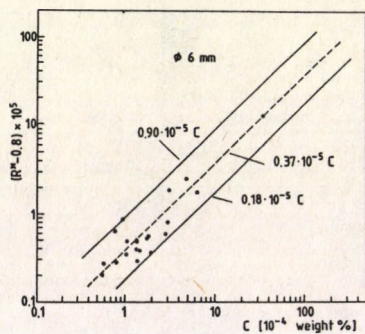


Fig. 4. Correlation diagram for Ga samples of various purity grades. R^x data obtained from ϕ 6 mm specimens [3]

STUDY OF NON-ISOTHERMAL PHASE TRANSFORMATION IN As_2Se_3 - As_2Te_3 GLASSESM.B. EL-DEN¹, M.K.EL MOUSLY²¹Physics Department, Faculty of Science, Ain Shams University
Cairo, Egypt²International Centre for Theoretical Physics
Trieste, Italy

The kinetic parameters for the crystallization processes of the system $AsSe_{1.5-x}Te_x$, with $0 \leq x \leq 1.15$ have been determined based upon the analysis of DTA plots. The compositional dependence of the crystallization activation energy, E_{cryst} , shows a minimum at equal ratio of selenium and tellurium, $E_{cryst} = 15$ kJ/mole for $AsSe_{0.75}Te_{0.75}$. The value of E_{cryst} is dependent on the ratio of selenium or tellurium.

Introduction

The last few years have shown a strong theoretical [1,2] and experimental [3] interest in the application of non-isothermal analysis techniques to the study of the phase transformations because of their several advantages. In the present paper the kinetics of the crystallization processes in the system As_2Se_3 - As_2Te_3 have been investigated.

Experimental technique

Seven compositions of the system $AsSe_{1.5-x}Te_x$ with $0 \leq x \leq 1.15$ have been prepared in evacuated quartz ampoules at $95^\circ C$ for 4-8h. The ampoules were shaken several times to ensure complete homogenization. The molten materials were quenched in air. The amorphous structure of the samples and the macroscopic homogeneity was checked [4-8].

DTA curves were carried out using Shimadzu DT-30 system [4].

Results and discussion

The DTA curves carried out at a constant heating rate $\Phi = 5$ deg/min. Fig.1 shows three DTA curves. These plots are characterized by the glass transition temperature (T_g), crystallization temperature (T_c) and melting temperature (T_m).

Table I summarizes the observed transition temperatures. The endothermic glass transition temperature T_g decreases upon introducing Te to the As_2Se_3 .

After passing T_g the thermograms show one or two exothermic peaks. The beginning of this exothermic peak has been considered as the temperature characterizing the starting of the crystallization process.

Kinetics of the crystallization process of the examined compositions has been investigated using a single-scan technique through an improved DTA model [5]. An estimation of the complex activation energy of crystallization E_{cryst} can be made by using Piloyan's method [9] and [10]. This leads to the following two equations [5]:

$$g(a) = \ln[-\ln(1-a)] = n \ln \left(\frac{K_0^{1/n} E_{cryst.}}{\Phi R} \right) - \frac{E_{cryst.}}{RT} \quad (1)$$

$$\ln[g(\alpha)] = \ln \left(\frac{K_0^{1/n} E_{\text{cryst.}}}{\phi R} \right) - \frac{E_{\text{cryst.}}}{nRT} \quad (2)$$

The transformed fraction α at any time t /or temperature T / was calculated by Borchard's equation [8]

$$\alpha(t) = a(t)/A, \quad (3)$$

where A is the peak area and $a(t)$ the area under the peak up to the time t /or the temperature T /.

Fig.2 illustrates the relations between $g(\alpha)$ vs. $1/T$. This relation yields a linear relationship for both of the crystallization peaks. For the two compositions $\text{AsSe}_{1.35}\text{Te}_{0.15}$ and $\text{AsSe}_{0.35}\text{Te}_{1.15}$ the relationship takes on two different slopes. The compositional dependence of E_{cryst} for the first peak shows a minimum of 15 kJ/mole for $x=0,75$. It reaches 25 kJ/mole for $x=0,15$ and 37,6 kJ/mole for $x=1,15$. The high value of E_{cryst} /44,7 kJ/mole/ for the second peak for $x=0,15$ might be due to some heterogeneous distribution with As_2Se_3 - As_2Te_3 matrix. This introduced barrier or boundaries act to retard the devitrification process in the material.

Comparing the experimental data with different kinetic equations [11] showed that the function $A_3(\alpha)$ where $-\ln(1-\alpha)^{1/3} = Kt$ is closely linear over the entire range of α [$0 < \alpha < 1$] with respect to the first and second peaks of the crystallization, (Fig.4). Accordingly, by using Eq.(2), the value of the order of crystal growth "n" which depends on the details of the nucleation and growth mechanism has been calculated /Table II/. "n" lies in the range 2.5-3.6 for all stages of crystallization of $\text{AsSe}_{1.5-x}\text{Te}_x$ glasses.

The glasses of the system $\text{As}_2/\text{SeTe}_3$ have a character of solid solution in the amorphous state [12], with heat treatment the crystallization processes may take place together with phase separation. There are two T_g / $\text{AsSeTe}_{0,5}$ /, two T_c / $\text{AsSe}_{1,35}\text{Te}_{0,15}$ / and two T_m / $\text{AsSe}_{0,75}\text{Te}_{0,75}$, $\text{AsSe}_{0,5}\text{Te}$ and $\text{AsSe}_{0,35}\text{Te}_{1,15}$ /, indicating two different processes. For the other compositions only one peak appears for the characteristics T_g , T_c and T_m indicating that the possibility of overlapping of the crystallization of the two phases or the crystallization of one phase is predominant.

X-ray measurements [13] have indicated a phase separation for system As_2Se_3 - As_2Te_3 at about 10 % Te content / $\text{AsSe}_{1,25}\text{Te}_{0,25}$ / . The phase separation is the result of a homogeneous melt being quenched through a region of subliquids liquid immiscibility [14]. The subliquids phase separation process often serves as processor reaction for the ultimate crystallization of the amorphous phase [15].

There is an eutectic point at 28,6 % Te content. For samples rich in Se, i.e. before the eutectic composition [12], it is difficult for the crystallization process to take place. This is clear from Table I for As_2Se_3 where there is no crystallization temperature T_c . Increasing the Te content, the crystallization process becomes more pronounced which may appear in one or two steps.

Acknowledgements

One of the authors /M.K.El-M./ would like to thank Professor Abdus Salam, the Internati-

onal Atomic Energy Agency and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste.

References

1. D.W.Henderson, *J. Non-Cryst. Solids*, **30**, 301, 1979.
2. L.Hodany, Central Research Institute for Physics, Budapest, KFKI preprint 24, 1981.
3. H.E.Kissinger, *J. Analytical Chem.*, **29**, No.11, 1957.
4. M.F. Kotkata, M.H. El-Fouly, A.Z. El-Behay and L.A. Wahab, *Mat. Sci. Eng.*, **60**, 163, 1983.
5. M.F. Kotkata and E.A. Mahmoud, *Mat. Sci. Eng.*, **54**, 163, 1982.
6. M.F. Kotkata, A.M. Shamah, M.B. El-Den and M.K. El-Mously, *Acta Phys. Hung.*, **54**, 47, 1983.
7. M.K. El-Mously, M.F. Kotkata and M.B. El-Den, *Egypt. J. Solids*, **1**, 166, 1980.
8. M.F. Kotkata and M.B. El-Den, ICTP, Trieste, Internal Report IC/83/99.
9. F.O. Piloyan, I.O. Ryabchikov and O.S. Novikova, *Nature*, **212**, 1229, 1966.
10. H.J. Borchard, *J. Inorg. Nucl. Chem.*, **12**, 252, 1960.
11. J.H. Sharp, G.V. Brindly and B.N. Achar, *J. Am. Ceram. Soc.*, **49**, 379, 1966.
12. Obrazov et al. *Bull. Acad. Sciences, USSR, Ing. Material*, **7**, 2166, 1971.
13. S.A. Saleh, M.F. Kotkata and M.K. El-Mously, *Proc. Math. and Physical Soc. of Egypt*, **47**, 73, 1976, **42**, 83, 1976.
14. M.B. Myers and J. Berkes, *J. Non-Cryst. Solids*, **8-10**, 804, 1972.
15. J.W. Cohn, *J. Am. Ceram. Soc.*, **52**, 118, 1969.

Table I
Characteristic temperatures of $AsSe_{1.05-x}Se_x$ glasses. The temperatures are given in $^{\circ}C$

Composition	T_g $^{\circ}C$	Crystallization exotherm		T_m $^{\circ}C$
		T_c $^{\circ}C$		
		First peak	Second peak	
$AsSe_{1.5}$	189	-	-	387
$AsSe_{1.35}Te_{0.15}$	131	170	272	324
$AsSeTe_{0.5}$	150	240	-	290
	196			314
$AsSe_{0.75}Te_{0.75}$	136	194	-	288
				309
$AsSe_{0.65}Te_{0.85}$	134	193		297
$AsSe_{0.5}Te$	131	187	-	310
				320
$AsSe_{0.35}Te_{1.15}$	129	160	-	323
				340

Table II
Crystallization kinetic parameters of $AsSe_{1.5-x}Te_x$ glasses

Composition	First peak		Second peak	
	n	E-kJ/mol	n	E kJ/mol
$AsSe_{1.35}Te_{0.15}$	3.6	251	3.5	448
$AsSeTe_{0.5}$	2.6	230	-	-
$AsSe_{1.75}Te_{0.75}$	2.8	151	-	-
$AsSe_{0.65}Te_{0.85}$	3.2	184	-	-
$AsSe_{0.5}Te$	2.9	209	-	-
$AsSe_{0.35}Te_{1.15}$	3.2	377	2.5	109

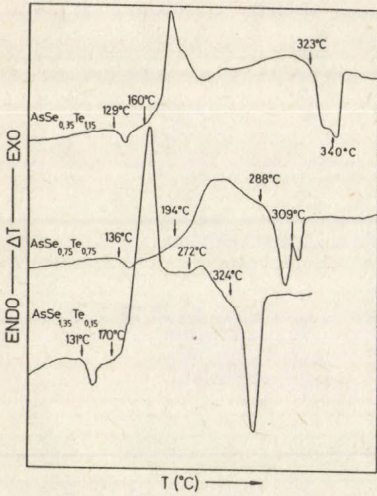


Fig.1. DTA thermograms for the compositions $AsSe_{1.35}Te_{0.15}$, $AsSe_{0.75}Te_{0.75}$ and $AsSe_{0.35}Te_{1.15}$

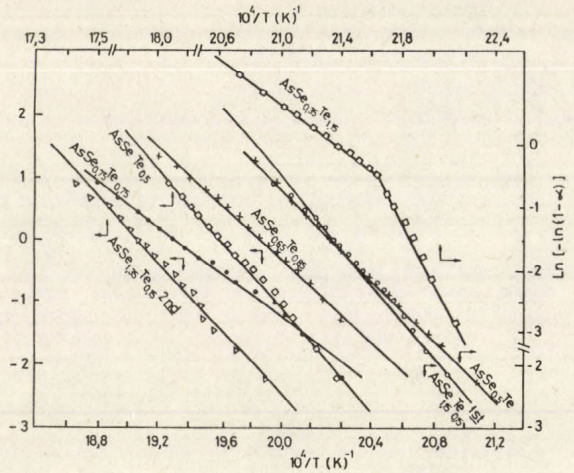


Fig.2. Plots of $\ln[-\ln(1-x)]$ vs $1/T$ for the first and second peaks of crystallization of $AsSe_{1.5-x}Te_x$ systems

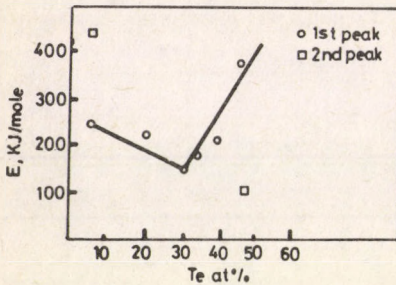


Fig.3. The variation of the activation energy of crystallization with tellurium concentration

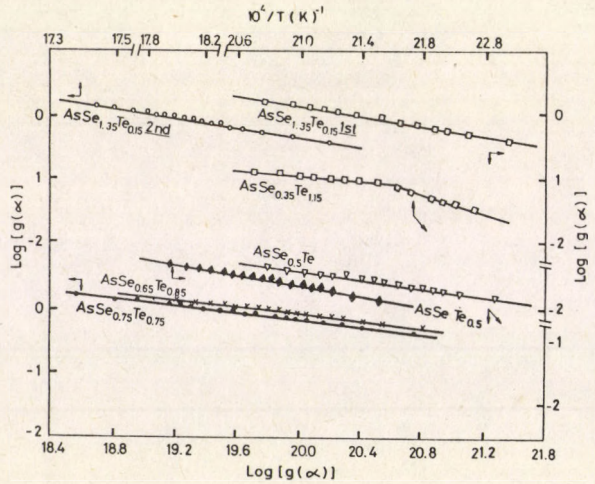


Fig.4. Plots of $\log g(\alpha)$ vs $1/T$ for the first and second peaks of crystallization of $AsSe_{1.5-x}Te_x$ systems

ELECTRON MOBILITY IN HIGH DENSITY NEON GAS

A.F.BORGHESANI, L.BRUSCHI, M.SANTINI and G.TORZO

Dipartimento di Fisica "G.Galilei", Università di Padova

35131 Padova, Italy, and

Gruppo Nazionale Struttura della Materia

Padova, Italy

The electron - atom scattering cross section in Ne is much smaller than that of other simple gases. Therefore the electron mobility is field-dependent even for very low fields. The zero-field electron mobility can be determined by extrapolation from $\mu(E)$ curves only.

Introduction

The density and temperature dependence of the electron drift mobility in high density Ne gas has been recently measured [1,2] using a "square wave" technique /SWT/. (Fig.1). This technique has been extensively described elsewhere [3].

For the current of the sample cell we get $i=0$ for $f > f_e$ and $i=(I_0/2)(1-f/f_e)$ for $f < f_e$. f_e and f are the constant and the variable square wave frequencies between G1-G and G2-C. I_0 is the current measured at zero frequency. The electron time of flight between G2 and C, $\tau_e=(2f_e)^{-1}$, is given by linear extrapolation to $i=0$. The mobility is then obtained as $\mu = d/(\tau_e E)$ /E is the electric field/.

When electron attachment to molecular impurities /usually oxygen/ is present, the same relation $i(f)$ holds, but I_0 is now replaced by $I_0 \exp(-\omega/f_e)$, where ω is the attachment frequency. An impurity concentration of 0.01 ppm can be detected with this method.

Measuring of the electron mobility

The SWT technique has been used to measure the electron mobility at very low values of the reduced field E/N , where N is the gas number density /typical values are $E/N \approx 10^{-25} \text{ m}^2 \text{ V}$ /. At such low E/N values the electrons are thermalized in the He, H_2 , Ar, N_2 , Kr, Xe, but not in Ne, where the electron-atom scattering cross section is very small.

In Ne, therefore, the electron mobility is field-dependent and the "zero-field" mobility μ_0 can only be extrapolated from the $\mu(E)$ curves /Fig.2/.

In low density gases the classical theory gives a formula for μ_0 as a function of temperature T , density N , and the thermal scattering cross section σ_{TH}

$$\mu_0 = \frac{4e}{3N\sigma_{TH}(2mk_B T)^{1/2}}, \quad (1)$$

$$\sigma_{TH}^{-1} = (k_B T)^{-2} \int_0^{\infty} d\epsilon \frac{\epsilon \exp(-\epsilon/k_B T)}{\sigma_{mt}(\epsilon)}, \quad (2)$$

where σ_{mt} is the energy-dependent momentum transfer cross section.

In the low density limit $\mu_0 N$ should be density-independent $\lim (\mu_0 N) = (\mu_0 N)_0 = A(T)$. The function $A(T)$, calculated from /1/ and /2/ for $Ne \rightarrow O$ is plotted in Fig.3. The agreement with the extrapolated $(\mu_0 N)_0$ data /closed circles/ is substantially good.

Multiple scattering effects

The classical theory, however, fails at finite density and several theories have been proposed [4,5,6] to explain the experimental data, which account for multiple scattering effects. All these theories /MST/ suggest, at moderate densities, a linear decrease of $(\mu_0 N)$ with N for gases with a positive scattering length, a , like He, H_2 and Ne:

$$\mu_0 N = (\mu_0 N)_0 (1 - \gamma N) = A(T) - \alpha N, \quad (3)$$

where $\gamma = h \cdot \sigma_{TH} / 2(2\pi mk_B T)^{1/2}$, and $\alpha = eh / (3\pi mk_B T)$.

The mobility behaviour predicted by /3/ is qualitatively /and for He and H_2 also quantitatively/ in agreement with the experimental results. For Ne a density dependence of $\mu_0 N$ much stronger than the expected one has been measured.

The $\alpha \cdot T$ vs. T functions for different gases are shown in Fig.4. The constant theoretical value $\alpha = eh / (3\pi mk_B T)$ is drawn by dashed line. α is derived from the experimental slope of $\mu_0 N$ vs. N .

Thomson discharge method

In order to check the strange results for Ne we repeated our measurements with the "Thomson discharge method" [9]. The TDM has been adapted to our low temperature and high pressure drift cell.

A pulsed photocathode has been used as electron source, and the electron time of flight has been measured by means of fast electronics and a transient recorder.

Special care has been also devoted to avoid systematic errors due to the finite electrodes geometry and to the electron attachment.

The "zero density" values $(\mu_0 N)_0$ obtained with TDM /open circles in Fig.3./ are now in perfect agreement with MST.

At high density, the density dependence of $\mu_0 N$ is still much stronger than predicted by MST /Fig.4, open circles/, leading us to the conclusion

that the multiple scattering theories by alone are inadequate to explain the electron mobility behavior in high density Ne.

An important peculiarity of Ne, with respect to other gases with $\alpha > 0$, is that $\sigma_{mt}(0)$ is much smaller, and moreover $\sigma_{mt}(\epsilon)$ is strongly dependent on the electron energy ϵ/a comparison with He and H_2 is shown in Fig.5/.

Conclusions

Starting from this consideration we suggest a rough model to fit the experimental results. Essentially we assume that the effective cross section is affected by density through the Fermi shift [10] of the electron ground state energy ϵ_0 .

Such an effect, in fact, becomes important in Ne owing to the steepness of the $\sigma_{mt}(\epsilon)$ curve at low energy /it is negligible in He and in H_2 where $\sigma_{mt}(\epsilon)$ is flatter/.

In Fig.6 the data obtained at three different temperatures are plotted as $\mu_0 N$ versus N . The full line is calculated following MST and the dashed line is calculated following our model without any adjustable parameters. In spite of the crudeness of our approach the agreement is good, suggesting that more theoretical effort in this direction could offer a correct interpretation of the electron mobility in dense gases.

References

1. L. Bruschi, M. Santini and G. Torzo, Phys. Lett., 102 A, 102, 1984.
2. A. Borghesani, L. Bruschi, M. Santini and G. Torzo, Phys. Lett., 108 A 255, 1985.
3. L. Bruschi, M. Santini and G. Torzo, J. Phys. E: Sci.Instrum, 18, 239, 1985.
4. G. L. Braglia and V. Dalla Casa, Phys. Rev., A26, 902, 1982.
5. T. F. O'Malley, J.Phys., B13, 1491, 1972.
6. V. M. Atrazhev and I. T. Yakubov, J.Phys. D10, 1977.
7. K. W. Schwarz, Phys.Rev., B21, 5125, 1980.
8. A. K. Bartels, Ph.D. Thesis, Hamburg, 1974.
9. A. F. Borghesani, L. Bruschi, M. Santini and G. Torzo, Z. Naturforsch., 41a, 912, 1986.
10. E. Fermi, Nuovo Cimento, 11, 537, 1934.

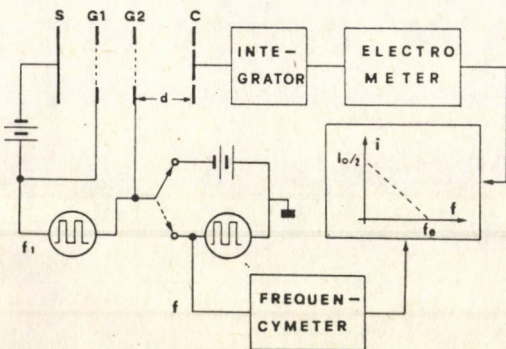


Fig.1. The square wave technique [1,2]

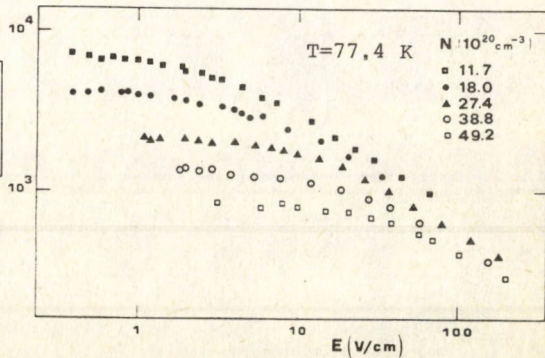


Fig.2. Electron mobility in Ne vs. electric field

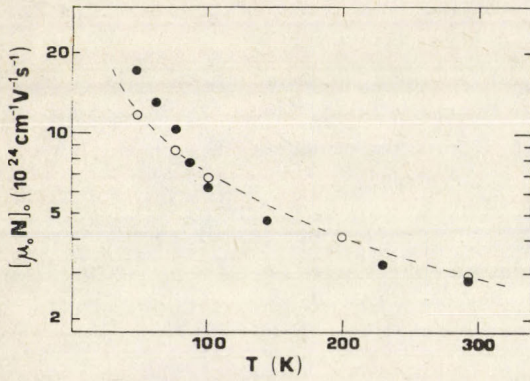


Fig. 3. $\mu_0 N$ extrapolated to zero density vs. temperature

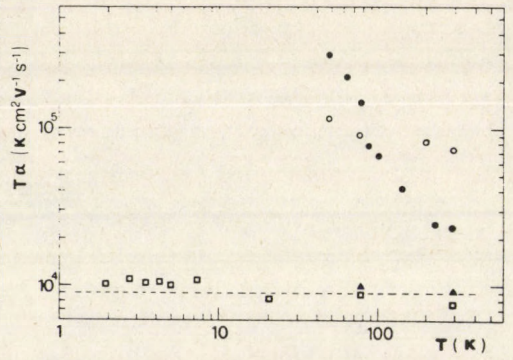


Fig. 4. αT vs. temperature

- He [7, 8]
- ▲ H₂ [8]
- Ne /SWT/
- Ne /TMD/ [1, 2, 9]

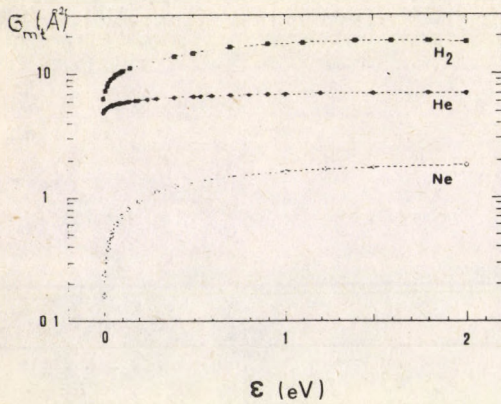


Fig. 5. Momentum transfer cross section of low energy electrons in Ne, He and H₂

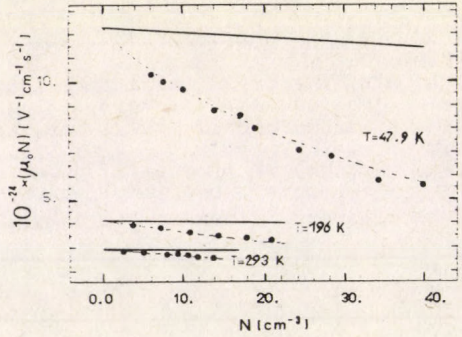


Fig. 6. Zero field mobility μ_0 times N in Ne vs. density at different temperatures

ARE THE STABILITY AND THE SUPERCONDUCTIVITY OF ALLOYS CONNECTED?

A. SZÁSZ¹, J. BÁNKUTI²

¹Institute for Solid State Physics, Roland Eötvös University
1088 Budapest, Hungary

²Department for Low Temperature Physics, Roland Eötvös University
1088 Budapest, Hungary

The electronic structure has a local stabilizing effect in metastable and amorphous phases. A model is given to explain the lower total energy in these cases. The DOS measurements confirm the role of the electronic structure in the stability. There is an explanation for the behaviour of the superconducting transition temperature T_c under the pressure and at heat treatment.

Introduction

The local stability of different alloys and their metastable phases have been continuously investigated from the beginning of solid state physics as an independent discipline. From the first moment it was clear that the electronic structure is one of the basic factors of this branch of problems. One of the largest results in the early period of these investigations were the Hume-Rothery-rules for metals [1]. Based on the Hume-Rothery's investigations many extended new results were reached [2,3,4]. The related problems have been raised again more widely and deeply in the last decade in connection with the amorphous materials (first of all the metallic glasses) and with the different problems of the surface and interface phenomena.

The role of the electronic structure in the formation of binary Ni-P metallic glasses has some interesting specialities, which is in a good agreement with the investigations of metastable phases of the aluminiumbased alloys [4].

On the basis of our soft X-ray emission investigations we have a conclusion that the electronic structure in the metastable state is definitely more stable than in the "equilibrium" structure. So the local minimum of the energy of the electronic structure is a stabilizing factor; the metastable structure is stabilized by the electronic structure. It is clear from the local investigations too that the changes in the chemical bond are responsible for the short-range order.

The model

The basic problem of all electronic structure investigations is the geometrical discrepancy of the Brillouin-zone BZ and Fermi-sphere FS. The Brillouin-zone reflects the geometry of crystal lattice and the Fermi-sphere has spherical type geometry, which is the smallest volume for the non-excited states. In such a simple case, when the Fermi-radius is smaller than the

smallest distance between the BZ centre and the boundary, we have no difficulties; the dispersion relation is parabolic, the density of electronic states DOS changes as the square root of electronic energy. Problems arise when the Fermi-sphere touches the boundary of Brillouin zone. This situation is given in the simple case of FCC lattice on the Fig. 1. After the first crossing of the BZ boundary the electrons need a gap energy to fill the second BZ. In the other region of the FS we have a continuous filling of the first BZ without gap energy. It means energetically that the minimum energy will be in the case of the full-filling of the first BZ, and the filling of the empty second BZ begins after this with defined energy gap. The more spherical BZ can secure much more electronic states in the first BZ, so the total energy can be much less than in a real, polyhedral BZ case. The deformation of the real lattices deforms the BZ too. The deformed BZ can form a more spherical shape, which holds more electronic states in the first BZ. In the case of random deformations the BZ will be a smeared sphere. It produces an elongated parabolic dispersion relation with a smaller gap, and a smoothed DOS. This effect can be seen in the DOS measurements of different metastable and amorphous alloys too [4,5]. In many cases of metastable phases the stabilizing factor is the electronic structure [4] and it can stabilize the amorphous structures. This is very similar to the Nagel-Tauc's assumption for the amorphous stability [6], where the DOS at the E_F has the local minimum (Fig. 2).

The microscopic stress-strain investigations of amorphous materials show a linear dependence, so the larger stress gives smaller real lattice parameters, consequently a larger BZ, which helps in the electronic stability too. This could be the reason of the pressure sensitivity of superconductivity [7], when the superconducting critical temperature T_c increases with the pressure. The more amorphous state ("smeared spherical" BZ, the higher internal stress) has a larger T_c in many cases [8] (Fig.3). At the annealing the stresses decrease, which is followed by the decrease of the T_c [9], (Fig.4).

References

1. W. Hume-Rothery, *Electrons, Atoms, Metals and Alloys*. ILIFFE et Sons. LTD. 1948.
2. L. Brewer, *Electronic Structure and Alloy Chemistry of the Transition Elements*. E.: P. Beck, Interscience, J. Wiley, N.Y., 1963.
3. D.G. Pettifer, *J. Phys. C. Sol. Stat. Phys.* **3**, 367, 1970.
4. A. Szász, L. Kertész, J. Hajdu, J. Kollár, *Aluminium (Düsseldorf)* **61**, 515, 1985.
5. I. Kübler, K.H. Bennemann, R. Lapka, F. Rösel, P. Oelhafen, H.J. Günterodt *Phys. Rev. B* **23**, 5176, 1981.
6. S.R. Nagel, J. Tauc, *Phys. Rev. Lett.*, **35**, 380, 1975.
7. I. Kovács, I. Kirschner, J. Bánkuti, L. Lászlóffy, *Proc. 12. Int. Conf. Low Temp. and Techn.*, Székesfehérvár, 1973, p. 99.
8. W.L. Johnson, *Glassy Metals I.*, Eds: H.J. Günterodt, H.Beck, Springer Verlag Berlin, 1981.
9. Qi-rui Zhang, H.U. Krebs, H.C. Freyhardt, in: *Proc. LT-17*, Eds: U. Eckern, A. Schmid, W. Weber, H. Wühl, Elsevier, 1984.

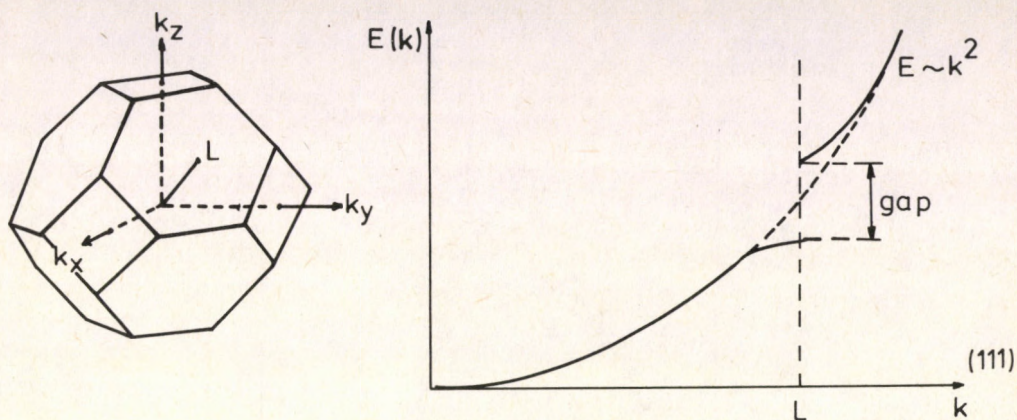


Fig. 1. The energy gap at the touching of the Fermi-sphere and the boundary of the Brillouin zone in the case of FCC lattice

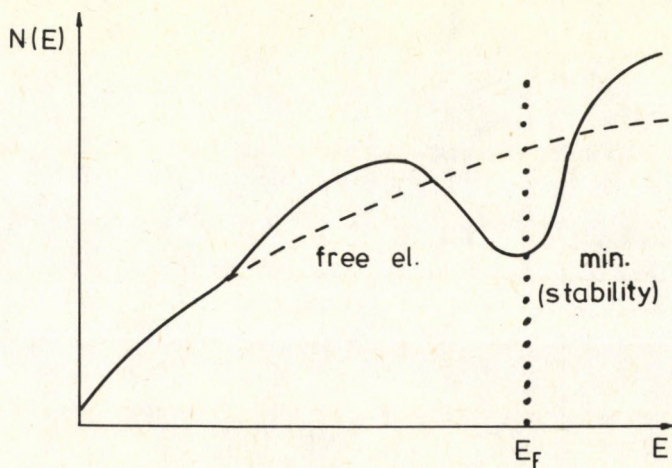


Fig. 2. The stability at the local minimum of the density of electronic states

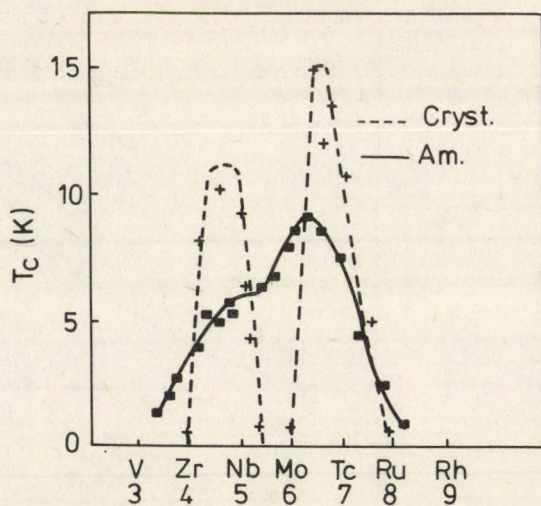


Fig. 3. The dependence of the T_c on the internal stress

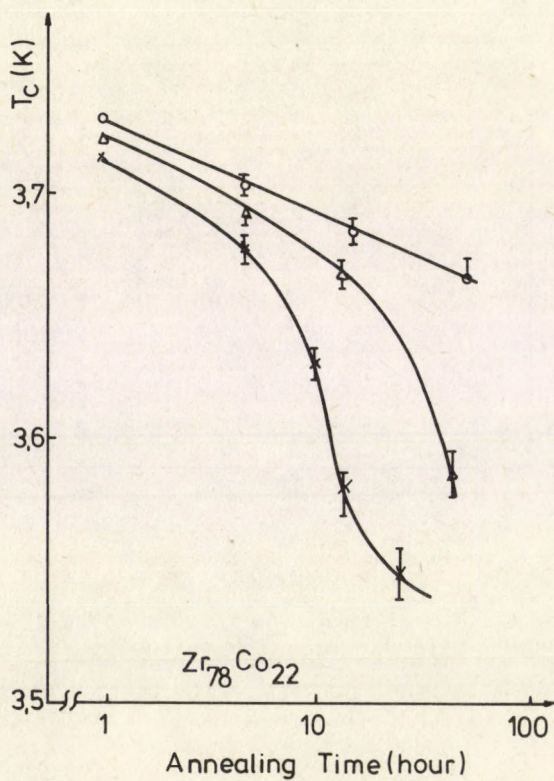


Fig. 4. The effect of the annealing on the T_c

MECHANICAL PROPERTIES OF TEFLONS AT LOW TEMPERATURES

L.LÁSZLÓFFY and J.BÁNKUTI

Department for Low Temperature Physics, Roland Eötvös University
1088 Budapest, Hungary

Many kinds of materials are used in the experimental investigations of superconducting or cryoresistive devices. Among them certain synthetics have important roles in the experimental and industrial applications as well. Earlier examinations show the very advantageous properties of teflon like materials in cryogenics.

The advantageous properties can be both mechanical and electrical. This is the reason why they are preferable for electric and thermal insulation and for some mechanical and stabilizing problems. In recent paper the mechanical properties of different teflons are studied.

The effect of the structure on the mechanical properties of teflons

The chemically homogeneous polymerized tetrafluorethylen is a crystalpolymer and its property is determined by the rate of crystallization [1]. The melting point of the crystallites is 600K and the temperature for the vitrification of non-crystallized parts is 153 K. In the applications at room temperatures the solid crystallites are mixed with the non-crystallized parts. The result is a soft, elastic material down to helium temperature. During the production of teflon the rate of crystallization can be controlled. By faster cooling we get a more elastic structure by the low rate of crystallization, and by slow cooling the crystallites can be so large, that the material can suffer only a small deformation without breaking.

The improvement of the properties of teflon can be made by adding different filling materials during the production. Our studies were concentrated to certain temperature dependent mechanical and thermal properties of some teflons filled in different ways.

Compressive strength measurements

We have made measurements on teflons filled by bronze + carbon, carbon, 15% glass, 40% glass and on pure Hungarian and GDR materials. The applied deformation was in the range of 5% - 25%. The temperatures of the measurements were 373 K, 296 K, 273 K and 77 K.

The experimental results are summarized in the Table I.

The measurements made at different temperatures unambiguously prove that the teflon doped by bronze + carbon is the most solid. Depending on the elastic property of material the accumulated mechanical energy causes an inner heating. The increase of the temperature results in a steep decrease in the solidity. If the heat originating inside the material can not be conducted

away so it causes a decrease in the solidity of the material. Supposing the better heat conductivity of the bronze + carbon filled teflon the better mechanical properties are evident.

Examination of the thermal expansion

The determination of the contraction caused by the decrease of the temperature is essential at low temperature application [2].

We have made measurements on the above mentioned materials. The values of the linear expansion $-\Delta l/l_0$ at the different temperatures of 473 K, 423 K, 373 K, 273 K, 223 K, 173 K, 77 K, are summarized in the Table II.

From the measurements it can be seen that the dilation of the teflon with glass is the least. It can be explained by the small thermal dilation of the glass.

Charpy impact value examination

The Charpy impact value examinations provide us informations how the materials become brittle at different temperatures. This time we have made measurements on two kinds of samples, the Hungarian and GDR teflons, at the temperatures of 373 K, 293 K, 273 K, and 77 K. The results of the measurements are given in the Table III.

It can be seen from the results, that the embrittlement of the teflon is not too high even at lower temperatures.

For the detailed evaluation we need some other measurements too, for instance the study of structure, the measurements of dielectric constant and loss and their frequency dependence, the determination of electric and thermal conductivity.

References

1. D.D. Csegodajev, Ftoroplaszti, Goszhimizdat, Moscow 1960.
2. L. Lászlóffy, I. Kirschner, J. Bánkuti and T. Porjesz, Examination of Contraction on Low Temperature, Industrial Study, Budapest, 1978.

Table I
Results of compressive strength measurements

Sort of Teflon	Bronze + Carbon	Carbon	15% Glass	40% Glass	H-Teflon	GDR-Teflon	
Temp (K) def.(%)	max. compressive force (N)						
373	5	720	480	440	380	270	
	10	1000	700	600	540	440	
	15	1360	890	760	670	610	
	20	1670	1060	910	800	780	
	25	2000	1260	1080	920	970	940
276	5	1250	900	800	700	620	
	10	1750	1150	1020	950	800	
	15	2200	1420	1240	1150	1000	
	20	2650	1660	1460	1320	1350	1200
	25	3150	1940	1720	1550	1600	1400
273	5	1600	1200	1020	1080	1000	1000
	10	2000	1440	1300	1300	1220	1220
	15	2400	1670	1520	1480	1420	1410
	20	2800	1940	1760	1660	1640	1590
	25	3200	2200	2090	1900	1900	1860
77	5	3100	2000	2650	2300	2620	2400
	10	3700	2300	2970	3000	2820	2700
	15	4000	2450	3150	3150	3020	2850
	20	4400	2620	3320	3300	3170	2980
	25	5000	2850	3570	3500	3370	3180

Table II
Results of thermal expansion measurements

Sort of Teflon Temp (K)	Bronze + Carbon	Carbon $10^3 \times \frac{\Delta l}{l_0}$	15% Glass	40% Glass	H-Teflon	GDR-Teflon
473	23.5	24.5	32	22.5	26	-
423	16.5	17.2	15.5	16	18	-
373	10.5	11	10	10	9.5	-
296	0	0	0	0	0	0
273	-4.5	-4	-7	-2	-5.5	-4.5
223	-9	-10	-12	-4	-12	-10.5
173	-10.5	-13	-14.5	-6	-15.5	-14
77	-13.5	-18	-17	-8	-18.5	-16

Table III
Results of Charpy impact value (C.I.V.) measurements

Teflons Temperature (K)	H-Teflon max.C.I.V.(cmkg/cm ²)	GDR-Teflon
373	27.5	>50
293	>50	14.5
273	12	13
77	13.5	7.5

NOTES TO CONTRIBUTORS

I. PAPERS will be considered for publication in *Acta Physica Hungarica* only if they have not previously been published or submitted for publication elsewhere. They may be written in English, French, German or Russian.

Papers should be submitted to

Prof. I. Kovács, Editor
Department of Atomic Physics, Technical University
1521 Budapest, Budafoki út 8, Hungary

Papers may be either articles with abstracts or short communications. Both should be as concise as possible, articles in general not exceeding 25 typed pages, short communications 8 typed pages.

II. MANUSCRIPTS

1. Papers should be submitted in three copies.
2. The text of papers must be of high stylistic standard, requiring minor corrections only.
3. Manuscripts should be typed in double spacing on good quality paper, with generous margins.
4. The name of the author(s) and of the institutes where the work was carried out should appear on the first page of the manuscript.

5. Particular care should be taken with mathematical expressions. The following should be clearly distinguished, e.g. by underlining in different colours: special founts (italics, script, bold type, Greek, Gothic, etc.); capital and small letters; subscripts and superscripts, e.g. x^2 , x_3 ; small l and I ; zero and capital O ; in expressions written by hand: e and l , n and u , v and v , etc.

A List of Symbols on a separate sheet should be attached to each paper.

6. References should be numbered serially and listed at the end of the paper in the following form: J. Ise and W. D. Fretter, *Phys. Rev.*, 76, 933, 1949.

For books, please give the initials and family name of the author(s), title, name of publisher, place and year of publication, e.g.: J. C. Slater, *Quantum Theory of Atomic Structures*, I. McGraw-Hill Book Company, Inc., New York, 1960.

References should be given in the text in the following forms: Heisenberg [5] or [5].

7. Captions to illustrations should be listed on a separate sheet, not inserted in the text.
8. In papers submitted to *Acta Physica* all measures should be expressed in SI units.

III. ILLUSTRATIONS AND TABLES

1. Each paper should be accompanied by three sets of illustrations, one of which must be ready for the blockmaker. The other sets attached to the copies of the manuscript may be rough drawings in pencil or photocopies.
2. Illustrations must not be inserted in the text.
3. All illustrations should be identified in blue pencil by the author's name, abbreviated title of the paper and figure number.
4. Tables should be typed on separate pages and have captions describing their content. Clear wording of column heads is advisable. Tables should be numbered in Roman numerals (I, II, III, etc.).

IV. RETURN OF MATERIAL

Owing to high postage costs, the Editorial Office cannot undertake to return *all* material not accepted for any reason for publication. Of papers to be revised (for not being in conformity with the above Notes or other reasons) only *one* copy will be returned. Material rejected for lack of space or on account of the Referees' opinion will not be returned to authors outside Europe.

Periodicals of the Hungarian Academy of Sciences are obtainable
at the following addresses:

AUSTRALIA

C.B.D. LIBRARY AND SUBSCRIPTION SERVICE
Box 4886, G.P.O., *Sydney N.S.W. 2001*
COSMOS BOOKSHOP, 145 Ackland Street
St. Kilda (Melbourne), Victoria 3182

AUSTRIA

GLOBUS, Höchstädtplatz 3, *1206 Wien XX*

BELGIUM

OFFICE INTERNATIONAL DES PERIODIQUES
Avenue Louise, 485, *1050 Bruxelles*
E. STORY-SCIENTIA P.V.B.A.
P. van Duyseplein 8, *9000 Gent*

BULGARIA

HEMUS, Bulvar Ruszki 6, *Sofia*

CANADA

PANNONIA BOOKS, P.O. Box 1017
Postal Station "B", *Toronto, Ont. M5T 2T8*

CHINA

CNPICOR, Periodical Department, P.O. Box 50
Peking

CZECHOSLOVAKIA

MAD'ARSKA KULTURA, Národní třída 22
115 66 Praha
PNS DOVOZ TISKU, Vinohradská 46, *Praha 2*
PNS DOVOZ TLAČE, *Bratislava 2*

DENMARK

EJNAR MUNKSGAARD, 35, Nørre Segade
1370 Copenhagen K

FEDERAL REPUBLIC OF GERMANY

KUNST UND WISSEN ERICH BIEBER
Postfach 46, *7000 Stuttgart 1*

FINLAND

AKATEEMINEN KIRJAKAUPPA, P.O. Box 128
00101 Helsinki 10

FRANCE

DAWSON-FRANCE S.A., B.P. 40, *91121 Palaiseau*
OFFICE INTERNATIONAL DE DOCUMENTATION ET
LIBRAIRIE, 48 rue Gay-Lussac
75240 Paris, Cedex 05

GERMAN DEMOCRATIC REPUBLIC

HAUS DER UNGARISCHEN KULTUR
Karl Liebknecht-Straße 9, *DDR-102 Berlin*

GREAT BRITAIN

BLACKWELL'S PERIODICALS DIVISION
Hythe Bridge Street, *Oxford OX1 2ET*
BUMPUS, HALDANE AND MAXWELL LTD.
Cowper Works, *Olney, Bucks MK46 4BN*
COLLET'S HOLDINGS LTD., Denington Estate,
Wellingborough, Northants NN8 2QT
WM DAWSON AND SONS LTD., Cannon House
Folkstone, Kent CT19 5EE
H. K. LEWIS AND CO., 136 Gower Street
London WC1E 6BS

GREECE

KOSTARAKIS BROTHERS INTERNATIONAL
BOOKSELLERS, 2 Hippokratous Street, *Athens-143*

HOLLAND

FAXON EUROPE, P.O. Box 167
1000 AD Amsterdam
MARTINUS NIJHOFF B. V.

Lange Voorhout 9-11, *Den Haag*
SWETS SUBSCRIPTION SERVICE
P.O. Box 830, *2160 Sz Lisse*

INDIA

ALLIED PUBLISHING PVT. LTD.
750 Mount Road, *Madras 600002*
CENTRAL NEWS AGENCY PVT. LTD.
Connaught Circus, *New Delhi 110001*
INTERNATIONAL BOOK HOUSE PVT. LTD.
Madame Cama Road, *Bombay 400039*

ITALY

D. E. A., Via Lima 28, *00198 Roma*
INTERSCIENTIA, Via Mazzé 28, *10149 Torino*
LIBRERIA COMMISSIONARIA SANSONI
Via Lamarmora 45, *50121 Firenze*
SANTO VANASIA, Via M. Macchi 58
20124 Milano

JAPAN

KINOKUNIYA COMPANY LTD.
Journal Department, P.O. Box 55
Chitose, *Tokyo 156*
MARUZEN COMPANY LTD., Book Department
P.O. Box 5050 Tokyo International, *Tokyo 100-31*
NAUKA LTD., Import Department
2-30-19 Minami Ikebukuro, *Toshima-ku, Tokyo 187*

KOREA

CHULPANMUL, *Phenjan*

NORWAY

TANUM-TIDSKRIFT-SENTRALEN A.S.
Karl Johansgata 43, *1000 Oslo*

POLAND

WĘGIERSKI INSTYTUT KULTURY
Marszałkowska 80, *00-517 Warszawa*
CKP I W. ul. Towarowa 28, *00-958 Warszawa*

ROMANIA

D. E. P., *Bucuresti*
ILEXIM, Calea Grivitei 64-66, *Bucuresti*

SOVIET UNION

SOYUZPECHAT — IMPORT, *Moscow*
and the post offices in each town
MEZH DUNARODNAYA KNIGA, *Moscow G-200*

SPAIN

DIAZ DE SANTOS Lagasca 95, *Madrid 6*

SWEDEN

ESSELTE TIDSKRIFTSSENTRALEN
Box 62, *101 20 Stockholm*

SWITZERLAND

KARGER LIBRI AG, Petersgraben 31, *4011 Basel*

USA

EBSCO SUBSCRIPTION SERVICES
P.O. Box 1943, *Birmingham, Alabama 35201*
F. W. FAXON COMPANY, INC.
15 Southwest Park, *Westwood Mass. 02090*
MAJOR SCIENTIFIC SUBSCRIPTIONS
1851 Diplomat, P.O. Box 819074,
Pallas, Tx. 75381-9074
READ-MORE PUBLICATIONS, INC.
140 Cedar Street, *New York, N. Y. 10006*

YUGOSLAVIA

JUGOSLOVENSKA KNJIGA, Terazije 27, *Beograd*
FORUM, Vojvode Mišića 1, *21000 Novi Sad*

Acta Physica Hungarica

VOLUME 62, NUMBERS 2-4, 1987

EDITOR-IN-CHIEF

I. KOVÁCS

EDITORIAL BOARD

**Z. BAY, R. GÁSPÁR, I. GYARMATI, N. KÜRTI,
K. NAGY, L. PÁL, A. SZALAY, P. SZÉPFALUSY, I. TARJÁN,
B. TELEGDI, L. TISZA, E. WIGNER**



Akadémiai Kiadó, Budapest

ACTA PHYS. HUNG. APAHAQ 62 (2-4) 101-384 (1987) HU ISSN 0231-4428

ACTA PHYSICA HUNGARICA

A JOURNAL OF THE HUNGARIAN ACADEMY
OF SCIENCES

EDITED BY
I. KOVÁCS

Acta Physica publishes original papers on subjects in physics. Papers are accepted in English, French, German and Russian.

Acta Physica is published in two yearly volumes (4 issues each) by

AKADÉMIAI KIADÓ
Publishing House of the Hungarian Academy of Sciences
H-1054 Budapest, Alkotmány u. 21

Subscription information

Orders should be addressed to

KULTURA Foreign Trading Company
1389 Budapest P.O. Box 149

or to its representatives abroad.

Acta Physica Hungarica is abstracted/indexed in Chemical Abstracts, Current Contents-Physical, Chemical and Earth Sciences, Mathematical Reviews, Science Abstracts, Physics Briefs, Risk Abstracts

© Akadémiai Kiadó, Budapest

CONTENTS

GENERAL PHYSICS

Correlations in a nonequilibrium steady state: Exact results for a generalized kinetic Ising model. <i>Z. Rácz and M. Plischke</i>	203
A symbolic algorithm for finding exactly soluble statistical mechanical models. <i>P. Ruján</i>	287
Unusual maps and their use to approach usual ones. <i>Z. Kaufmann, P. Szépfalussy and T. Tél</i>	321

ELEMENTARY PARTICLES AND FIELDS

Infrared asymptotics of the quark propagator in a Bloch–Nordsieck-type model. <i>G. Pócsik and T. Torma</i>	101
Infrared asymptotics of the quark propagator in nonabelian gauge theories I. <i>G. Pócsik and T. Torma</i>	107
Remarks on the bilocal field theories. <i>Z. Fodor and E. Regős</i>	115
Energy-energy correlations for the $p\bar{p}$ collider. <i>F. Csikor and G. Pócsik</i>	155
Scalar Bethe–Salpeter equation and the relativistic bound state problem. <i>K. Ladányi</i>	161
On a problem of spontaneous compactification. <i>P. Forgács, Z. Horváth and L. Palla</i>	177
The nature of the deconfining phase transition in the SU(2) gauge theory in $1 + \epsilon$ dimension. <i>A. Patkós</i>	185
Expansions at the parameter space boundary in the standard Higgs model. <i>I. Montvay</i>	191
Weak interactions of heavy quarks. <i>H. Pietschmann</i>	211
The possibility of avoiding the axion in the Peccei–Quinn mechanism. <i>T. Nagy</i>	237
Yang–Feldman formalism of dipole fields. <i>K. L. Nagy</i>	257
Canonical quantization of the relativistic theory of the Dirac monopole. <i>A. Frenkel</i>	307

NUCLEAR PHYSICS

Fragment yields in the microcanonical model of nuclear disassembly. <i>G. Fai</i>	277
Nucleon–nucleon interaction with ultra short range attraction. <i>I. Lovas, Erika Lovas and K. Sailer</i>	347
A quantitative analysis of the collective outward motion of nuclear matter in central Ca + Ca and Nb + Nb collisions. <i>L. P. Csernai and G. Fai</i>	355
A three-fluid model of heavy ion collisions. <i>J. Németh</i>	365
The effect of correlations on the entropy and hadrochemical composition in heavy ion reactions. <i>H-W. Barz, T. S. Biró, B. Lukács and J. Zimányi</i>	371

ATOMIC AND MOLECULAR PHYSICS

Spin orbitals and total energy calculated by the $X\alpha$ method including ab initio self-consistent exchange parameters α_{SCF} . <i>R. Gáspár and Á. Nagy</i>	131
--	-----

FLUIDS, PLASMAS AND ELECTRIC DISCHARGES

Hamilton principle for the vortex flow of an ideal fluid in special relativity. <i>I. Abonyi</i>	199
Inversion method to test radial structures in cylindrical plasmas. <i>K. G. Antal, J. F. Bitó and F. Halász</i>	225

CONDENSED MATTER

Zero modes and reparametrization invariance in the Ising spin glass. <i>I. Kondor and R. Németh</i>	219
The glass transition feedback: Review and comments. <i>T. Geszti</i>	267

ASTROPHYSICS

Entropy in the universe. <i>G. Marx</i>	139
Density peaks and large scale velocities. <i>Alexander S. Szalay and Lars Gerhard Jensen</i>	263

INTERDISCIPLINARY

Global-scale changes of the environment: Observations from space. <i>K. Ya. Kondratyev</i>	243
--	-----

The final manuscript was received
by the Editorial Office on 20 February 1986

INFRARED ASYMPTOTICS OF THE QUARK PROPAGATOR IN A BLOCH-NORDSIECK-TYPE MODEL*

G. PÓCSIK

*Institute for Theoretical Physics, Roland Eötvös University
1088 Budapest, Hungary*

and

T. TORMA

*Computing Centre, Roland Eötvös University
1088 Budapest, Hungary*

Assuming a k^{-4} singularity for the infrared behaviour of the gluon propagator in QCD, in a Bloch–Nordsieck-type model it is shown that the quark propagator is an entire function or vanishing in the infrared limit depending on the gauge used. In special cases one can restore the free propagator, too.

1. Introduction

The absolute confinement of quarks can manifest itself in the lack of singularities of the quark propagator in the infrared limit, and a pole or branch point in this limit can go with the existence of free quark states. As for the first possibility, in various gauges and approximations it has been shown that the quark propagator is vanishing on the mass shell (e.g. [1, 2, 3]). In other approaches the quark propagator turns out to be consistent with the free one in the infrared limit [4, 5]. In all of these considerations the infrared behaviour of the gluon propagator was described by a more singular term than the free one, k^{-2} , namely, in several cases a k^{-4} behaviour was used corresponding to a linear confining potential.

Considering the importance of the above problems, in the present paper we calculate the infrared asymptotics of the quark propagator S_F' in QCD in a Bloch–Nordsieck-type model where the Dirac matrices are considered as c -numbers. The infrared behaviour of the electron propagator in QED has been successfully treated in this model in [6]. Our calculations are carried out in a kind of vacuum saturation approximation leading to an exponential form of S_F' in the coordinate space. For the gluon propagator a k^{-4} behaviour is assumed.

* Dedicated to Prof. K. Nagy on his 60th birthday

In axial gauges $n^2 \neq 0$ we get that $S'_F(k)$ is an entire function which is finite on the mass shell. In the timelike axial gauge, $n^2 = 1$, one can still find $S'_F = S_F$ as a special case by choosing n_μ as the constant Dirac matrices. In Feynman and Landau gauges S'_F is again an entire function in $4 + \varepsilon$ dimension going to zero for vanishing ε .

In Section 2 we describe the model considered and our calculations of $S'_F(k)$, while Section 3 contains a discussion of our results.

2. Model calculations for the quark propagator

The model we consider consists of taking the Dirac matrices as c -numbers u_ν , $u^2 = 1$, and we may assume $u_0 > 0$. Thus, quark-antiquark pairs cannot be produced and effects of closed quark loops are vanishing. This follows from the fact that $S_F(x)$ is a retarded function, indeed

$$S_F(x) = \begin{cases} 0, & x_0 < 0, \\ -\frac{i}{u_0} \exp\left(-im \frac{x_0}{u_0}\right) \delta^{(3)}\left(x_j - u_j \frac{x_0}{u_0}\right), & x_0 \geq 0. \end{cases} \quad (1)$$

(1) is the Fourier transform of $(uk - m + i\varepsilon)^{-1}$.

We are working in axial gauges $n^2 \neq 0$ [7] where ghost loops are absent. The quark propagator can be expressed by functional derivatives in the following form

$$S'_F(x-y) = N \left(G \left(x, y \left| \frac{1}{i} \frac{\delta}{\delta J} \right. \right) Z(J) \right)_{J=0}, \quad (2)$$

where

$$Z(J) = \exp \left(i \int d^4x L_I \left(\frac{1}{i} \frac{\delta}{\delta J} \right) \right) \cdot \exp \left(-\frac{i}{2} \int d^4x d^4y J_{\mu a}(x) G_{0ab}^{\mu\nu}(x-y) J_{\nu b}(y) \right), \quad (3)$$

$$N^{-1} = (Z(J))_{J=0}. \quad (4)$$

The Lagrangian L_I contains the self-couplings of gluons, $J_a^\mu(x)$ is an external colour current, $G_{0ab}^{\mu\nu}(x-y)$ is the free gluon propagator in axial gauge and $G(x, y|A)$ is the Green's function of the quark moving in the external gluon field $A_{\mu a}$. Under the above assumptions (2) is exact, which can be verified by solving the Schwinger's equations for the vacuum functional by functional derivatives. The gluon Green's functions are given by the functional derivatives of $Z(J)$ at $J=0$ multiplied by N .

The Green's function $G(x, y|A)$ satisfies the equation

$$\left(iu_\nu \partial^\nu - m + \frac{1}{2} g\lambda_a u_\nu A_a^\nu(x) \right) G(x, y|A) = \delta^{(4)}(x-y). \tag{5}$$

m means the mass parameter, λ_a is the colour matrix.

First, we find a solution to (5). Therefore, Fourier transform $G(x, y|A)$ with respect to $x-y$ leading to a function $G(k, x)$ which can be represented as

$$G(k, x) = -i \int_0^\infty dv U(v, x|A) \exp[-i(m-uk-i\varepsilon)v], \tag{6}$$

and for $U(v, x|A)$ we have

$$\left[i \frac{\partial}{\partial v} + u_\rho \left(i\partial^\rho + \frac{1}{2} g\lambda_a A_a^\rho(x) \right) \right] U(v, x|A) = 0 \tag{7}$$

with $U(v=0, x|A) = 1$. Being the gluon Green's functions translational invariant, we can take $U(v, x|A)$ at $x-uv=0$ on the subspace of the functionals $Z(J)$ in (2). This new quantity $U(v|A)$ obeys the equation

$$\frac{dU(v|A)}{dv} = \frac{1}{2} ig\lambda_a u^\rho A_{a\rho}(uv)U(v|A), \tag{8}$$

$$U(0|A) = 1.$$

The solution of (8) should be taken at $A = i^{-1} \delta/\delta J$ applied to $Z(J)$, and $J=0$ in (2). This yields for the quark propagator in momentum space

$$S'_F(k) = -i \int_0^\infty \langle 0|U(v|A)|0\rangle \exp[-i(m-uk-i\varepsilon)v] dv \tag{9}$$

and the operator $U(v|A)$ satisfies (8) where now $A_{a\rho}(uv)$ is the interacting gluon field operator. We evaluate the vacuum expectation value in (9) in a vacuum saturation approximation, so that terms with only an even number of A 's remain, corresponding to

$$U(v|A) = 1 + \frac{1}{4} (ig)^2 u_\alpha u_\beta \lambda_a \lambda_b \cdot \int_0^v dv_1 \int_0^{v_1} dv_2 A_a^\alpha(uv_1) A_b^\beta(uv_2) U(v_2|A). \tag{10}$$

An approximate solution of (10) is

$$U(v|A) = \exp \left(\frac{1}{4} (ig)^2 u_\alpha u_\beta \lambda_a \lambda_b \cdot \int_0^v dv_1 \int_0^{v_1} dv_2 A_a^\alpha(uv_1) A_b^\beta(uv_2) \right). \tag{11}$$

Terms of order $-g^4$ are neglected in the exponent. Taking (11) in the vacuum saturation approximation leads to

$$\langle 0|U(v|A)|0\rangle = \exp\left(-\frac{1}{8}g^2i(\lambda_a)^2 \int_0^v dv_1 dv_2 u_\alpha G^{\alpha\beta}(u(v_1 - v_2))u_\beta\right) \quad (12)$$

with $G^{\alpha\beta}$ the two-point gluon Green's function, and $(\lambda_a)^2 = 16/3$.

Exactly the result (12) follows by assuming commutable λ_a 's in (8) and replacing the many-point gluon Green's functions by their disconnected pieces. For free gluons (12) is still an approximation but in QED it holds exactly [6].

In the examples considered below $u_\alpha u_\beta G^{\alpha\beta}(u(v_1 - v_2))$ is v -independent and imaginary, thus (12) is of the form $\exp(bv^2)$, $b \leq 0$. For $b=0$ (9) reproduces the free quark propagator, while for $b < 0$ $S'_F(k)$ is an entire function of $m - uk$:

$$S'_F(k) = -\frac{i}{2} \sqrt{-\frac{\pi}{b}} \exp\left(\frac{(m-uk)^2}{4b}\right) \left[1 - \Phi\left(\frac{i(m-uk)}{2\sqrt{-b}}\right)\right], \quad (13)$$

Φ means the Gauss' error integral. On the contrary, the free propagator in (12) leads to an $\exp(b \ln v)$ -type function assuring a branch point for $S'_F(k)$ at $uk = m$ [6].

In what follows we assume the form

$$G_{\alpha\beta}(k) = -\frac{\Omega^2}{k^4} [g_{\alpha\beta} - (kn)^{-1}(k_\alpha n_\beta + n_\alpha k_\beta) + (kn)^{-2} k_\alpha k_\beta n^2], \quad n^2 \neq 0 \quad (14)$$

for the Fourier transform of $G_{\alpha\beta}(x)$ in axial gauge, Ω is a constant. Without the dimensionless term Ω^2/k^2 , (14) is just the free propagator. Ω^2/k^2 [1, 2, 3] modifies the potential into a linear one. From (14)

$$u_\alpha u_\beta G^{\alpha\beta}(u(v_1 - v_2)) = -i \frac{\Omega^2}{8\pi^2} \left(1 - \frac{(un)^2}{n^2}\right), \quad (15)$$

$$b = -\frac{1}{12} \frac{g^2 \Omega^2}{\pi^2} \left(1 - \frac{(un)^2}{n^2}\right). \quad (16)$$

We have calculated (12) also in Feynman and Landau gauges using the confining term Ω^2/k^2 . Here (13) is still valid in the approximation of neglecting ghost loops with

$$b_F = -\frac{1}{12} \frac{g^2 \Omega^2}{\pi^2} \Gamma(\varepsilon), \quad b_L = \frac{3}{4} b_F, \quad (17)$$

where the calculation has been carried out in $n = 4 + \varepsilon$ dimension. For $\varepsilon \rightarrow 0$ $S'_F(k)$ is vanishing.

3. Discussion

In the present paper we have calculated the quark propagator in a model with constant Dirac matrices. We feel that the main features of (8) and (9) will remain valid also in the relativistic case in the infrared limit. The exponential form (12) emerged in a vacuum saturation approximation. In spacelike axial gauges $n^2 < 0$, $b < 0$ and $S'_F(k)$ tends to the constant $-\frac{i}{2}\sqrt{-\frac{\pi}{b}}$ in the infrared limit $uk \rightarrow m$, even if we started with massless quarks, provided we use the confining gluon propagator (14). For timelike axial gauges $n^2 = 1$ this is still true except the choice $u_\alpha = n_\alpha$ when $b = 0$ and S'_F becomes the free propagator irrespective of the factor Ω^2/k^2 . This follows immediately from $u_\alpha u_\beta G^{\alpha\beta}(k)$ in (14). If one adds a term like $(g_{\alpha\beta} - n_\alpha n_\beta/n^2)$ to (14) [5], in general, S'_F is still an entire function except one choice which cancels (16) and $b = 0$ [5]. In Feynman and Landau gauges the approximate $S'_F(k)$ is vanishing in four-dimension.

The appearance of a finite $S'_F(k)$ in the mass shell limit is a nonperturbative phenomenon. Indeed, the dependence of $S'_F(k)$ on the coupling constant g is $\sim g^{-1}$ on the mass shell, and for $m - uk \neq 0$

$$S'_F(k) \sim g^{-1} \exp\left(-\frac{c_1}{g^2}\right) \left[1 - \Phi\left(\frac{ic_2}{g}\right)\right], \tag{18}$$

c_1, c_2 are finite. For $g \rightarrow 0$ (and $g \rightarrow \infty$) $S'_F \rightarrow 0$ in (18).

Finally we remark that (13) leads to the following propagator in coordinate space

$$S'_F(x) = S_F(x) \exp\left(b \frac{x_0^2}{u_0^2}\right). \tag{19}$$

S'_F does not generate any dynamical singularity.

In conclusion we have considered a common approximation to S'_F in various gauges and shown that in most cases the quark propagator does not possess singularities in the infrared limit.

References

1. H. Pagels, Phys. Rev., *D14*, 2747, 1976; *D15*, 2991, 1977.
2. J. S. Ball and F. Zachariasen, Phys. Lett., *106B*, 133, 1981.
3. K. Harada, Progr. Theor. Phys., *68*, 1324 (1982).
4. R. L. Stuller, Phys. Rev., *D13*, 513, 1976.
5. B. A. Arbuzov, Phys. Lett., *125B*, 497, 1983.
6. N. N. Bogoliubov and D. V. Shirkov, Introduction to the Theory of Quantized Fields, Interscience, New York, 1959, p. 503.
7. W. Kummer, Acta Phys. Austriaca, *41*, 315, 1975.

INFRARED ASYMPTOTICS OF THE QUARK PROPAGATOR IN NONABELIAN GAUGE THEORIES I*

G. PÓCSIK

*Institute for Theoretical Physics, Roland Eötvös University
1088 Budapest, Hungary*

and

T. TORMA

*Computing Centre, Roland Eötvös University
1088 Budapest, Hungary*

Assuming a k^{-4} singularity for the infrared behaviour of the gluon propagator, it is shown that the quark propagator is an entire function of p^2 in the infrared region and it has an essential singularity at $g^2=0$ for $p^2 \neq m^2$ in axial gauges. A similar behaviour is shown in a Bloch–Nordseick-type model, too.

1. Introduction

The absolute confinement of quarks can manifest itself in the lack of singularities of the quark propagator in the infrared limit. In various gauges and approximations it has been shown that the quark propagator is vanishing on the mass shell (e.g. [1, 2, 3]), while in other approaches the quark propagator turns out to be consistent with the free one in the infrared limit [4, 5]. In all of these considerations the infrared behaviour of the gluon propagator was described by a more singular term than k^{-2} , namely, in several cases a k^{-4} behaviour was used corresponding to a linear confining potential.

Considering the complexity of the problem whose different approaches have emphasized various remarkable aspects of the confinement we calculate the infrared asymptotics of the unrenormalized quark propagator S_F by functional methods [6]. Our approximation corresponds to a resummation of quark lines with many dressed gluon propagators, both ends of which are attached to the quark line. For the gluon propagator a k^{-4} singularity is assumed. This approximation can be interpreted as a result of an effective bilinear gluon Lagrangian with an inverse propagator of the type k^4 . Because of algebraic complications due to colour pairings, the resummation is carried out for SU (2).

* Dedicated to Prof. K. Nagy on his 60th birthday

In Section 2 we describe our calculations of S'_F . In axial gauges $n^2 \neq 0$, $S'_F(p)$ turns out to be an entire function of p^2 for $0 < p^2 \approx m^2$, $p^2 \neq \frac{(pn)^2}{n^2}$, and for $p^2 \neq m^2$ it has an essential singularity at $g^2 = 0$. This disappears in the mass shell limit. In Feynman and Landau gauges S'_F is again an entire function of p^2 in $4 + \varepsilon$ dimensions and it remains finite and independent of g^2 for $\varepsilon \rightarrow 0$.

In Section 3 it is shown that similar conclusions are valid also in a Bloch-Nordsieck-type model.

Section 4 contains a discussion of the results.

2. Infrared limit of the quark propagator

We are working in axial gauges $n^2 \neq 0$ [7] where ghost loops are absent. The quark propagator can be expressed by functional derivatives in the following form

$$S'_F(x-y) = N \left(G \left(x, y \left| \frac{1}{i} \frac{\delta}{\delta J} \right. \right) Z(J) \right)_{J=0}, \quad (1)$$

where

$$Z(J) = \exp \left(i \int d^4x L_I \left(\frac{1}{i} \frac{\delta}{\delta J} \right) \right) \times \\ \times \exp \left(- \frac{i}{2} \int d^4x d^4y J_{\mu a}(x) G_{0ab}^{\mu\nu}(x-y) J_{\nu b}(y) \right) \quad (2)$$

and

$$N^{-1} = (Z(J))_{J=0}. \quad (3)$$

The Lagrangian L_I contains the self-couplings of gluons, $J_a^\mu(x)$ is an external colour current, $G_{0ab}^{\mu\nu}(x-y)$ is the free gluon propagator in axial gauge, and $G(x, y|A)$ means the Green's function of the quark moving in the external gluon field A . Under the above assumptions (1) is exact which can be verified by solving the Schwinger's equations for the vacuum functional by functional derivatives (see e.g. [6]). The dressed gluon Green's functions are given by the functional derivatives of $Z(J)$ at $J=0$ multiplied by N .

The Green's function $G(x, y|A)$ satisfies the equation

$$\left[i\gamma_\mu \left(\partial_x^\mu - ig \frac{\lambda_a}{2} A_a^\mu(x) \right) - m \right] G(x, y|A) = \delta^{(4)}(x-y), \quad (4)$$

m means the mass parameter, λ_a is the colour matrix. Let us introduce the functional $H(x, y|A)$ by the definition

$$G(x, y|A) = \left[i\gamma_\mu \left(\partial_x^\mu - ig \frac{\lambda_a}{2} A_a^\mu(x) \right) + m \right] H(x, y|A) \tag{5}$$

and from (4) we get

$$\begin{aligned} & (-\partial_x^2 - m^2 + ig \frac{\lambda_a}{2} \gamma_\mu \gamma_\nu \partial_x^\mu A_a^\nu(x) + ig \lambda_a A_{a\mu}(x) \partial_x^\mu + \\ & + \frac{g^2}{4} (\gamma_\mu \lambda_a A_a^\mu(x))^2) H(x, y|A) = \delta^{(4)}(x-y). \end{aligned} \tag{6}$$

The Fourier transform of $H(x, y|A)$, $\tilde{H}(p, q|A)$, determines the quark propagator in momentum space as follows

$$(2\pi)^4 S'_F(p) \delta^{(4)}(p-q) = S^I + S^{II},$$

$$S^I = (\gamma_\mu p^\mu + m) (\tilde{H}(p, q|A) NZ(J))_{J=0}, \tag{7}$$

$$S^{II} = \frac{g}{2(2\pi)^4} \gamma_\mu \lambda_a \left(\int d^4 q' \tilde{A}_a^\mu(q') \tilde{H}(p-q', q|A) NZ(J) \right)_{J=0};$$

here $\tilde{A}_a^\mu(q')$ is the Fourier transform of $A_a^\mu(x) \rightarrow \frac{1}{i} \frac{\delta}{\delta J_{a\mu}(x)}$ and \tilde{H} satisfies the equation

$$\begin{aligned} & (p^2 - m^2) \tilde{H}(p, q|A) + \frac{1}{(2\pi)^4} \int d^4 k [g \lambda_a p^\mu \tilde{A}_{a\mu}(k) - \\ & - g \frac{\lambda_a}{2} (g_{\mu\nu} + i\sigma_{\mu\nu}) k^\mu \tilde{A}_a^\nu(k) + \frac{g^2}{4(2\pi)^4} \gamma_\mu \gamma_\nu \lambda_a \lambda_b \times \\ & \times \int d^4 k' \tilde{A}_a^\mu(k') \tilde{A}_b^\nu(k-k')] \tilde{H}(p-k, q|A) = (2\pi)^4 \delta^{(4)}(p-q). \end{aligned} \tag{8}$$

Following [6] we represent $\tilde{H}(p, q|A)$ as the integral

$$\tilde{H}(p, q|A) = -i \int_0^\infty dv U(p, q; v|A) \exp(iv(p^2 - m^2 + i\epsilon)), \tag{9}$$

where the new functional $U(p, q; v|A)$ obeys the normalisation

$$U(p, q; 0|A) = (2\pi)^4 \delta^{(4)}(p-q). \tag{10}$$

This can be shown by taking the Fourier transform of (6) in $x-y$ and writing the corresponding Fourier transform of H as the integral of the exponential function of its inverse with respect to v .

Substituting (9) into (8), using (10), leads to the definition equation of $U(p, q; v|A)$

$$\begin{aligned} \frac{d}{dv} U(p, q; v|A) - \frac{i}{(2\pi)^4} \int d^4k [g\lambda_a p^\mu \tilde{A}_{a\mu}(k) - \\ - g \frac{\lambda_a}{2} k^\mu (g_{\mu\nu} + i\sigma_{\mu\nu}) \tilde{A}_a^\nu(k) + \frac{g^2}{4(2\pi)^4} \gamma_\mu \gamma_\nu \lambda_a \lambda_b \cdot \\ \cdot \int d^4k' \tilde{A}_a^\mu(k') \tilde{A}_b^\nu(k-k')] \exp(iv(k^2 - 2kp)) U(p-k, q; v|A) = 0. \end{aligned} \quad (11)$$

We need the functional U for p^2 's around m^2 . Since $\tilde{A}_{a\mu}(k)$ and the brackets in the integral act as gluon operators in the gluon Green's functions in $S^{I,II}$, we can approximate the integrand in (11) by its value at $k \rightarrow 0$. Actually, assuming $\tilde{A}_{a\mu}(k) \sim k^{-l}$, $l < 4$ for $k \rightarrow 0$, the most singular term is $p\tilde{A}$ at $k \ll p \approx m$ in the bracket. Finally, we replace (11) in the infrared limit as follows

$$i \frac{d}{dv} U(p, q; v|A) + g\lambda_a p^\mu A_{a\mu}(2pv) U(p, q; v|A) = 0. \quad (12)$$

(12) has the usual time ordered operator solution which, making use of (7), (9), (10) yields

$$\begin{aligned} S^I = -(2\pi)^4 i \delta^{(4)}(p-q) (\gamma_\mu p^\mu + m) \int_0^\infty dv \exp(iv(p^2 - m^2 + i\varepsilon)) \times \\ \times \left[1 + \sum_{n=1}^\infty (ig)^n \prod_{i=1}^n \lambda_{a_i} p^{\mu_i} \int_0^v dv_1 \dots \right. \\ \left. \dots \int_0^{v_{n-1}} dv_n \langle TA_{a_1\mu_1}(2pv_1) \dots A_{a_n\mu_n}(2pv_n) \rangle_0 \right], \end{aligned} \quad (13)$$

where $A_{a\mu}(2pv)$ is the interacting gluon field.

For S^{II} one obtains in the above approximation

$$\begin{aligned} S^{II} = -\frac{ig}{2} \int_0^\infty dv \exp(iv(p^2 - m^2 + i\varepsilon)) \int d^4x \exp(i(p-q)x) \cdot \\ \cdot \gamma^\mu \lambda_a [\langle A_a^\mu(x) \rangle_0 + \sum_{n=1}^\infty (ig)^n \prod_{i=1}^n \lambda_{a_i} p^{\mu_i} \int_0^v dv_1 \dots \\ \dots \int_0^{v_{n-1}} dv_n \langle TA_{a\mu}(x) A_{a_1\mu_1}(2pv_1) \dots A_{a_n\mu_n}(2pv_n) \rangle_0]. \end{aligned} \quad (14)$$

Now, we calculate (13) and (14) in such an approximation where gluons starting from the quark line are absorbed by the same line, corresponding to keeping the propagators in the dressed gluon Green's functions in (13), (14). For the gluon propagator a k^{-4} singularity is assumed. The same approximation follows by using an

effective bilinear gluon Lagrangian with an inverse propagator of the type k^4 . In this way we get

$$\begin{aligned} \frac{1}{(2\pi)^4} \int d^4q S^I &= -i(\gamma_\mu p^\mu + m) \int_0^\infty dv \exp(iv(p^2 - m^2 + i\varepsilon)) \cdot \\ &\cdot \left\{ 1 + \sum_{n=1}^\infty \frac{1}{(2n)!} (-ig^2 v^2 p_\alpha p_\beta G^{\alpha\beta}(2p(v_1 - v_2)))^n \cdot \right. \\ &\cdot \left. \left[\prod_{i=1}^{2n} \lambda_{a_i} \sum_{\text{d.p.}} \delta_{a_1 a_2} \cdots \delta_{a_{2n-1} a_{2n}} \right] \right\}, \end{aligned} \tag{15}$$

where $G^{\alpha\beta}$ denotes the dressed gluon propagator without the colour $-\delta_{ab}$ and $\sum_{\text{d.p.}}$ goes over all the distinct $a_i a_j$ -pairings. We have employed that in the cases considered below $p_\alpha p_\beta G^{\alpha\beta}$ is a v -independent quantity,

$$p_\alpha p_\beta G^{\alpha\beta}(2p(v_1 - v_2)) = i \frac{2b}{g^2}, \quad b \leq 0. \tag{16}$$

In QCD the summation $\left[\sum_{\text{d.p.}} \right]$ is extremely complicated, therefore, we confine ourselves to the SU (2) gauge group for which $\left[\sum_{\text{d.p.}} \right]$ is given by $(2n + 1)! (n! 2^n)^{-1}$.

Hence

$$\frac{1}{(2\pi)^4} \int d^4q S^I = -i(\gamma_\mu p^\mu + m) \int_0^\infty dv (1 + 2bv^2) \exp(iv(p^2 - m^2 + i\varepsilon) + bv^2). \tag{17}$$

This type of exponential v -dependence has been shown in an Abelian gauge model with a dipole gluon field, too [8].

In order to fix b in (17), assume for the Fourier transform of the gluon propagator in axial gauges $n^2 \neq 0$ in d dimensions [5]:

$$\begin{aligned} \tilde{G}_{\alpha\beta}(k) &= -\frac{\Omega^2}{k^4} \left[g_{\alpha\beta} - \frac{k_\alpha n_\beta + k_\beta n_\alpha}{kn} + \frac{k_\alpha k_\beta}{(kn)^2} n^2 + \right. \\ &\left. + (1 + \delta)(4 - d) \left(g_{\alpha\beta} - \frac{n_\alpha n_\beta}{n^2} \right) \right], \end{aligned} \tag{18}$$

with Ω a constant and δ is a parameter. The choice $\delta = 0$ is used in [5], $\delta = -1$ corresponds to the usual axial gauge. From (18) one gets in four dimensions [5]

$$b = \frac{\delta \Omega^2 g^2}{16\pi^2} \left(p^2 - \frac{(pn)^2}{n^2} \right). \tag{19}$$

For first orientation we consider b in Feynman and Landau gauges too, where (17) remains valid in the approximation of neglecting ghost loops. Making use of the term Ω^2/k^4 , we find in $4 + \varepsilon$ dimensions

$$b_F = -\frac{\Omega^2 g^2}{16\pi^2} \Gamma(\varepsilon) p^2, \quad b_L = -\frac{3\Omega^2 g^2}{64\pi^2} \Gamma(\varepsilon) p^2. \quad (20)$$

In the above infrared approximation S^{II} can also be calculated with the result

$$\frac{1}{(2\pi)^4} \int d^4 q S^{II} = \frac{ig^2}{2} \gamma^\alpha p^\beta G_{\alpha\beta}(2pv) \int_0^\infty dv (3 + 2bv^2) v \exp(iv(p^2 - m^2 + i\varepsilon) + bv^2). \quad (21)$$

Here

$$\frac{g^2}{2} \gamma^\alpha p^\beta G_{\alpha\beta}(2pv) = ib(\gamma p - (\gamma n)(pn)n^{-2}) \cdot (p^2 - (pn)^2 n^{-2})^{-1} \quad (22)$$

for axial gauges (18), and

$$\frac{g^2}{2} \gamma^\alpha p^\beta G_{\alpha\beta}(2pv) = ib_{F,L} \frac{\gamma p}{p^2} \quad (23)$$

for Feynman and Landau gauges.

For $b=0$ (e.g. $\delta=0$ in (19)) (21) is vanishing and (17) reproduces the free quark propagator [5]. For $b>0$ (17) and (21) are nonexistent, but for $b<0$ one can carry out the integrations [9] and obtains

$$\frac{1}{(2\pi)^4} \int d^4 q S^I = -(\gamma p + m) \left[\frac{i\pi^{1/2} a^2}{4(-b)^{3/2}} \cdot \exp\left(\frac{a^2}{4b}\right) \cdot \left(1 - \Phi\left(\frac{ia}{2(-b)^{1/2}}\right)\right) + \frac{a}{2b} \right], \quad (24)$$

$$\frac{1}{(2\pi)^4} \int d^4 q S^{II} = \frac{g^2}{2} \gamma^\alpha p^\beta G_{\alpha\beta}(2pv) \left[\frac{a^3}{8} \left(\frac{\pi}{(-b)^5}\right)^{1/2} \cdot \exp\left(\frac{a^2}{4b}\right) \left(1 - \Phi\left(\frac{ia}{2(-b)^{1/2}}\right)\right) + \frac{i}{2b} \left(\frac{a^2}{2b} - 1\right) \right], \quad (25)$$

with $a = m^2 - p^2$, Φ means the Gauss' error integral. The sum of (24) and (25) provides $S'_F(p)$ in the infrared limit, $p^2 \approx m^2$.

In spacelike axial gauges $n^2 = -1$, for $\delta < 0$ and $p^2 > 0$ around m^2 , the quark propagator S'_F is finite; at $p^2 \rightarrow m^2$, (24) tends to zero, (25) remains nonsingular. Thus S'_F is an entire function of p^2 in the infrared region. This conclusion holds also in timelike axial gauges $n^2 = 1$, $\delta > 0$, $p \neq (pn)^2$. In the mass shell limit S'_F is independent of $g^2 \delta \Omega^2 \neq 0$.

In Feynman and Landau gauges (in the approximation without ghost loops) the above conclusions are still valid in $4 + \epsilon$ dimensions and no g -dependence survives the limit $a \rightarrow 0$. In four dimensions S'_F is an entire function of p^2 ((24) vanishes and (25) is $\gamma p/2p^2$) which is completely independent of g^2 .

3. The Bloch—Nordsieck model

In what follows we would like to show that the result (24) can be derived in a Bloch—Nordsieck-type model too [11].

As is known [10] this model corresponds to taking the Dirac matrices as c -numbers u_ν , $u^2 = 1$, and we may assume $u_0 > 0$. It follows that contributions of closed quark loops are vanishing.

First, find a solution to (4). Fourier transform $G(x, y|A)$ with respect to $x - y$ leading to a function $G(p, x)$ and write

$$G(p, x) = -i \int_0^\infty dv U(v, x|A) \exp(-iv(m - up - i\epsilon)). \tag{26}$$

For $U(v, x|A)$ we obtain

$$\left[i \frac{\partial}{\partial v} + u_\rho \left(i \partial^\rho + g \frac{\lambda_a}{2} A_a^\rho(x) \right) \right] U(v, x|A) = 0. \tag{27}$$

It is easy to see that $U(0, x|A) = 1$. Being the gluon Green's functions translation invariant, we can take $U(v, x|A)$ at $x - uv = 0$ on the subspace of the functionals in (1). This new quantity $U(v|A)$ obeys the equation

$$\begin{aligned} \frac{d}{dv} U(v|A) &= ig \frac{\lambda_a}{2} u^\rho A_{a\rho}(uv) U(v|A), \\ U(0|A) &= 1. \end{aligned} \tag{28}$$

The solution of (28) should be taken at $A = \frac{1}{i} \frac{\delta}{\delta J}$ applied to $Z(J)$ and $J = 0$ in (1). This yields

$$S'_F(p) = -i \int_0^\infty dv \langle 0|U(v|A)|0\rangle \exp(-iv(m - up - i\epsilon)) \tag{29}$$

for the quark propagator in momentum space where the operator $U(v|A)$ satisfies (28). Here the vacuum expectation value agrees with the bracket [] in (13) with p_{μ_i} replaced by $u_{\mu_i}/2$ and $2pv \rightarrow uv$. Applying the same treatment as in Section 2, for the SU(2) gauge group we obtain (24) as $S'_F(p)$ without $\gamma p + m$, where $a = m - up$ and instead of (19), (20):

$$b = \frac{\delta \Omega^2 g^2}{12\pi^2} \left(1 - \frac{(un)^2}{n^2} \right),$$

$$b_F = -\frac{g^2 \Omega^2}{12\pi^2} \Gamma(\varepsilon), \quad b_L = \frac{3}{4} b_F. \quad (30)$$

Therefore, $S'_F(p)$ in the infrared limit, $up \approx m$, has identical features with those of (24) in the region $p^2 \approx m^2$. In particular, $S'_F(p) \rightarrow 0$ for $up \rightarrow m$, $g \neq 0$.

In an Abelian gauge model with a dipole gluon field the gauge invariant fermion fields have zero asymptotic limits, as it has been in [8].

4. Discussion

In the present paper we have calculated the quark propagator in the infrared limit by functional methods for a confining gluon propagator. The results (24), (25) emerged in a vacuum saturation approximation where due to algebraic complications an SU (2) gauge group was assumed.

As we have seen, in axial gauges $S'_F(p)$ is an entire function of p^2 in the infrared region $p^2 \approx m^2$, ($\delta, n^2 \neq 0$). This seems to be true also in four-dimensional Feynman and Landau gauges. In order to reach the above conclusion in timelike axial gauges $n^2 = 1$, one must assume $p^2 \neq (pn)^2$ otherwise unphysical singularities would appear in (24), (25). Likewise (17) and (21) do not exist for $b > 0$.

Under the ν -integrals in (17) and (21) exclusively analytic expressions of g^2 take place. After the ν -integrations this is no longer true, $S'_F(p)$ cannot be expanded around $g^2 = 0$. In axial gauges $n^2 \neq 0$ for $g^2 \rightarrow 0$ ($b \rightarrow -0$) and $a \neq 0$, $S'_F(p)$ is finite. In Feynman and Landau gauges S'_F is finite in $4 + \varepsilon$ dimensions. In four dimensions S'_F is independent of g^2 .

References

1. H. Pagels, Phys. Rev., *D14*, 2747, 1976; *D15*, 2991, 1977.
2. J. S. Ball and F. Zachariasen, Phys. Lett., *106B*, 133, 1981.
3. K. Harada, Progr. Theor. Phys., *68*, 1324, 1982.
4. R. L. Stuller, Phys. Rev., *D13*, 513, 1976.
5. B. A. Arbuzov, Phys. Lett., *125B*, 497, 1983.
6. E. S. Fradkin, Nucl. Phys., *76*, 588, 1966.
7. W. Kummer, Acta Phys. Austriaca, *41*, 315, 1975.
8. E. d'Emilio and M. Mintchev, Phys. Rev., *D27*, 1840, 1983.
9. I. S. Gradshteyn and I. M. Ryzhik, Tables of Integrals, Sums, Series and Products (Phys.-Mat. Lit., Moscow, 1963).
10. N. N. Bogoliubov and D. V. Shirkov, Introduction to the Theory of Quantized Fields, Interscience, New York, 1959.
11. A preliminary version of the calculation can be seen in G. Pócsik and T. Torma: Former paper in this volume.

REMARKS ON THE BILOCAL FIELD THEORIES*

Z. FODOR and E. REGŐS

*Institute for Theoretical Physics, Roland Eötvös University
1088 Budapest, Hungary*

Bilocal field theories of spin-0 and 1 mesons as permanently bound quark systems are studied. Internal consistency of the field equations frequently imposes strong restrictions on the potentials. An anharmonic oscillator model is described leading to a hadron mass spectrum that bends back from the linear one. In the framework of bilocal field theories we work out the classical field theory of baryons consisting of three quarks, and differential as well as integral conservation laws are derived.

1. Introduction

Bilocal field theories (e.g. [1]) are space-time models and they have been constructed for describing hadrons as permanently bound quark systems. In these theories the subsidiary conditions play an important role since they fix the dependence on the relative time in such a way that oscillations (as in the Bethe-Salpeter approach) are absent. In one of these models [2] the subsidiary condition allowed a harmonic oscillator potential leading to a linear hadron mass spectrum. The classical and quantised conservation laws for hadrons were studied in [3] and [4].

In this work we look for certain classes of potentials which are allowed in bilocal field theories by the field equations and the subsidiary condition. This problem was partly discussed in [5] for three-quark systems of baryons.

After describing the problem for spin-0 mesons (Section 2), we study the influence of the condition of compatibility on the potential and corresponding differential equations are derived in Section 3. An anharmonic oscillator model is constructed. Its mass spectrum deviates from the linear one, so that the hadrons lie on Regge trajectories bending back from the linear one. This is allowed by quantum mechanics. In Section 4 the bilocal theory of spin-1 vector-mesons is described in the Lorentz gauge. Assuming the symmetry of the x and r -spaces we get such models which do not give quantised mass spectrum for the vector-meson, otherwise quite sensible.

In the last part of the work (Section 5) we derive the classical conservation laws for the three-quark systems of baryons and show how to construct conserved quantities for the baryon as a whole using field quantities depending on the three quarks.

* Dedicated to Prof. K. Nagy on his 60th birthday

2. The compatibility of the field equations

A spin-0 meson is described by a $\psi(x_1, x_2)$ classical scalar function (field quantity), x_1 and x_2 are the coordinates of the two pointlike constituents. As in usual field theories, $\psi(x_1, x_2)$ is determined by the equations of motion

$$\begin{aligned}(\partial_{1\mu}^2 - V(x_1, x_2))\psi(x_1, x_2) &= 0, \\(\partial_{2\mu}^2 - V(x_2, x_1))\psi(x_1, x_2) &= 0.\end{aligned}\tag{1}$$

The quark 1 is moved by the effective potential $V(x_1, x_2)$, etc. V is an invariant scalar which is a real function of x_i and $\frac{\partial}{\partial x_i}$. To decompose the centre-of-mass and relative motions, we define $x_\mu = \frac{1}{2}(x_{1\mu} + x_{2\mu})$, $r_\mu = x_{1\mu} - x_{2\mu}$. If p_μ is the four momentum of the centre-of-mass, a free meson is represented as follows

$$\psi(x_1, x_2) = e^{-ipx} \cdot \psi(p, r),\tag{2}$$

where $\psi(p, r)$ is responsible for the internal motions. Substituting into (1) gives

$$\left(\partial_r^2 - \frac{1}{4}p^2 - \text{Re } V(r, \partial_r, ip)\right)\psi(p, r) = 0,\tag{3}$$

$$(p\partial_r + \text{Im } V(r, \partial_r, ip))\psi(p, r) = 0.\tag{4}$$

Here $p^2 = M^2$, M means the meson mass. A useful necessary condition for the compatibility of these equations is

$$([\partial_r^2, \text{Im } V] - [\text{Re } V, p\partial_r + \text{Im } V])\psi(p, r) = 0,\tag{5}$$

where $\partial_r = \frac{\partial}{\partial r_\mu}$. Now we answer the question if (5) is sufficient for the compatibility of the equations (3), (4). Let $\{\psi_M(p, r)\}$ be the solution of (4). Using (4) we have

$$\left(\partial_r^2 - \frac{1}{4}p^2 - \text{Re } V\right)(p\partial_r + \text{Im } V)\psi(p, r) = 0,\tag{6}$$

substituting into (5) gives

$$(p\partial_r + \text{Im } V)\left(\partial_r^2 - \frac{1}{4}p^2 - \text{Re } V\right)\psi(p, r) = 0.\tag{7}$$

If $\psi_0 \in \{\psi_M(p, r)\}$ and (5) is satisfied, one has only

$$\left(\partial_r^2 - \frac{1}{4}p^2 - \text{Re } V\right)\psi_0(p, r) \in \{\psi_M(p, r)\},$$

so, the condition (5) is necessary but not sufficient.

It is easy to provide an example when (5) is not sufficient. Write

$$\operatorname{Im} V = \operatorname{Re} V = 0, \quad (8)$$

then (5) is trivially satisfied. From the relativistic invariance

$$\psi(p, r) = \psi(pr, r^2, p^2). \quad (9)$$

Using (4)

$$p_\mu \frac{\partial}{\partial r_\mu} \psi(pr, r^2, p^2) = p^2 \frac{\partial \psi}{\partial (pr)} + 2pr \frac{\partial \psi}{\partial (r^2)} = 0 \quad (10)$$

and solving (10)

$$\psi = g((pr)^2 - p^2 r^2). \quad (11)$$

Put this form into (3):

$$\left(\partial_r^2 - \frac{1}{4} p^2 \right) g = 2p^2 [2g'' - 3g'], \quad (12)$$

here the comma means the differentiation with respect to the argument. It seems that this is an element of $\{\psi_M(p, r)\}$ but it is 0 only if $g(u) = -2u^{-1/2}$.

Furthermore, employing the relativistic invariance of $\psi(p, r)$, it will be useful to write the equations (3), (4) in the more explicit form

$$\begin{aligned} & \frac{\partial^2 \psi}{\partial (r^2)^2} \cdot 4 \left(r^2 - \frac{(pr)^2}{p^2} \right) + 2 \left(2 \operatorname{Im} V \cdot \frac{pr}{p^2} - 2 \frac{pr}{p^2} \operatorname{Im} V - 1 \right) \frac{\partial \psi}{\partial (r^2)} - \\ & - \left[\frac{2pr}{p^2} \cdot \frac{\partial \operatorname{Im} V}{\partial (r^2)} + \frac{\partial \operatorname{Im} V}{\partial (pr)} - \frac{(\operatorname{Im} V)^2}{p^2} + \frac{1}{4} p^2 + \operatorname{Re} V \right] \psi = 0, \end{aligned} \quad (13)$$

$$\frac{\partial \psi}{\partial (pr)} = - \frac{2pr}{p^2} \cdot \frac{\partial \psi}{\partial (r^2)} - \frac{\operatorname{Im} V \cdot \psi}{p^2}. \quad (14)$$

3. Discussion of the compatibility

The case of $\operatorname{Re} V \neq 0, \operatorname{Im} V = 0$

Let $\operatorname{Re} V$ be the most general function consisting of quadratic invariant quantities:

$$\operatorname{Re} V = V(r^2, r\partial_r, (pr)^2, \partial_r^2, (p\partial_r)^2), \quad (15)$$

$\operatorname{Re} V$ contains $r\partial_r, \partial_r^2, (p\partial_r)^2$ only as a finite polynomial. Using (5)

$$\operatorname{Re} V p \partial_r \psi - p \partial_r (\operatorname{Re} V \psi) = 0, \quad (16)$$

hence

$$p_\mu \left(\frac{\partial}{\partial r_\mu} \operatorname{Re} V \right) \psi(p, r) = 0. \quad (17)$$

Differentiating the function (15) and substituting into (17) we get

$$\frac{\partial \operatorname{Re} V}{\partial(r^2)} \cdot pr\psi + p^2 \cdot \frac{\partial \operatorname{Re} V}{\partial(pr)^2} \cdot pr\psi = 0. \quad (18)$$

(Hence $\operatorname{Im} V = 0$, $p\partial_r\psi = 0$.) The only solution of this differential equation is

$$\operatorname{Re} V = f(p^2r^2 - (pr)^2) \cdot g(r\partial_r, \partial_r^2, (p\partial_r)^2), \quad (19)$$

where g is a polynomial.

A special case of this is solved in [6]. Now we solve the equations (3), (4) for $g = 1$ in (19) and $\operatorname{Im} V = 0$, that is

$$V = f\left(r^2 - \frac{(pr)^2}{p^2}\right). \quad (20)$$

In the centre-of-mass frame $p = (p_0, 0, 0, 0)$ and ψ is independent of r_0 ,

$$r^2 - \frac{(pr)^2}{p^2} = -\mathbf{r}^2, \quad (21)$$

where $\mathbf{r} = (r_1, r_2, r_3)$ a three-vector. Furthermore, (3) has the form

$$\left[\Delta_{\mathbf{r}} + \frac{1}{4} M^2 - f(\mathbf{r}^2) \right] \psi = 0. \quad (22)$$

Separate the internal motion into angle and distance variables

$$\psi = Y_l^m(\vartheta, \varphi) \cdot \frac{\chi(r)}{r}, \quad (23)$$

then

$$\frac{d^2\chi}{dr^2} + \left(\frac{M^2}{4} - f(r^2) - \frac{l(l+1)}{r^2} \right) \chi = 0. \quad (24)$$

In general, we are interested in the regular solutions of this equation since $\psi(p, r)$, describing the motions in the hadron, decreases quickly with increasing r . (The quarks are permanently bound.) Here we consider the model

$$f(r^2) = \sigma^2 r^2 + ar^4, \quad (25)$$

where

$$\frac{a}{\sigma^3} \ll 1. \quad (26)$$

This model is a perturbed harmonic oscillator.

The eigenfunctions of the normal harmonic oscillator are known from the quantum mechanics:

$$\psi_{n,l,m}^{(0)} = A \cdot r^l \cdot e^{-\frac{\sigma r^2}{2}} \cdot Y_l^m(\vartheta, \varphi) \cdot F\left(-\frac{n-l}{2}, l + \frac{3}{2}, \sigma r^2\right), \quad (27)$$

where F is a degenerated hypergeometric function and the mass spectrum is

$$M_n^{2(0)} = 8\sigma \left(n + \frac{3}{2} \right); \quad n=0, 1, 2, \dots \quad (28)$$

and

$$|m|=0, 1, \dots, l, \quad l = \begin{cases} 0, 2, \dots, n \\ 1, 3, \dots, n \end{cases} \frac{n-l}{2} = s \quad \text{is } 0, 1, 2, \dots \quad (29)$$

The mass spectrum of the unperturbed oscillator gets a correction which is in first order perturbation calculation

$$M^{2(1)} = 4 \int \psi_{n,l,m}^{(0)*} \cdot (ar^4) \cdot \psi_{n,l,m}^{(0)} \cdot r^2 dr \sin \vartheta d\vartheta d\varphi. \quad (30)$$

Calculating this from the integrals of the hypergeometric functions, we arrive at the correction

$$M_{n,l}^{2(1)} = \frac{2a}{\sigma^2} (3n^2 + 9n - l^2 - l + 7.5). \quad (31)$$

The degeneration of the spectrum disappeared. For a fixed n the internal angular momentum causes a fine structure in the mass spectrum. The relative correction to the unperturbed spectrum is as follows

$$\frac{M_{n,l}^{2(1)}}{M_n^{2(0)}} = \frac{a(3n^2 + 9n - l^2 - l + 7.5)}{4\sigma^3 \left(n + \frac{3}{2} \right)}, \quad (32)$$

this correction is real for $n \gg \frac{\sigma^3}{3a}$, (see Fig. 1). For a fixed n the mass spectrum has a

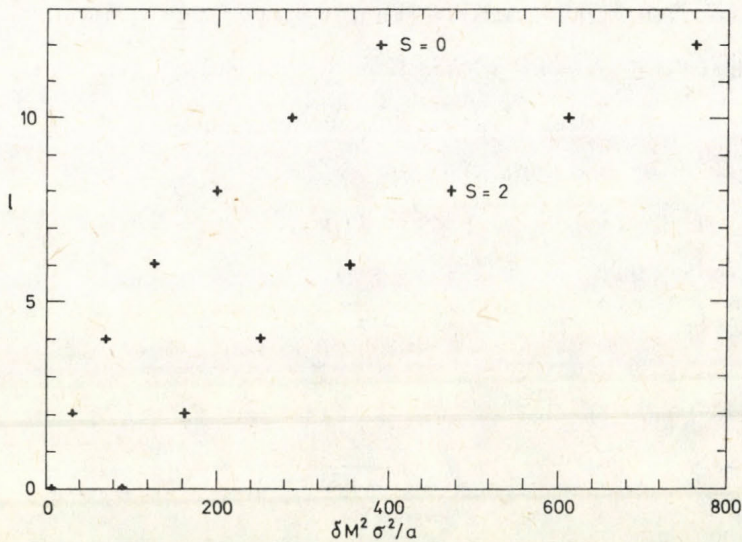


Fig. 1. Mass spectrum of an anharmonic oscillator vs l at $s = \frac{n-1}{2} = 0,2$

structure of a Regge trajectory bending back from the linear one. For a fixed n there always exists a maximal angular momentum.

The case of $\text{Re } V = \text{constant}$, $\text{Im } V \neq 0$

The most general form of $\text{Im } V$ is

$$\begin{aligned} \text{Im } V = & p\partial_r V^{(1)}(r^2, r\partial_r, (pr)^2, \partial_r^2, (p\partial_r)^2) + \\ & + pr \cdot V^{(2)}(r^2, r\partial_r, (pr)^2, \partial_r^2, (p\partial_r)^2). \end{aligned} \quad (33)$$

Using (5) one can write

$$(\partial_\mu \partial_\mu \text{Im } V + 2\partial_\mu \text{Im } V \partial_\mu) \psi = 0. \quad (34)$$

This leads to the following partial differential equation ($x = r^2$, $y = (pr)^2$, $z = r\partial_r$):

$$\begin{aligned} \{ & (p\partial_r) [2V_x^{(1)} + 4V_{xx}^{(1)}x + 4V_{yy}^{(1)}yp^2 + V_{zz}^{(1)}\partial_r^2 + 4V_{xy}^{(1)}y + 2V_{xz}^{(1)}(z+4) + \\ & + 2V_{yz}^{(1)}(p^2 + prp\partial_r)] + pr[2V_x^{(2)} + 4V_{xx}^{(2)}x + 4V_{yy}^{(2)}yp^2 + \\ & + V_{zz}^{(2)}\partial_r^2 + 4V_{xy}^{(2)}y + 2V_{xz}^{(2)}(z+4) + 2V_{yz}^{(2)}(p^2 + pr + p\partial_r)] + \\ & + 2V_x^{(2)}pr + 2V_y^{(2)}prp^2 + V_z^{(2)}p\partial_r + 2p\partial_r V_x^{(1)}z + 2p\partial_r V_y^{(1)} \cdot \\ & \cdot prp\partial_r + p\partial_r V_z^{(1)}\partial_r^2 + V^{(2)}p\partial_r + pr[2V_x^{(2)}r\partial_r + 2V_y^{(2)}pr \cdot \\ & \cdot p\partial_r + V_z^{(2)}\partial_r^2] \} \psi = 0. \end{aligned} \quad (35)$$

In the special case of $V^{(1)} = V^{(1)}(r^2)$ and $V^{(2)} = V^{(2)}(r^2)$, clearly

$$\text{Im } V = pr f(r^2) + g(r^2) p\partial_r. \quad (36)$$

Substituting into the condition (34) we get

$$6f'pr + 2f''r^2pr + 4g'p\partial_r + 2r^2g'' + 2f'prr\partial_r + fp\partial_r + 2g'r\partial_r p\partial_r = 0. \quad (37)$$

From (4) it follows that

$$p\partial_r \psi = -\frac{f}{g+1} pr \psi. \quad (38)$$

Comparing (37) and (38):

$$f' = g' \frac{f}{g+1} \quad (39)$$

which has the solution

$$p\partial_r \psi = -cpr \psi. \quad (40)$$

Making use of (40) in (37) and (38) we have as the only solution

$$f = 0, \quad (41)$$

or in other words $c = 0$.

The case of $\text{Re } V = \text{constant}$, $\text{Im } V = 0$

If we suppose (40) for f and g , the parts $r\hat{\partial}_r$ vanish in $[\partial_r^2, \text{Im } V]$, and

$$[\partial_r^2, \text{Im } V] = -2c^2(g+1)pr. \quad (42)$$

Substituting into (5) and using (42) we obtain

$$\frac{\partial \text{Re } V}{\partial r^2} + p^2 \frac{\partial \text{Re } V}{\partial (pr)^2} = -2c^2. \quad (43)$$

Solving this partial differential equation yields

$$\text{Re } V = -c^2 \left(r^2 + \frac{(pr)^2}{p^2} \right) + h \left(r^2 - \frac{(pr)^2}{p^2} \right), \quad (44)$$

where h is an arbitrary function, and c is determined by (40).

4. Bilocal field theory of spin-1 vector mesons

We would like to extend the bilocal description of mesons with 0 spin (consisting of two quarks) to 1 spin. In this case the vector potential has 4 components, $A_\nu(1, 2)$; $\nu = 0, 1, 2, 3$. The field equations are assumed in the form

$$[\partial_{1\mu}^2 - V(1, 2)]A_\nu(1, 2) = 0, \quad (45)$$

$$[\partial_{2\mu}^2 - V(2, 1)]A_\nu(1, 2) = 0, \quad (46)$$

where V is an invariant scalar potential, as in Section 2. Separating the motion of the centre-of-mass we can write

$$A_\nu = e_\nu(p) \cdot e^{-ipx} \cdot \psi(p, r), \quad (47)$$

where $e_\nu(p)$ is the polarisation unit-vector. We can always add the Lorentz condition to (45), (46):

$$\partial_\nu^x A^\nu = 0, \quad (48)$$

or in momentum-space: $p_\nu A^\nu = 0$, that is $e_\nu p^\nu = 0$. For this reason the results concerning the 0 spin are valid also for the internal $\psi(p, r)$ space in the present case. There is,

however, a possibility, namely the possible symmetry of x - and r -spaces in the subsidiary condition, that is whether or not

$$\partial_r^\nu A^\nu = h(x, p, r) \quad (49)$$

can be assumed where h is vanishing or nonvanishing. From (47) and (49)

$$h(x, p, r) = e^{-ipx} \cdot g(p, r), \quad (50)$$

where

$$g(p, r) = e_\nu(p) \cdot \partial_r^\nu \psi(p, r), \quad (51)$$

so that to a given $\psi(p, r)$ there exists an $h(x, p, r)$. Applying ∂_r^ν , using (49) we can write

$$\partial_{1\mu}^2 h(x, p, r) - \partial_r^\nu [V(1, 2)A_\nu] = 0 \quad (52)$$

$$\partial_{2\mu}^2 h(x, p, r) - \partial_r^\nu [V(2, 1)A_\nu] = 0, \quad (53)$$

and because of $\partial_{1\mu}^2 = \left(-\frac{1}{2}ip + \partial_r\right)^2$, $\partial_{2\mu}^2 = \left(-\frac{1}{2}ip - \partial_r\right)^2$, and (47), (50) one gets

$$\left(-\frac{1}{2}ip + \partial_r\right)^2 g(p, r) - e_\nu(p) \partial_r^\nu [V(r, \partial_r, ip)\psi(p, r)] = 0, \quad (54)$$

$$\left(-\frac{1}{2}ip - \partial_r\right)^2 g(p, r) - e_\nu(p) \partial_r^\nu [V(-r, -\partial_r, ip)\psi(p, r)] = 0. \quad (55)$$

Taking the real and imaginary parts:

$$\left(-\frac{1}{4}p^2 + \partial_r^2\right) g(p, r) = e_\nu \partial_r^\nu (\text{Re } V \cdot \psi), \quad (56)$$

$$p\partial_r g(p, r) = e_\nu \partial_r^\nu (\text{Im } V \cdot \psi). \quad (57)$$

In general, if $V = f(r^2, pr, p^2) \cdot (p\partial_r)^l \cdot (\partial_r^2)^k$; substituting into (56), (57) we obtain

$$\left(-\frac{1}{4}p^2 \partial_r^2 - \text{Re } f \cdot (p\partial_r)^l (\partial_r^2)^k\right) g(p, r) = 2 \frac{\partial \text{Re } f}{\partial(r^2)} (e_\nu r_\nu) (p\partial_r)^l (\partial_r^2)^k \psi, \quad (58)$$

$$(p\partial_r + \text{Im } f (p\partial_r)^l (\partial_r^2)^k) g(p, r) = 2 \frac{\partial \text{Im } f}{\partial(r^2)} (e_\nu r_\nu) (p\partial_r)^l (\partial_r^2)^k \psi. \quad (59)$$

It follows that without knowing the ψ solution, we do not get natural solutions for g , since in general $\frac{\partial \text{Re } V}{\partial(r^2)} \neq \frac{\partial \text{Im } V}{\partial(r^2)}$. However, if V is independent of r^2 ,

$$\left(\partial_r^2 - \frac{1}{4}p^2 - \text{Re } V\right) g(p, r) = 0, \quad (60)$$

$$(p\partial_r + \text{Im } V) g(p, r) = 0, \quad (61)$$

and similarly to (3) and (4), the compatibility under (5) is necessary for $g(p, r)$. Using (9), (48), (51)

$$g = e_\nu \partial_r^\nu \psi = 2e_\nu r_\nu \cdot \frac{\partial \psi}{\partial (r^2)}, \quad (62)$$

therefore $g=0$ is equivalent with

$$\psi = \psi(pr, p^2). \quad (63)$$

Let us discuss now in more detail the case $g=0$. From (56) and (57)

$$e_\nu \partial_r^\nu (V \cdot \psi) = 0, \quad (64)$$

partially differentiating and since $e_\nu r_\nu \neq 0$, we get

$$\frac{\partial V}{\partial (r^2)} = 0. \quad (65)$$

Assume now

$$V = f(p^2, (pr)^2) \cdot (\partial_r^2)^k \cdot (p \partial_r)^m, \quad (66)$$

and define

$$f_1 = -\frac{\frac{1}{4}p^2 - \operatorname{Re} f}{p^2}, \quad (67)$$

$$f_2 = \frac{\operatorname{Im} f}{p^2}. \quad (68)$$

Substituting into the field equations (3), (4):

$$\psi'' + f_1 \cdot \psi = 0, \quad (69)$$

$$\psi' + f_2 \cdot \psi = 0, \quad (70)$$

consequently

$$\psi = A \cdot e^{-\int f_2 d(pr)}, \quad (71)$$

and f_1 is also fixed by

$$f_1 = \frac{-\psi''}{\psi}. \quad (72)$$

Since the wave functions of the type (71) are regular, they do not provide mass-spectra in these models motivated by symmetrical subsidiary conditions. The vector mesons are compound, but the mass is not quantised, it remains a continuous parameter.

It is, however, evident that the harmonic (or anharmonic) oscillator models described in Section 2 can be maintained for vector mesons, too, and in this case there exist mass spectra.

5. Conservation laws for compound baryons

The classical field theories and conservation laws of the bilocal fields are worked out (e.g. [3]). In what follows we shall show that this can be done for baryons consisting of three quarks, too. We suppose that ψ vanishes quickly at the boundary of the relative coordinate-space (for instance the hadron bag is described by a Gaussian). The baryon consists of three components, whose coordinates are x_1, x_2, x_3 . Let us introduce centre-of-mass and relative coordinates:

$$X = \frac{1}{3}(x_1 + x_2 + x_3), \quad x = \frac{1}{\sqrt{2}}(x_2 - x_3), \quad y = \sqrt{\frac{2}{3}}\left(x_1 - \frac{x_2 + x_3}{2}\right). \quad (73)$$

We assume that there exists a real, invariant Lagrangian

$$L = L(\psi, \psi^*, \partial_i \psi, \partial_i \psi^*, x_\mu, y_\mu); \quad \mu = (1, 2, 3, 4) \quad (74)$$

$$\xi_i = \{X_\mu, x_\mu, y_\mu\}; \quad \partial_i = \frac{\partial}{\partial \xi_i}. \quad (75)$$

The field equations are derived by the usual method from the Euler-Lagrange equations

$$\frac{\partial L}{\partial \psi} = \partial_i \frac{\partial L}{\partial \partial_i \psi} \equiv \partial_\mu \frac{\partial L}{\partial \partial_\mu \psi} + \bar{\partial}_\mu \frac{\partial L}{\partial \bar{\partial}_\mu \psi} + \bar{\partial}_\mu \frac{\partial L}{\partial \bar{\partial}_\mu \psi}; \quad (76)$$

$$\frac{\partial L}{\partial \psi^*} = \partial_i \frac{\partial L}{\partial \partial_i \psi^*} \equiv \partial_\mu \frac{\partial L}{\partial \partial_\mu \psi^*} + \bar{\partial}_\mu \frac{\partial L}{\partial \bar{\partial}_\mu \psi^*} + \bar{\partial}_\mu \frac{\partial L}{\partial \bar{\partial}_\mu \psi^*}; \quad (77)$$

$$\bar{\partial}_\mu = \frac{\partial}{\partial x_\mu}, \quad \bar{\partial}_\mu = \frac{\partial}{\partial y_\mu}. \quad (78)$$

In practice these are completed by subsidiary conditions (examples in [5]).

Demanding the invariance of the Lagrangian under infinitesimal translations in ξ -space we get a conserved quantity. Introducing

$$T^{ij} = \frac{\partial L}{\partial \partial_i \psi} \cdot \partial^j \psi + \frac{\partial L}{\partial \partial_i \psi^*} \cdot \partial^j \psi^* - L g^{ij} \quad (79)$$

as the energy-momentum tensor in the ξ -space, it follows

$$\partial_i T^{ij} = -\partial_{\text{expl}}^j L. \quad (80)$$

(80) can be verified directly from the field equations too.

The tensor (79) can be decomposed into the tensors:

$$T^{\mu\nu} = \frac{\partial L}{\partial \partial_\mu \psi} \partial^\nu \psi + \frac{\partial L}{\partial \partial_\mu \psi^*} \partial^\nu \psi^* - g^{\mu\nu} L, \quad (81)$$

$$\bar{T}^{\mu\nu} = \frac{\partial L}{\partial \bar{\partial}_\mu \psi} \bar{\partial}^\nu \psi + \frac{\partial L}{\partial \bar{\partial}_\mu \psi^*} \bar{\partial}^\nu \psi^* - g^{\mu\nu} L, \quad (82)$$

$$\bar{\bar{T}}^{\mu\nu} = \frac{\partial L}{\partial \bar{\bar{\partial}}_\mu \psi} \bar{\bar{\partial}}^\nu \psi + \frac{\partial L}{\partial \bar{\bar{\partial}}_\mu \psi^*} \bar{\bar{\partial}}^\nu \psi^* - g^{\mu\nu} L, \quad (83)$$

$$T_{12}^{\mu\nu} = \frac{\partial L}{\partial \bar{\partial}_\mu \psi} \partial^\nu \psi + \frac{\partial L}{\partial \bar{\partial}_\mu \psi^*} \partial^\nu \psi^*, \quad (84)$$

$$T_{23}^{\mu\nu} = \frac{\partial L}{\partial \bar{\bar{\partial}}_\mu \psi} \bar{\partial}^\nu \psi + \frac{\partial L}{\partial \bar{\bar{\partial}}_\mu \psi^*} \bar{\partial}^\nu \psi^*, \quad (85)$$

$$T_{31}^{\mu\nu} = \frac{\partial L}{\partial \partial_\mu \psi} \bar{\partial}^\nu \psi + \frac{\partial L}{\partial \partial_\mu \psi^*} \bar{\partial}^\nu \psi^*, \quad (86)$$

$$T_{21}^{\mu\nu} = \frac{\partial L}{\partial \partial_\mu \psi} \bar{\partial}^\nu \psi + \frac{\partial L}{\partial \partial_\mu \psi^*} \bar{\partial}^\nu \psi^*, \quad (87)$$

$$T_{32}^{\mu\nu} = \frac{\partial L}{\partial \bar{\partial}_\mu \psi} \bar{\partial}^\nu \psi + \frac{\partial L}{\partial \bar{\partial}_\mu \psi^*} \bar{\partial}^\nu \psi^*, \quad (88)$$

$$T_{13}^{\mu\nu} = \frac{\partial L}{\partial \bar{\bar{\partial}}_\mu \psi} \partial^\nu \psi + \frac{\partial L}{\partial \bar{\bar{\partial}}_\mu \psi^*} \partial^\nu \psi^*. \quad (89)$$

Here $T^{\mu\nu}$, $\bar{T}^{\mu\nu}$, $\bar{\bar{T}}^{\mu\nu}$ are attached to the centre-of-mass and internal motion, respectively. These tensors satisfy the conservation equations:

$$\partial_\mu T^{\mu\nu} + \bar{\partial}_\mu T_{12}^{\mu\nu} + \bar{\bar{\partial}}_\mu T_{13}^{\mu\nu} = 0, \quad (90)$$

$$\partial_\mu T_{21}^{\mu\nu} + \bar{\partial}_\mu \bar{T}^{\mu\nu} + \bar{\bar{\partial}}_\mu T_{23}^{\mu\nu} = -\bar{\partial}_{\text{expl}}^\nu L, \quad (91)$$

$$\partial_\mu T_{31}^{\mu\nu} + \bar{\partial}_\mu T_{32}^{\mu\nu} + \bar{\bar{\partial}}_\mu \bar{\bar{T}}^{\mu\nu} = -\bar{\bar{\partial}}_{\text{expl}}^\nu L. \quad (92)$$

Divergences of $T_{12}^{\mu\nu}$, $T_{23}^{\mu\nu}$, $T_{31}^{\mu\nu}$, $T_{21}^{\mu\nu}$, $T_{32}^{\mu\nu}$, $T_{13}^{\mu\nu}$ act as force-densities and they show correlations between the three constituents of the baryon.

Since $T_{12}^{\mu\nu} = 0$ at the boundary of the x -space, $T_{13}^{\mu\nu} = 0$ at the boundary of the y -space:

$$\partial_\mu \int d^4 x d^4 y T^{\mu\nu} = 0. \quad (93)$$

This integral can be interpreted as the energy-momentum tensor of the baryon attached to the centre-of-mass motion. Similarly, because of quickly decreasing $\bar{T}^{\mu\nu}$, $T_{23}^{\mu\nu}$, $T_{32}^{\mu\nu}$, $\bar{\bar{T}}^{\mu\nu}$:

$$\partial_\mu \int d^4 x d^4 y T_{21}^{\mu\nu} = 0, \quad (94)$$

$$\partial_\mu \int d^4 x d^4 y T_{31}^{\mu\nu} = 0. \quad (95)$$

It is, however, probable that these integrals do not play any role because they include mixed derivatives. Indeed, for example the Lagrangian of the free baryon without spin depends bilinearly on the derivatives, and in the state of P momentum

$$\int d^4x T_{21}^{\mu\nu} \alpha i p^\mu \int d^4x \psi(p, x, y) \bar{\delta}^\nu \psi(p, x, y) = 0 \quad (96)$$

in case of a real internal wave function. Alternative forms of (90), (91), (92), expressing with the variables of the constituents, are

$$\begin{aligned} \partial_{1\mu} \left(3T_{21}^{\mu\nu} + \sqrt{\frac{3}{2}} T_{23}^{\mu\nu} \right) + \partial_{2\mu} (3T_{21}^{\mu\nu} + \sqrt{2} \bar{T}^{\mu\nu} - \sqrt{6} T_{23}^{\mu\nu}) + \\ + \partial_{3\mu} (3T_{21}^{\mu\nu} - \sqrt{2} \bar{T}^{\mu\nu} - \sqrt{6} T_{23}^{\mu\nu}) = -\bar{\delta}_{\text{expl}}^\nu L, \end{aligned} \quad (97)$$

$$\begin{aligned} \partial_{1\mu} \left(3T_{31}^{\mu\nu} + \sqrt{\frac{3}{2}} \bar{T}^{\mu\nu} \right) + \partial_{2\mu} (3T_{31}^{\mu\nu} + \sqrt{2} T_{32}^{\mu\nu} - \sqrt{6} \bar{T}^{\mu\nu}) + \\ + \partial_{3\mu} (3T_{31}^{\mu\nu} - \sqrt{2} T_{32}^{\mu\nu} - \sqrt{6} \bar{T}^{\mu\nu}) = -\bar{\delta}_{\text{expl}}^\nu L, \end{aligned} \quad (98)$$

$$\begin{aligned} \partial_{1\mu} \left(3T^{\mu\nu} + \sqrt{\frac{3}{2}} T_{13}^{\mu\nu} \right) + \partial_{2\mu} (3T^{\mu\nu} + \sqrt{2} T_{12}^{\mu\nu} - \sqrt{6} T_{13}^{\mu\nu}) + \\ + \partial_{3\mu} (3T^{\mu\nu} - \sqrt{2} T_{12}^{\mu\nu} - \sqrt{6} T_{13}^{\mu\nu}) = 0, \end{aligned} \quad (99)$$

as well as

$$\partial_{1\mu} \int d^4x_2 d^4x_3 \left(3T^{\mu\nu} + \sqrt{\frac{3}{2}} T_{13}^{\mu\nu} \right) = 0, \quad (100)$$

$$\partial_{2\mu} \int d^4x_1 d^4x_3 (3T^{\mu\nu} + \sqrt{2} T_{12}^{\mu\nu} - \sqrt{6} T_{13}^{\mu\nu}) = 0, \quad (101)$$

$$\partial_{3\mu} \int d^4x_1 d^4x_2 (3T^{\mu\nu} - \sqrt{2} T_{12}^{\mu\nu} - \sqrt{6} T_{13}^{\mu\nu}) = 0, \quad (102)$$

$$\partial_{1\mu} \int d^4x_2 d^4x_3 \left(3T_{21}^{\mu\nu} + \sqrt{\frac{3}{2}} T_{23}^{\mu\nu} \right) = 0, \quad (103)$$

$$\partial_{2\mu} \int d^4x_1 d^4x_3 (3T_{21}^{\mu\nu} + \sqrt{2} \bar{T}^{\mu\nu} - \sqrt{6} T_{23}^{\mu\nu} + \sqrt{2} \delta_{\mu\nu} \cdot L) = 0, \quad (104)$$

$$\partial_{3\mu} \int d^4x_1 d^4x_2 (3T_{21}^{\mu\nu} - \sqrt{2} \bar{T}^{\mu\nu} - \sqrt{6} T_{23}^{\mu\nu} - \sqrt{2} \delta_{\mu\nu} \cdot L) = 0. \quad (105)$$

We can introduce the current vector, too. Demanding the invariance of the Lagrangian under gauge transformations of the first kind,

$$\psi \rightarrow \psi' = \psi \cdot e^{i\alpha}, \quad \psi^* \rightarrow \psi'^* = \psi^* \cdot e^{-i\alpha}, \quad (106)$$

we get the conservation equation for the currents:

$$j^\mu = ie \left(\frac{\partial L}{\partial \partial_\mu \psi} \cdot \psi - \frac{\partial L}{\partial \partial_\mu \psi^*} \cdot \psi^* \right), \quad (107)$$

$$\bar{j}^\mu = ie \left(\frac{\partial L}{\partial \bar{\partial}_\mu \psi} \cdot \psi - \frac{\partial L}{\partial \bar{\partial}_\mu \psi^*} \cdot \psi^* \right), \quad (108)$$

$$\bar{\bar{j}}^\mu = ie \left(\frac{\partial L}{\partial \bar{\bar{\partial}}_\mu \psi} \cdot \psi - \frac{\partial L}{\partial \bar{\bar{\partial}}_\mu \psi^*} \cdot \psi^* \right), \quad (109)$$

$$\partial_\mu j^\mu + \bar{\partial}_\mu \bar{j}^\mu + \bar{\bar{\partial}}_\mu \bar{\bar{j}}^\mu = 0. \quad (110)$$

Integrating again over the internal space:

$$\partial_\mu \int d^4 x d^4 y j^\mu = 0. \quad (111)$$

$\int d^4 x d^4 y j^\mu$ can be interpreted as the conserved current vector of the whole system. For a free baryon $\bar{j}^\mu = \bar{\bar{j}}^\mu = 0$ in the case of a real internal wave-function.

An alternative form of the current conservation:

$$\begin{aligned} \partial_{1\mu} \left(3j^\mu + \sqrt{\frac{3}{2}} \bar{\bar{j}}^\mu \right) + \partial_{2\mu} (3j^\mu + \sqrt{2} \bar{j}^\mu - \sqrt{6} \bar{\bar{j}}^\mu) + \\ + \partial_{3\mu} (3j^\mu - \sqrt{2} \bar{j}^\mu - \sqrt{6} \bar{\bar{j}}^\mu) = 0, \end{aligned} \quad (112)$$

whose double integrals are

$$\partial_{1\mu} \int d^4 x_2 d^4 x_3 \left(3j^\mu + \sqrt{\frac{3}{2}} \bar{\bar{j}}^\mu \right) = 0, \quad (113)$$

$$\partial_{2\mu} \int d^4 x_1 d^4 x_3 (3j^\mu + \sqrt{2} \bar{j}^\mu - \sqrt{6} \bar{\bar{j}}^\mu) = 0, \quad (114)$$

$$\partial_{3\mu} \int d^4 x_1 d^4 x_2 (3j^\mu - \sqrt{2} \bar{j}^\mu - \sqrt{6} \bar{\bar{j}}^\mu) = 0 \quad (115)$$

in terms of constituent coordinates.

Finally let us introduce the angular-momentum. Demanding the invariance of the Lagrangian under infinitesimal rotations

$$\xi'_i = \xi_i + \varepsilon_{ij} \xi_j; \quad \varepsilon_{ij} = -\varepsilon_{ji}, \quad (116)$$

we get the conservation equation concerning the angular-momentum. Define the angular-momentum tensor in ξ space as

$$J_{i,lj} = \xi_j T_{il} - \xi_l T_{ij} - \frac{\partial L}{\partial \partial^i \psi} \Sigma_{lj} \psi - \frac{\partial L}{\partial \partial^i \psi^*} \Sigma_{lj} \psi^* \quad (117)$$

we obtain

$$\partial^i J_{i,lj} = 0, \quad (118)$$

where $l, j = 1, 2, 3, 4$, or $5, 6, 7, 8$, or $9, 10, 11, 12$. ε_{ij} consists of blocks,

$$\varepsilon^{ij} = \begin{array}{|c|c|c|} \hline \varepsilon_1^{\mu\nu} & & \\ \hline & \varepsilon_2^{\mu\nu} & \\ \hline & & \varepsilon_3^{\mu\nu} \\ \hline \end{array} \quad (119)$$

$J_{i,lj}$ can be decomposed into tensors of the four-space

$$\begin{array}{ll} J_{\mu, \nu\sigma}; & i, l, j = 1, 2, 3, 4, \\ \bar{J}_{\mu, \nu\sigma}; & i, l, j = 5, 6, 7, 8, \\ \bar{\bar{J}}_{\mu, \nu\sigma}; & i, l, j = 9, 10, 11, 12, \\ J_{12\mu, \nu\sigma}; & i = 1, 2, 3, 4, \quad l, j = 5, 6, 7, 8, \\ J_{i,lj} = J_{23\mu, \nu\sigma}; & i = 5, 6, 7, 8, \quad l, j = 9, 10, 11, 12, \\ J_{31\mu, \nu\sigma}; & i = 9, 10, 11, 12, \quad l, j = 1, 2, 3, 4, \\ J_{21\mu, \nu\sigma}; & i = 5, 6, 7, 8, \quad l, j = 1, 2, 3, 4, \\ J_{32\mu, \nu\sigma}; & i = 9, 10, 11, 12, \quad l, j = 5, 6, 7, 8, \\ J_{13\mu, \nu\sigma}; & i = 1, 2, 3, 4, \quad l, j = 9, 10, 11, 12. \end{array} \quad (120)$$

These satisfy the the conservation equations:

$$\partial^\mu J_{\mu, \nu\sigma} + \bar{\partial}^\mu J_{12\mu, \nu\sigma} + \bar{\bar{\partial}}^\mu J_{13\mu, \nu\sigma} = 0, \quad (121)$$

$$\partial^\mu J_{21\mu, \nu\sigma} + \bar{\partial}^\mu \bar{J}_{\mu, \nu\sigma} + \bar{\bar{\partial}}^\mu J_{23\mu, \nu\sigma} = 0, \quad (122)$$

$$\hat{\partial}^\mu J_{31\mu, \nu\sigma} + \bar{\partial}^\mu J_{32\mu, \nu\sigma} + \bar{\bar{\partial}}^\mu \bar{\bar{J}}_{\mu, \nu\sigma} = 0 \quad (123)$$

and

$$\partial_\mu \int d^4x d^4y J_{\mu, \nu\sigma} = 0. \quad (124)$$

$\int d^4x d^4y J_{\mu, \nu\sigma}$ can be interpreted as the angular-momentum tensor of the baryon. Furthermore

$$\partial_\mu \int d^4x d^4y J_{21\mu, \nu\sigma} = 0, \quad (125)$$

$$\partial_\mu \int d^4x d^4y J_{31\mu, \nu\sigma} = 0. \quad (126)$$

Write (121), (122), (123) in terms of constituents:

$$\partial_{1\mu} \left(3J_{\mu, \nu\sigma} + \sqrt{\frac{3}{2}} J_{13\mu, \nu\sigma} \right) + \partial_{2\mu} (3J_{\mu, \nu\sigma} + \sqrt{2} J_{12\mu, \nu\sigma} - \sqrt{6} J_{13\mu, \nu\sigma}) +$$

$$+\partial_{3\mu}(3J_{\mu,\nu\sigma}-\sqrt{2}J_{12\mu,\nu\sigma}-\sqrt{6}J_{13\mu,\nu\sigma})=0, \quad (127)$$

$$\begin{aligned} \partial_{1\mu}\left(3J_{31\mu,\nu\sigma}+\sqrt{\frac{3}{2}}\bar{J}_{\mu,\nu\sigma}\right)+\partial_{2\mu}(3J_{31\mu,\nu\sigma}+\sqrt{2}J_{32\mu,\nu\sigma}-\sqrt{6}\bar{J}_{\mu,\nu\sigma})+ \\ +\partial_{3\mu}(3J_{31\mu,\nu\sigma}-\sqrt{2}J_{32\mu,\nu\sigma}-\sqrt{6}\bar{J}_{\mu,\nu\sigma})=0, \end{aligned} \quad (128)$$

$$\begin{aligned} \partial_{1\mu}\left(3J_{21\mu,\nu\sigma}+\sqrt{\frac{3}{2}}J_{23\mu,\nu\sigma}\right)+\partial_{2\mu}(3J_{21\mu,\nu\sigma}+\sqrt{2}\bar{J}_{\mu,\nu\sigma}-\sqrt{6}J_{23\mu,\nu\sigma})+ \\ +\partial_{3\mu}(3J_{21\mu,\nu\sigma}-\sqrt{2}\bar{J}_{\mu,\nu\sigma}-\sqrt{6}J_{23\mu,\nu\sigma})=0. \end{aligned} \quad (129)$$

Their integrals

$$\partial_{1\mu}\int d^4x_2d^4x_3\left(3J_{\mu,\nu\sigma}+\sqrt{\frac{3}{2}}J_{13\mu,\nu\sigma}\right)=0, \quad (130)$$

$$\partial_{2\mu}\int d^4x_1d^4x_3(3J_{\mu,\nu\sigma}+\sqrt{2}J_{12\mu,\nu\sigma}-\sqrt{6}J_{13\mu,\nu\sigma})=0, \quad (131)$$

$$\partial_{3\mu}\int d^4x_1d^4x_2(3J_{\mu,\nu\sigma}-\sqrt{2}J_{12\mu,\nu\sigma}-\sqrt{6}J_{13\mu,\nu\sigma})=0. \quad (132)$$

Similarly,

$$\partial_{1\mu}\int d^4x_2d^4x_3\left(3J_{31\mu,\nu\sigma}+\sqrt{\frac{3}{2}}\bar{J}_{\mu\nu}\right)=0, \quad (133)$$

$$\partial_{2\mu}\int d^4x_1d^4x_3(3J_{31\mu,\nu\sigma}+\sqrt{2}J_{32\mu,\nu\sigma}-\sqrt{6}\bar{J}_{\mu,\nu\sigma})=0, \quad (134)$$

$$\partial_{3\mu}\int d^4x_1d^4x_2(3J_{31\mu,\nu\sigma}-\sqrt{2}J_{32\mu,\nu\sigma}-\sqrt{6}\bar{J}_{\mu,\nu\sigma})=0, \quad (135)$$

$$\partial_{1\mu}\int d^4x_2d^4x_3\left(3J_{31\mu,\nu\sigma}+\sqrt{\frac{3}{2}}\bar{J}_{\mu,\nu\sigma}\right)=0, \quad (136)$$

$$\partial_{2\mu}\int d^4x_1d^4x_3(3J_{31\mu,\nu\sigma}+\sqrt{2}J_{32\mu,\nu\sigma}-\sqrt{6}\bar{J}_{\mu,\nu\sigma})=0, \quad (137)$$

$$\partial_{3\mu}\int d^4x_1d^4x_2(3J_{31\mu,\nu\sigma}-\sqrt{2}J_{32\mu,\nu\sigma}-\sqrt{6}\bar{J}_{\mu,\nu\sigma})=0. \quad (138)$$

In conclusion, the conserved quantities attached to the motion of the centre-of-mass (four-momentum, current and angular-momentum) have been described by integrating over the internal motion in the baryon consisting of three pointlike quarks.

References

1. H. Yukawa, *Phys. Rev.*, 77, 219, 1950; *ibid.* 80, 1047, 1950; *ibid.* 91, 415, 416, 1953; J. Rayski, *Proc. Phys. Soc.*, 64A, 957, 1951; *Nuovo Cimento*, 2, 255, 1955; M. Markov, *Nuovo Cimento Suppl.* III. 760, 1956; G. Pócsik, *Acta Phys. Hung.*, 9, 245, 1959; T. Takabayashi, *Phys. Rev. B*, 139, 1381, 1965; *Prog. Theor.*

- Phys., 42, 423, 1969; *ibid.* 48, 1718, 1972; Ishida and J. J. Otokoza, Progr. Theor. Phys., 47, 2117, 1972. A discussion of conservation laws is contained in G. Pócsik ITP-Rep. No. 365, 1976.
2. A. Z. Capri and C. C. Chiang, Nuovo Cimento, 36A, 331, 1976;
 3. G. Pócsik, Acta Physica Austriaca, 49, 47, 1978;
 4. I. Farkas and G. Pócsik, Lett. Nuovo Cimento, 23, 444, 1978;
 5. G. Pócsik, Nuovo Cimento, 54A, 413, 1979.
 6. G. Pócsik, ITP-Budapest Report No. 367, November 1976.

SPIN ORBITALS AND TOTAL ENERGY
CALCULATED BY THE $X\alpha$ METHOD INCLUDING
AB INITIO SELF-CONSISTENT EXCHANGE
PARAMETERS α_{SCF} *

R. GÁSPÁR and Á. NAGY

*Institute of Theoretical Physics, Kossuth Lajos University
4010 Debrecen, Hungary*

A way of calculating ab initio self-consistent exchange parameters in the $X\alpha$ method is presented. It provides a self-contained theory, having no need of any external parameter. Expectation values of powers of radius and the total energy for some atoms and ions are calculated with this self-consistent exchange parameter α_{SCF} and compared with results obtained from other $X\alpha$ calculations and the Hartree-Fock theory. It is pointed out that the $X\alpha$ results with exchange parameters α_{SCF} are nearly as good as the Hartree-Fock values and the method greatly simplifies the many-body problem.

Introduction

The $X\alpha$ method, which has become extremely popular for the past few years, greatly simplifies the many-body problem, as only integrals on density functionals are calculated instead of two-electron integrals. Since the $X\alpha$ method was first suggested by Slater [1], enormous development has occurred in the theory and in its applications. One of the main aims of the researches in the $X\alpha$ method is to develop it into an ab initio method. The $X\alpha$ method has an adjustable parameter α . Generally, the parameter α used is exterior in the method, e.g. the most commonly applied parameter α , the so-called α_{HF} , has been determined by Schwarz [2] so that the Hartree-Fock total energy and the $X\alpha$ total energy be equal. This way of determining the parameter α_{HF} seems to be very fruitful in a large number of applications, but it has the shortcoming that it can be applied consistently to atoms only as few Hartree-Fock calculations for molecules or solids are available nowadays. On the other hand, it seems to have little sense doing an $X\alpha$ calculation if we do a Hartree-Fock one beforehand to calculate the exchange parameter α_{HF} of the $X\alpha$ method.

There have been several attempts [3] to determine the value of the exchange parameter α of the $X\alpha$ method. The most interesting ones seem to be those that determine the value of the parameter α without adjustments. Gopinathan, Whitehead and Bogdanovic [4] calculated theoretical exchange parameters. However, they started out from the assumption that the density of the Fermi-hole varies linearly with

* Dedicated to Prof. K. Nagy on his 60th birthday

the radial distance and the exchange parameters $\alpha(\alpha_{\text{GWB}})$ they obtained were not good enough so they had to introduce a set of scaled parameters $\alpha(\alpha'_{\text{GWB}})$ to get agreement with the free-electron gas value. Gazquez and Keller [5] applying the method of Gopinathan, Whitehead and Bogdanovic with another expression for the exchange-hole density obtained theoretical exchange parameters. However, they also had to apply an adjusting parameter.

In this paper a review of a method proposed by one of the authors [6] is given. This method provides an ab initio self-consistent method for the calculation of the exchange parameters α . In this method there is no need of any external parameter, the exchange parameter is determined in the method itself, in a self-consistent manner. This method can be used for molecules and solids, too. It is shown here that the expectation values of powers of the radius and the total energy calculated by the X α method with parameters α_{SCF} are nearly as good as the Hartree-Fock ones. The ab initio exchange parameters α_{SCF} are fairly close to the exchange parameters α_{HF} and show a similar behaviour vs atomic number.

The method

This way of theoretical determination of the exchange parameter α in the X α method starts out from the free-electron gas theory. The exchange potential of an electron with spin up in the electron gas is given by [7]

$$V_{x\uparrow}(1) = -8F(\eta) \left[\frac{3}{4\pi} \rho_{\uparrow}(1) \right]^{1/3}, \quad (1)$$

where

$$F(\eta) = \frac{1}{2} + \frac{1-\eta^2}{4\eta} \ln \left| \frac{1+\eta}{1-\eta} \right| \quad (2)$$

and

$$\eta = \frac{p}{p_F}, \quad (3)$$

where p is the momentum of the electron considered and p_F is the Fermi momentum. The total density of electrons having spin up is

$$\rho_{\uparrow} = \sum_i n_i u_i^* u_i \quad (4)$$

and a similar formula holds for electrons with spin down. n_i and u_i are the occupation numbers and spin orbitals of the electrons, respectively.

The averaged exchange potential is given by

$$\overline{V_{x\uparrow}(1)} = -8 \left[\frac{3}{4\pi} \rho_{\uparrow}(1) \right]^{1/3} \frac{\int_{\eta_1}^{\eta_2} F(\eta) \eta^2 d\eta}{\int_{\eta_1}^{\eta_2} \eta^2 d\eta}, \quad (5)$$

which is an X α exchange potential

$$V_{X\alpha\uparrow} = -6\alpha \left(\frac{3}{4\pi} \rho_{\uparrow} \right)^{1/3} \quad (6)$$

with an exchange parameter

$$\alpha = \left\{ \frac{1}{2}(\eta^3 + \eta) - \frac{1}{4}(\eta^2 - 1)^2 \ln \left| \frac{1 + \eta}{1 - \eta} \right| \right\}_{\eta_1}^{\eta_2} \left[\left\{ \eta^3 \right\}_{\eta_1}^{\eta_2} \right]^{-1} \quad (7)$$

(Energy is in Rydbergs.)

If the average is taken over all the occupied states, i.e. $\eta_1 = 0$ and $\eta_2 = p_F/p_F = 1$, the exchange parameter $\alpha = 1$ is given. As it has already been shown by Slater [8] an average of the exchange potential (1) for the total Fermi sphere leads to the exchange parameter $\alpha = 1$. If the averaging is done over a thin shell near p_F , i.e. $\eta_1 = (p_F - \varepsilon)/p_F$ and $\eta_2 = p_F/p_F = 1$ and $\varepsilon \rightarrow 0$, $\alpha = 2/3$ is obtained. This is the exchange parameter α suggested by Gáspar [9] and later by Kohn and Sham [10].

Obviously, the reality is between these two extreme cases. On the one hand, there is no need to average for the whole Fermi sphere to determine the exchange potential of the one electron considered. On the other hand, the average must be taken for a shell of finite thickness. Thus, the exchange potential is constructed by an averaging process near the Fermi surface for a layer containing v_{\uparrow} electrons in the unit volume, i.e. $\eta_1 = (1 - v_{\uparrow}/\rho_{\uparrow})^{1/3}$ and $\eta_2 = 1$. v_{\uparrow} is the density of the electron considered. The exchange potential is given by

$$V_{X\alpha\uparrow} = -6\alpha_{\text{shell}} \left[\frac{3}{4\pi} \rho_{\uparrow} \right]^{1/3} \quad (8)$$

where

$$\alpha_{\text{shell}} = \frac{\rho_{\uparrow}}{v_{\uparrow}} \left\{ 1 - \frac{1}{2}\eta^3 - \frac{1}{2}\eta + \frac{1}{4}(\eta^2 - 1)^2 \ln \left| \frac{1 + \eta}{1 - \eta} \right| \right\} \quad (9)$$

and

$$\eta = \left(1 - \frac{v_{\uparrow}}{\rho_{\uparrow}} \right)^{1/3} \quad (10)$$

Here we introduced a subscript "shell" emphasizing that the exchange potential (8) is different for different spin orbitals, i.e. for different shells. To get an exchange potential that can easily be compared with the original X α exchange potential, the exchange potentials (8) are averaged over the shells

$$V_{X\alpha\uparrow(\vec{r})} = -6\alpha_{\uparrow}(\vec{r}) \left[\frac{3}{4\pi} \rho_{\uparrow}(\vec{r}) \right]^{1/3} \quad (11)$$

$$\alpha_{\uparrow}(r) = \frac{\sum_{i\uparrow} n_i \alpha_{\text{shell}}^i}{\sum_{i\uparrow} n_i} \quad (12)$$

where α_{shell}^i is the exchange parameter of the shell i according to the formula (9). As the expression (12) contains the spin orbitals, this $\alpha_i(\vec{r})$ is not a constant, but it is a function of the position \vec{r} . Of course, it is possible to introduce a constant exchange parameter α_i so that the mean squared deviation of the original $X\alpha$ exchange potential

$$V_{X\alpha i}(\vec{r}) = -6\alpha_i \left[\frac{3}{4\pi} \rho_i(\vec{r}) \right]^{1/3} \quad (13)$$

and this modified exchange potential (11) be a minimum.

These exchange parameters α_i and α_j can be calculated self-consistently. Starting out from a trial exchange parameter α , the $X\alpha$ one-electron equations are solved self-consistently. The spin orbitals obtained are used to calculate the new exchange parameter α applying the formulae (9) and (10). With the new exchange parameter α the $X\alpha$ equations are solved again self-consistently, and so on. The procedure is carried on until the self-consistent exchange parameter α_{SCF} is obtained.

It can be emphasized that it is not necessary to apply the constant exchange parameter α_i . Instead, it is possible to use the exchange potentials (11) or even (8). In the calculations presented in this paper the constant α has been applied as the results are compared with those of other $X\alpha$ calculations having constant exchange parameters.

It can be seen that this method is self-contained. The exchange parameter α_{SCF} is determined in the course of the calculation without any need of an adjusting procedure or a parameter exterior to the method.

This method is not restricted to atoms or atomic ions. It can be straightforwardly applied for molecules and solids, too.

Results and discussion

A test of spin orbitals is offered by the expectation values of powers of the radius. $\langle r^2 \rangle$, $\langle r \rangle$ and $\langle r^{-1} \rangle$ calculated for a few atoms and ions using self-consistent parameters α_{SCF} and these are presented in Table I and II. The total electron density here is normalized to 1. For comparison Hartree-Fock [11] and CI [12] values are presented in Table I. Table II contains $X\alpha$ results with α_{HF} , $\alpha=1$ and $\alpha=2/3$, too. Hartree-Fock values and experimental data for $\langle r^2 \rangle$ obtained from diamagnetic susceptibility [13] are also presented in Table II. The expectation values of $\langle r^2 \rangle$, $\langle r \rangle$ and $\langle r^{-1} \rangle$ reflect the goodness of the spin orbitals and the total electron density. The $\langle r^{-1} \rangle$ values show how well the electron densities behave in the domain of inner spin orbitals nearer to the nucleus. The values $\langle r \rangle$ and $\langle r^2 \rangle$ show the behaviour of spin orbitals in the middle and outer regions of the atom. In a recent article of the authors [14], spin orbitals of Ne, Ar and Kr calculated by the $X\alpha$ method with several parameters α (among others α_{SCF}) have been compared. It was pointed out that the $X\alpha$ spin orbitals with α_{SCF} agree well with those of the Hartree-Fock method. Studying Table I and II we can come to the same conclusion. The $X\alpha$ results with parameters α_{SCF}

Table I

Expectation values of powers of the radius for a few atoms and ions calculated by the Hartree-Fock [11], CI [12] and $X\alpha$ methods with self-consistent α_{SCF} parameters

	HF	CI	$X\alpha_{\text{SCF}}$
Li			
$\langle r^2 \rangle$	6.21038	6.12404	6.99186
$\langle r \rangle$	1.67325	1.66474	1.75168
$\langle r^{-1} \rangle$	1.90516	1.90587	1.87662
Be			
$\langle r^2 \rangle$	4.32971	4.08138	4.37394
$\langle r \rangle$	1.53221	1.49702	1.53807
$\langle r^{-1} \rangle$	2.10219	2.10615	2.08613
C^{2+}			
$\langle r^2 \rangle$	1.15549	1.11799	1.10809
$\langle r \rangle$	0.81968	0.80836	0.80495
$\langle r^{-1} \rangle$	3.35567	3.35947	3.37625
Ne			
$\langle r^2 \rangle$	0.9372	0.9448	0.95914
$\langle r \rangle$	0.78911		0.79221
$\langle r^{-1} \rangle$	3.11134		3.11677
Mg			
$\langle r^2 \rangle$	2.4676	2.3965	2.36169
$\langle r \rangle$	1.02148		1.00746
$\langle r^{-1} \rangle$	3.32672		3.32870
Ar			
$\langle r^2 \rangle$	1.4464		1.45700
$\langle r \rangle$	0.89282		0.89276
$\langle r^{-1} \rangle$	3.87364		3.87412
Ca			
$\langle r^2 \rangle$	2.8283		2.70531
$\langle r \rangle$	1.06229		1.04900
$\langle r^{-1} \rangle$	4.00801		4.00615
Ca^{2+}			
$\langle r^2 \rangle$	0.8573		0.85301

are very close to the Hartree-Fock data. In several cases the $X\alpha_{\text{SCF}}$ values are as good as the Hartree-Fock ones or sometimes they are even better. A comparison with the other values of α show that the α_{SCF} and α_{HF} give better results than the $\alpha = 1$ or $\alpha = 2/3$.

Table III presents the total energy of some light atoms. In addition to $X\alpha$ results with exchange parameters α_{SCF} , Table III contains the Hartree-Fock data [15], the values obtained from the Hartree-Fock plus correlation energies [16], and the LSD (local-spin-density) energies [17]. For Ne and Ar atoms $X\alpha$ total energies calculated using the parameters α_{GWB} and α'_{GWB} are also available [4]. The $X\alpha$ total energy

Table II

Expectation values of powers of the radius for Ne and Ar atoms calculated by the Hartree-Fock [15] and $X\alpha$ methods with α_{SCF} , α_{HF} , $\alpha=1$ and $\alpha=2/3$ parameters. Experimental values $\langle r^2 \rangle$ are obtained from diamagnetic susceptibility [13]

	HF	Exp	$X\alpha$			
			α_{SCF}	α_{HF}	$\alpha=1$	$\alpha=2/3$
Ne						
$\langle r^2 \rangle$	0.9372	0.84–0.98	0.9591	0.9701	0.8509	1.0038
$\langle r \rangle$	0.7891		0.7922	0.7959	0.7539	0.8070
$\langle r^{-1} \rangle$	3.1113		3.1168	3.1113	3.1782	3.0953
Ar						
$\langle r^2 \rangle$	1.4464	1.34–1.38	1.4570	1.4530	1.2948	1.4894
$\langle r \rangle$	0.8928		0.8928	0.8918	0.8519	0.9004
$\langle r^{-1} \rangle$	3.8734		3.8741	3.8753	3.9277	3.8648

Table III

Total energy (in Ry) of some light atoms calculated by Hartree-Fock [15], Hartree-Fock plus correlation [16], the LSD [17] and $X\alpha$ methods with α_{SCF} parameters

Z	HF	HF + correlation	LSD	$X\alpha(\alpha_{\text{SCF}})$
10	-257.095	-257.855	-256.427	-257.825
11	-323.720	-324.490	-322.847	-322.725
12	-399.231	-400.085	-398.234	-399.263
13	-483.755	-484.671	-482.587	-483.107
14	-577.710	-578.696	—	-577.418
15	-681.438	-682.479	-679.937	-680.968
16	-795.012	-796.199	—	-794.567
17	-918.966	-920.296	—	-918.508
18	-1053.638	-1055.096	-1051.792	-1053.117

determined applying the parameters α_{GWB} is -257.459 Ry for Ne and -1054.657 Ry for Ar. The $X\alpha$ total energy calculated using the parameters α'_{GWB} is -256.862 Ry and -1052.947 Ry for Ne and Ar, respectively. It can be seen that the $X\alpha$ method with parameters α_{SCF} generally gives nearly as good results as the Hartree-Fock method and always better results than the $X\alpha$ method with α_{GWB} and α'_{GWB} . The $X\alpha$ total energies with parameter α_{SCF} are often closer to the exact results than the LSD total energies that explicitly contain correlation.

In conclusion, it can be found that the $X\alpha$ method with ab initio self-consistent exchange parameters is nearly as good as the Hartree-Fock method.

References

1. J. C. Slater, *Phys. Rev.*, *81*, 385, 1951.
2. K. Schwarz, *Phys. Rev.*, *B5*, 2466, 1972.
3. J. C. Slater, *The Self-Consistent Field for Molecules and Solids*, Vol. 4, McGraw-Hill, Inc., New York, 1974.
4. M. S. Gopinathan, M. A. Whitehead and R. Bogdanovic, *Phys. Rev.*, *A14*, 1, 1976.
5. J. L. Gazquez and J. Keller, *Phys. Rev.* *A16*, 1358, 1977.
6. R. Gáspár, *Acta Phys. Hung.*, *35*, 213, 1974; *Acta Phys. et Chim. Debr.*, *19*, 7, 1974; *Acta Phys. et Chim. Szeged*, *20*, 321, 1974.
7. P. Gombás, *Die statistische Theorie des Atoms und ihre Anwendungen*, Springer Verlag, Wien, 1949; N. H. March, *Self-Consistent Fields in Atoms*, Pergamon, Oxford, 1975.
8. J. C. Slater, *Quantum Theory of Atomic Structure*, Vol. II, McGraw-Hill, Inc., New York, 1960.
9. R. Gáspár, *Acta Phys. Hung.*, *3*, 263, 1954.
10. W. Kohn and L. J. Sham, *Phys. Rev.*, *140A*, 1133, 1965.
11. E. Clementi and C. Roetti, *At. Data Nucl. Data Tables* *14*, 177, 1974; A. Gupta and R. J. Boyd, *J. Chem. Phys.*, *68*, 1951, 1978.
12. A. W. Weiss, *Phys. Rev.*, *122*, 1826, 1961; S. H. Vosko and L. Wilk, *J. Phys.*, *B16*, 3687, 1983.
13. Landolt-Börnstein, *Zahlenwerte and Funktionen aus Physik, Chemie, Astronomie, Geophysik und Technik*, Springer-Verlag, Berlin, 1950, 6. Auflage, Band 1.
14. R. Gáspár and Á. Nagy, *Acta Phys. Hung.* *53*, 247, 1982.
15. C. Froese-Fischer, *Atomic Data*, *4*, 301, 1972.
16. H. Hartmann and E. Clementi, *Phys. Rev.*, *133A*, 1295, 1964; T. J. Tseng, S. H. Hong and M. A. Whitehead, *J. Comp. Chem.*, *1*, 88, 1980.
17. J. P. Perdew and A. Zunger, *Phys. Rev.*, *B23*, 5048, 1981.

ENTROPY IN THE UNIVERSE*

G. MARX

*Department of Atomic Physics, Roland Eötvös University
1088 Budapest, Hungary*

The relic microwave radiation indicates that the universe was in the state of thermal equilibrium (heat death) in the past. The question arises: how could it happen that the universe departed from the thermal equilibrium? According to the models presented here this was the outcome of an interplay among dynamic instability, matter differentiation and weakening of some interactions.

*Past and future are so markedly
different because the universe
is still very young. (F. Hund)*

Dynamical controversy

The success of the Newtonian programme has been due to the fact that the laws of motion have been formulated in the form of time dependent differential equations like

$$\dot{q} = F(q), \quad (1)$$

where q is a set of variables describing the state. The function F describes the specific interaction, it can be obtained from local observations. In order to get the present state $q(t)$, beside the equation (1) of motion one needs also the initial condition

$$q(0). \quad (2)$$

But this cannot be learned by experimenting in laboratories! One tried to borrow it from elsewhere, but scientists of the last century were worried about such an alien input. To be able to explain the material world by itself, a steady state universe was postulated (at least on large scale):

$$\dot{q} = 0.$$

In this way physicists hoped to get q by solving time independent equilibrium conditions:

$$F(q) = 0.$$

* Dedicated to Prof. K. Nagy on his 60th birthday

At the turn of the century, however, theoretical investigations led to inherent logical and empirical contradictions:

Olbers' paradoxon is well known: The intensity of the light reaching Earth from a star being at distance r is proportional to r^{-2} . The number of stars in a layer of unit thickness and radius r is proportional to r^2 , so in a homogeneous universe, being in steady state, each layer contributes to the brightness of the sky with the same amount. In an infinite universe this would give a brightly glowing sky, so why is it dark at night? — Clausius argued that in an eternal universe all differences would have levelled up, the universe would have reached its state of maximum entropy long ago. How can it be hot and cold, dense and empty at different places? — Seeliger stressed that the known equations of motion cannot describe a static homogeneous universe, the potential equation

$$\nabla^2\varphi = 4\pi G\rho$$

does not have an acceptable solution for a mass density constant in space and time. (G is the Newtonian constant of gravity.) How comes that the sky does not fall onto our head?

The greatest minds were ready to modify the empirical equations in order to force a steady state solution. Finally Alex Friedman took the courage to acknowledge that the physical equations of motion did not have static solution at all. He used the equations of general relativity and postulated a homogeneous and isotropic universe.

$$ds^2 = R(t)^2 \left[\frac{dr^2}{1+2Kr} + r^2(d\vartheta^2 + \sin^2\vartheta d\varphi^2) \right] - c^2 dt^2.$$

The length scale $R(t)$ can be obtained from Einstein's equation

$$\frac{1}{2} \dot{R}^2 - \frac{G\rho}{R} \left(\frac{4\pi}{3} R^3 \right) = K. \quad (3)$$

General relativity also gives

$$dE + pdV = 0$$

for an adiabatically closed world, in our case

$$d \left(\frac{4\pi}{3} R^3 \rho c^2 \right) + pd \left(\frac{4\pi}{3} R^3 \right) = 0. \quad (4)$$

If the equation $p = p(R)$ of state is known, Eq. (4) gives $\rho(R)$ and the Eq. (3) enables us to compute $R(t)$. E.g. in a dust (nonrelativistic gas) dominated universe $p \ll \rho c^2$, consequently $4\pi R^3 \rho / 3 = M = \text{const}$, and Eq. (3) gives

$$\dot{R}^2 = 2GM/R + 2K \quad (5)$$

and all of its solutions show a singularity $R=0$. There is a combined empirical and theoretical indication that K cannot be large. E.g. by putting $K=0$ and by choosing $R=0$ as the origin of time, one gets

$$R(t) = (4.5GM)^{1/3} t^{2/3} \quad \text{and} \quad \rho(t) = 1/6\pi G t^2. \quad (6)$$

(Hawking and Penrose confirmed later that even if one dropped the specific restrictions about p and ρ , the realistic solutions of Einstein's equations could not be extended to infinity in both directions of the time axis. Only empty universe could be static.)

The theoretical conclusion has been confirmed empirically by the observation of the runaway of galaxies (the overall expansion of the universe). This has invalidated the objections cited above.

Thermodynamical controversy

If the universe was dense in the past, how comes that the present cool world is not dominated by medium-heavy nuclei having the deepest binding (like Fe) but it is abundant in light elements (H, He, O, C)? Gamow offered the answer that the early universe was hot. The entropy of the black body radiation at temperature T is

$$S = \frac{4\pi^2 k^4 T^3}{45 \hbar^3 c^3} V = \frac{16\pi^3 k^4}{135 \hbar^3 c^3} (TR)^3.$$

If radiation has reached its thermal equilibrium,

$$dS = (dE + p dV)/T = 0,$$

then $S = \text{const}$ gives $T \sim R^{-1}$. This results in a radiation (mass) density

$$\rho_{\text{rad}} = (\pi^2/15) k^4 T^4 / \hbar^3 c^3 \sim R^{-4}, \quad (7)$$

to be compared to the (nonrelativistic) gas density

$$\rho_{\text{gas}} = 3M/4\pi R^3 \sim R^{-3}. \quad (8)$$

By approaching the limit $t \rightarrow 0$, $R \rightarrow 0$, ρ_{rad} diverges faster and exceeds ρ_{gas} . In the early era of radiation dominance $\rho \approx \rho_{\text{rad}} \sim R^{-4}$ so

$$\dot{\rho}/\rho = -4(\dot{R}/R).$$

Eq. (3) gives (for $K=0$)

$$\dot{R}/R = (8\pi G/3)\rho. \quad (9)$$

The two equations combined offer the solution

$$\rho_{\text{rad}} = (\pi^2/15) k^4 T^4 / \hbar^3 c^5 = 3/32\pi G t^2 \quad (10)$$

and from this one gets $T \sim t^{-1/2}$. In this way the thermal history of the past can be reconstructed as far as the behaviour of matter is explored in high energy laboratories.

The prediction of the early hot universe has been confirmed by the discovery of the cosmic microwave background noise. Radioastronomical observation has shown that this radiation is highly isotropic (up to 4–5 decimals), its spectrum has a Planck shape corresponding to a temperature $T=2.7$ K. This means that at present the number of (microwave) photons exceeds the number of atomic particles by a factor of 10^9 , so the overall entropy of the universe is mainly due to photons, nowadays atomic contribution is negligible.

By reconstructing the past of the universe by using the formulas $R \sim t^{2/3}$ (while $\rho_{\text{gas}} > \rho_{\text{rad}}$) and $T_{\text{rad}} \sim R^{-1}$, one obtains that space cooled below 10^4 K within the first million years. Before that it was $kT > 1$ eV, charged matter was completely ionized, interacting immediately with electromagnetic waves. After that ions formed neutral atoms, which became practically invisible for electromagnetic waves. Since that the mean collision time of photons got larger than the age of the universe, so the relic radiation depicts the state of the universe within the first million years! (The picture is, however, red shifted.) The Planck shape of its spectrum tells us that the universe was in thermal equilibrium 10^5 years after its origin! (No wonder: gravity was able to thermalize the very early — hot and dense — universe within Planck time.) The relic radiation we observe is not a bliss of creation but the glow of hell!

Our world certainly is not in thermal equilibrium now. How can a closed system get out of thermal equilibrium? This question is a burning one since decades [3]. In the next chapters models will be presented to show that the rejuvenation of our world is the outcome of an interplay among dynamical instability, matter differentiation and weakening of some interactions [11, 12].

Gas model

Let us consider two equal containers, one filled with argon gas, the other with nitrogen gas; both at 2 atm pressure and 200°C (Fig. 1A). The equal temperatures are consequences of a thermal contact between the two pistons. In this equilibrium state there is no arrow of time. — Remove now half of the weights! Both gases expand quickly. By raising the pistons, they perform the same work against gravity at the cost of their internal thermal motion. The monoatomic argon cools to 76°C , the diatomic nitrogen cools to 121°C (Fig. 1B). Nitrogen molecules store energy in the form of molecular rotations as well, so in order to perform the same work, nitrogen gas has to cool less.) Expansion against gravity created temperature difference! (One could even drive a steam engine for a while.) But heat conduction starts immediately to level up the temperatures, the appearance of an arrow of time is only a transient phenomenon. Due to the weak (thermal) coupling between the two pistons, they will soon reach a new thermal equilibrium at a common temperature of 104°C .

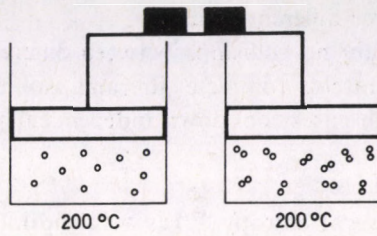


Fig. 1A

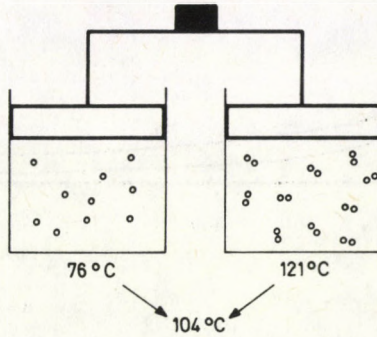


Fig. 1B

To build up this model to a closed system, let us assume that Einstein's world is filled with two gases: one is monoatomic, the other is diatomic (with molecular masses m_1, m_2 and with temperatures T_1, T_2 ; each of them containing $N/2$ molecules within a sphere of radius R). The energy, pressure and entropy of the gas mixture are:

$$E = (N/2)(m_1 c^2 + 1.5 k T_1) + (m_2 c^2 + 2.5 k T_2), \quad (11)$$

$$p = N(k T_1 + k T_2)/2V, \quad (12)$$

$$S = 0.75 k N \ln T_1 + 1.25 k N \ln T_2 + k N \ln V. \quad (13)$$

The Einstein's equation (3) reads for $K = 0$:

$$\dot{R}^2 = 2GE/Rc^2 \quad (14)$$

or by introducing dimensionless variables

$$r = R mc/\hbar, \quad x = kT_1/2(m_1 + m_2)c^2, \quad y = kT_2/2(m_1 + m_2)c^2 \quad (15)$$

one gets

$$\dot{r} = g[N(1 + 3x + 5y)/r]^{1/2}. \quad (16)$$

The entropy (13) can be expressed by the same variables:

$$S = (kN/4) [\ln(x^3 r^6) + \ln(y^5 r^6)]. \quad (17)$$

Now one can imagine three different scenarios:

Scenario A: There are no collisions between different molecules ($T_1 = T_2$, $S_1 = \text{const}$, $S_2 = \text{const}$ separately, complete thermal isolation). As space expands according to Eq. (16), both gases cool down independently and adiabatically:

$$x \sim r^{-2}, \quad y \sim r^{-1.2}, \quad (18)$$

evidently conserving the overall entropy (17) as well. The different molar heats result in different cooling rates (18), so temperature difference will be created, mainly in the fast initial phase of the expansion (Fig. 2A).

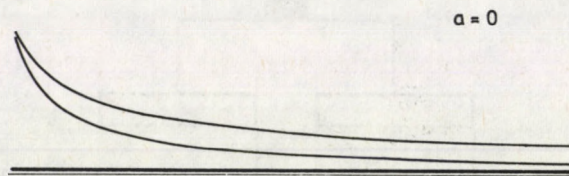


Fig. 2A

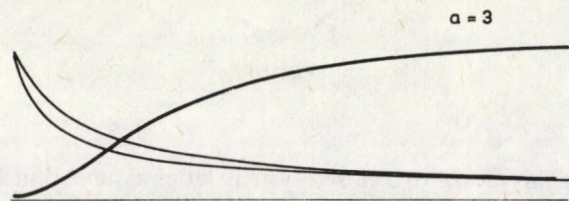


Fig. 2B

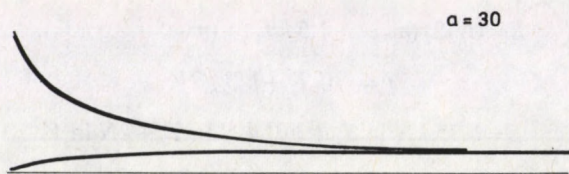


Fig. 2C

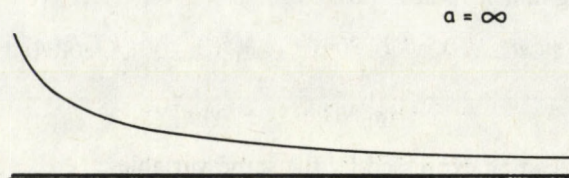


Fig. 2D

Scenario D: There are frequent collisions between all sorts of molecules ($T_1 = T_2$). During expansion both gases cool at the same rate:

$$x = y \sim r^{-1.5}. \quad (19)$$

Matter is in thermal equilibrium all the time (Fig. 2D):

$$S = 2N \ln(xr^{1.5}) = \text{const.}$$

Scenario B: There are frequent collisions among monoatomic molecules, frequent collisions among diatomic molecules, but rare collisions among different molecules. The thermal coupling of the two gases is not enough to equalize their temperatures, but there is a sort of modest heat conduction between them:

$$\dot{x} = -2x\dot{r}/r + A(y-x)/3, \quad (20)$$

$$\dot{y} = -1.2y\dot{r}/r - A(y-x)/5. \quad (21)$$

(The first terms on the right hand side take the expansion cooling into account, $r \sim \lambda$, where λ is the de Broglie wave length of molecules. The second terms describe the heat conduction. Here the factors are chosen so that Eq. (4) is satisfied all the time.) Now the behaviour of this system is described by the equations (16), (20), (21). As indicated on Figures 2B–C, the early fast expansion creates temperature difference, but the temperatures equalize in the later — more quiet — phase, as a consequence of heat conduction. Any temperature difference is a transient phenomenon, its duration depends on the value of the conduction coefficient A . Heat conduction is surely irreversible, producing an increase in the entropy S (Eq. (13), solid lines on Figures 2B–C):

$$dS > (dE + p dV)/T = 0. \quad (22)$$

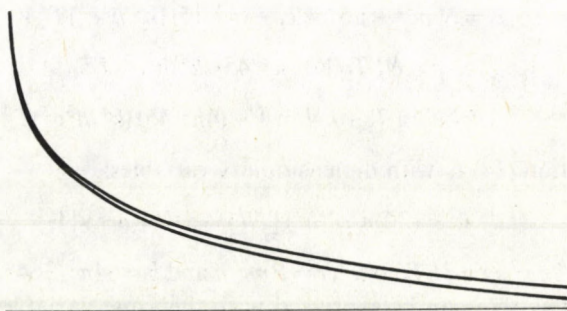


Fig. 3

This last remark deserves some attention. In scenario D the expansion started from and went through equilibrium states, so the entropy was constant (because it had already its maximum value). In scenario B one experienced a sequence of non-equilibrium states with irreversible process going on, consequently the entropy started from the same value as in scenario D, but later it increased! How is it possible for S to increase above the maximum value? The straightforward answer is that after the first moment different scenarios (different conductivity coefficients) led to different universes. According to scenario A the pressure behaves (Eqs (12), (18), 19)) like

$$p = p_0(q^5 + q^{4.2})/2, \quad (23)$$

where $q = R(0)/R(t)$ is the inverse expansion ratio. In scenario D it behaves like

$$P = p_0 q^{4.5}, \quad (24)$$

being smaller than in the earlier case (Fig. 3). In scenario B the pressure is somewhere between the values (23) and (24). If one starts the integrations from identical initial conditions, at a later time t the sizes of $R(t)$ will be different in different scenarios. (Smaller conductivity A , larger entropy S leads to faster expansion.)

Model of gas with radiation

Let us consider an expanding Einstein's world filled with ionized plasma and radiation. At a certain temperature neutral atoms will be formed, the (monoatomic) gas decouples from thermal radiation, both expand adiabatically, but independently of each other. The gas temperature T_1 deviates from the radiation temperature T_2 . This leads to irreversible phenomena.

This happens in our actual universe.

Our second model deals with such a two-component system, expanding against its own gravitational attraction, consequently cooling. The energy, pressure and entropy of the system are:

$$E = Nmc^2 + 1.5NkT_1 + (\pi^2/15)(k^4/\hbar^3c^3)T_2^4V, \quad (25)$$

$$p = NkT_1/V + (\pi^2/45)(k^4/\hbar^3c^3)T^3, \quad (26)$$

$$S = 1.5kN \ln T_1 + kN \ln V + (4\pi^2/45)(k^4/\hbar^3c^3)T_2^3V. \quad (27)$$

Einstein's equation (14) is with dimensionless variables:

$$\dot{r}^2 = (2Gm^3c^3/\hbar^3) [N(1+x)/r + br^2y^4], \quad (28)$$

where $x = 3kT_1/2mc^2$, $y = 3kT_2/2mc^2$, $r = Rmc/\hbar$ and $b = 64\pi^3/3645$. The entropy of this two-component system can be expressed with the same variables:

$$S/k = 1.5N \ln(xr^2) + 2b(yr)^3. \quad (29)$$

Now let us consider again three scenarios:

Scenario A: The neutral atoms do not interact with radiation, the two components cool adiabatically ($\lambda \sim R$), but due to their different scaling behaviour (energy quantum $= hc/\lambda$ for photons and $= h^2/8m\lambda^2$ for atoms) the nonrelativistic component cools at a much faster rate (Fig. 4A):

$$x \sim r^{-2}, \quad y \sim r^{-1}. \quad (30)$$

Each term of the entropy (29) remains constant.

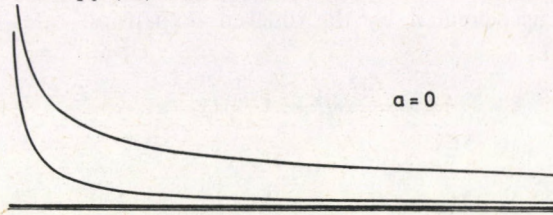


Fig. 4A

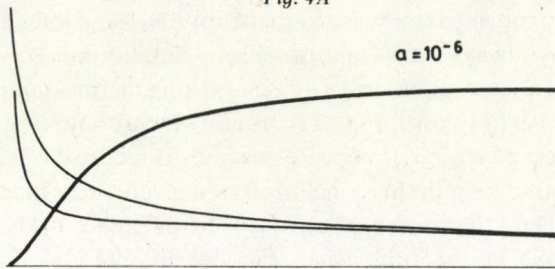


Fig. 4B

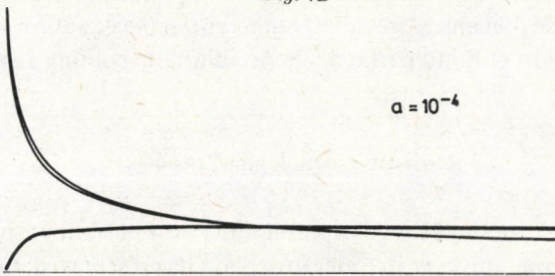


Fig. 4C

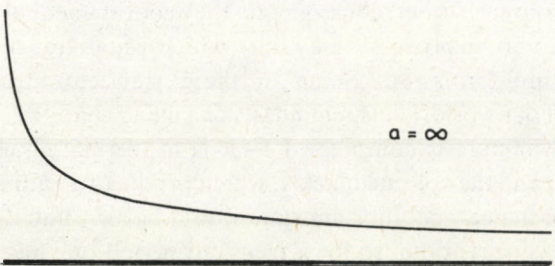


Fig. 4D

Scenario D: There are frequent collisions between the atoms and photons, so they have a common temperature ($x=y$). The $x(r)$ function can be obtained from the equation

$$S/k = 1.5N \ln(xr^2) + 2b(xr)^3 = \text{const.} \quad (31)$$

By substituting $x(r)$ into Eq. (28) and by integrating it numerically, one gets an expansion through equilibrium states (Fig. 4D).

Scenario B: There is now a weak heat transfer which tries to level up the temperature difference, created by the different expansion rates of the adiabatic cooling. If one writes

$$\dot{x} = -2xr\dot{r}/r + (A/N)(y-x), \quad (32)$$

$$\dot{y} = -yr\dot{r}/r - (A/4br^3y^3)(y-x), \quad (33)$$

the condition (4) of the Einstein's equations is satisfied, but the irreversible heat transfer from radiation to the cooler gas increases the entropy (29) (solid lines on Figures 4B–C). The integration may always start from isothermal equilibrium (let us say, from the formation of neutral atoms out of ions). In case of strong thermal coupling equilibrium states follow each other ($S = \text{const}$, Fig. 4D). In case of weak coupling non-equilibrium states emerge, leading to irreversible processes and to increase of entropy (Fig. 4C). Again, there is no paradoxon in this conclusion: at a given time t the actual size $R(t)$ of the universe depends on the chosen value of A . The difference in the expansion rates becomes considerable at the time when the densities of gas and radiation are comparable. No wonder: the escape from thermal equilibrium is a consequence of an interplay between the dynamical instability inherent in the equation of motion (3), (28) and the differentiation of matter (the different adiabatic cooling rates (30)).

Realistic cosmological models

In this Chapter we shall shortly recapitulate, how current cosmological models depict the history of our universe, in order to explain its present structure. We shall do it in order to be able to show that in certain time periods two components of cosmic material decouple from each other, consequently they cool at unequal rates, in this way irreversible non-equilibrium processes may start, which lead to transient processes and may enable structure formation. Some of these transient structures give the characteristic face of our world. (This reminds us of a cloud chamber. At the beginning the air–alcohol–gas mixture is compressed — it is in thermal equilibrium at room temperature. We expand the volume quickly, which produces an adiabatic cooling. At the low temperature the new equilibrium state would be air + liquid alcohol. But the formation of droplets needs time, so for a transient period one has air + overcooled alcohol vapour. This is just the short sensitivity period of the cloud chamber: if an

ionizing particle passes through the chamber, its trajectory will be made visible by the chain of droplets formed around the ions.)

In order to be specific, we use the standard cosmological model, complemented with the inflationary scenario of the Grand Unified Theory (GUT).

Bottleneck 1: Let us recall the formula (5). If $K < 0$ then $R(t) \rightarrow \infty$ is excluded. After a finite time expansion stops and then the world will collapse. The condition for the universe to be practically infinite in space and time can be taken from Eq. (5):

$$\rho = \rho_{\text{crit}}, \quad \text{where} \quad \rho_{\text{crit}} = (3/8\pi G) (\dot{R}/R)^2 = 3/8\pi G \tau^2. \quad (34)$$

The present expansion rate is known empirically: $\tau^{-1} = \dot{R}(t)/R(t) = 50$ km/s per megaparsec, leading to the value $\rho_{\text{crit}} = 10^{26}$ kg/m³. Astronomical evidence indicates that the actual value of ρ cannot differ from this critical value by orders of magnitude. Using formula (34) in Eq. (5), taking the solution (6) into account one gets

$$1 - \rho/\rho_{\text{crit}} = 2K/\dot{R}^2 \sim t.$$

If there was $|1 - \rho/\rho_{\text{crit}}| > 10^{-12}$, say, at 10^{-40} s when thermal equilibrium was first established, this deviation would have increased (proportionally to the age t of the world) to 10^{45} till today! If $\rho \approx \rho_{\text{crit}}$, there would be no galaxies on the sky. If $\rho \gg \rho_{\text{crit}}$, the world would have collapsed long ago. What could be the explanation that the universe was extremely fine tuned to $\rho = \rho_{\text{crit}}$, i.e. to $K = 0$? [4].

According to GUT the early universe thermalized itself by gravitational interactions within Planck time. At such high energies all particles were massless, all gauge interactions were of the same strength, matter behaved like black body radiation. (Its Gibbs free energy was zero.) The differentiation of matter into light and heavy particles, differentiation of forces to weak and strong interactions can be attributed to a phase transition: the freezing of the vacuum resulted in an order parameter $W > 0$. As a consequence different particles picked up different rest energies: $m_i c^2 = g_i W > 0$. During the vacuum phase transition a latent heat (Q per unit volume) was liberated which warmed the world to a temperature $T(Q \sim T^4)$. This heat may be the origin of the huge entropy what we observe now by radio telescopes in the form of microwave noise.

The cooling of the early universe was very fast, due to the rapid expansion (for small r values $\dot{r} \sim r^{-1}$ in Eq. (28)). But freezing of vacuum (to start formatting domains with non-vanishing order parameters) needed a certain time. So the universe was actually overcooled for a considerable period [6]. An energy density Q remained latent in the 'liquid vacuum' in spite of the low temperature of its particle content. If Q dominates over the Stefan Boltzmann energy density then Eq. (9) takes the form

$$(\dot{R}/R)^2 = 8\pi G Q / 3c^2 = \text{const},$$

i.e.

$$\dot{r}/r = 1/\tau_0 = \text{const}$$

with

$$\tau_0 = (3c^2/8\pi G Q),$$

giving the exponential solution

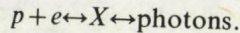
$$r(t) = r(t_0) \exp [(t - t_0)/\tau_0].$$

If the duration $t - t_0$ of overcooling is large with respect to the constant τ_0 , the universe might be inflated by many orders of magnitude. The inflation of r makes the last term negligible in Einstein's equation

$$(\dot{R}/R)^2 = 8\pi G\rho/3 + 2K/R,$$

which is equivalent of saying that after inflation K may practically be set zero. Conclusion: space and time enough for our world has been created by the thermodynamical disequilibrium between the (still liquid) vacuum and the (already cool) particles.

Bottleneck 2: Radiation has higher entropy than hydrogen gas of the same energy. Annihilation of protons plus electrons into photons would increase disorder in an irreversible way. According to GUT such transitions may go via the hypothetical X particle:



There are still protons around: one for 10^9 photons. (Luckily not too many. One per million would have made the universe recollapsed long ago. One per trillion would be not enough to start galaxy buildup.) How comes that in our cool world atoms did not annihilate long ago?

At $kT > m_x c^2$ quark matter and antiquark matter were in thermal equilibrium with X particles and radiation, which meant equal number of quarks and antiquarks. When kT dropped below $m_x c^2$, the equilibrium number $N_x \sim \exp(-m_x c^2/kT)$ of X particles became practically zero. The actual decay of X particles was, however, weak and slow, so they survived for a while in the cool environment. When they decayed in an irreversible way, they produced a tiny excess of quarks over antiquarks, due to the observed slight inherent asymmetry of weak interactions. If we assume that at this time X was one sort out of 10^3 sorts of particles, and if N was the overall number of particles, then $N_x = 10^{-3} N$. If the semileptonic branching ratio is assumed to be 10^{-3} and if the asymmetry of weak decay is 10^{-3} , then the number of quark excess produced by the decay of X particles is $B = 10^{-3} \cdot 10^{-3} \cdot N_x = 10^{-9} N$. Later on antiquarks annihilated with quarks, only the excess quarks survived. Almost all particles decayed into photons, their entropy has been transferred to what is called today photon entropy. This means that the present photon number is $\approx N$. Thus the present ratio of baryonic charge to photon number can be estimated to be $B/N \sim 10^{-9}$, in acceptable agreement with the empirical evidence. The universe has cooled down so fast, that a tiny fraction of its material got trapped in the quark state (of low entropy). Quarks may decay into radiation (of high entropy) only by tunnelling slowly through the X barrier (Fig. 5). Their decay lifetime is estimated to be longer than 10^{30} years. This is why our present world (stuffed mainly by photons) still contains some quarks (condensed into protons and neutrons) as overcooled transient relics from the hot $kT > m_x c^2$ era. It contains

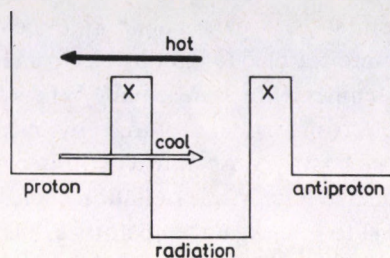


Fig. 5

protons enough to make stars and planets, but not enough to terminate world history too early by a collapse.

Bottleneck 3: After quark–antiquark pair annihilation one had mainly photons and leptons in equilibrium due to electromagnetic and weak interactions. When the densities and energies became low, weak collision time became longer than the age of the universe, so neutrinos got decoupled from the electron–positron–photon plasma. Both components cooled adiabatically according to the law $T \sim \lambda^{-1} \sim R^{-1}$.

When kT dropped below $2m_e c^2 = 1$ MeV (after about 1 s), electrons and positrons annihilated into photons, in this way they transferred their entropy to the electromagnetic radiation, increasing the photon temperature:

$$S = S_{\text{rad}} + S_{e^-} + S_{e^+} = \frac{4\pi^2}{45} \left(1 + \frac{7}{8} + \frac{7}{8} \right) \frac{k^4 T_{\text{old}}^3}{\hbar^3 c^3} V = \frac{4\pi^2}{45} \frac{k^4 T_{\text{new}}^3}{\hbar^3 c^3} V$$

giving $T_{\text{new}} = T_{\text{old}} (11/4)^{1/3}$. Since this era the neutrinos became a bit cooler than photons: $T_\nu = (4/11)^{1/3} T_{\text{rad}}$. The present temperature of weakly interacting massless particles must be lower than the measured temperature of photons, $T_{\text{rad}} = 2.7$ K. ‘Cold’ particles are not warmed up because weak collisions like $\nu + e \rightarrow e + \nu$ are very rare.

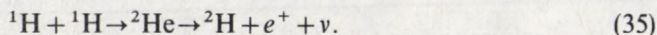
Bottleneck 4: In the cooling universe the quark excess condensed to neutrons and protons. Their number ratio was given by Boltzmann’s formula:

$$N(n)/N(p) = \exp [-(M_n - M_p)c^2/kT],$$

i.e. practically one while $kT > (M_n - M_p)c^2 = 0.75$ eV. The temperature dropped below this value after the first second. At lower temperatures the equilibrium number of neutrons is practically zero, but their actual decay into protons $n \rightarrow p^+ + e^- + \nu$ is a weak and slow process, taking about 10 minutes. Free neutrons survived for a few minutes in a cool world where nuclear binding already became possible. This is again a non-equilibrium situation! Some of the neutrons got captured by protons, making light nuclei, others decayed. When the free neutrons disappeared, the nuclear buildup stopped before establishing an equilibrium distribution of chemical elements. Heavy hydrogen and lithium are excellent nuclear (fusion) fuels, so they could not originate from hot stars, only in the cool outer space, when free neutrons were still present. Their further fusion is prevented by the mutual Coulomb repulsion of the positive nuclei.

Bottleneck 5: In the first 10^5 years, above $kT > 1$ eV gases were completely ionized. The charged ions were coupled to light by electromagnetic interaction, so ion temperature and photon temperature were equal. When the world cooled below ionization temperature, electrons were captured by nuclei. The neutral atoms decoupled from electromagnetic waves. As shown in our second model, nonrelativistic gas cools like R^{-2} , radiation like R^{-1} . While radiation cooled from 10 000 K to 100 K, the gas cooled from 10 000 K to 1 K. Again temperature difference was created! In the cool hydrogen gas (slightly contaminated by light elements) statistical fluctuations produced mass concentrations. If these clouds were large enough, their excess gravity helped them to survive against thermal motion. Unluckily this happened too late: the gas density in the universe was too low to make galaxies. The galaxy formation might start earlier if there were cold (weakly interacting) nonrelativistic particles (like massive neutrinos [7]) around. Because they were not coupled to electromagnetic radiation, they started the faster nonrelativistic cooling $T \sim R^{-2}$ much earlier, reaching a very low temperature till today. These slow particles (affected by gravity) might build up the first astronomical objects (neutrino superstars?) before the formation of neutral atoms. The gravitational fields of these (neutrino?) superstars were already present when the neutral hydrogen gas appeared. The neutral atoms cooled and fell into the gravitational wells of (neutrino?) superstars, forming the first hydrogen clouds, which later differentiated into galaxies and stars [2, 8, 9, 14].

Bottleneck 6: The present atomic universe contains today 75% hydrogen, these simplest nuclei, in spite of the fact that iron offers the deepest binding energy per particle. Nuclear material was trapped in hydrogen (characteristic for a hot thermal universe) because the fast cooling and the electric repulsion among positive nuclei prevented reaching nuclear equilibrium appropriate for a cold world. Later the contraction work of gravity heated up some fragments of matter (called stars) to several million degrees, offering a new chance for nuclear fusion. The first step would be the ${}^1\text{H} + {}^1\text{H} \rightarrow {}^2\text{He}$ reaction, but ${}^2\text{He}$ isotope does not exist. (This is why the Sun does not explode like an H bomb.) The only slight chance for lasting fusion is a simultaneous weak decay:



From now on fusion may run via strong interactions, but the 'weak' bottleneck at the first step (35) slows down the nuclear fusion chain from microseconds to billions of years, offering ample time for biological evolution.

The Ultimate Bottlenecks: Our universe is about 16 billion years old. It is made mainly of thermal radiation, but some surviving protons, light elements, stars and planets give it the shape we know. There exists a thermodynamical arrow of time. How long will these transient disequilibrium phenomena last? This is under discussion in recent literature [1, 5].

Nuclear material will be transformed into iron, stars will become black dwarfs, neutron stars or black holes in 10^{10} or 10^{12} years.

Protons may decay into positrons and radiation, stars and planets may crumble away through X tunnelling in 10^{30} to 10^{32} years.

Positrons (of proton decay) and electrons (left over from atoms), massive neutrinos and antineutrinos will annihilate into radiation in 10^{40} to 10^{80} years (if leptons remain bound in (neutrino?) superstars).

Even black holes may annihilate in 10^{70} — 10^{90} years.

In about 10^{100} years matter may turn 'back' into radiation (without rest mass, possessing 'maximum' entropy), if the expansion is not too fast ($K=0$). (Some fast expanding scenarios with $K < 0$ prolong the existence of some massive particles ad infinitum [5].)

Conclusion

The discussion of some simplified and other more realistic models has shown that in order to be able to explain the (transient) deviation of the universe from thermal equilibrium, beside the inherent instability of Einstein's equation one has to take into account the (transient) differentiation of matter: the emergence of 'weak' forces, which can make the rate of thermalization among different components of matter slower than the creation of thermal differences by adiabatic expansion.

Gravity created stars and planets. Strong interactions created protons, neutrons, nuclei: they stabilized the chemical elements. Electromagnetic interactions created and stabilized atoms, molecules, solids. And weak interactions created the arrow of time. All together have shaped our world as we watch it today.

Appendix: Computations

The equations of the Gas Model (16), (20), (21) were rewritten for computation as

$$\begin{aligned} r_1 &= r_0 + d[N(1 + 3x_0 + 5y_0)/r_0]^{1/2}, \\ x_1 &= x_0(r_0/r_1)^2, \quad y_1 = y_0(r_0/r_1)^{1/2}, \\ x_2 &= x_1 + (ad/3)(y_1 - x_1), \quad y_2 = y_1 - (ad/5)(y_1 - x_1), \end{aligned}$$

where $d = g \cdot \Delta t$. By starting with e.g. $N = 1$, $x = y = 0.4$, $d = 0.03$ one gets at $t = 10$ the following values:

$r = 7.032$	and	$S - S(0) = (0)$	if	$a = 0$
$r = 7.008$		222		1
$r = 6.991$		92		4
$r = 6.983$		39		10
$r = 6.978$		8		40

showing how different heat conductivity coefficients lead to different universes of different (actual and 'maximum') entropy values at a given later time. (As a rule, higher

entropy leads to faster expansion and larger size.) These initial conditions were used in Fig. 2, the two temperatures (thin lines) and the entropy change (solid curve) were plotted for the time interval 1–40 (139 steps). Fig. 2A is for $a=0$, Fig. 2B for $a=3$, Fig. 2C for $a=30$, Fig. 2D for $a=\infty$.

The equations of the Gas + Radiation Model were rewritten for computation as

$$r_1 = r_0 + d[N(1 + x_0)/r_0 + by_0^4 r_0^2]^{1/2},$$

$$x_1 = x_0(r_0/r_1)^2, \quad y = y_0 r_0/r_1,$$

$$x_2 = x_1 - (ad/N)(y_1 - x_1), \quad y_2 = y_1 + (ad/4br_1 y_1)(y_1 - x_1)$$

(with $ad = A\Delta t$). In our computation we used the initial conditions $N = 0.0001$, $x = y = 0.3$, $d = 5$, $r = 1$. (At these initial conditions the masses of gas and radiation were comparable, so the interplay of the two components could manifest itself in creating temperature difference and irreversibility.) At $t = 1000$ (after 200 steps) this gave the following values:

$r = 12.096$	and	$S - S(0)$	if	$a = 0$
$r = 12.094$		173		10
$r = 12.087$		108		10
$r = 12.089$		27		10
$r = 12.052$		1		10

It can be seen how higher entropy is related to faster expansion. The same initial conditions were used in Fig. 4, when the two temperatures and the entropy curve was plotted in the time interval 0–700 (in 139 steps). Fig. 4A has used $A = 0$, Fig. 4B has used $A = 10$, Fig. 4C has used $A = 10$, Fig. 4D has used $A = \infty$.

References

1. J. Balog and G. Marx, *Acta Phys. Hung.*, 58, 33, 1985.
2. R. Cowsik and J. McClelland, *Phys. Rev. Lett.*, 29, 669, 1972.
3. P. C. V. Davies in "Entropy and Information", Acad. Prague 1, 1975. T. Gold, *Am. Journal of Phys.*, 30, 403, 1962; J. V. Marlikar and F. Hoyle, *Nature*, 222, 1040, 1969; P. T. Landsberg, *Nature*, 225, 1204, 1969; David Layzer, *Scientific American*, 233/6, 56, 1975; E. L. Schucking and E. A. Spiegel, *Comm. Astrophys. Space Sci.*, 11, 121, 1970; P. C. V. Davies, *The Runaway Universe*, J. M. Dent, London, 115, 1978; R. U. Sexl, "Origin of Order in the Universe", ICUS, Miami, 1985.
4. P. J. Peebles et al., *Astrophys. J.*, 184, 467, 1943.
5. F. Dicus et al., *Astrophys. J.*, 252, 1, 1982.
6. A. Gueth, *Phys. Rev. Letters*, 44, 631, 963, 1980.
7. V. Lubimov et al., *Physics Lett.*, 948, 226, 1980.
8. G. Marx and A. S. Szalay, *Proc. Neutrino '72*, Budapest, 191, 1972.
9. G. Marx, *Proc. Neutrino '76*, Aachen, 643, 1976.
10. G. Marx, *Proc. Electroweak Interactions at DESY*, World Sci., 159, 1983.
11. G. Marx, *Fortschritte der Physik*, 32, 185, 1984.
12. G. Marx, *Proc. EPS General Assembly, JEDNOTA*, Prague, 247, 1983.
13. P. J. Peebles, *Physical Cosmology*, Princeton University Press, 1971.
14. A. S. Szalay and G. Marx, *Astron. and Astrophys.*, 49, 437, 1976.
15. Ya. B. Zeldovich, *Astron. and Astrophys.*, 5, 84, 1980.

ENERGY-ENERGY CORRELATIONS FOR THE $p\bar{p}$ COLLIDER*

F. CSIKOR and G. PÓCSIK

*Institute for Theoretical Physics, Roland Eötvös University
1088 Budapest, Hungary*

Hadronic energy-energy correlations (EEC) for the $p\bar{p}$ collider are defined and calculated in QCD in the colliding parton center of mass frame of each event. Normalized EEC calculated from only those events for which the parton and the $p\bar{p}$ center of mass frames coincide is also investigated. It is shown that in both cases the normalized EEC are quite insensitive to the choice of parton density parametrizations and may thus serve as reasonable tests of QCD.

Jet physics for the $p\bar{p}$ collider has been a very successful field in the last years [1]. Experimentally determined inclusive jet yields reasonably agree with the QCD predictions. The clear emergence of two jet events makes possible a direct study of the underlying parton-parton interaction, too. On the other hand multijet final states are not yet studied in detail. Only well separated three jet systems were studied by measuring p_{out} distributions (where p_{out} is the momentum perpendicular to the plane defined by the colliding $p\bar{p}$ and the highest transverse momentum jet.) To study more complicated final states it is natural to take over experience obtained in e^+e^- annihilation.

In this note we discuss the ways energy-energy correlations (EEC) — which have proven to be very useful in e^+e^- annihilation — may be defined for the $p\bar{p}$ collider. Though next to leading order calculations are not expected to be available in the near future, we mention that EEC are properly defined, “infrared safe” cross sections, which are finite in higher order QCD [2]. EEC are sensitive to two jets only through jet fragmentation, while the first real contribution comes from three jet final states. On the other hand, measuring EEC it is not necessary to set up a specific jet definition, EEC is simply defined in terms of observed hadron energies. The EEC cross section is not a Lorentz invariant quantity. To be as close as possible to the e^+e^- annihilation case, we shall always define EEC in the colliding parton center of mass frames of each event. A different definition of EEC the so called transverse EEC is given in [3].

We define the normalized energy-energy correlation function as

$$F(\cos \chi) = \frac{1}{N} \sum_{A=1}^N \frac{1}{\Delta\chi \sin \chi} \sum \frac{E_{Aa} E_{Ab} (2 - \delta_{ab})}{(E_A)^2}, \quad (1)$$

* Dedicated to Prof. K. Nagy on his 60th birthday

where the label A specifies the event $A = 1, \dots, N$ and the energies and angles are measured in the center of mass frames of the colliding partons. This frame will be identified with the rest frame of the (hard scattered) final state hadrons. The first sum is over the accepted events (see later), while the second sum is over all pairs of accepted hard scattered hadrons a, b , whose momenta are at relative angles between χ and $\chi + \Delta\chi$. E_{Aa}, E_{Ab} are the energies of particles a, b and E_A is the total energy of the (hard scattered) final state hadrons in event A . $F(\cos \chi)$ satisfies the usual normalization condition: $\int_{-1}^1 F(\cos \chi) \alpha \cos \chi = 1$.

The accepted events/particles may be defined in several ways resulting in different EEC functions. For obvious reasons, we impose the usual pseudorapidity cuts, i.e. $|y| < y_{\max}$ on the accepted hadrons. Second, we require that the total transverse energy E_T of the accepted hadrons be larger than E_{\min} . This requirement should ensure that the event is a hard scattering one. The EEC function defined with these requirements we shall call $F_1(\cos \chi)$. Thinking on the parton level, it is the three parton final state, which first contributes. It might happen e.g. that in a given three parton final state only two partons lie in the accepted phase space region. Such events will contribute only at $\chi = 180^\circ$ (and by self correlation at 0°), since the Lorentz transformation to the rest frame of the detected partons makes the event back to back. Since the fixed order QCD prediction is reliable only for $\delta < \chi < 180^\circ - \delta$ ($\delta \approx 30^\circ$) such events do not contribute in the interesting χ range. Alternatively, one might exclude events, which have a large total transverse momentum.

Determining $F_1(\cos \chi)$ one has to make a Lorentz transformation for each accepted particle into the rest frame of the final state hadrons. To avoid this one may accept only those events for which the colliding parton center of mass frame coincides with the $p\bar{p}$ center of mass frame. The EEC function defined in this way will be called $F_2(\cos \chi)$.

The lowest order (leading log) calculation of the EEC function $F_1(\cos \chi)$ is in principle straightforward. We have to calculate the convolution of the structure functions with the $2 \rightarrow 3$ subprocess cross sections (known from [4]), insert a δ function to keep the relative angle of the detected partons fixed, multiply by the appropriate scaled energy factors, and sum up for all pairs of final partons. To lowest order the normalization cross section (σ_{norm}) is given by a similar expression involving the $2 \rightarrow 2$ subprocess cross sections [5]. Schematically

$$F_1(\cos \chi) = \frac{\sum_{a_i b_i} G_{a_1|p}(x_1, Q^2) * G_{a_2|\bar{p}}(x_2, Q^2) * \hat{\sigma}^{a_1 a_2 \rightarrow b_1 b_2 b_3}}{\sum_{a_i b_i} G_{a_1|p}(x_1, Q^2) * G_{a_2|\bar{p}}(x_2, Q^2) * \hat{\sigma}^{a_1 a_2 \rightarrow b_1 b_2}} \quad (2)$$

To calculate the EEC function $F_2(\cos \chi)$ one has to proceed similarly, keeping however the momenta of the colliding partons equal in magnitude.

Since the normalized EEC functions $F_i(\cos \chi)$ are ratios of integrals both of which contain the parton densities $G_{a_1|p}$, $G_{a_2|\bar{p}}$, it is expected that the sensitivity to the parton density choice will be less than e.g. for the inclusive jet cross section case (where it is very large, see e.g. [6]). To study this expectation (following [7]) we have used two different parametrizations: the Owens-Reya (OR) [8] with $\Lambda=0.5$ GeV and the Glück-Hofmann-Reya (GHR) [9] parametrization with $\Lambda=0.4$ GeV. Following [7], in the OR case we have put $Q=p_T$ and in the GHR case $Q=2p_T$ in the denominator of Eq. (2) (i.e. σ_{norm}), while for the numerator (involving the $2 \rightarrow 3$ subprocesses) we have used $Q=\bar{p}_T$ (OR) and $Q=2\bar{p}_T$ (GHR). (\bar{p}_T is the average transverse momentum of the three final partons.) Our results are shown on Figs 1–4 for $\sqrt{s}=540$ GeV.

Fig. 1 shows $F_1(\cos \chi)$ for different values of y_{max} and different parton density parametrizations. The sensitivity to the parton density choice is 10–20%, small indeed. Fig. 2 shows the E_{min} dependence of $F_1(\cos \chi)$ at $\chi=90^\circ$. It is remarkable that F_1 is quite insensitive to E_{min} . The reason is again that both the numerator and the denominator in Eq. (2) depend on E_{min} in approximately the same way. Fig. 3 shows the E_{min} dependence of the normalizing cross section σ_{norm} (the denominator in Eq. (2)), which of course decreases fast with increasing E_{min} .

Fig. 4 shows $F_2(\cos \chi)$ for different values of y_{max} and different parton density parametrizations. The sensitivity to the parton density choice is again 10–20%.

Figs 1–4 show the lowest order (leading log) QCD results. The pure QCD predictions will be modified, when parton fragmentation is taken into account. Experience obtained with e^+e^- annihilation shows that fragmentation corrections are not negligible even at energies comparable to the Z^0 mass [10]. EEC are sensitive to the choice of fragmentation models applied. This sensitivity — though a great, important challenge — is not very large, it is usually less than 20%. Referring to the e^+e^- annihilation experience we do not expect larger sensitivity in the $p\bar{p}$ EEC case either. Also it is clear that the EEC asymmetry, (defined as $A(\cos \chi)=F(-\cos \chi)-F(\cos \chi)$) is less sensitive to the 2 parton final state fragmentation, than the EEC itself. To take into account fragmentation corrections in the $p\bar{p}$ case (and check the above expectations) the only possibility is to perform a complete M.C. calculation with one (or several) of the existing fragmentation models. This is, however, outside the scope of the present note.

Another source of correction is that our definitions of accepted events/particles will necessarily include some low energy hadrons belonging to the spectator jets. Because of the presence of the energy factors in the definition, we estimate that this correction may be small [11].

In summary, we have defined two kinds of parton center of mass frame normalized energy–energy correlation functions. Both of them (and $F_2(\cos \chi)$ more easily) could be determined from existing collider data. A precise comparison with QCD predictions (Figs 1–4) will be possible if fragmentation corrections are properly included. Since fragmentation corrections change both the numerator and denominator of Eq. (2), we expect that the net effect will be smaller than in the e^+e^-

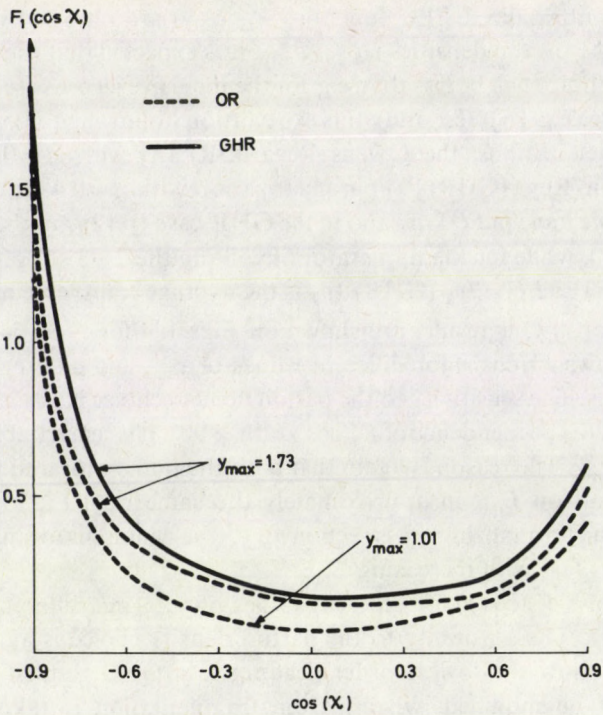


Fig. 1. The EEC function $F_1(\cos \chi)$ for different y_{\max} cuts and parton density parametrizations for $E_{\min} = 40$ GeV

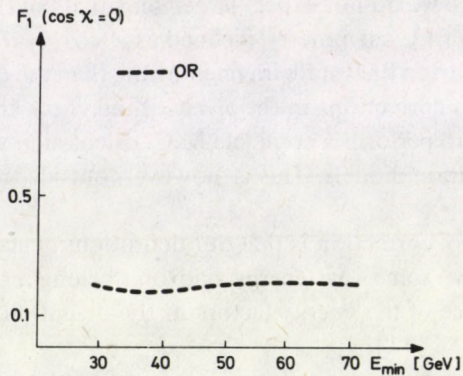


Fig. 2. E_{\min} dependence of $F_1(\cos \chi = 0)$ with OR [8] parton density parametrization and $y_{\max} = 1.73$

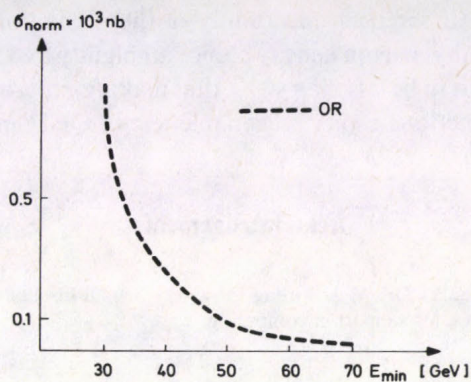


Fig. 3. E_{\min} dependence of the normalizing cross section of σ_{norm} of $F_1(\cos \chi)$ with OR [8] parton density parametrization and $y_{\text{max}} = 1.73$

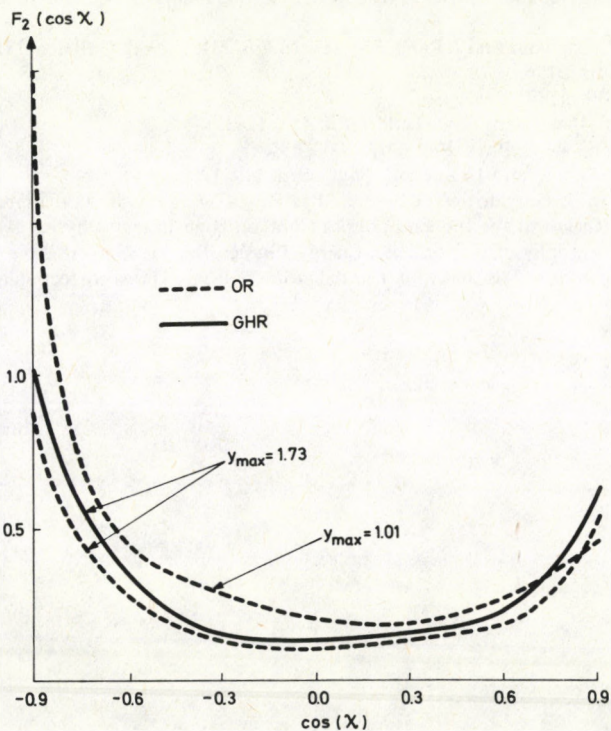


Fig. 4. The EEC function $F_2(\cos \chi)$ for different y_{max} cuts and parton density parametrizations for $E_{\min} = 40$ GeV

annihilation case. The theoretical uncertainty of the prediction will be quite small (including scale ambiguity, parton density choice ambiguity and fragmentation model sensitivity, we estimate it to be less than 40%), thus making experimental determination of normalized EEC functions a very reasonable test of QCD in $p\bar{p}$ collisions.

Acknowledgement

One of us (F. Cs.) thanks Dr. W. J. Stirling for a very helpful discussion. We thank the Central Research Institute for Physics, for support in computing.

References

1. M. Jacob, Proc. of the XXII. Int. Conf. on High Energy Physics, Leipzig 16–25 July, 1984, Vol II. p. 150.
2. C. L. Basham, L. S. Brown, S. D. Ellis and S. T. Love, Phys. Rev. Lett., *41*, 1585, 1978; Phys. Rev., *D17*, 2298, 1978; *D19*, 2018, 1979; L. S. Brown and S. D. Ellis, Phys. Rev., *D24*, 2383, 1981.
3. A. Ali, E. Pietarinen and W. J. Stirling, Phys. Lett., *141B*, 447, 1984.
4. J. Kripfganz and A. Schiller, Phys. Lett., *79B*, 317, 1978; Z. Kunszt and E. Pietarinen, Nucl. Phys., *B164*, 45, 1980; T. Gottschalk and D. Sivers, Phys. Rev., *D21*, 102, 1980; F. A. Berends et al., Phys. Lett., *103B*, 124, 1981.
5. B. Combridge, J. Kripfganz and J. Ranft, Phys. Lett., *70B*, 234, 1977; R. Cutler and D. Sivers, Phys. Rev., *D16*, 679, 1977; *D17*, 196, 1978.
6. B. Humpert, Z. Phys., *C27*, 257, 1985.
7. Z. Kunszt and E. Pietarinen, Phys. Lett., *132B*, 453, 1983.
8. J. F. Owens and E. Reya, Phys. Rev., *D17*, 3003, 1978.
9. M. Glück, E. Hoffmann and E. Reya, Z. Phys., *C13*, 119, 1982.
10. F. Csikor, G. Eszes, L. Garrido and G. Pócsik, Phys. Rev., *D34*, 129, 1986; a preliminary version has been presented by F. Csikor at the Int. Europhysics Conf. on High-Energy Physics, Bari, Italy, 18/24 July 1985, Proc. Int. Europhys. Conf. on High-Energy Physics, Bari, 1985, p. 182.
11. Dr. W. J. Stirling informed us that with their definition of EEC [3] this correction is fairly large despite the energy factors.

SCALAR BETHE–SALPETER EQUATION AND THE RELATIVISTIC BOUND STATE PROBLEM*

K. LADÁNYI

*Institute of Theoretical Physics, Roland Eötvös University
1088 Budapest, Hungary*

The structure of the scalar Bethe–Salpeter equation is studied in the coordinate space. Expansion methods are applied to derive a generalized matrix eigenvalue problem. The standard truncation procedure is modified by using a least-squares variational method.

1. Introduction

The Bethe–Salpeter (BS) equations [1–3] offer an efficient theoretical tool to solve relativistic bound state problems. Important progress has been made by using the ladder approximation along the lines suggested by Wick [4], Cutkosky [5], Schwartz [6], and Schwartz and Zemach [7]. (For reviews and early developments see [8–10].) Generalizations of the standard ladder approximation have been presented in order to analyse the asymptotic behaviour of the spinor BS wave functions at short distances [11–14]. The short-distance properties of the BS wave functions lead to a number of predictions for the covariant observables, such as structure functions, distribution amplitudes, form factors, anomalous magnetic moments, and other hadronic properties [15–19]. In addition, the short-distance analysis is a crucial ingredient in bound-state calculations which involve the selection of the admissible asymptotic solutions by using the BS normalizability condition [8, 20].

In the present paper we investigate nonperturbative expansion methods which can be applied to the solution of relativistic bound-state problems. The BS equation will be formulated in the space of a complete set of normalized basis functions. In this way, we arrive at a generalized matrix eigenvalue problem which is ideally suited for machine computations, since it gives a systematic procedure for generating a sequence of higher approximations. Applications of the Rayleigh–Ritz method [21] will be discussed. In addition, we suggest a least-squares variational method based on a convenient error functional [22]. A similar procedure has been already applied to avoid the spurious solutions which may appear in the application of the standard basis set expansion methods to the Dirac eigenvalue problem [23]. (The least-squares method can also be extended to the scattering theory [24–27].) As an illustrative step, we next consider the scalar BS equation of unequal-mass systems in the bound-state region.

* Dedicated to Prof. K. Nagy on his 60th birthday

2. Ladder approximation

Let us start by writing down the simple interaction Lagrangian

$$L_I = -\frac{g'}{2} [:\phi_1^2\varphi: + :\phi_2^2\varphi:]. \quad (2.1)$$

Here the quantum fields $\phi_1(x)$, $\phi_2(x)$, and $\varphi(x)$ generate scalar particles of rest masses m_1 , m_2 , and κ , respectively. The definition of the covariant BS wave function τ_P is the following

$$\tau_P(x_1, x_2) = \langle 0 | T \phi_1(x_1) \phi_2(x_2) | \Phi_P \rangle, \quad (2.2)$$

where $|\Phi_P\rangle$ is the Heisenberg state of the composite system which can be characterized by total four-momentum P_ν . The ladder approximation of the wave function τ_P satisfies the BS equation

$$\begin{aligned} (\square_1 + m_1^2) \tau_P(x_1, x_2) (\square_2 + m_2^2) = \\ = g'^2 \frac{1}{i} D^c(z; \kappa^2) \tau_P(x_1, x_2), \end{aligned} \quad (2.3)$$

with $z = x_1 - x_2$ and

$$D^c(z; \kappa^2) = \frac{1}{(2\pi)^4} \int d^4q \frac{\exp(iq^\mu z_\mu)}{\kappa^2 - q^2 - i\epsilon}. \quad (2.4)$$

Translational invariance implies

$$\tau_P(x_1, x_2) = \varphi_P(z) \exp[-iP_\nu(\mu_1 x_1^\nu + \mu_2 x_2^\nu)] \quad (2.5)$$

with the restriction $\mu_1 + \mu_2 = 1$. The calculations will be carried out in the center-of-mass (c.m.) coordinate frame by choosing

$$P_0 = E, \quad P_j = 0 \quad (j = 1, 2, 3), \quad (2.6)$$

where E is the total c.m. energy of the system.

The Wick-rotated relative BS wave function $\psi(x)$ is given by [4, 8]

$$\psi(x) = \varphi_P(-ix_4, x_j). \quad (2.7)$$

Here the vectors z^μ have been rotated to the real Euclidean values $x_j = z^j$ ($j = 1, 2, 3$) and $x_4 = iz^0$.

We shall apply the following notations

$$\square = \partial_\nu \partial_\nu = \sum_{\nu=1}^4 \partial^2 / \partial x_\nu^2, \quad (2.8)$$

$$R = (x_\nu x_\nu)^{1/2}, \quad x_\nu x_\nu = \sum_{\nu=1}^4 x_\nu^2. \quad (2.9)$$

The BS equation of the Wick-rotated wave function $\psi(x)$ can be derived by using Eqs (2.3)–(2.9). One obtains

$$[F(x) - g^2 V(R; \kappa^2)]\psi(x) = 0, \quad (2.10)$$

where $g = g'/4\pi$,

$$F(x) = \left[-\square + 2\mu_1 E \frac{\partial}{\partial x_4} - \mu_1^2 E^2 + m_1^2 \right] \times \\ \times \left[-\square - 2\mu_2 E \frac{\partial}{\partial x_4} - \mu_2^2 E^2 + m_2^2 \right], \quad (2.11)$$

and

$$V(R; \kappa^2) = \frac{4\kappa}{R} K_1(\kappa R). \quad (2.12)$$

(K_1 is the first-order modified Bessel function.) In the short-distance limit we get

$$V(R; \kappa^2) \rightarrow 4R^{-2} + 2\kappa \ln R \quad \text{as } R \rightarrow 0. \quad (2.13)$$

The numerical calculations can be simplified by applying the Vosko approximation of the interaction kernel $V(R; \kappa^2)$ [28]. This approximation is of the form

$$V(R; \kappa^2) \sim \frac{4}{1+\bar{a}} R^{-2} [(1+\kappa R)e^{-\kappa R} + \bar{a}(1+2\kappa R)e^{-2\kappa R}], \quad (2.14)$$

where $\bar{a} = 0.66746$. In the subsequent part of this paper we shall choose the convention

$$\mu_1 = \mu_2 = \frac{1}{2}. \quad (2.15)$$

In addition, the mass scale will be fixed by

$$m_1 = 1 + \bar{\delta}, \quad m_2 = 1 - \bar{\delta}, \quad |\bar{\delta}| < 1. \quad (2.16)$$

The BS equation of equal-mass systems ($\bar{\delta} = 0$) becomes

$$[F_0(x) - g^2 V(R; \kappa^2)]\psi(x) = 0, \quad (2.17)$$

where

$$F_0(x) = \left(-\square + 1 - \frac{E^2}{4} \right)^2 - E^2 \frac{\partial^2}{\partial x_4^2}. \quad (2.18)$$

Notice that Eq. (2.17) leads to the Wick–Cutkosky model [4, 5, 8] if the exchange mass κ is zero.

3. Separation of the angular variables

a) Expansion of the scalar BS wave function in partial wave amplitudes of $O(4)$

We shall use the angular variables θ , ϑ , and φ as defined by

$$x_1 = R \sin \theta \sin \vartheta \sin \varphi, \quad x_2 = R \sin \theta \sin \vartheta \cos \varphi, \quad (3.1)$$

$$x_3 = R \sin \theta \cos \vartheta, \quad x_4 = R \cos \theta. \quad (3.2)$$

One obtains

$$r = R \sin \theta, \quad r = (x_1^2 + x_2^2 + x_3^2)^{1/2}. \quad (3.3)$$

The BS equation (2.10) is invariant under rotations in the three-dimensional subspace, and the wave functions $\psi(x)$ can be characterized by the total angular momentum quantum numbers J and M . Thus, we have

$$\psi(x) = \psi_{JM}(x) = \Phi_J(r, x_4) Y_{JM}(\vartheta, \varphi). \quad (3.4)$$

Here Y_{JM} is the three-dimensional spherical harmonic, and the wave function Φ_J can be written as

$$\Phi_J(r, x_4) = \tilde{\Phi}_J(R, \theta). \quad (3.5)$$

As a first step, we consider the BS wave functions at vanishing c.m. energy E . In this limit the BS equation (2.10) is invariant under the transformations of the four-dimensional rotation group $O(4)$. Consequently, the BS wave functions $\psi(x)$ have the following form [29]

$$\psi(x) = f_{N-J}(R) Y_{NJM}(\Omega) \quad \text{for } E \rightarrow 0, \quad (3.6)$$

$$N - J = n = 0, 1, 2, \dots \quad (3.7)$$

Here $f_n(R)$ is a radial BS wave function, and the four-dimensional scalar spherical harmonics $Y_{NJM}(\Omega)$ involve the Gegenbauer polynomials [30] $C_k^\lambda(\cos \theta)$ as given by

$$Y_{NJM}(\Omega) = Y_{NJM}(\theta, \vartheta, \varphi) = G_N^{(J)}(\theta) Y_{JM}(\vartheta, \varphi), \quad (3.8)$$

$$G_N^{(J)}(\theta) = \left[\frac{2^{2J+1} (N+1) (N-J)!}{\pi (N+J+1)!} \right]^{1/2} J! C_{N-J}^{J+1}(\cos \theta) \sin^J \theta. \quad (3.9)$$

Let us turn to the analysis of the BS wave functions at finite c.m. energies E . In this case, the $O(4)$ invariance of the BS equation is broken according to Eq. (2.11). Therefore, we shall expand the wave functions $\psi(x)$ in partial wave amplitudes of $O(4)$. The truncated version of this expansion can be written as [6, 29]

$$\psi(x) = \psi_{JM}(x) = \sum_{N=J}^{K+J} f_{N-J}(R) Y_{NJM}(\Omega). \quad (3.10)$$

Relations (3.4)–(3.5) together with Eqs (3.8) and (3.10) imply

$$\tilde{\Phi}_J(R, \theta) = \sum_{N=J}^{K+J} f_{N-J}(R) G_N^{(J)}(\theta). \quad (3.11)$$

In summary, the approximate solutions of the BS equation (2.10) can be obtained by calculating a finite set of radial BS wave functions $f_n(R)$.

b) Relations

It is convenient to introduce the following notations

$$d(-1; N, R) = \frac{d}{dR} - \frac{N}{R}, \quad (3.12)$$

$$d(1; N, R) = \frac{d}{dR} + \frac{N+2}{R}, \quad (3.13)$$

$$d(-2; N, R) = \frac{d^2}{dR^2} - \frac{2N+1}{R} \frac{d}{dR} + \frac{N(N+2)}{R^2}, \quad (3.14)$$

$$d(0; N, R) = \frac{d^2}{dR^2} + \frac{3}{R} \frac{d}{dR} - \frac{N(N+2)}{R^2}, \quad (3.15)$$

$$d(2; N, R) = \frac{d^2}{dR^2} + \frac{2N+3}{R} \frac{d}{dR} + \frac{N(N+2)}{R^2}. \quad (3.16)$$

Integration by parts gives

$$\int_0^\infty dRR^3 f_m^*(R) d(-1; N, R) f_n(R) = - \int_0^\infty dRR^3 f_n(R) d(1; N+1, R) f_m^*(R), \quad (3.17)$$

$$\int_0^\infty dRR^3 f_m^*(R) d(-2; N, R) f_n(R) = \int_0^\infty dRR^3 f_n(R) d(2; N+2, R) f_m^*(R), \quad (3.18)$$

$$\int_0^\infty dRR^3 f_m^*(R) d(0; N, R) f_n(R) = \int_0^\infty dRR^3 f_n(R) d(0; N, R) f_m^*(R), \quad (3.19)$$

provided that the contributions vanish at $R=0$ and $R \rightarrow \infty$. Thus, we may write

$$d^\dagger(-1; N, R) = -d(1; N+1, R), \quad (3.20)$$

$$d^\dagger(-2; N, R) = d(2; N+2, R), \quad (3.21)$$

$$d^\dagger(0; N, R) = d(0; N, R). \quad (3.22)$$

Simple recursion formulae involving the Gegenbauer polynomials imply [6, 29, 30]

$$\square f(R) Y_{NJM}(\Omega) = Y_{NJM}(\Omega) d(0; N, R) f(R), \quad (3.23)$$

$$\begin{aligned} \frac{\partial}{\partial x_4} f(R) Y_{NJM}(\Omega) &= A(N+1, J) Y_{N+1, JM}(\Omega) d(-1; N, R) f(R) + \\ &+ A(N, J) Y_{N-1, JM}(\Omega) d(1; N, R) f(R), \end{aligned} \quad (3.24)$$

where

$$A(N, J) = \frac{1}{2} \left[\frac{(N-J)(N+J+1)}{N(N+1)} \right]^{1/2}, \quad A(0, 0) = 0. \quad (3.25)$$

We shall also use the notations

$$A^{(2)}(2; N, J) = A(N+1, J) A(N+2, J), \quad (3.26)$$

$$A^{(2)}(0; N, J) = [A(N+1, J)]^2 + [A(N, J)]^2. \quad (3.27)$$

c) Radial BS equations

Let us substitute the expansion (3.10) into the BS equation (2.10). By using Eqs (3.23)–(3.25), we arrive at the truncated system of radial BS equations which can be written as

$$\sum_{n=0}^K \hat{B}(m|n) f_n(R) = 0, \quad m = 0, 1, 2, \dots, K, \quad (3.28)$$

where

$$\hat{B}(m|n) = \hat{F}(m|n) - g^2 V(R; \kappa^2) \delta_{mn}. \quad (3.29)$$

The diagonal matrix element $\hat{F}(n|n)$ becomes

$$\begin{aligned} \hat{F}(n|n) &= \left[-d(0; n+J, R) + 1 + \delta^2 - \frac{E^2}{4} \right]^2 - \\ &- E^2 A^{(2)}(0; n+J, J) d(0; n+J, R) - 4\delta^2, \end{aligned} \quad (3.30)$$

and the nondiagonal matrix elements are given by

$$\hat{F}(n|n+p) = F^{(0)}(n|n+p) d(p; n+J+p, R), \quad (3.31)$$

$$p = -2, -1, 1, 2, \quad (3.32)$$

$$F^{(0)}(n|n-2) = -E^2 A^{(2)}(2; n+J-2, J), \quad (3.33)$$

$$F^{(0)}(n|n-1) = -4\delta EA(n+J, J), \quad (3.34)$$

$$F^{(0)}(n|n+1) = -4\delta EA(n+J+1, J), \quad (3.35)$$

$$F^{(0)}(n|n+2) = -E^2 A^{(2)}(2; n+J, J). \quad (3.36)$$

All the other nondiagonal matrix elements of $\hat{F}(m|n)$ vanish.

Taking into account Eqs (3.21)–(3.22), we observe that

$$\hat{F}^\dagger = \hat{F} \quad \text{for} \quad \delta = 0. \quad (3.37)$$

Thus, in the equal-mass case, the operators \hat{F} and $V\delta_{mn}$ are self-adjoint, the latter being positive definite. On the other hand, Eq. (3.20) implies

$$\hat{F}^\dagger \neq \hat{F} \quad \text{for } \delta E \neq 0. \quad (3.38)$$

In other words, the operator \hat{F} of an unequal-mass system is not self-adjoint on account of the matrix elements $\hat{F}(n|n \pm 1)$.

The radial BS equations (3.28) can also be written as

$$\sum_{n=0}^K \hat{B}_H(m|n) f_n(R) = 0, \quad m = 0, 1, 2, \dots, K, \quad (3.39)$$

$$\hat{B}_H(m|n) = \hat{F}_H(m|n) - g^2 \hat{V}_H(m|n), \quad (3.40)$$

where

$$\hat{F}_H(m|n) = (-1)^m \hat{F}(m|n), \quad (3.41)$$

and

$$\hat{V}_H(m|n) = (-1)^m V(R; \kappa^2) \delta_{mn}. \quad (3.42)$$

It follows from Eqs (3.20)–(3.22) and (3.31)–(3.36) that the operators \hat{F}_H and \hat{V}_H are self-adjoint due to a convenient definition of the scalar product [31, 32]. However, \hat{F}_H and \hat{V}_H are no longer positive definite operators (see Eqs (3.41) and (3.42)). One should note that, at $\delta E \neq 0$, the truncated radial BS equations may lead to ghost problems and trajectory-mixing phenomena [8, 31–34].

4. Methods of solution

a) Asymptotic solutions

Since the leading short-distance behaviour of the solutions is independent of E and δ , we next consider the radial BS equations at $E = 0$ and $\delta = 0$. In this case, Eq. (3.28) becomes

$$\{[-d(0; N, R) + 1]^2 - g^2 V(R; \kappa^2)\} f_{N-J}(R) = 0. \quad (4.1)$$

We may now proceed to investigate the solution $f_n(R)$ in the short-distance limit $R \rightarrow 0$. Taking into account Eq. (2.13), standard theory of linear differential equations tells us that there exist powerlike asymptotic solutions:

$$f_{N-J}(R) \rightarrow a_{(N)} R^\rho \quad \text{as } R \rightarrow 0. \quad (4.2)$$

By substituting Eq. (4.2) into the radial BS equation (4.1), we obtain the homogeneous linear equation

$$\hat{B}^{(2)}(N, \rho) a_{(N)} = 0, \quad (4.3)$$

with

$$\hat{B}^{(2)}(N, \rho) = [\rho - (N + 2)] (\rho - N) (\rho + N) (\rho + N + 2). \quad (4.4)$$

To guarantee nonzero coefficients $a_{(N)}$, the indices ρ must satisfy the indicial equation

$$\hat{B}^{(2)}(N, \rho) = 0. \quad (4.5)$$

The solutions of Eq. (4.5) are the following indices

$$\rho_1 = N + 2, \quad \rho_2 = N, \quad (4.6)$$

$$\rho_3 = -N, \quad \rho_4 = -(N + 2). \quad (4.7)$$

As $R \rightarrow 0$, all the asymptotic solutions $f_{N-J}(R)$ are of the powerlike form (4.2) provided that the indices ρ_i ($i = 1, \dots, 4$) are nondegenerate. Otherwise, according to well-known theorems, the short-distance behaviour of the solutions may be governed by logarithmic terms.

The crucial ingredient of the bound-state calculations is the selection of the admissible solutions which satisfy the BS normalizability condition [8, 20]. Simple calculation yields two admissible asymptotic solutions:

$$f_{N-J}^{(1)}(R) \xrightarrow{R \rightarrow 0} a_{(N)}^{(1)} R^{N+2}, \quad (4.8)$$

$$f_{N-J}^{(2)}(R) \xrightarrow{R \rightarrow 0} a_{(N)}^{(2)} R^N. \quad (4.9)$$

In addition, there are two nonadmissible solutions which can be ruled out by means of the BS normalizability condition. These solutions are given by

$$f_{N-J}^{(3)}(R) \xrightarrow{R \rightarrow 0} a_{(N)}^{(3)} R^{-N} \quad \text{for } N > 0, \quad (4.10)$$

$$f_{N-J}^{(4)}(R) \xrightarrow{R \rightarrow 0} a_{(N)}^{(4)} R^{-(N+2)} \quad \text{for } N > 0. \quad (4.11)$$

At $N = J = 0$, the nonadmissible asymptotic solutions are of the form

$$f_0^{(3)}(R) \xrightarrow{R \rightarrow 0} a_0^{(3)} \ln R, \quad (4.12)$$

$$f_0^{(4)}(R) \xrightarrow{R \rightarrow 0} a_0^{(4)} R^{-2}. \quad (4.13)$$

We next discuss the large-distance ($R \rightarrow \infty$) behaviour of the solutions in the bound-state region. For simplicity, let us consider Yukawa-like interaction terms $V(R; \kappa^2)$ involving nonzero exchange masses κ . In this case, the BS equation becomes simply the free equation as $R \rightarrow \infty$. One can verify that, at large values of R , the four independent asymptotic solutions consist of two exponentially decreasing (admissible) solutions, and two exponentially increasing ones. Of course, the nonnormalizable (exponentially increasing) solutions must be rejected.

We now are in a position to investigate the eigenvalue problem. In the short-distance region ($R \rightarrow 0$) the radial wave function $f_{N-j}(R)$ is a linear combination of two admissible asymptotic solutions. Since the BS equation is linear and homogeneous, we are left with *one free parameter* ($a_{(N)}^{(2)}/a_{(N)}^{(1)}$) as we integrate out to the large-distance region $R \rightarrow \infty$, where *two* exponentially increasing asymptotic solutions must be absent. These conditions may be satisfied only at discrete eigenvalues of some other parameter (e.g. g^2) in the BS equation.

b) *Rayleigh-Ritz method*

To begin with let us expand the radial BS wave functions $f_n(R)$ in terms of a complete set of normalizable basis functions $\varphi_{nj}(R)$. The truncated version of this expansion is given by

$$f'_n(R) = \sum_{j=0}^Q a_{nj} \varphi_{nj}(R). \quad (4.14)$$

The leading short-distance behaviour of the approximate radial wave functions $f'_n(R)$ will be fitted to that of the exact solutions by using Slater-type basis functions $\varphi_{nj}(R)$. According to Eq. (4.9), we choose

$$\varphi_{nj}(R) = A_{nj} R^{n+J+j} e^{-\alpha R}, \quad (4.15)$$

with a positive nonlinear scale parameter α . (A_{nj} is a normalization factor.)

The BS eigenvalue problem (3.28)–(3.29) can also be formulated in the space of the basis functions $\varphi_{nj}(R)$. By substituting the expansion (4.14) into Eq. (3.28), and using the standard procedure, we obtain a homogeneous system of linear algebraic equations for the coefficients a_{nj} . The truncated version of these equations can be written as

$$\sum_{n=0}^K \sum_{j=0}^Q [\bar{F}(m, h|n, j) - g^2 \bar{V}(m, h|n, j)] a_{nj} = 0, \quad (4.16)$$

$$m = 0, 1, 2, \dots, K; \quad h = 0, 1, 2, \dots, Q, \quad (4.17)$$

where

$$\bar{F}(m, h|n, j) = \langle \varphi_{mh} | \hat{F}(m|n) | \varphi_{nj} \rangle, \quad (4.18)$$

$$\bar{V}(m, h|n, j) = \langle \varphi_{mh} | \hat{V}(m|n) | \varphi_{nj} \rangle, \quad (4.19)$$

and

$$\hat{V}(m|n) = V(R; \kappa^2) \delta_{mn}. \quad (4.20)$$

All the matrix elements of the type (4.18)–(4.19) are defined by

$$\langle \varphi_{mh} | \hat{D}(m|n) | \varphi_{nj} \rangle = \int_0^\infty dR R^3 \varphi_{mh}(R) \hat{D}(m|n) \varphi_{nj}(R). \quad (4.21)$$

In addition, we shall use the notation

$$\bar{B}(m, h|n, j) = \bar{F}(m, h|n, j) - g^2 \bar{V}(m, h|n, j). \quad (4.22)$$

Nontrivial solutions of Eq. (4.16) exist if

$$\text{Determinant} |\bar{B}(m, h|n, j)| = 0. \quad (4.23)$$

In other words, the approximate eigenvalues are prescribed by Eq. (4.23). The matrix elements $\bar{B}(m, h|n, j)$ involve the following integrals

$$I_m^{(1)}(h|j) = \langle \varphi_{mh} | V(R; \kappa^2) | \varphi_{mj} \rangle, \quad (4.24)$$

$$I_m^{(2)}(h|j) = \langle \varphi_{mh} | \varphi_{mj} \rangle, \quad (4.25)$$

$$I_m^{(3)}(h|j) = \langle \varphi_{mh} | [d(0; m+J, R)]^2 | \varphi_{mj} \rangle, \quad (4.26)$$

and

$$I_m^{(4)}(m, h|n, j) = \langle \varphi_{mh} | d(p; n+J, R) | \varphi_{nj} \rangle, \quad (4.27)$$

with

$$m \geq 0, \quad n = m + p \geq 0, \quad p = 0, \pm 1, \pm 2. \quad (4.28)$$

It is practical to replace the interaction term $V(R; \kappa^2)$ by the Vosko approximation (2.14). In this case, all the integrals (4.24)–(4.27) can be easily evaluated by using the formula

$$\int_0^\infty dR R^\Delta e^{-\omega R} = \frac{\Gamma(\Delta+1)}{\omega^{\Delta+1}}. \quad (4.29)$$

For further insight, let us investigate the application of the Rayleigh–Ritz method to the solution of the equal-mass BS equation ($\bar{\delta} = 0$). This method is based on the variational functional [6, 28]

$$g^2 = \frac{\int d^4x \psi^*(x) F(x) \psi(x)}{\int d^4x \psi^*(x) V(R; \kappa^2) \psi(x)}. \quad (4.30)$$

In accord with Eqs (3.10), (3.23)–(3.24), and Eq. (4.14) the functional (4.30) has the explicit form [35]

$$g^2 = \frac{\sum_{m=0}^K \sum_{h=0}^Q \sum_{n=0}^K \sum_{j=0}^Q a_{mh}^* \bar{F}(m, h|n, j) a_{nj}}{\sum_{m=0}^K \sum_{h=0}^Q \sum_{n=0}^K \sum_{j=0}^Q a_{mh}^* \bar{V}(m, h|n, j) a_{nj}}. \quad (4.31)$$

By variation of the expression (4.31) with respect to the linear coefficients a_{nj} , we are led to the matrix eigenvalue problem (4.16) which can be solved by a computer. Thus, in this way, one can calculate a sequence of approximate eigenvalues $[g^{(K, Q)}]^2$ (and the corresponding linear coefficients a_{nj}) by choosing finite and increasing values of K and Q . Since the matrices \bar{F} and \bar{V} are Hermitean and positive definite (in the equal-mass

case), the Hylleraas–Undheim theorem [36] ensures that, for $\delta=0$, the sequence of successive approximations monotonically approaches the true eigenvalues from above. In addition, the “optimal” value of the nonlinear scale parameter α can be selected by using standard variational procedures. Here we consider the c.m. energy E as given and g^2 to be found. By solving the BS eigenvalue problem (4.16) at many values of E , we obtain the function $g^2(E)$ from which $E(g^2)$ can be recovered without loss of accuracy.

We next discuss the eigenvalue problem (4.16) in the unequal-mass case ($\delta E \neq 0$). According to Eqs (3.20)–(3.22) and (3.30)–(3.36), we have

$$\bar{F}(m, h|m+2, j) = \bar{F}(m+2, j|m, h), \quad (4.32)$$

and

$$\bar{F}(m, h|m+1, j) = -\bar{F}(m+1, j|m, h). \quad (4.33)$$

Consequently, the matrix \bar{F} is non-Hermitian at nonzero values of δE . Of course, we may use Eqs (3.39)–(3.42) to write the eigenvalue problem (4.16) in the Hermitian form

$$\sum_{n=0}^K \sum_{j=0}^Q [\bar{F}_H(m, h|n, j) - g^2 \bar{V}_H(m, h|n, j)] a_{nj} = 0, \quad (4.34)$$

$$m=0, 1, 2, \dots, K; \quad h=0, 1, 2, \dots, Q, \quad (4.35)$$

where

$$\bar{F}_H(m, h|n, j) = (-1)^m \bar{F}(m, h|n, j), \quad (4.36)$$

and

$$\bar{V}_H(m, h|n, j) = (-1)^m \bar{V}(m, h|n, j). \quad (4.37)$$

We now observe that Eqs (4.19)–(4.20) and (4.32)–(4.37) imply

$$\bar{F}_H(m, h|n, j) = \bar{F}_H(n, j|m, h), \quad (4.38)$$

and

$$\bar{V}_H(m, h|n, j) = \bar{V}_H(n, j|m, h). \quad (4.39)$$

However, \bar{F}_H and \bar{V}_H are no longer positive definite operators. On account of these properties (including Eqs (4.32)–(4.33)), the Hylleraas–Undheim theorem cannot be applied in the unequal-mass case $\delta E \neq 0$. Therefore, in general, one should not expect a sequence of monotonically convergent approximations of the eigenvalues g^2 (from above) by choosing finite and increasing values of K and Q . Furthermore, one cannot automatically use the Rayleigh–Ritz variational principle to select optimal values for the nonlinear scale parameter α . This lack of the control of the classical variational methods may result in unpleasant convergence problems (and instabilities) which restrict the applicability of the standard truncation procedures to the solution of the BS eigenvalue problem (4.16). Finite basis-set expansion methods may also lead to spurious “approximate” solutions of the Dirac equation [23]. Consequently, the standard variational methods should be replaced by other efficient procedures in order to select convenient basis sets and acceptable numerical solutions.

c) *Least-squares variational method*

Let us consider the approximate BS wave function $\psi'(x)$ that can be expressed as

$$\psi'(x) = \sum_{n=0}^K f'_n(R) Y_{n+J, JM}(\Omega), \quad (4.40)$$

in accord with Eqs (3.10) and (4.14). By substituting Eq. (4.40) into the BS equation (2.10), we may write

$$\sum_{n=0}^K \hat{B}(m|n) f'_n(R) = \Delta_m(R), \quad m=0, 1, 2, \dots, \bar{K}. \quad (4.41)$$

Here the operators $\hat{B}(m|n)$ are defined by Eqs (3.29)–(3.36), and the deviations $\Delta_m(R)$ are related to the error of the approximate BS wave function $\psi'(x)$. Of course, the truncated system of radial BS equations (3.28) can be recovered by choosing $\Delta_m(R) \equiv 0$ and $\bar{K} = K$. In order to generalize the standard truncation procedure, we shall prescribe

$$\bar{K} > K + 1. \quad (4.42)$$

In this case, Eq. (4.41) may be regarded as an overdetermined system of differential equations. Therefore, in general, the deviations $\Delta_m(R)$ are different from zero, and also depend on the linear parameters a_{nj} :

$$\Delta_m(R) = \Delta_m(R; a_{00}, \dots, a_{KQ}). \quad (4.43)$$

We now introduce a complete set of normalizable test functions $\chi_{mh}(R)$. According to Eq. (4.9), we choose

$$\chi_{mh}(R) = A'_{mh} R^{m+J+h} e^{-\gamma R} \quad (h=0, 1, 2, \dots), \quad (4.44)$$

where γ is a positive scale parameter, and the normalization factor is denoted by A'_{mh} .

As a first step, we consider the components of the deviation vectors $\Delta_m(R)$ as given by

$$\langle \chi_{mh} | \Delta_m \rangle = \int_0^\infty dR R^3 \chi_{mh}(R) \Delta_m(R). \quad (4.45)$$

A measure of the error of the approximate wave function $\psi'(x)$ will be defined as

$$\lambda[\psi'(x)] = \frac{\sum_{m=0}^{\bar{K}} \sum_{h=0}^{\bar{Q}} w_{mh} |\langle \chi_{mh} | \Delta_m \rangle|^2}{\sum_{m=0}^K \sum_{h=0}^Q \sum_{j=0}^Q a_{mh}^* \tilde{W}(m, h|m, j) a_{mj}}, \quad (4.46)$$

where \tilde{W} is a convenient (positive definite) "norm operator", and the w_{mh} 's are positive weighting factors. We shall use a sufficiently large set of test functions by imposing the requirements (4.42) and

$$\bar{Q} > Q + 1. \quad (4.47)$$

In order to simplify the calculations, one may use

$$\gamma = \alpha, \quad \tilde{W}(m, h | m, j) = \langle \varphi_{mh} | \varphi_{mj} \rangle, \quad (4.48)$$

and, in addition,

$$w_{mh} = 1 \quad (4.49)$$

for all possible values of m and h .

By variation of the expression (4.46) with respect to the linear coefficients a_{mj} (and using the notation (4.22)) we obtain the eigenvalue problem

$$\sum_{n=0}^K \sum_{j=0}^Q [\bar{L}(m, h | n, j) - \lambda \tilde{W}(m, h | n, j)] a_{nj} = 0, \quad (4.50)$$

$$m = 0, 1, 2, \dots, K; \quad h = 0, 1, 2, \dots, Q, \quad (4.51)$$

where

$$\bar{L}(m, h | n, j) = \sum_{m'=0}^{\bar{K}} \sum_{h'=0}^{\bar{Q}} w_{m'h'} \bar{B}^*(m', h' | m, h) \bar{B}(m', h' | n, j) \quad (4.52)$$

and

$$\tilde{W}(m, h | n, j) = \tilde{W}(m, h | m, j) \delta_{mn}. \quad (4.53)$$

The matrix \bar{L} is Hermitean and positive semidefinite, and the acceptable solutions a_{nj} belong to the lowest eigenvalue λ_1 which depends, of course, on the size of the basis set ($\lambda_1 = \lambda_1^{(K, Q)}$). According to the requirements (4.42) and (4.47), we shall choose

$$\bar{K} = K + k, \quad k > 1, \quad (4.54)$$

$$\bar{Q} = Q + q, \quad q > 1, \quad (4.55)$$

where the integers k and q are fixed.

We next consider g^2 and J as given and the bound-state energies E to be found. By solving Eq. (4.50) at many values of E and α , we obtain the function

$$\lambda_1 = \lambda_1^{(K, Q)}(E, \alpha). \quad (4.56)$$

The approximate eigenvalues of the c.m. energy E (and the "optimal" scale parameters α) can be calculated by using the least-squares variational requirement

$$\lambda_1^{(K, Q)}(E, \alpha) = \text{minimum}. \quad (4.57)$$

In other words, the eigenvalue $\lambda_1^{(K, Q)}(E, \alpha)$ may be regarded as a reasonable measure of the error of the approximate BS wave function $\psi'(x)$, and the minima of the function $\lambda_1^{(K, Q)}(E, \alpha)$ are at the approximate bound-state energies $E_b^{(K, Q)}(\alpha)$ ($b = 1, 2, \dots$). Thus, at fixed values of α , we apply Eq. (4.57) to compute a sequence of the approximate eigenvalues $E_b^{(K, Q)}(\alpha)$ by choosing finite and increasing values of K and Q .

Let us first discuss the approximate ground-state energies $E_1^{(K, Q)}(\alpha)$. The completeness of the basis set (4.15) implies that $E_1^{(K, Q)}(\alpha)$ is almost constant in a finite region of the nonlinear scale parameter α , if the values of K and Q are sufficiently large.

(These stability properties are, in general, less pronounced for the excited states ($b > 1$)). Of course, the "optimal value" of the scale parameter α may also be computed by using the least-squares variational requirement (4.57).

Illustrative numerical calculations [22, 23] indicate that the least-squares variational method can be successfully applied to the approximate solution of the BS equation, if the basis functions (4.15) are chosen according to the following requirements. (i) There must exist a region of the nonlinear scale parameter α where a set of the computed energy eigenvalues is stable (i.e., in that region one has $E_b^{(K, Q)}(\alpha) \sim \text{constant}$). (ii) An apparent convergence of these eigenvalues should be observed at increasing basis sizes. (iii) The measure of the error λ_1 must be sufficiently small. Similar criteria have been already used to select acceptable numerical solutions both in the bound-state and in the scattering region [23, 27].

5. Discussion

At nonvanishing c.m. energies, the BS equation leads to an infinite system of coupled ordinary differential equations which are referred to as the radial BS equations. In the equal-mass case, these equations can be solved by the Rayleigh–Ritz method which is based on the standard truncation procedure. However, in general, the unequal-mass BS equation is beyond the current scope of the classical variational methods and, indeed, the standard truncation procedure may result in unpleasant convergence problems and instabilities [23]. Consequently, the classical variational methods should be replaced by other efficient procedures. In the present paper we have suggested a particular least-squares method which involves a nontrivial generalization of the standard truncation procedure.

References

1. E. E. Salpeter and H. A. Bethe, *Phys. Rev.*, *84*, 1232, 1951.
2. M. Gell-Mann and F. Low, *Phys. Rev.*, *84*, 350, 1951.
3. J. Schwinger, *Proc. Nat. Acad. Sci. USA*, *37*, 453, 1951.
4. G. C. Wick, *Phys. Rev.*, *96*, 1127, 1954.
5. R. E. Cutkosky, *Phys. Rev.*, *96*, 1135, 1954.
6. C. Schwartz, *Phys. Rev.*, *137*, B717, 1965.
7. C. Schwartz and C. Zemach, *Phys. Rev.*, *141*, 1454, 1966.
8. N. Nakanishi, *Progr. Theor. Phys. Suppl.*, *43*, 1, 1969.
9. A. Pagnamenta, "Solutions of the Bethe–Salpeter Equations" in *Lectures in Theoretical Physics*, Eds. K. T. Mahanthappa and W. E. Brittin, Gordon and Breach Science Publishers, New York, 1969.
10. M. Böhm, H. Joos and M. Krammer, *Acta Phys. Austriaca Suppl.*, *11*, 3, 1973.
11. K. Ladányi, *Phys. Rev.*, *D11*, 2320, 1975.
12. K. Ladányi, *Phys. Rev.*, *D14*, 2648, 1976.
13. K. Ladányi, *Acta Phys. Austriaca*, *49*, 113, 1978.
14. K. Ladányi, *Annals of Physics N. Y.*, *130*, 427, 1980.

15. S. D. Drell and T. D. Lee, *Phys. Rev.*, *D5*, 1738, 1972.
16. T. Appelquist and E. Poggio, *Phys. Rev.*, *D10*, 3280, 1974.
17. M. L. Goldberger, A. H. Guth and D. E. Soper, *Phys. Rev.*, *D14*, 2633, 1976.
18. G. P. Lepage and S. J. Brodsky, *Phys. Rev.*, *D22*, 2157, 1980.
19. S. J. Brodsky, Chueng-Ryong Ji and M. Sawicki, SLAC-PUB-3382, 1984.
20. S. Mandelstam, *Proc. Roy. Soc.*, *A233*, 248, 1955.
21. M. A. Abdel-Raouf, *Physics Reports*, *84*, 163, 1982.
22. K. Ladányi, *Nuovo Cimento*, *A70*, 405, 1970.
23. B. Gazdy and K. Ladányi, *J. Chem. Phys.*, *80*, 4333, 1984.
24. K. Ladányi, *Nuovo Cimento*, *A61*, 173, 1969.
25. K. Ladányi and T. Szondy, *Nuovo Cimento*, *B5*, 70, 1971.
26. M. A. Abdel-Raouf, *Physics Reports*, *108*, 1, 1984.
27. B. Apagyí and K. Ladányi, *Phys. Rev.*, *A33*, 182, 1986.
28. S. H. Vosko, *J. Math. Phys.*, *1*, 505, 1960.
29. G. Domokos and P. Surányi, *Nucl. Phys.*, *54*, 529, 1964.
30. W. Magnus, F. Oberhettinger and R. P. Soni, *Formulas and Theorems for Special Functions of Mathematical Physics*, Springer, New York, 1966.
31. R. E. Cutkosky and B. B. Deo, *Phys. Rev. Lett.*, *19*, 1256, 1967.
32. W. B. Kaufmann, *Phys. Rev.*, *187*, 2051, 1969.
33. R. N. Madan, R. W. Haymaker and R. Blankenbecler, *Phys. Rev.*, *172*, 1788, 1968.
34. E. zur Linden, *Nuovo Cimento*, *A63*, 181, 1969.
35. K. Ladányi, *Nuovo Cimento*, *A56*, 173, 1968.
36. E. A. Hylleraas and B. Undheim, *Z. Phys.*, *65*, 759, 1930.

ON A PROBLEM OF SPONTANEOUS COMPACTIFICATION*

P. FORGÁCS**

*Department of Physics, The University, Southampton
SO9 5NH England*

Z. HORVÁTH and L. PALLA

*Institute for Theoretical Physics, Roland Eötvös University
1088 Budapest, Hungary*

We investigate how the equations governing spontaneous compactification to non-symmetric coset spaces can be solved for two particular classes of internal spaces.

1. Introduction

In the last few years we witnessed a revival of interest in using theories based on the old assumption that the physical space-time has more than four dimensions to explain the observed properties of elementary particles. The spontaneously compactified theories [1] form a class of these models: they are theories of gravity coupled to Yang–Mills and spinor fields in a $4 + d$ dimensional space-time which are reduced — via a *solution* of the Einstein–Yang–Mills (EYM) field equations — to a 4 dimensional low energy theory by compactifying the extra d dimensions into a compact coset space S/H with an appropriately small size R_0 . Recent investigations showed that it is rather difficult to find a model explaining all details of unified gauge theories [2]. In these models mostly symmetric coset spaces were used to describe the extra compact dimensions [3]. Quite recently the equations governing spontaneous compactification to non-symmetric coset spaces (NSCS) have been derived [4, 5] and it was shown [4] that in contrast to the case of symmetric spaces they cannot be solved for all NSCS. The aim of this paper is to determine a class of NSCS for which the compactification equations do admit a solution.

In the next Section we briefly outline the derivation of the equations governing spontaneous compactification to NSCS while in Section 3 we show how these equations can be solved in two special cases, namely for S/H 's having the form of

$SU(k+l)/SU(k) \times \overbrace{U_1 \times \dots \times U_1}^l$ or $SO(2[k+l-1])/SU(k) \times \overbrace{U_1 \times \dots \times U_1}^l$ where k and l are arbitrary integers ($k \geq 1, l \geq 2$).

* Dedicated to Prof. K. Nagy on his 60th birthday

** On leave from Central Research Institute for Physics, H-1525 Budapest 114, Hungary

2. Spontaneous compactification to NSCS

We start with the standard EYM action for a coupled system of gravity and gauge fields with gauge group G_{YM} in $4+d$ dimensions

$$S = \int dz^{4+d} (-g)^{1/2} \left(\frac{R}{\kappa^2} - \frac{1}{4g^2} F_{AB}^\alpha F^{\alpha AB} - \Lambda \right), \quad (1)$$

where κ and g are the $4+d$ -dimensional gravitational and gauge coupling constants and Λ is a cosmological constant. The equations following from (1) are $4+d$ dimensional Einstein equations and the Yang–Mills equations coupled to $4+d$ dimensional gravity. Spontaneous compactification means that we find a solution of this coupled system of equations where the $4+d$ dimensional space-time is the direct product of the flat 4 dimensional Minkowski space (M_4) and a compact coset space S/H ($\dim S - \dim H = d$). To obtain $M_4 \times S/H$ as a solution we must have non-vanishing gauge fields on the internal space S/H , and as this internal space admits S as a symmetry group we also require that the explicit gauge fields be S symmetric in the sense of [6]. In this paper we assume that S is a simple Lie group and S/H is a NSCS.

To describe $M_4 \times S/H$ we split the coordinates $z^M = (x^m, y^\mu)$ $m=0, 1, \dots, 3; \mu=1, 2, \dots, d$ denoting Minkowski versus internal ones. Furthermore, we choose from each coset an element labelled by y^μ denoted as $L(y)$. Using this quantity the vielbeins e^a on S/H are given by [7]

$$L^{-1}(y)dL = e^a Q_a + \omega^i Q_i = (e_\mu^a(y) Q_a + \omega_\mu^i(y) Q_i) dy^\mu. \quad (2)$$

Here Q_A , $A=1, \dots, \dim S$, are the generators of S satisfying $[Q_A, Q_B] = f_{ABC} Q_C$ with f_{ABC} being the completely antisymmetric structure constants of S . Furthermore, we denote the indices corresponding to H and to the coset S/H by i and a , respectively (so $A = (i, a)$). As H is a closed subgroup of S , $f_{ijb} = 0$ and f_{ijk} are the structure constants of H . The fact that S/H is not a symmetric space means that $f_{abc} \neq 0$. Since e^a and ω^i satisfy the Maurer–Cartan equations — as a consequence of Eq. (2) — we determine the torsion free (“natural”) connection on S/H as

$$B_b^a = -\frac{1}{2} f_{bc}^a e^c - f_{bt}^a \omega^t. \quad (3)$$

It is straightforward to determine the Riemann and Ricci tensors for this connection, in fact for the latter we obtain

$$R_{ab} = \frac{1}{4} (f_{cbe} f_{eca} + 4 f_{cai} f_{icb}). \quad (4)$$

To make an ansatz for the compactifying gauge fields we must choose appropriate G_{YM} admitting the possibility of constructing S symmetric gauge configurations. The

minimal choice is $G_{YM} = H$ [6]. If H consists of several factors $H = H_1 \times H_2 \times \dots \times H_p$ then our ansatz is

$$A = \sum_{j=1}^p \frac{1}{g_j} \omega^{i_j} Q_{i_j} = \sum_{j=1}^p A^{(j)};$$

$$F = -\frac{1}{2} \sum_{j=1}^p f_{abi_j} Q^{ij} e^a \Lambda e^b \frac{1}{g_j} = \frac{1}{2} \sum_{j=1}^p F_{ab}^{(j)} e^a \Lambda e^b, \quad (5)$$

where i_j runs over the indices corresponding to the j -th factor, H_j , of H and g_j denotes the gauge coupling constant referring to this particular factor.

Using the Jacobi identity for $f_{i_j ab}$, the antisymmetry of f_{abc} and also that $f_{adc} f_{ci_j a} = 0$ one can readily show that $F_{ab}^{(j)}$ in Eq. (5) indeed solves the YM equations

$$\partial_a F_{ad}^{(j)} - \Gamma_{aa}^b F_{bd}^{(j)} - \Gamma_{ad}^c F_{ac}^{(j)} - g_j [A_d^{(j)}, F_{ad}^{(j)}] = 0, \quad j = 1, \dots, p,$$

where $e^c \Gamma_{ca}^b = B_a^b$ is the torsion free connection and $A_a^{(j)} = e_a^\mu A_\mu^{(j)}$ with e_a^μ being the inverse of e_μ^a .

Before turning to the Einstein equations we introduce a length scale, R_0 , fixing the size of S/H . This implies that both R_{ab} and $F_{ab}^{(j)}$ in Eq. (4) resp. (5) get multiplied by R_0^2 . The $4+d$ -dimensional Einstein equations split into a four-dimensional and a d -dimensional part. The 4-dimensional part — expressing that we assume here a flat Minkowski space — determines Λ in terms of the other parameters. Fine tuning Λ to this value the d -dimensional part has the form

$$\frac{1}{4} C_2^S(Adj) \delta_{ac} = \sum_j \left(\frac{\kappa a^2}{g_j^2} - \frac{1}{2} \right) f_{i_j ab} f_{i_j cb}, \quad (6)$$

where $C_2^S(Adj)$ is the value of the quadratic Casimir of S on its adjoint representation: $f_{adc} f_{cDc} = -C_2^S(Adj) \delta_{ac}$. If for a NSCS, S/H , there is a choice of the g_j coupling constants guaranteeing that the expression on the r.h.s. of Eq. (6) is proportional to δ_{ac} then this coset space can be used for compactification in our framework. In [4] we analyzed briefly Eq. (6) and concluded that in the case of NSCS physically interesting 4-dimensional models may be obtained only if H is a *regular subgroup* of S satisfying in addition $\text{rank } H = \text{rank } S$. Therefore, in the following we consider only such coset spaces, S/H , that satisfy these two conditions.

3. Some solutions of Eq. (6)

First, we note that $f_{ai_j c}$ ($i_j = 1 \dots \dim H_j$) form the representation matrices on the tangent space (T) of the generators of the H_j factor of H thus $f_{i_j ac} f_{i_j bc}$ is proportional to the value of the second Casimir of H_j taken on this representation. Rewriting the

expression on the r.h.s. of Eq. (6) as

$$\left(\frac{\kappa a^2}{g_1^2} - \frac{1}{2}\right) \left\{ f_{i_1 a b} f_{i_1 c b} + \sum_{j=2}^p \lambda_j f_{i_1 a b} f_{i_1 c b} \right\};$$

$$\lambda_j = \left(\frac{\kappa a^2}{g_j^2} - \frac{1}{2}\right) / \left(\frac{\kappa a^2}{g_1^2} - \frac{1}{2}\right) \quad (7)$$

we see that Eq. (6) may have a solution if we can introduce the λ_j factors that change the relative values of the second Casimirs of the various H_j 's in such a way that the sum of all of the modified Casimirs for all irrep components of

$$T \left(T = \sum_{\{s_1 \dots s_p\}} \oplus T_{s_1}^{(H_1)} \otimes \dots \otimes T_{s_p}^{(H_p)} \right)$$

becomes identical.

We conjecture that Eq. (6) can be solved in this way for all H 's satisfying the two conditions mentioned above if all the roots of S have equal length; i.e. for all regular subgroups with maximal rank H , of $SU(n)$, $SO(2n)$ and E_n . Even if this conjecture is right it does not mean that all solutions of Eq. (6) are found as e.g. the G_2/SU_3 solution (discussed in details in [4]) does not belong to this class.

To motivate our conjecture we note that if we take H to be the maximal torus of S i.e. when the tangent space consists of all the roots of S then the expression in the curly bracket of (7) in the root basis of S has the form

$$\text{diag} \left(\sum_{i=1}^p \lambda_i (\alpha_1^i)^2, \sum_{i=1}^p \lambda_i (\alpha_2^i)^2, \dots, \sum_{i=1}^p \lambda_i (\alpha_d^i)^2 \right),$$

where α_k^i ($k=1 \dots d$) are the roots of S ($\dim S = d+p$). If all the roots are of the same length then choosing equal λ_i 's in this expression it gets proportional to δ_{ab} . On the other hand if there are roots of different lengths then a simple inspection revealed that this expression cannot be made of this form.

To substantiate our conjecture we show how the λ_j factors must be chosen in order to solve Eq. (6) for $SU_{k+l}/SU_k \times \overbrace{U_1 \times \dots \times U_1}^l$ and $SO(2[k+l-1])/SU_k \times \overbrace{U_1 \times \dots \times U_1}^l$ respectively. For $k \geq 1$; $l \geq 2$ these are NSCS and belong to the conjectured class of coset spaces.

For $SU_{k+l}/SU_k \times \overbrace{U_1 \times \dots \times U_1}^l$ we use the $k+l$ mutually orthogonal unit vectors e_i ($i=1 \dots k+l$; $e_i e_j = \delta_{ij}$) to describe the roots of SU_{k+l} as well as the weights

in the tangent space. In terms of these vectors the positive roots of SU_{k+l} and SU_k have the form [8]

$$\begin{array}{ll}
 e_1 - e_2, \dots, e_1 - e_{k+l}, & e_1 - e_2, \dots, e_1 - e_k, \\
 e_2 - e_3, \dots, e_2 - e_{k+l}, & e_2 - e_3, \dots, e_2 - e_k, \\
 \dots & \dots \\
 e_{k+l-1} - e_{k+l}, \quad (SU_{k+l}), & e_{k-1} - e_k, \quad (SU_k),
 \end{array} \tag{8}$$

while the roots that (together with their negatives) constitute the tangent space, T , are

$$e_i - e_{k+n}, \quad i = i, \dots, k, \quad n = 1, \dots, l; \tag{9}$$

$$e_{k+n} - e_{k+m}, \quad n < m, \quad n = 1, \dots; \quad m = n + 1, \dots, l. \tag{10}$$

Clearly for all fixed n the roots in Eq. (9) form the weights of the fundamental irreducible representation of SU_k , while those in Eq. (10) are SU_k singlets. The normalized U_1 generators are expressed in terms of e_i as

$$Q_m = \frac{e_1 + \dots + e_{k+m-1} - (k+m-1)e_{m+m}}{\sqrt{(k+m)(k+m-1)}}, \quad m = 1, \dots, l. \tag{11}$$

For $SU_k \times \overbrace{U_1 \times \dots \times U_1}^l$ the sum of second Casimirs is nothing but the sum of the SU_k Casimir and of the squares of the U_1 charges. The value of the second SU_k Casimir is $\left(k - \frac{1}{k}\right)$ for the fundamental irrep and is 0 for the singlet. The U_1 charges of $e_i - e_{k+n}$ and $e_{k+n} - e_{k+m}$ are obtained by simple multiplication

$$Q_m(e_i - e_{k+n}) = \begin{cases} \frac{1}{\sqrt{(k+m)(k+m-1)}}, & m < n; \\ \frac{k+n}{\sqrt{(k+n)(k+n-1)}}, & m = n; \\ 0, & m > n; \end{cases} \tag{12}$$

$$Q_p(e_{k+n} - e_{k+m}) = \begin{cases} 0, & p < n; \quad \frac{1}{\sqrt{(k+p)(k+p-1)}}, & n < p < m; \quad 0, & m < p; \\ -\frac{k+n-1}{\sqrt{(k+n)(k+n-1)}}, & p = n; \quad \frac{k+m}{\sqrt{(k+m)(k+m-1)}}, & p = m. \end{cases} \tag{13}$$

Using these expressions in the curly bracket of Eq. (7) one gets that the sum of all terms for $e_i - e_{k+n}$ becomes

$$k - \frac{1}{k} + \sum_{m=1}^{n-1} \frac{m}{(k+m)(k+m-1)} + \lambda_n \frac{k+n}{k+n-1}, \quad (14)$$

while for $e_{k+n} - e_{k+m}$ it has the form

$$\lambda_n \frac{k+n-1}{k+n} + \lambda_m \frac{k+m}{k+m-1} + \sum_{p=n+1}^{m-1} \frac{\lambda_p}{(k+p)(k+p-1)}. \quad (15)$$

Eq. (6) can be solved if there is a choice of λ_m guaranteeing that the expression in (14) is independent of n and is identical to that of (15) for all n and m . Demanding the equality of the expressions in (14) for n and $n+1$ leads to $\lambda_n = \lambda_{n+1}$, i.e. all the U_1 -s must be shifted uniformly with respect to the SU_k . Setting all $\lambda_m = \lambda$ in both Eqs (14) and (15) we get

$$k - \frac{1}{k} + \lambda(k+1), \quad (14')$$

$$2\lambda. \quad (15')$$

The equality of these two expressions finally determines λ ; $\lambda = k+1$.

To describe the roots of $SO(2[k+l-1])$ and the tangent space for $SO(2[k+l-1])/SU_k \times \overbrace{U_1 \times \dots \times U_1}^l$ we again use the e_i vectors ($i=1, \dots, k+l-1$; $e_i e_j = \delta_{ij}$). The positive roots of $SO(2[k+l-1])$ can be expressed as

$$\begin{array}{ll} e_1 - e_2, e_1 - e_3, \dots, e_1 - e_{k+l-1}, & e_1 + e_2, \dots, e_1 + e_{k+l-1}, \\ \dots\dots\dots & \dots\dots\dots \\ e_{k+l-2} - e_{k+l-1}, & e_{k+l-2} + e_{k+l-1}. \end{array} \quad (16)$$

The root system of SU_k remains identical to that of given in (8) while the roots constituting the tangent space are now

$$e_i - e_{k+n}; \quad e_i + e_{k+n}; \quad i=1, \dots, k; \quad n=1, \dots, l-1, \quad (17)$$

$$e_i + e_j; \quad i, j=1, \dots, k; \quad i \neq j, \quad (18)$$

$$e_{k+n} - e_{k+m}; \quad e_{k+n} + e_{k+m}; \quad n=1, \dots, l-2; \quad m=n+1, \dots, l-1. \quad (19)$$

Clearly the expressions in (17) describe the weights of two fundamental SU_k irreps, while those in (19) are SU_k singlets. The novel feature of this coset space is the appearance of the completely antisymmetric two tensor irrep of SU_k in (18). For $m=1, \dots, l-1$ the normalized U_1 generators are still given by (11), however, for $m=l$ the U_1 generator has the form

$$Q_l = \frac{e_1 + \dots + e_{k+l-1}}{\sqrt{k+l-1}}. \quad (20)$$

It is easy to see that for $e_i - e_{k+n}$ and $e_{k+n} - e_{k+m}$ the l -th U_1 charge vanishes while for $m < l$ they are identical to those in (12, 13), thus repeating the discussion of $SU_{k+l}/SU_k \times U_1 \times \dots \times U_1$ case we conclude that all λ_m -s ($m=1, \dots, l-1$) must be equal to $\lambda = k+1$ and the sum of all (modified) Casimirs must be $2(k+1)$. Thus we must check if it is possible to choose a λ_l guaranteeing that this latter holds for the other irreps in (17-19) as well. The U_1 charges of the $e_{k+n} + e_{k+m}$ SU_k singlets are

$$Q_l(e_{k+n} + e_{k+m}) = \frac{2}{\sqrt{k+l-1}},$$

$$Q_p(e_{k+n} + e_{k+m}) = \begin{cases} 0, & p < n; \frac{1}{\sqrt{(k+p)(k+p-1)}}, & n < p < m; \frac{2}{\sqrt{(k+p)(k+p-1)}}, & m < p; \\ -\frac{k+n-1}{\sqrt{(k+n)(k+n-1)}}, & p = n; \frac{2-(k+m)}{\sqrt{(k+m)(k+m-1)}}, & p = m; \end{cases}$$

$$p = 1, \dots, l-1. \quad (21)$$

Using that $\lambda_m = (k+1)$ ($m=1, \dots, l-1$) the sum of the squares of the modified U_1 charges for $e_{k+n} + e_{k+m}$ is thus

$$2(k+1) \left\{ 1 - \frac{2}{k+l-1} \right\} + 4 \frac{\lambda_l}{k+l-1}.$$

Equating this expression to $2(k+1)$ yields $\lambda_l = k+1$. The completely antisymmetric two

tensor irrep of SU_k (18) has a Young tableaux $\begin{array}{|c|} \hline \square \\ \hline \end{array}$ thus the value of the second Casimir on this irrep is $[9] \frac{2}{k}(k^2 - k - 2)$. The U_1 charges of this irrep are given by

$$Q_m(e_i + e_j) = \frac{2}{\sqrt{(k+m)(k+m-1)}}, \quad m = 1, \dots, l-1;$$

$$Q_l(e_i + e_j) = \frac{2}{\sqrt{k+l-1}}.$$

Since all λ_m -s ($m=1, \dots, l$) are equal to $k+1$ we finally get for the (modified) sum of the Casimirs

$$4(k+1) \frac{1}{k} + \frac{2}{k}(k^2 - k - 2) = 2(k+1),$$

thus the criterion is indeed satisfied for this irrep. As the U_1 charges of the $e_i + e_{k+n}$ irrep

are

$$Q_m(e_i + e_{k+n}) = \begin{cases} \frac{2}{\sqrt{(k+m)(k+m-1)}}, & n < m; \\ \frac{1}{\sqrt{(k+m)(k+m-1)}}, & m < n; \\ \frac{2-(k+n)}{\sqrt{(k+n)(k+n-1)}}, & m = n; \end{cases} \quad m = 1, \dots, l-1;$$

$$Q_l(e_i + e_{k+n}) = \frac{2}{\sqrt{k+l-1}}$$

it is not difficult to verify that with all λ_m -s set to $(k+1)$ we get the needed $2(k+1)$ when computing the curly bracket of (7) for this irrep. This then completes the proof that $SO(2[k+l-1])/SU_k \times \overbrace{U_1 \times \dots \times U_1}^l$ is indeed a solution of the compactification problem.

References

1. E. Cremmer and J. Scherk, Nucl. Phys., *B103*, 399, 1976, *B108*, 409, 1976; *B118*, 61, 1977; Z. Horváth, L. Palla, E. Cremmer and J. Scherk, Nucl. Phys., *B127*, 57, 1977; Z. Horváth and L. Palla, Nucl. Phys., *B142*, 327, 1978; J. F. Luciani, Nucl. Phys., *B135*, 111, 1978.
2. E. Witten, Princeton preprint, Oct. 1983; Ch. Wetterich, Bern Univ. preprint BUTP — 85/4; D. Olive and P. West, Nucl. Phys., *B217*, 248, 1983; S. Randjbar-Daemi, A. Salam and J. Strathdee, Trieste preprint IC/84/15; P. Forgács and G. Zoupanos, Phys. Lett., *B148*, 99, 1984; F. A. Bais et al., CERN preprint TH 4045, 1984.
3. S. Randjbar-Daemi and R. Percacci, Phys. Lett., *B117*, 41, 1982; S. Randjbar-Daemi, A. Salam and J. Strathdee, Nucl. Phys., *B214*, 491, 1983; *B242*, 447, 1984; P. H. Frampton, P. J. Moxhay and K. Yamamoto, Phys. Lett., *144B*, 354, 1984; Ch. Wetterich, Nucl. Phys., *B244*, 355, 1984; *B260*, 402, 1985; A. N. Schellekens, Nucl. Phys., *B248*, 706, 1984.
4. P. Forgács, Z. Horváth and L. Palla, KFKI-report-1985-51 (rev.) to be published in Z. Phys. C.
5. A. N. Schellekens, Nucl. Phys., *B256*, 109, 1985.
6. P. Forgács and N. Manton, Comm. Math. Phys., *72*, 15, 1980.
7. A. Salam and J. Strathdee, Ann. Phys., *141*, 316, 1986; L. Palla, Z. Phys. C. *24*, 195, 1984.
8. J. Gilmore, Lie Groups, Lie Algebras, and Some of Their Applications, John Wiley and Sons, New York-London-Sidney-Toronto, 1974.
9. K. Pilch and A. N. Schellekens, Stony Brook preprint ITP-SB-84-20.

THE NATURE OF THE DECONFINING PHASE TRANSITION IN THE SU(2) GAUGE THEORY IN $1 + \varepsilon$ DIMENSION*

A. PATKÓS

*Department of Atomic Physics, Roland Eötvös University
1088 Budapest, Hungary*

The effective theory of SU(2)-coloured charges at finite temperature is shown to undergo a deconfining phase transition in $D = 1 + \varepsilon (\varepsilon > 0)$ dimension and to belong to the Ising universality class by explicit computation of the coupling renormalisation to subleading order.

Theoretical approach to deconfining phase transitions starts from the important universality arguments of Svetitsky and Yaffe [1]. They state, *if* the transition is continuous and *if* a single fixed point exists to describe the breakdown of the symmetry, then the indices of the deconfining transition are determined by the characteristics of that fixed point of the renormalisation group transformations.

Yet, an eventual first order transition cannot be excluded. Actually, Monte Carlo simulations indicate continuous phase transition only for the pure SU(2) gauge theory among the SU(N) groups. There the deconfinement breaks a global Z(2) Ising-like symmetry. Numerical “measurements” of the critical magnetization index in 2 and 3 spatial dimensions support that the transition belongs to the Ising universality class [2, 3].

The only analytical approach to these phenomena to date is the mean field treatment of the effective theory for the order parameter P [4, 5]:

$$S_{\text{eff}} = -\frac{1}{2} J \sum_{\mathbf{x}, \mathbf{e}} [P(\mathbf{x})P^+(\mathbf{x} + \mathbf{e}) + h.c.]. \quad (1)$$

P takes its values among the characters of the group in the fundamental representation; its expectation value gives the excess free energy of static coloured charges in the fundamental representation of the gauge group over the ground state. The coupling J is related to the gauge coupling and the temperature. For the strong coupling regime

* Dedicated to Prof. K. Nagy on his 60th birthday

($g^2 \gg 1$) the following functional form can be derived

$$J(g^2, T) = \begin{cases} \left(\frac{1}{g^2}\right)^{1/Ta}, & G = \text{SU}(2), \\ 2\left(\frac{1}{Ng^2}\right)^{1/Ta}, & G = \text{SU}(N) \quad N > 2, \end{cases} \quad (2)$$

where a is the lattice spacing.

In this note we shall establish two results concerning the nature of the phase transition in the SU(2) model defined by Eq. (1). First, we show that its lower critical dimension is 1, secondly we compute its critical point and critical thermal index in $D = 1 + \varepsilon$ dimension and find that the model belongs to the Ising universality class. These results although very plausible, are not self-evident as the model (1) is not an Ising model by its definition. It has a much wider symmetry, than Z(2) as its partition function

$$Z = \int_0^{4\pi} \prod_{\mathbf{x}} \left[d\varphi_{\mathbf{x}} \sin^2 \frac{\varphi_{\mathbf{x}}}{2} \right] \exp \left\{ 4J \cos \frac{\varphi_{\mathbf{x}}}{2} \cos \frac{\varphi_{\mathbf{x}+\mathbf{e}}}{2} \right\} \quad (3)$$

is defined by integrating over all the elements of the gauge group. Our results complement the evidence collected by numerical simulations for the nature and mechanism of the deconfining transition in the finite temperature pure SU(2) gauge model.

In order to prove the above statements we employ the Migdal-Kadanoff renormalisation group transformation [6, 7]. As for 1D global symmetric (spin-) models this transformation is exact the first conclusion will be proven rigorously. In $D = 1 + \varepsilon$ dimensions the experience shows that this version of the renormalisation transformations yields critical exponents which are exact to leading order in ε [6]. We shall use the bond moving variant of this transformation due to Kadanoff [7], with a scale factor $b = 2$. The first step of the transformation is bond-moving where along each direction every second bond is suppressed and the corresponding strength is redistributed uniformly among the remaining bonds:

$$J' = 2^{D-1} J. \quad (4)$$

In the second step a 1-dimensional integration is performed leading to the renormalized Boltzmann factor:

$$w_{\text{ren}}(P, P') = \int dP'' \sqrt{1 - \frac{P''^2}{4}} \exp \{ J' P'' (P + P') \}. \quad (5)$$

We shall explicitly compute (5) for $J \gg 1$, because in $D = 1 + \varepsilon$ we find that the critical point $J_c \sim O(1/\varepsilon)$. In this asymptotic region we shall see that

$$w_{\text{ren}}(P, P') = \text{const} \cdot \exp \{ J_{\text{ren}} P P' \}, \quad (6)$$

that is the transformation is closed with a single coupling (for infinitesimal ε).

The generic form of the 1D decimation at large J will be

$$J' = J \left(1 - \frac{c}{J^\beta} \right). \quad (7)$$

This means that to leading order the coupling is marginal in 1D, and subleading terms in the renormalisation make the fixed point $J = \infty$ unstable. This is exactly the way we identify the lower critical dimension. The bond moving step in $D = 1 + \varepsilon$ dimension introduces the opposite tendency by Eq. (4) in the coupling flow leading to a summary coupling renormalisation

$$J_{\text{ren}} = 2^\varepsilon J \left(1 - \frac{c}{2^\varepsilon J^\beta} \right). \quad (8)$$

Expanding the right hand side of (8) to leading order in ε (assume that $J^\beta \sim O(\varepsilon^{-1})$) one finds the shifted location of the fixed point:

$$J_c^\beta = \frac{c}{\varepsilon \ln 2}. \quad (9)$$

Then one linearizes the transformation (8) around the point (9) and finds the expression

$$y_J = 1 + \varepsilon \beta \ln 2, \quad v_J = \frac{\ln 2}{\ln y_J} \approx \frac{1}{\beta \varepsilon} \quad (10)$$

for the "thermal" exponent of the transition. From Eqs (9) and (10) we see that it is the subleading term in the 1D renormalisation transformation formula (7) which determines the critical data to leading order in ε .

The parameters c and β are what we determine for large J next.

In a previous publication [8] we have introduced the double character expansion for the Boltzmann factor of the theory (1):

$$\exp \{ J \text{Tr } g \text{Tr } g' \} = \sum_{i,j} F_{ij}(J) d_i d_j \chi_i(g) \chi_j(g'), \quad (11)$$

where χ_j is the character of the spin- j representation of $SU(2)$ and d_j is the dimension of the representation:

$$F_{ij}(J) = \frac{1}{d_i d_j} \{ I_{j-i}(x) [I_{i+j+2}(x) + I_{i+j}(x)] - \\ - I_{i+j+1}(x) [I_{i-j+1}(x) + I_{j-i+1}(x)] \}, \quad x \equiv 8J. \quad (12)$$

In Eq. (12) $I_n(x)$ is the n -th modified imaginary argument Bessel function. Also, both indices of F_{ij} are either integers or half-integers.

In terms of the expansion coefficients the 1D decimation with scale factor $b=2$ takes the form

$$F'_{ij}(J) = \sum_k d_k^2 F_{ik}(J) F_{kj}(J). \quad (13)$$

For large J we make use of the asymptotic saddle point expansion of the Bessel functions to order x^{-3} :

$$I_n(x) \underset{x \rightarrow \infty}{\cong} \frac{e^x e^{-\frac{n^2}{2x}}}{\sqrt{2\pi x}} \left[1 + \frac{1}{8x} + \frac{1}{x^2} \left(-\frac{n^2}{4} + \frac{9}{128} \right) + \frac{1}{x^3} \left(\frac{n^4}{24} - \frac{29n^2}{96} + \frac{75}{1024} \right) \right]. \quad (14)$$

Substituting this form into Eq. (12) one finds

$$F_{ij}(J) \underset{x \rightarrow \infty}{\cong} \frac{1}{d_i d_j} \left(F_{(1)i} F_{(1)j} + \frac{1}{x^3} F_{(2)i} F_{(2)j} \right) \frac{e^{2x}}{2\pi x}, \quad x \equiv 8J$$

$$F_{(\alpha)i} \equiv f_{(\alpha)i} - f_{(\alpha)i+1},$$

$$f_{(1)i} = e^{-i^2/x} \left[1 + \frac{1}{8x} + \frac{1}{x^2} \left(\frac{9}{128} - \frac{i^2}{2} \right) + \frac{1}{x^3} \left(\frac{75}{1024} - \frac{\kappa^2}{2} + \frac{i^2}{2} \left(\kappa - \frac{29}{24} \right) + \frac{i^4}{12} \right) \right],$$

$$f_{(2)i} = \frac{1}{\sqrt{2}} (i^2 + \kappa) e^{-i^2/x}. \quad (15)$$

Here the separation is not unique. We exploit this fact by introducing into (15) a constant κ to be determined later.

Now the decimation transformation (13) takes the form

$$F'_{ij} = \frac{e^{4x}}{(2\pi x)^2} \frac{1}{d_i d_j} a_1 \left[F_{(1)i} F_{(1)j} + \frac{a_2}{x^3} (F_{(1)i} F_{(2)j} + F_{(2)i} F_{(1)j}) \right],$$

$$a_1 = \sum_k F_{(1)k}^2, \quad a_1 a_2 = \sum_k F_{(1)k} F_{(2)k}. \quad (16)$$

It should be noted that the summations run over integer or halfintegers depending on the values of i and j . To the accuracy implied in Eq. (16) one evaluates the constants a_1 and a_2 with help of Poisson's resummation formula:

$$\sum_{k=-\infty}^{\infty} f(k) = \sum_{m=-\infty}^{\infty} \int d\xi e^{2\pi i m \xi} f(\xi). \quad (17)$$

In order to be able to apply (17) one has to reshuffle the summations over $k=1/2, 3/2, \dots$ or $k=0, 1, 2, \dots$ into summations over all integers with help of evenness

properties of the summands. It turns out finally that to leading order in J a_1 and a_2 are the same for both classes of the indices:

$$a_1 = \left(\frac{\pi}{8x}\right)^{1/2} + O(x^{-3/2}), \quad a_2 = -\frac{1}{4\sqrt{2}}x - \frac{1}{8\sqrt{2}} + O(x^{-1}) \quad x \gg 1. \quad (18)$$

Then the use of (18) in (16) shows that actually that recursion formula is $O(J^{-2})$ accurate.

Now we would like to find a rescaled value of J which fulfils the renormalisation group invariance equation:

$$F'_{ij}(J) = F_{ij}(J'). \quad (19)$$

The only possibility is that the leading piece in (15) when evaluated for J' having the form of (7) and expanded around J , gives rise to the second term in the square bracket of (16). We proceed by comparing the two expressions in order to find the constants c , β and κ . The result is

$$c = \frac{1}{32}, \quad \beta = 1, \quad \kappa = -\frac{1}{8}. \quad (20)$$

Finally using these values in (9) and (10) we have for the critical data:

$$J_c = \frac{1}{32\varepsilon \ln 2}, \quad \nu_J = \frac{1}{\varepsilon}. \quad (21)$$

The location of the fixed point is sensitive to the value of the scale ($b=2$) that is to the definition of the renormalisation group transformation. The "thermal" index ν_J is however universal in the Migdal-Kadanoff framework too.

The fact that we found a consistent solution to the renormalisation group equation (19) with the Ansatz (7) proves that the coupling J is marginal in $D=1$, which is therefore the lower critical dimension of the system (1). This means that the symmetry broken in the transition should be discrete.

The critical index ν_J found in (21) is the same as calculated by Migdal for the $D = 1 + \varepsilon$ dimensional Ising model in his original paper [6]. This fact establishes that the theory (1) belongs to the Ising universality class.

Along the same lines, in principle, one should be able to demonstrate that the SU(3) effective theory belongs to the Potts universality class in $D = 1 + \varepsilon$ dimensions.

References

1. B. Svetitsky and L. G. Yaffe, Nucl. Phys., *B210* [FS6] 423, 1982.
2. J. Kuti, J. Polónyi and K. Szlachányi, Phys. Lett., *98B*, 198, 1981.
R. V. Gavai and H. Satz, Phys. Lett., *145B*, 248, 1984.
3. J. Engels, E. Kehl, H. Satz and B. Wlitt, Bielefeld-preprint 85-10.
4. J. Polónyi and K. Szlachányi, Phys. Lett., *110B*, 395, 1982.
5. F. Green and F. Karsch, Nucl. Phys., *B238*, 297, 1984.
6. A. A. Migdal, ZsETF, *69*, 810, 1457, 1975.
7. L. P. Kadanoff, Ann. Phys. (N. Y.), *100*, 359, 1976.
8. P. H. Damgaard and A. Patkós, CERN-TH. 4322/85; Nucl. Phys. (to be published).

EXPANSIONS AT THE PARAMETER SPACE BOUNDARY IN THE STANDARD HIGGS MODEL*

I. MONTVAY

*Deutsches Elektronen-Synchrotron DESY
2000 Hamburg, FRG*

The strong self-coupling expansion and weak gauge coupling expansion are discussed in the lattice regularized SU(2) Higgs model with scalar doublet field.

1. Introduction

A basic ingredient of the Glashow–Weinberg–Salam SU(2)⊗U(1) theory of electro-weak interactions is the Higgs mechanism due to the non-zero vacuum expectation value of the scalar field. Its consequence is a non-zero mass for the gauge vector bosons and, after introducing the appropriate Yukawa couplings, the chiral symmetry breaking mass of the leptons and quarks. Both the SU(2) and U(1) gauge couplings are experimentally known to be small, therefore can be taken into account by perturbation theory. The starting point of the usual dimensionally regularized perturbation theory is, however, the *assumption* that the vacuum state is identical to the semi-classical one, given by the minimum of the classical potential. This fact alone could justify a non-perturbative investigation of the Higgs mechanism. The need for a non-perturbative approach to the electro-weak theory would become, of course, acute if it would turn out that the presently unknown scalar Higgs sector is strongly interacting. In addition, non-perturbative results in lattice regularized scalar ϕ^4 models, without the gauge couplings, strongly suggest that such models have a trivial continuum limit with vanishing physical interaction [1]. In order to establish the relation of the triviality of ϕ^4 models to the standard SU(2)⊗U(1) perturbation theory, a better understanding of lattice regularized Higgs models is required.

The lattice regularization [2] allows for a variety of different approaches: besides the powerful exact theorems it is also possible to perform approximate numerical calculations or different sorts of analytic expansions. The general strategy of the analytic expansions is to reduce the number of coupling parameters by sending some of them to the boundary of the coupling parameter space. The small parameter in the expansion is the distance from the boundary in some appropriately chosen metric. In

* Dedicated to Prof. K. Nagy on his 60th birthday

the example of the standard Higgs model (i.e. SU(2) gauge field coupled to a complex scalar doublet), which will be considered throughout this paper, there are three couplings: the scalar self-coupling λ , the gauge coupling g (or $\beta \equiv 4g^{-2}$) and the hopping parameter κ which represents in the lattice regularization the mass parameter for the scalar field. A possible expansion in this model is the *hopping parameter expansion* in powers of κ at $\kappa = 0$. The expansion coefficients are, in general, functions of the other two parameters (λ , β) involving non-trivial integrations over the gauge variables. A similar hopping parameter expansion is often used in gauge models with spin- $\frac{1}{2}$ fermion matter fields [3].

Another possibility, among others, is to expand at some point of the $\lambda = \infty$ plane. The resulting *strong self-coupling expansion* [5, 6] will be discussed in Section 2 of the present paper. As we shall see, the expansion coefficients depend also in this case non-trivially on the gauge variables. Such expansions can be rather useful, but for a completely analytic calculation one has to choose some particularly simple expansion point on the parameter space boundary, where all the integrations can be done analytically. Another important aspect is that one would like to control analytically the continuum limit of the lattice regularized theory, therefore the expansion point should also be a critical point with infinite correlation lengths. The classical example for an expansion fulfilling both these requirements is usual lattice perturbation theory, which is a double expansion in powers of λ and g at $\left(\lambda = 0, \beta = \infty, \kappa = \frac{1}{8}\right)$. The action is gaussian at this point, therefore the gauge and scalar field integrations can be easily performed. An important feature of perturbation theory is that it is done at a point with vanishing gauge coupling, therefore the expansion coefficients do not involve non-trivial gauge field integrations. Expanding only in the gauge coupling at some arbitrary point of the $\beta = \infty$ plane one obtains the more general *weak gauge coupling expansion*, which will be briefly discussed in Section 3 of this paper.

The advantage of the general weak gauge coupling expansion is that it allows for a formulation of the constraints imposed on the Higgs model by the expected triviality of the pure ϕ^4 model at $\beta = \infty$. The triviality of ϕ^4 means that the limit of the physical Green's functions at the critical line $(\lambda, \kappa_{cr}(\lambda))$ in the $\beta = \infty$ plane is proportional to the physical Green's functions of a free scalar field theory. Although the scalar field action looks non-trivial, at the critical line it describes a free field theory "in disguise". Therefore, if ϕ^4 is trivial indeed, then there exists an infinity of points $(\lambda = \text{arbitrary}, \beta = \infty, \kappa = \kappa_{cr}(\lambda))$ on the boundary of the parameter space, where in a weak gauge coupling expansion the remaining theory is free. One can also combine the weak gauge coupling expansion at the point $(\lambda = \infty, \beta = \infty, \kappa = \kappa_{cr}(\infty))$ with the strong self-coupling expansion. The resulting "strong-weak coupling expansion" is in principle similar to the usual double expansion at $\left(\lambda = 0, \beta = \infty, \kappa = \frac{1}{8}\right)$.

2. Strong self-coupling expansion

In this Section the general formulation of the strong self-coupling expansion [5, 6] will be summarized. The notations and conventions will in general be the same as in [7, 5, 6]. Therefore, the SU(2) gauge link-variables will be denoted by $U(x, \mu) \in \text{SU}(2)$, and the Higgs field will be described by the length ($\rho_x \geq 0$) and angular variable ($\alpha_x \in \text{SU}(2)$). x denotes lattice points, $\mu = \pm 1, \pm 2, \pm 3, \pm 4$ are link directions and (x, μ) is the link from the point x to the neighbouring point $(x + \hat{\mu})$ in direction μ . The lattice action in these variables can be written like

$$S_{\lambda, \beta, \kappa} = \beta \sum_P \left(1 - \frac{1}{2} \text{Tr} U_P \right) + \sum_x [\rho_x^2 - 3 \log \rho_x + \lambda(\rho_x^2 - 1)^2] - \kappa \sum_{(x\mu)} \rho_{x+\hat{\mu}} \rho_x \text{Tr} (\alpha_{x+\hat{\mu}}^+ U(x, \mu) \alpha_x). \quad (1)$$

Here Σ_P stands for a summation over positively oriented plaquettes, and $\Sigma_{(x\mu)} \equiv \Sigma_{x, \mu > 0}$ is a sum over positive links. The integration measure in the path integral is $d\rho_x d^3 \alpha_x d^3 U(x, \mu)$ (where $d^3 g$ denotes the Haar-measure in SU(2)). The peculiarity of the SU(2) doublet scalar field is that its angular part is equivalent to the local gauge degree of freedom. Therefore, at any finite β it is possible to introduce, instead of the SU(2) link- and site-variables, a gauge invariant link variable

$$V(x, \mu) \equiv \alpha_{x+\hat{\mu}}^+ U(x, \mu) \alpha_x. \quad (2)$$

In terms of this, the lattice action is

$$S_{\lambda, \beta, \kappa} = \beta \sum_P \left(1 - \frac{1}{2} \text{Tr} V_P \right) + \sum_x [\rho_x^2 - 3 \log \rho_x + \lambda(\rho_x^2 - 1)^2] - \kappa \sum_{(x\mu)} \rho_{x+\hat{\mu}} \rho_x \text{Tr} V(x, \mu). \quad (3)$$

After performing the trivial integration over α_x , the integration measure for Eq. (3) is $d\rho_x d^3 V(x, \mu)$.

In the limit $\beta \rightarrow \infty$ the variable change in Eq. (2) is inappropriate, because the gauge part of the action vanishes (the link-variables become gauge equivalent to unity). Therefore, for the $\beta \rightarrow \infty$ action one has to use

$$S_{\lambda, \beta = \infty, \kappa} = \sum_x [\rho_x^2 - 3 \log \rho_x + \lambda(\rho_x^2 - 1)^2] - \kappa \sum_{(x\mu)} \rho_{x+\hat{\mu}} \rho_x \text{Tr} (\alpha_{x+\hat{\mu}}^+ \alpha_x). \quad (4)$$

This defines a four-component ϕ^4 -model with global $\text{SU}(2) \otimes \text{SU}(2)$ - (or $\text{O}(4)$ -) symmetry.

In the $\lambda \rightarrow \infty$ limit the length of the Higgs-field is frozen to $\rho_x = 1$, and the action in Eq. (3) goes over into

$$S_{\lambda=\infty, \beta, \kappa} = \beta \sum_P \left(1 - \frac{1}{2} \text{Tr} V_P \right) - \kappa \sum_{(x\mu)} \text{Tr} V(x, \mu). \quad (5)$$

The corresponding limit of the action in Eq. (4) defines the $SU(2) \otimes SU(2)$ -symmetric non-linear σ -model on the lattice:

$$S_{\lambda=\infty, \beta=\infty, \kappa} = -\kappa \sum_{(x\mu)} \text{Tr} (\alpha_x^+ \hat{\mu} \alpha_x). \quad (6)$$

The derivation of the strong self-coupling expansion at the point ($\lambda = \infty$, β , $\bar{\kappa} \equiv s^2 \kappa$) starts from the relation

$$S_{\lambda, \beta, \kappa} = S_{\lambda=\infty, \beta, \bar{\kappa}} + \sum_x \left[-3 \log \rho_x + \lambda \left(\rho_x^2 - 1 + \frac{1}{2\lambda} \right)^2 \right] - \kappa \sum_{(x\mu)} (\rho_x + \hat{\mu} \rho_x - s^2) \text{Tr} V(x, \mu). \quad (7)$$

The integration over the length variables can be explicitly performed. Introducing the λ -dependent expectation value $\langle \dots \rangle_\lambda$, for an arbitrary ρ -dependent function $f(\rho)$, by

$$\langle f \rangle_\lambda \equiv \frac{\int [d\rho] f(\rho) \exp \left\{ - \sum_x \left[-3 \log \rho_x + \lambda \left(\rho_x^2 - 1 + \frac{1}{2\lambda} \right)^2 \right] \right\}}{\int [d\rho] \exp \left\{ - \sum_x \left[-3 \log \rho_x + \lambda \left(\rho_x^2 - 1 + \frac{1}{2\lambda} \right)^2 \right] \right\}} \quad (8)$$

the necessary integrals are of the form:

$$\langle \rho_y^k \rangle_\lambda = \frac{\int_0^\infty d\rho_y \rho_y^{3+k} \exp \left[-\lambda \left(\rho_y^2 - 1 + \frac{1}{2\lambda} \right)^2 \right]}{\int_0^\infty d\rho_y \rho_y^3 \exp \left[-\lambda \left(\rho_y^2 - 1 + \frac{1}{2\lambda} \right)^2 \right]} \equiv \frac{I_k(\lambda)}{I_0(\lambda)} \equiv i_k. \quad (9)$$

Here ρ_y is the length variable for some site y . An explicit form of these integrals can be given in terms of the parabolic cylinder functions (see [6], where numerical values are also listed in a few representative cases).

Let us denote a positively oriented link shortly by $l \equiv (x\mu)$, and the trace of the corresponding gauge invariant link variable by $T_l \equiv \text{Tr} V(x, \mu)$. Due to Eq. (7), the required λ -dependent expectation value is:

$$\begin{aligned} & \left\langle \exp \left[\kappa \sum_l T_l (\rho_x + \hat{\mu} \rho_x - s^2) \right] \right\rangle_\lambda = \\ & = \exp \left\{ \kappa \sum_l T_l (i_1^2 - s^2) + \sum_{n=2}^\infty \frac{\kappa^n}{n!} \sum_{l_1 \dots l_n} T_{l_1} \dots T_{l_n} c_n(l_1 \dots l_n) \right\}. \end{aligned} \quad (10)$$

Here the "connected i -product" c_n is defined by

$$c_n(l_1 \dots l_n) \equiv c(l_1 \dots l_n) \equiv \langle (\rho_{x_1} + \hat{\mu}_1 \rho_{x_1}) \dots (\rho_{x_n} + \hat{\mu}_n \rho_{x_n}) \rangle_{\lambda}^c. \quad (11)$$

The superscript c on $\langle \dots \rangle_{\lambda}^c$ denotes "connected" (λ -dependent) expectation value. In the definition of the connected parts, the products within parentheses have to be considered as a single entity. Identities similar to Eq. (10) are the basic tools for the derivation of the expansions considered in this paper. They convert integrals over some group of lattice variables into exponentials of connected expectation values of some functions of these variables.

The strong self-coupling coefficients of, say, the generating function of connected correlation functions can directly be obtained by applying identities of the type Eq. (10) also to the integral over link variables. Some explicit expressions for the expansion of different connected correlation functions are given in [6]. As a representative example, let us consider the connected correlation functions of the link trace T_l :

$$\langle T_{l_1} T_{l_2} \dots T_{l_m} \rangle_{\lambda, \beta, \kappa}^c = \sum_{n=0}^{\infty} \kappa^n \langle F^{(n)} T_{l_1} \dots T_{l_m} \rangle_{\lambda = \infty, \beta, \bar{\kappa}}^c. \quad (12)$$

The function $F^{(n)}$ can be written as a sum over partitions:

$$F^{(n)} = \sum_{n\{n_1 \dots n_k\}} f_{\{n_1 \dots n_k\}}^n T_{\{n_1 \dots n_k\}}^n. \quad (13)$$

The "partition" $n\{n_1 \dots n_k\}$ of the positive integer n is defined as a set of positive integers n_1, \dots, n_k satisfying

$$n = n_1 + 2n_2 + \dots + kn_k. \quad (14)$$

In Eq. (13) the factor $f_{\{n_1 \dots n_k\}}^n$ belonging to the partition $n\{n_1 \dots n_k\}$ is given by

$$f_{\{n_1 \dots n_k\}}^n = (i_1^2 - s^2)^{n_1} \prod_{i=1}^k \frac{1}{n_i! (i!)^{n_i}}. \quad (15)$$

The function $T_{\{n_1 \dots n_k\}}^n$ is the product of n_1 single-link variables, n_2 two-link variables, \dots , n_k k -link variables, summed over all the links. The first few $F^{(n)}$ in Eq. (13) are given explicitly by

$$\begin{aligned} F^{(1)} &= (i_1^2 - s^2) \sum_{l_1} T_{l_1}, \\ F^{(2)} &= \sum_{l_1 l_2} \left[\frac{1}{2} (i_1^2 - s^2)^2 T_{l_1} T_{l_2} + \frac{1}{2} c(l_1 l_2) (T_{l_1} T_{l_2}) \right], \\ F^{(3)} &= \sum_{l_1 l_2 l_3} \left[\frac{1}{6} (i_1^2 - s^2)^3 T_{l_1} T_{l_2} T_{l_3} + \frac{1}{2} (i_1^2 - s^2) c(l_2 l_3) T_{l_1} (T_{l_2} T_{l_3}) + \right. \\ &\quad \left. + \frac{1}{6} c(l_1 l_2 l_3) (T_{l_1} T_{l_2} T_{l_3}) \right], \end{aligned} \quad (16)$$

where let us recall that in the connected expectation values the variables put in parentheses count as a single entity.

The explicit form of the expansion coefficients can be used for the study of convergence of the strong self-coupling expansion series. As it was shown in [6], the series is essentially always convergent on a finite lattice, since it can be written as an $\alpha = 1$ ratio of two entire functions of some appropriate auxiliary variable α . The eventually appearing zeros of the denominator for $|\alpha| \leq 1$ can always be transformed away by a suitable conformal mapping in the α -plane. This implies a finite reordering of the series such that the k^{th} term of the new series is a linear combination of the first k terms of the old one. After this eventually necessary resummation the series will be convergent.

The extension of the convergence proof seems straightforward also to the case of infinite lattices. The only additional question is the convergence of the lattice link sums, but this can probably be controlled for appropriately chosen expansion points ($\lambda = \infty$, β , $\bar{\kappa}$) and λ -dependent curves $\kappa = s(\lambda)^{-2} \bar{\kappa}$.

In [6] the numerical values of some low order expansion coefficients were also determined by Monte Carlo integration at the point ($\lambda = \infty$, β , $\bar{\kappa}$). The calculated third order series are sufficient for λ -values in the range of $\lambda \simeq 10$. For $\lambda \simeq 0.1-1.0$, however, only a qualitative description of the direct Monte Carlo results is possible, for suitably chosen s^2 . A quantitative agreement in this λ -range requires higher than third order.

3. Weak gauge coupling expansion

The relation between the actions, instead of Eq. (7), is now

$$S_{\lambda, \beta, \kappa} = S_{\lambda, \beta = \infty, \kappa_0} + S_g - \sum_{(x\mu)} \{ \kappa \rho_x + \hat{\mu} \rho_x \text{Tr} [\alpha_x^+ \hat{\mu} U(x, \mu) \alpha_x] - \kappa_0 \rho_x + \hat{\mu} \rho_x \text{Tr} (\alpha_x^+ \hat{\mu} \alpha_x) \}. \quad (17)$$

The hopping parameter of the expansion point is denoted here by κ_0 , and S_g is the Wilson action for the SU(2) gauge field:

$$S_g \equiv \beta \sum_P \left(1 - \frac{1}{2} \text{Tr} U_P \right). \quad (18)$$

Introducing the notations

$$A(x, \mu) \equiv U(x, \mu) - 1 \equiv -a_{x\mu} + i\tau_r a_{rx\mu}, \quad \varphi_x \equiv \rho_x \alpha_x \quad (19)$$

the above equation can also be written as

$$S_{\lambda, \beta, \kappa} = S_{\lambda, \beta = \infty, \kappa_0} + S_g - \sum_{(x\mu)} \{ \kappa \text{Tr} [\varphi_x^+ \hat{\mu} A(x, \mu) \varphi_x] + (\kappa - \kappa_0) \text{Tr} (\varphi_x^+ \hat{\mu} \varphi_x) \}. \quad (20)$$

We are interested in the generating function Z of the gauge invariant connected correlation functions:

$$Z[r, k]_{\lambda, \beta, \kappa} \equiv \log \left\langle \exp \left\{ \sum_x r_x \rho_x + \sum_{r(x\mu)} k_{rx\mu} \text{Tr} [\tau_r \varphi_x^+ \hat{\mu} U(x, \mu) \varphi_x] \right\} \right\rangle_{\lambda \beta \kappa}. \quad (21)$$

Here τ_r , ($r = 1, 2, 3$) is a weak-isospin Pauli matrix. The derivatives of Z with respect to r give the connected correlation functions of the gauge invariant Higgs-boson variable ρ (weak isospin zero), whereas derivatives with respect to k produce the connected correlation functions of the isospin 1 gauge invariant W-boson variable $\text{Tr} [\tau \varphi^+ U \varphi]$. (Note that one could use, in principle, also other interpolating fields, for instance, in the Higgs-boson channel $\text{Tr} (\varphi^+ U \varphi)$ or in the W-boson channel $\text{Tr} [\tau \alpha^+ U \alpha]$, but this would not change anything essential.)

The path integral needed in Eq. (21) can be written as

$$\int [d\rho d^3 \alpha d^3 U] \exp(-S_{\lambda, \beta = \infty, \kappa_0} - S_g) \cdot \exp \left\{ \sum_x r_x \rho_x + \sum_{(x\mu)} [k_{rx\mu} v_{rx\mu} + (\kappa - \kappa_0) v_{x\mu} + a_{x\mu} j_{x\mu} + a_{rx\mu} j_{rx\mu}] \right\}, \quad (22)$$

where a summation over repeated isospin indices ($r, s, t, \dots = 1, 2, 3$) is always understood. The definition of the currents appearing in Eq. (22) is the following:

$$\begin{aligned} v_{x\mu} &\equiv \text{Tr} (\varphi_x^+ \hat{\mu} \varphi_x), \\ v_{rx\mu} &\equiv \text{Tr} (\varphi_x^+ \hat{\mu} \varphi_x \tau_r), \\ u_{rx\mu} &\equiv \text{Tr} (\varphi_x^+ \hat{\mu} \tau_r \varphi_x), \\ u_{rsx\mu} &\equiv \text{Tr} (\varphi_x^+ \hat{\mu} \tau_r \varphi_x \tau_s), \\ j_{x\mu} &\equiv -\kappa v_{x\mu} - k_{rx\mu} v_{rx\mu}, \\ j_{rx\mu} &\equiv i\kappa u_{rx\mu} + ik_{sx\mu} u_{rsx\mu}. \end{aligned} \quad (23)$$

The integration over the gauge field variables in Eq. (22) can be performed in the same way as in usual perturbation theory. The result is, of course, an infinite series in g^2 . The final form of the generating function Z can be obtained from an identity of the same type as Eq. (10) (for more details see [8]). The expansion point ($\lambda, \beta = \infty, \kappa_0$) can be chosen everywhere in the $\beta = \infty$ plane. If one wants, however, to study the questions related to the possibility of a continuum limit at ($\lambda, \beta = \infty, \kappa_c(\lambda)$) [7], then κ_0 has to be identified with $\kappa_c(\lambda)$. It is also possible to combine the weak gauge coupling expansion with the strong self-coupling expansion at $\lambda = \infty$. In this case the expansion point has only a single parameter (κ), and the expansion coefficients are expressed by expectation values in the non-linear σ -model defined by the action in Eq. (6).

Note added in proof: The lowest order weak gauge coupling expansion near the $\beta = \infty$ critical line implies that the continuum limit of the standard Higgs model is trivial [9].

References

1. For a review see, for instance, J. Fröhlich, in Progress in gauge field theory, Cargèse lecture 1983, ed. G.'t Hooft et al., Plenum Press New York, 1984.
2. K. G. Wilson, Phys. Rev., *D10*, 2445, 1974.
3. K. G. Wilson, in New phenomena in subnuclear physics, Erice 1975, edited by A. Zichichi, Plenum Press, New York, 1976; for a recent review on the application of the hopping parameter expansion to the numerical study of QCD see [4].
4. I. Montvay, Rev. Mod. Phys., *59*, 263, 1987.
5. I. Montvay, in Proceedings of the conference "Advances in Lattice Gauge Theory", April 1985, Tallahassee, Florida, ed. D. W. Duke, J. F. Owens.
6. K. Decker, I. Montvay, P. Weisz, Nucl. Phys. B. *268*, 362, 1986.
7. I. Montvay, Nucl. Phys. B. *269*, 170, 1986.
8. I. Montvay, Phys. Lett., *172*, 71, 1986.
9. I. Montvay, Nucl. Phys. B, to appear in 1987.

HAMILTON PRINCIPLE FOR THE VORTEX FLOW OF AN IDEAL FLUID IN SPECIAL RELATIVITY*

I. ABONYI

*Institute for Theoretical Physics, Roland Eötvös University
1088 Budapest, Hungary*

The fundamental equations of an ideal — entropy current conserving — fluid are obtained from a Hamilton Principle. The Lagrangian is formulated in the usual way. The physical situation, however, requires the introduction of some constraints in the variational principle to restrict the variations of the original field quantities. The Lagrange multipliers entering in such a way are shown to lead to the possibility of vortices in the velocity field.

1. Variational principles based on a convenient action integral play an important role in classical mechanics as well as in classical and quantum field theory. This role — because of its homogeneity in mathematical methods, its inherent insensitivity to the coordinate system used in the description — offers a unique possibility to exhibit the common features of the phenomena that belong to a branch of physics and so to decrease the number of first principles in the theory and may help to formulate the laws of the branch in question in a coordinate independent manner. The procedure called canonical quantization, based on an action principle offers a way to pass from classical to quantum theory.

In the case of the hydrodynamics of the classical ideal fluid an action principle has been used for a long time [1] which, however, has been found to be insufficient since it accounted only for the vortex-free flows of the fluid.

The possibility of the extension of the variational procedure to account for vortex flow has been shown by Herivel [2]. He realized that when one takes into account consequently requirements formulating the substance character and the ideal thermodynamic behaviour of the fluid, i.e. adds the continuity equation for the density and that of entropy density with multipliers to the Lagrange density, the variational principle with constraints will lead to a sufficient number of equations, one for each field quantity and multiplier, the flow velocity \mathbf{v} being determined in a Clebsch form

$$\mathbf{v} = -\nabla\alpha + \beta\nabla s,$$

* Dedicated to Prof. K. Nagy on his 60th birthday

where α and β are the Lagrange multipliers and s is the specific entropy. Evidently, if $\nabla s \neq 0$, then $\nabla \times \mathbf{v} \neq 0$. Herivel has shown that

$$\frac{\partial \beta}{\partial t} + (\mathbf{v} \cdot \nabla) \beta = \left(\frac{\partial u}{\partial s} \right)_\rho = T,$$

where $u(\rho, s)$ is the specific internal energy and T the absolute temperature.

2. The relativistic fluid will be described by the rest mass density (in the proper volume) μ , the specific entropy (per unit rest mass) s and the four-vector u_k of the flow velocity field. Use will be made of the specific internal energy which is related to the internal degrees of freedom of the fluid and so is quite different from rest energy already specified by μ . The specific internal energy e is considered as a function of μ and s , and by definition will be a scalar quantity, which satisfies the First and Second Law of Thermodynamics written in the form

$$de(\mu, s) = \left(\frac{\partial e}{\partial \mu} \right)_s d\mu + \left(\frac{\partial e}{\partial s} \right)_\mu ds = \frac{p}{\mu^2} d\mu + T ds, \quad (1)$$

where p and T stand for pressure and absolute temperature, both scalars; especially, when (1) is put in the form

$$T ds = dw - \frac{1}{\mu} dp, \quad (2)$$

where $w = e + \frac{p}{\mu}$ is the specific enthalpy.

We have to require that the flow velocity u_k be a four-vector of bound length, namely

$$u_r u_r = -c^2 \quad (3)$$

(summation convention understood for doubly occurring Latin indices) to make the physical interpretation possible.

The fluid material should be considered as a substance which is guaranteed by the continuity equation

$$\partial_r(\mu u_r) = 0. \quad (4)$$

The relevant thermodynamical model in the frame of which we want to describe the fluid is now chosen to be the ideal fluid, where no dissipation effects are considered throughout the fluid. So along the world line of any fluid element

$$u_r \partial_r s = \frac{ds}{d\tau} = 0. \quad (5)$$

Let us remark that Eq. (5) combined with Eq. (4) gives

$$\partial_r(\mu s u_r) = 0, \quad (6)$$

μs being the entropy density of the fluid, then the entropy current four-vector $\mu s u_k$ is divergence-free in our case.

The Hamilton principle will be announced for a four-dimensional space-time region Ω using a Lagrange-density L , a Lorentz-scalar of the dimension proper energy per proper volume. The fundamental part of the Lagrangian starts with a scalar rest mass density of the fluid (being the lowest order available nontrivial construction). The flow vector as a four-vector of bound length is not interesting for this fundamental part. The internal (or external) interaction of the fluid will be considered by the product of μ and the specific internal energy e (or specific potential φ), both being lowest order nontrivial scalars. Then the fundamental part of the Lagrangian L_F has the form:

$$L_F = \mu c^2 - \mu (e(\mu, s) + \varphi(x_k)). \quad (7)$$

Since the flow is not free, but subject to the constraints (3), (4) and (5), we take them into account by Lagrange multipliers using $K\mu$, α , $\beta\mu$, respectively.

After all this introduction the actual Lagrangian L for this variational problem with constraints is the following:

$$L = \mu c^2 - \mu (e(\mu, s) + \varphi(x_k)) + K\mu (u_k u_k + c^2) - \alpha \partial_r (\mu u_r) - \beta u_r \partial_r s. \quad (8)$$

The variation has to be executed in such a way — as usually — that the variations $\delta\mu$, δs , δu_k , δK , $\delta\alpha$, $\delta\beta$ vanish at the boundaries of the region of Ω .

Evidently, the Euler equations give back the constraints for the variations δK , $\delta\alpha$ and $\delta\beta$:

$$\delta K: \quad u_k u_k = -c^2, \quad (9)$$

$$\delta\alpha: \quad \partial_r (\mu u_r) = 0, \quad (10)$$

$$\delta\beta: \quad u_r \partial_r s = 0. \quad (11)$$

For the other variations, the calculation is not so trivial. So we follow it giving a short outline.

The variation of μ gives

$$\delta\mu: \quad c^2 + u_r \partial_r \alpha - \left(e + \mu \frac{\partial e}{\partial \mu} + \varphi \right) = 0. \quad (12)$$

When varying u_k , we obtain

$$\delta u_k: \quad 2K\mu u_r + \mu \partial_r \alpha - \mu \beta \partial_r s = 0. \quad (13)$$

Now, since $\mu \neq 0$, we omit this factor, and observe that the factor $2K$ does not play any important role, so it can be considered as incorporated into α and β , permitting us to write

$$u_k = -\partial_k \alpha + \beta \partial_k s. \quad (14)$$

When carrying out the variation of s , we obtain

$$\delta s: \quad \partial_r(\beta\mu u_r) - \mu \left(\frac{\partial e}{\partial s} \right)_\mu = 0. \quad (15)$$

Here, taking (10) and (1) into account, we have

$$u_r \partial_r \beta = \left(\frac{\partial e}{\partial s} \right)_\mu = T. \quad (16)$$

In the following, we transform Eq. (12) by taking its gradient, which gives

$$\partial_i(u_r \partial_r \alpha) - \partial_i \left(e + \varphi + \mu \frac{\partial e}{\partial \mu} \right) = 0. \quad (17)$$

When substituting here $\partial_r \alpha$ from (14), we conclude, that (17) leads to

$$-u_r \partial_r u_i + \left(\frac{\partial e}{\partial s} \right)_\mu \partial_i s - \partial_i \left(e + \frac{p}{\mu} \right) + \partial_i \varphi = 0,$$

which in virtue of the thermodynamic relations (1) and (2) can be put in the form

$$u_r \partial_r u_i = T \partial_i s - \partial_i w - \partial_i \varphi, \quad (18)$$

or

$$u_r \partial_r u_i = -\frac{1}{\mu} \partial_i p - \partial_i \varphi. \quad (19)$$

This, however, is equivalent with

$$\partial_r(\mu u_i u_r) = -\partial_i(p + \varphi), \quad (20)$$

or

$$\partial_r(\mu u_i u_k + \delta_{ir}(p + \varphi)) = 0. \quad (21)$$

Equations (18)–(21) represent one of the usual forms of the equations of motion for a barotropic fluid in special relativity (we use the term barotropic to underline that in this case $p = p(\mu)$, i.e. the pressure can be envisaged as a function of the rest mass density only).

References

1. F. London, *Superfluids*, J. Wiley, New York, 1955.
2. J. W. Herivell, *Proc. Cambridge Phil. Soc.*, 51, 344, 1955.

CORRELATIONS IN A NONEQUILIBRIUM STEADY STATE: EXACT RESULTS FOR A GENERALIZED KINETIC ISING MODEL*

Z. RÁCZ** and M. PLISCHKE

Department of Physics, Simon Fraser University Burnaby
British Columbia Canada V5A 1S6

A one-dimensional kinetic Ising model with domain wall dynamics corresponding to diffusion-limited annihilation in the presence of particle sources is considered. The spin-spin correlation function C_r is calculated exactly for a nonequilibrium steady state which can be brought close to a critical point of the system. We find that in contrast to equilibrium situations, C_r contains two distinctly diverging correlation lengths $\xi_1 \ll \xi_2$. While scaling appears for $r \sim \xi_1$, where ξ_1 is an equilibrium type correlation length characterizing the cooperativity of the spins, no simple scaling is observed for $r \gtrsim \xi_2$, where ξ_2 is a length closely related to the nonequilibrium kinetics of the system.

1. Introduction

The description of systems in thermodynamic equilibrium is enormously simplified by the fact that the probability for a given equilibrium state to occur is independent of the details of the underlying dynamics; it depends only on the additive constants of the motion which, in practice, means the energy of the state. It is an intriguing and quite important question whether some aspects of the equilibrium description could be generalized to *far-from-equilibrium* steady states. The attempts to answer this question seem to go along the following two lines.

i) One may try to generalize thermodynamics. Usually, this means that a set of deterministic equations are written down for the macroscopic variables and the influence of fluctuations arising from the microscopic degrees of freedom are taken into account by adding noise terms to those equations. Then, if possible, the equations are transformed into an equivalent Fokker-Planck equation and the stationary solution of the Fokker-Planck equation gives the steady-state distribution of the macroscopic variables and thus one finds the non-equilibrium thermodynamic potential of the system [1]. The aim of this approach is to examine simple cases and try to find the rules of construction of the nonequilibrium potential and thus bypass the tedious and often unsolvable details of the dynamics. Unfortunately no general rules have been found so far.

* Dedicated to Prof. K. Nagy on his 60th birthday

** On leave from the Institute for Theoretical Physics, Roland Eötvös University, H-1088 Budapest, Hungary.

ii) Another approach is the building and solving of simple models which incorporate the main features of concrete nonequilibrium systems [2, 3, 4]. As in the case of generalized thermodynamics, one hopes that by solving a sufficient number of these models, an understanding of the distinguishing features of far-from-equilibrium steady states will emerge. The advantage of the second method is that it does not aim at calculating the nonequilibrium potential which might not exist or may have strange analytical properties [5]; thus it is more likely to lead to the discovery of features which are novel and not present in the equilibrium case. The disadvantage of this approach, on the other hand, is that exactly solvable, nontrivial models are hard to come by and most approximation methods are biased by equilibrium notions thus making the results questionable. With the advent of computers, however, it seems that quite complicated models can be simulated successfully [2, 6, 7, 8] and thus there is a real possibility of accumulating a wealth of information about simple nonequilibrium situations.

In this paper a contribution to the second approach is presented by an exact calculation of the two-point correlations in a far-from-equilibrium steady state of a generalized kinetic Ising model. The model in question originates in studies of one-dimensional diffusion-limited annihilation in the presence of particle sources [4]. This model may be relevant for understanding both the dynamics of solitons in systems like trans-polyacetylene when solitons are generated by photons and chemical reactions on low dimensional substrates.

As has been shown [4], one-dimensional diffusion-limited annihilation can be described as the domain-wall motion and annihilation in the one-dimensional kinetic Ising model [9] at zero temperature ($T=0$). The particle creation can be incorporated by introducing an extra spin-flip process and the resulting generalized kinetic Ising model can be solved exactly [4]. For example, for the steady state in which the annihilation of neighboring pairs of particles is balanced by the particle creation, one obtains that the average density \bar{n} is given by $\bar{n} \sim h^{1/3}$ where h is the rate of creation of the particles. Furthermore, the relaxation towards the steady state can be characterized by a relaxation time τ which scales with h as $\tau \sim h^{-2/3}$, i.e. one can observe critical slowing down in the $h \rightarrow 0$ limit. The scaling of \bar{n} and τ is similar to that found in equilibrium systems near a critical point. This is not entirely surprising since $h=0$ is the equilibrium critical point of the one-dimensional Ising model. It should be pointed out, however, that particle creation is a nonequilibrium process (there is, e.g., no detailed balance) and thus the analogies with an equilibrium system are not obvious even if the system is close to an equilibrium critical point. The aim of this paper is to demonstrate how the nonequilibrium aspects of the process show up in the correlation function of the system.

Near an equilibrium critical point of a homogeneous system one expects that the two-point correlation function $C(r)$ obeys scaling

$$C(r) \sim \frac{1}{r^x} \varphi\left(\frac{r}{\xi}\right). \quad (1)$$

Here x is a critical exponent characterizing the scale invariance at the critical point where the correlation length diverges ($\xi \rightarrow \infty$) and $\varphi(y)$ is just a cutoff function ($\varphi(y) \ll 1$ for $y \ll 1$) which describes the fact that the correlations diminish for $r \gg \xi$. We shall find that in contrast to equilibrium systems, the correlation function in the steady state is characterized by two correlation lengths. One of them ξ_1 is similar in nature to the equilibrium correlation length since it is determined by the steady-state properties of the system. Namely, ξ_1 is just the average distance between the particles (i.e. the average length of the aligned spin domains) and thus scales as $\xi_1 \sim \bar{n}^{-1} \sim h^{-1/3}$. The other length ξ_2 is kinetic in nature, it is the average distance between the particles created in a given time interval. Since this distance is inversely proportional to the rate of creation of the particles, one has $\xi_2 \sim h^{-1}$ and thus $\xi_1 \ll \xi_2$ in the $h \rightarrow 0$ limit. Our results for $C(r)$ can be summarized in terms of ξ_1 and ξ_2 as follows. For $r \sim \xi_1$, the correlation function assumes a scaling form (1) with $x=0$ and $\xi \sim \xi_1$ being the "steady-state" correlation length. For $r \gtrsim \xi_2 \gg \xi_1$, the particle creation process introduces extra effects and, as a result, the asymptotics of $C(r)$ changes and it cannot be written any more in a simple scaling form. Since the correlations are practically negligible in the region $r \sim \xi_2$ it appears that the extra correlations do not affect the properties of the system significantly. The presence of the second diverging lengthscale might nevertheless be important since it may turn out to be a distinguishing feature of nonequilibrium steady states near a critical point. This idea is supported by a growing body of observations that second diverging lengthscales appear in various growth models which display nonequilibrium steady states if the growth is restricted in all but one dimension [6-8, 10].

The physical content of this paper is described above. What remains is the derivation of the results which will be done in the next Section.

2. The model and the calculation of the correlation function

Consider a system with states $\{\sigma\} \equiv \{\dots, \sigma_i, \sigma_{i+1}, \dots\}$ specified by stochastic spin variables $\sigma_i(t) = \pm 1$ which are assigned to lattice sites labeled by an integer $-\infty < i < \infty$. The dynamics of this spin system is described by the probability distribution $P(\{\sigma\}, t)$ satisfying the following master equation:

$$\dot{P}(\{\sigma\}, t) = \sum_{i=-\infty}^{\infty} \sum_{\alpha=1}^2 [w_i^{(\alpha)}(\{\sigma\}_i^\alpha) P(\{\sigma\}_i^\alpha, t) - w_i^{(\alpha)}(\{\sigma\}) P(\{\sigma\}, t)]. \quad (2)$$

Here the state $\{\sigma\}_i^1$ differs from $\{\sigma\}$ by flipping the i -th spin and the rate of flipping is given by

$$w_i^{(1)}(\{\sigma\}) = \frac{\Gamma}{2} \left[1 - \frac{1}{2} \sigma_i (\sigma_{i-1} + \sigma_{i+1}) \right], \quad (3)$$

where Γ is just a parameter setting the timescale. If no other spin-flip processes were present ($w^{(2)}=0$) then (2) and (3) would define the kinetic Ising model [9] at zero temperature. A little reflection on the form (3) shows that only spins neighboring a domain wall can flip and, as a result, the domain walls move to the left or right with equal rate $\Gamma/2$ while two nearest-neighbor domain walls annihilate each other at a rate Γ . Thus identifying the domain walls with particles, we can see that this model describes diffusion limited annihilation as discussed in the Introduction. In order to have single particle sources homogeneously distributed in the lattice (Model I in [4]), one assumes that $\{\sigma_i\}_i^2$ in (2) differs from $\{\sigma\}$ by the simultaneous flipping of all the spins σ_i , $l \leq i$ and that the rate of this process is independent of the state of the system

$$w_i^{(2)}(\{\sigma\}) = \Gamma h. \quad (4)$$

In this way one creates domain walls (i.e. particles) at a rate Γh per site and this process also obeys the rule that particle creation at an already occupied site is equivalent to emptying that site. Thus equation (2) with (3) and (4) defines a generalized kinetic Ising model which describes diffusion limited annihilation in the presence of particle sources.

Although the particle interpretation of this model is more important for practical purposes, the underlying spin model is equally interesting since it is an example of a simple but nontrivial model which displays a nonequilibrium steady state. In this paper we shall concentrate on the spin aspects of the model and calculate the steady state spin-spin correlation function

$$C(i, j) = \lim_{t \rightarrow \infty} \langle \sigma_i(t) \sigma_j(t) \rangle, \quad (5)$$

where the brackets denote averaging over the distribution function $P(\{\sigma\}, t)$. Since there is no mechanism in this model which could produce an inhomogeneous state, the calculation of $C(i, j)$ involves only translationally invariant states. Thus we can assume that the initial distribution $P(\{\sigma\}, 0)$ and consequently $P(\{\sigma\}, t)$ at any t is translationally invariant. It follows then that $\langle \sigma_i(t) \sigma_j(t) \rangle$ and $C(i, j)$ depend only on $l = i - j$ and the following notation can be introduced

$$C_l = \lim_{t \rightarrow \infty} C_l(t) = \lim_{t \rightarrow \infty} \langle \sigma_i(t) \sigma_{i+l}(t) \rangle. \quad (6)$$

We can obtain a closed set of linear differential equations for C_l by multiplying (2) by $\sigma_i \sigma_{i+l}$ and summing over all configurations $\{\sigma\}$. Following the steps of Glauber's work [9] where the equation determining $C_l(t)$ for the kinetic Ising model is derived, one obtains for $l > 0$

$$\Gamma^{-1} \dot{C}_l(t) = C_{l+1}(t) + C_{l-1}(t) - 2(1 + hl)C_l(t), \quad (7)$$

while $C_l(t) = C_{-l}(t)$ for $l < 0$ and $C_{l=0}(t) = \langle \sigma_i^2 \rangle = 1$.

Equation (7) has already been investigated in [4] since the nearest-neighbor correlation $C_1(t)$ determines the particle density and the time evolution of homo-

geneous density fluctuations in the particle version of the model ($C_l(t)$ was denoted as r_l and C_l as \bar{r}_l in [4]). It has been established in [4] that for $t \rightarrow \infty$, all the solutions of Eq. (7) decay exponentially to a time independent solution C_l determined from

$$C_{l+1} + C_{l-1} - 2(1 + hl)C_l = 0. \quad (8)$$

The solution of this equation which is finite in the $l \rightarrow \infty$ limit (note that $|C_l| \leq 1$) and satisfies the condition $C_{l=0} = 1$ is obtained by noting that the Bessel functions $J_\nu(z)$ of the first kind of order ν satisfy the recursion relation [11]

$$J_{\nu+1}(z) + J_{\nu-1}(z) - \frac{2\nu}{z} J_\nu(z) = 0. \quad (9)$$

Comparing (8) and (9) one obtains the steady-state correlation function [4]

$$C_l = J_{l+h^{-1}}(h^{-1}) / J_{h^{-1}}(h^{-1}). \quad (10)$$

The properties of the correlation function can now be found by studying various limits of the Bessel function. Since we are interested in the form of C_l near the critical point ($h \ll 1$), we have $l + h^{-1} \gg 1$, i.e. the properties of the Bessel functions of large order are needed. They can be found in standard monographs [11, 12], and we shall quote only the results.

There are two main regions to be distinguished in the large order asymptotics of $J_\nu(z)$ (Note that only the case $\nu \geq z$ has to be examined since $\nu = l + h^{-1}$ and $z = h^{-1}$ in (10)). The first one is the region where $\nu - z \sim \nu^{1/3}$ while the second one is the limit $\nu \gg z$. The first case corresponds to $l \sim h^{-1/3}$ while the second one to $l \gg h^{-1}$ and this is how the two correlation lengths $\xi_1 \sim h^{-1/3}$ and $\xi_2 \sim h^{-1}$ arise from the properties of the Bessel function. The physical interpretation of ξ_1 and ξ_2 is quite obvious since $\xi_1 - h^{-1/3} \sim \bar{n}^{-1}$ is proportional to the average distance between the domain walls [4], i.e. it is the average length of the aligned spin domains while $\xi_2 \sim h^{-1} \sim [w_i^{(2)}]^{-1}$ is proportional to the average distance between domain walls created in a given time interval.

In order to have the actual form of the correlation function in the $l \sim \xi_1 \sim h^{-1/3}$ region one can use the following limit [11]

$$\lim_{\nu \rightarrow \infty} J_\nu(\nu + z\nu^{1/3}) \approx 2^{1/3} \nu^{-1/3} Ai(-2^{1/3}z) + O(\nu^{-1}), \quad (11)$$

where $Ai(z)$ is the Airy function [11]. Substituting (11) into (10) one obtains

$$C_l \approx \frac{Ai(2^{1/3}lh^{1/3})}{Ai(0)} = \frac{Ai(l/\xi_1)}{Ai(0)}, \quad (12)$$

where we defined $\xi_1 = (2h)^{-1/3}$ (Note that all quantities which have the dimension of length are measured in units of the lattice spacing). Comparing (12) with (1) one can see that C_l has a scaling form in the $l \sim \xi_1$ region. The particular value of the critical

exponent $x=0$ is not surprising since x seems to be zero for many equilibrium models in one dimension (simple examples are the Ising and Potts models).

Within the region $l \sim h^{-1/3}$ one can still consider limits $l/\xi_1 \ll 1$ and $l/\xi_1 \gg 1$ and obtain more explicit forms for C_l by using the properties of the Airy function [11]. For $l/\xi_1 \ll 1$ we find

$$C_l \approx 1 + \frac{Ai'(0)}{Ai(0)} \frac{l}{\xi_1} + \dots = 1 - \frac{3^{1/3} \Gamma(2/3)}{\Gamma(1/3)} \frac{l}{\xi_1} + 0 \left(\left(\frac{l}{\xi_1} \right)^2 \right), \quad (13)$$

where $\Gamma(z)$ is the Gamma function [11]. For $l/\xi_1 \gg 1$, on the other hand, the correlations decay faster than exponentially

$$C_l \approx \frac{1}{2\sqrt{\pi} Ai(0)} \frac{\exp \left[-\frac{2}{3} (l/\xi_1)^{3/2} \right]}{(l/\xi_1)^{1/4}} \left[1 + 0 \left(\frac{\xi_1}{l} \right) \right], \quad (14)$$

where $Ai(0) = [3^{2/3} \Gamma(2/3)]^{-1}$. Of course, the above form is valid only if not only $1 \ll lh^{1/3}$ but also $lh \ll 1$. For $l \gg h^{-1} \sim \xi_2$ one must use another asymptotic expansion of $J_\nu(z)$ [11, 12]:

$$J_\nu(z) \approx \frac{1}{\sqrt{2\pi}} \exp \left[\nu + \nu \ln \frac{z}{2} - \left(\nu + \frac{1}{2} \right) \ln \nu \right] \cdot \left[1 + 0 \left(\frac{1}{\nu} \right) \right]. \quad (15)$$

Substituting $\nu = l + h^{-1}$ and $z = h^{-1}$ one can see that no simple scaling form (1) appears. The most important factor in the decay of correlations, however, can easily be identified to be $(\xi_2 = (2h)^{-1})$:

$$C_l \sim \exp \left[-l \left(\ln \frac{l}{\xi_2} - 1 \right) \right], \quad (16)$$

i.e. the decay of correlations is again faster than exponential.

The reason for the fact that there is scaling in the region $l \sim h^{-1/3}$ but no scaling appears for $l > h^{-1}$ is quite simple. Scaling is usually the result of some cooperative phenomena and the lengthscale on which scaling appears is the lengthscale of cooperativity, i.e. the lengthscale on which some degrees of freedom in the system are strongly correlated. In our case the length of domains of aligned spins are of the order of ξ_1 , i.e. the spins are strongly correlated on the lengthscale of ξ_1 and this is why it can be expected without any calculation that there is a scaling region around $l \sim \xi_1$. In contrast to ξ_1 the other correlation length ξ_2 is not connected with cooperativity. It is actually related to a process (creation of domain walls) which destroys correlations thus, as has been found, no scaling exists in the region $l \gtrsim \xi_2$.

Finally, it should be noted that the correlation function C_l has not been investigated in the transitional region

$$h^{-1/3} \ll l \sim h^{-\sigma} \ll h^{-1}, \quad (17)$$

where $1/3 < \sigma < 1$. No simple representation of the Bessel function is known to us in this regime. On the basis of the above remarks about cooperativity, however, we believe that nothing but the matching of the forms (14) and (16) happens in the transitional region.

3. Remarks

Although the two-point correlations contain a great deal of information about the system, it would clearly be valuable if one could calculate the steady-state distribution for the simple but non-trivial model discussed in this paper. At this moment we do not see any straightforward way of doing it but the following considerations may help in making the first steps towards constructing the steady-state distribution.

The model described by Eqs (2), (3) and (4) can be generalized to a finite temperature T so that detailed balance is observed not only for the single spin-flip process but also for the domain-wall creation and annihilation. In order to achieve this (3) and (4) must be changed to

$$w_i^{(1)}(\{\sigma\}) = \frac{\Gamma}{2} \left[1 - \frac{\gamma}{2} \sigma_i (\sigma_{i+1} + \sigma_{i-1}) \right], \quad (18)$$

$$w_i^{(2)}(\{\sigma\}) = \Gamma h (1 - v \sigma_i \sigma_{i+1}), \quad (19)$$

where $\gamma = \tan h(2J/k_B T)$ and $v = \tan h(J/k_B T)$ with J being the nearest-neighbor coupling in the Ising model. In this new model the steady state is an equilibrium state and the equilibrium distribution is known. Now one can consider the parameters γ and v as independent quantities and introduce for example $v' = v - \varepsilon$ and carry out systematic expansion in ε in order to approach the $v' = 0$ limit. Although this will be an expansion around an equilibrium state it might still provide a clue of how to construct the far-from-equilibrium ($\gamma = 1, v = 0$) steady-state distribution.

Acknowledgement

This work was supported by the Natural Sciences and Engineering Research Council of Canada.

References

1. R. Graham, in *Stochastic Processes in Non-equilibrium Systems*, edited by L. Garrido, P. Seglar, and P. J. Sheppard, Lecture Notes in Physics, Vol. 84, Springer, New York, 1978, p. 82.
2. S. Katz, J. L. Lebowitz and H. Spohn, *Phys. Rev.*, B28, 1655, 1983; *J. Stat. Phys.*, 34, 497, 1984.
3. H. van Beijeren and L. S. Schulman, *Phys. Rev. Lett.*, 53, 806, 1984.
4. Z. Rácz, *Phys. Rev. Lett.*, 55, 1707, 1985.

5. R. Graham and T. Tél, *Phys. Rev. Lett.*, *52*, 9, 1984.
6. P. Meakin, *Phys. Rev.*, *A27*, 2616, 1983; Z. RÁCZ and T. Vicsek, *Phys. Rev. Lett.*, *51*, 2382, 1983.
7. F. Family and T. Vicsek, *J. Phys.*, *A18*, L75, 1985.
8. M. Plischke and Z. RÁCZ, *Phys. Rev.*, *A32*, 3825, 1985.
9. R. J. Glauber, *J. Math. Phys.*, *4*, 294, 1963.
10. M. Plischke and Z. RÁCZ, *Phys. Rev. Lett.*, *53*, 415, 1984.
11. *Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun, Dover, New York, 1965.
12. G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge, 1962.

WEAK INTERACTIONS OF HEAVY QUARKS* **

H. PIETSCHMANN

Institute for Theoretical Physics, Vienna University
Vienna, Austria

The status of our knowledge about weak interactions of heavy quarks is reviewed within the framework of the Standard Theory. Several approaches to connect the basic Lagrangian to experimental measurements are discussed.

1. Introduction

After the spectacular discovery of the W and Z bosons [1] it has become clear that the so-called "Standard Model" is the paradigm of our days for the description of elementary particles and their interactions. As far as the leptons are concerned, we can directly derive predictions from the standard Lagrangian and observe spectacular agreement with experimental measurements. On the other hand, in the quark sector things are not as easy; this is due to the fact that quarks can never be observed as free particles, rather their interactions have to be connected to their behaviour in bound states. Hence, phenomenology is still a very important tool in deriving these connections and leaves ample room for theoretical work.

The term "heavy quarks" can either be restrictedly used for the third generation (i.e. top and bottom quarks) or one can more widely use the term also for the charmed quark. Due to the lower mass of the charmed quark, corrections may be bigger and therefore predictions of the standard Lagrangian not so clean. We shall therefore mainly deal with the bottom quark, but refer to the charmed quark also whenever it is of particular interest.

The part of the standard model Lagrangian [2] which is relevant for our considerations is given in Eq. (1),

$$L_q = \sum_{qq'} \left\{ \frac{g}{2\sqrt{2}} [\bar{q}\gamma_\lambda(1 + \gamma_5)U_{qq'}q' \cdot W^\lambda + h.c.] + \frac{g}{4 \cos \theta_w} [\bar{q}\gamma_\lambda(C_V + \gamma_5)q - \bar{q}'\gamma_\lambda(C_V' + \gamma_5)q']Z^\lambda \right\}, \quad (1)$$

* Dedicated to Prof. K. Nagy on his 60th birthday.

** Work supported in part by "Fonds zur Förderung der wissenschaftlichen Forschung in Österreich", Project Nr. P5444.

where

$$q \in (u, c, t), \quad (2)$$

$$q' \in (d, s, b)$$

and

$$C_V = 1 - \frac{8}{3} \sin^2 \theta_W, \quad (3)$$

$$C'_V = 1 - \frac{4}{3} \sin^2 \theta_W.$$

$U_{qq'}$ is the Kobayashi–Maskawa matrix [3]. Eq. (1) also serves to establish notation. In the forthcoming, we shall only use the charged current part of the Lagrangian equation (1), but let us notice in passing, that the neutral current part predicts that there are no flavour changing transitions in the neutral current sector. This has indeed been tested experimentally and we shall here just mention one of the results [4],

$$\text{B.R.}(\overline{B}^0 \rightarrow \mu^+ \mu^-) < 2 \cdot 10^{-4}, \quad 90\% \text{ C.L.} \quad (4)$$

$$\text{B.R.}(\overline{B}^0 \rightarrow e^+ e^-) < 3 \cdot 10^{-4}.$$

2. Semi-leptonic decays

As is good old tradition in the physics of weak interactions, semileptonic processes provide the best means to test weak interactions of hadrons. This is again the case for heavy quarks, since we can use the spectator approximation as a very accurate tool.

The semi-leptonic decay rate for a bottom quark (with obvious modification also for charmed or top quarks) is given in Eq. (5):

$$\Gamma(b \rightarrow ql\bar{\nu}_l) = \frac{G^2 m_b^5}{192\pi^3} \eta_{SL} |U_{bq}|^2 c_{ql}, \quad (5)$$

where η_{SL} is a semi-leptonic correction due to real and virtual gluon emission and c_{ql} is a phase space factor. The value of η_{SL} is 0.88–0.90 for the QCD cut-off Λ_{QCD} between 0.1–0.2 GeV. For the electronic decay mode, in which the leptonic mass can be neglected, the phase space factor is given in Eq. (6):

$$c_{ql} = (1 - a^2)(1 - 8a + a^2) - 12a^2 \ln a, \quad (6)$$

where

$$a = \left(\frac{m_q}{m_b}\right)^2. \quad (7)$$

For the muonic mode, the phase space factor is given by

$$c_{q\mu} = 12 \int_{(a+b)^2}^1 dx \frac{(1-x)^2}{x} (x-a-b) \sqrt{x - (\sqrt{a} + \sqrt{b})^2} \sqrt{x - (\sqrt{a} - \sqrt{b})^2}, \quad (8)$$

where a is given by Eq. (7) and

$$b = \left(\frac{m_\mu}{m_b} \right)^2. \quad (9)$$

In this way, it was for instance possible to obtain information on the ratio of the Kobayashi-Maskawa matrix elements [5]

$$\left| \frac{U_{bu}}{U_{bc}} \right|^2 < 0.08. \quad (10)$$

3. Non-leptonic decays

The diagrams contributing to non-leptonic decays are shown in Fig. 1. The spectator approximation may still be good, but the presence of other diagrams is directly visible in certain exclusive decay channels, in which only non-spectator diagrams can contribute. An example is the process

$$D^0 \rightarrow \bar{K}^0 \Phi, \quad (11)$$

which can only occur via the W -exchange diagram. Its branching ratio has been determined by the ARGUS-collaboration [6] to be $(0.99 \pm 0.32 \pm 0.17)\%$.

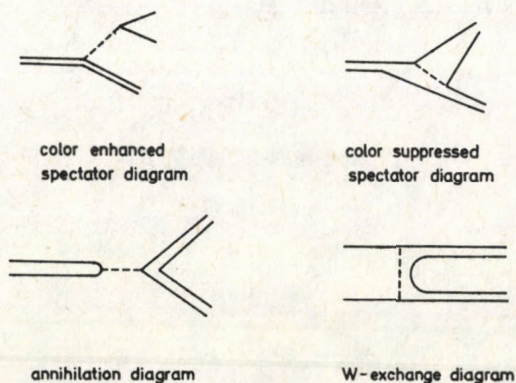


Fig. 1

Also the color suppressed spectator diagram has been established in the decay

$$B \rightarrow \psi + \text{anything}, \quad (12)$$

which has been seen with a branching ratio of $(1.37 \pm 0.55)\%$ by the ARGUS-collaboration and $(1.10 \pm 0.21 \pm 0.23)\%$ by the CLEO-collaboration [5] in fair agreement with theoretical expectations [7].

The validity of the spectator approximation can be tested by its prediction that the decay rate for charged and neutral hadrons involving a heavy quark should be identical. Quite obviously, this is strongly violated in the case of D particles, but it might be more satisfactory in the case of B particles. Another way of testing the spectator approximation is the observation of the decay particle energy distribution. This has been done in satisfactory agreement with spectator predictions [8].

The non-leptonic decay rate for bottom quark decay (and similarly for top or charm quark decay) in the spectator approximation can be written as

$$\Gamma_{NL}^b = \frac{G^2 m_b^5}{192\pi^3} \sum_{q,q'} |U_{bq}|^2 c_{qq'} \eta_{\text{QCD}} = \frac{G^2 m_b^5}{192\pi^3} \{ |U_{bu}|^2 C_u^b + |U_{bc}|^2 C_c^b \}. \quad (13)$$

$c_{qq'}$ is again a phase space factor which is either similar to Eqs (6) and (8) or — in the more general case — can be taken from [9].

η_{QCD} is the enhancement factor stemming from box diagrams in which a gluon goes parallel to the W from one quark leg to the other. They vanish in lowest order and have to be taken in leading log approximation [9]

$$\eta_{\text{QCD}} = \frac{1}{3} \left\{ 2 \left[\frac{\alpha_s(m_W^2)}{\alpha_s(m_b^2)} \right]^{12/23} + \left[\frac{\alpha_s(m_b^2)}{\alpha_s(m_W^2)} \right]^{24/23} \right\}. \quad (14)$$

In principle, there is also a correction similar to η_{SL} of Eq. (5) stemming from real and virtual one gluon emission but in the non-leptonic case it is very close to unity and we have dropped it from Eq. (13).

“Canonical” values for C_u^b and C_c^b have been derived by Gaillard and Maiani [11]. For the quark masses given in Eqs (15)

$$\begin{aligned} m_b &= 5 \text{ GeV}, \\ m_c &= 1.75 \text{ GeV}, \\ m_s &= 0.5 \text{ GeV}, \\ m_\tau &= 1.8 \text{ GeV}, \\ \eta_{\text{QCD}} &= 1.3 \end{aligned} \quad (15)$$

they obtain

$$C_u^b = 7.69, \quad C_c^b = 2.75. \quad (16)$$

These values are still used in most of the experimental analysis to obtain Kobayashi–Maskawa matrix elements (for example [12]).

In order to get some feeling for the range, in which the parameters of Eq. (16) can vary due to theoretical model assumptions, one can use a semi-phenomenological

approach [13] and express the decay rate, Eq. (13), by means of spectral functions in the following way [2]

$$\Gamma(b, q) = \frac{G^2 |U_{bq}|^2 m_b^3}{16\pi} \int_0^{(1-\sqrt{a})^2} \frac{dx}{x} w(1, a, x) \cdot \{ [(1-a)^2 + (1+a)x - 2x^2] \rho_b(x) + [(1-a)^2 - (1+a)x] \rho'_b(x) \}, \quad (17)$$

where a is again given by Eq. (7) and the spectral functions are defined by

$$(2\pi)^3 \sum_n \delta^{(4)}(P-k) \langle 0 | j_\mu^\dagger(0) | n(k) \rangle \langle n(k) | j_\nu(0) | 0 \rangle = \left(\frac{P_\mu P_\nu}{P^2} - g_{\mu\nu} \right) \rho(P^2) + \frac{P_\mu P_\nu}{P^2} \rho'(P^2) \quad (18)$$

and

$$\rho_b(x) = \rho(x m_b^2) = \rho[(p_b - p_q)^2]. \quad (19)$$

Following [13], one can now use empirical informations on the spectral functions either from τ -decay or from the semi-theoretical relation of the spectral functions to the ratio

$$R = \frac{\sigma(e^+ e^- \rightarrow \text{hadrons})}{\sigma(e^+ e^- \rightarrow \mu^+ \mu^-)} = \frac{12\pi^2}{s} \rho_{\text{el.m.}}(s). \quad (20)$$

The latter method has successfully been used in deriving decay rates for the Z -boson [14]. Using the exact lepton masses and varying the quark masses within a reasonable range with respect to Eq. (15), one finds in this way that a reliable range for the coefficients, Eq. (16), is

$$C_u^b \simeq 7.4 - 8.0, \quad C_c^b \simeq 2.7 - 3.2. \quad (21)$$

(QCD corrections for the semi-phenomenological method of [13] have been calculated by Klein and Rupertsberger [15].)

For the top decay, the variation of the parameters is much limited due to the high mass of the top. On the other hand, the top mass is not yet known, but for a range of 35–45 GeV the equation analogous to Eq. (17) becomes [16]

$$\tau_t = \frac{\tau_\mu (m_\mu/m_t)^5}{(7.5 - 8.1) |U_{tb}|^2 + 8.9 |U_{ts}|^2 + 8.9 |U_{td}|^2}, \quad (22)$$

where the lower end of the range of the coefficient of U_{tb} corresponds to lower mass values of the top.

4. Branching ratios

Combining the results of the two previous sections, the semi-leptonic branching ratio (per generation) turns out to be 13–15%. Experimentally, however, the average value seems to be rather smaller ($11.6 \pm 0.6\%$) [17]. Some authors attribute this to a rather large contribution of non-spectator diagrams [18]. The matter should be cleared up, as soon as the lifetime of charged and neutral B -mesons can be disentangled.

Another interesting quantity which has been measured is the ratio of B -mesons decaying into D^* versus D -mesons; it is connected to the ratio of spin flip and non-flip decay, i.e. to the old nuclear physics questions of the ratio of Gamow–Teller versus Fermi matrix element. Measurements [19] are still rather crude and theoretical predictions somewhat uncertain [20]. In the spirit of the old nuclear physics question, the static quark model should give a first indication [21].

5. Conclusion

In spite of the fact that the standard theory provides us with a paradigm Lagrangian for heavy quarks which should describe all the interactions, interesting theoretical work is still necessary to arrive at firm predictions. This is due to the fact that quarks are not available as free particles. Much credit should be given in particular to experimentalists who are able to isolate relevant phenomena at energies in which background problems as well as precision problems are a fantastic challenge, which nevertheless has been overcome.

References

1. UA1 Collaboration, CERN, G. Arnison et al., Phys. Lett., *122B*, 103, 1983; UA2 Collaboration, M. Banner et al., Phys. Lett., *122B*, 476, 1983.
2. For a complete review of the Standard Model see for example H. Pietschmann, Weak Interactions, Formulae, Results and Derivations, Springer Verlag, Wien–New York, 1983.
3. M. Kobayashi and K. Maskawa, Progr. Theor. Phys., *49*, 652, 1973.
4. CLEO-Collaboration, Phys. Rev., *D30*, 2279, 1984.
5. E. H. Thorndike, Int. Symp. on Lepton and Photon Interactions at High Energies, Kyoto, Japan, 1985.
6. ARGUS-Collaboration, DESY 85-048, 1985.
7. J. H. Kuhn and R. Rückl, Phys. Lett., *135B*, 477, 1984; P. H. Cox, S. Horvater, S. T. Jones and L. Clavelli, Phys. Rev., *D32*, 1157, 1985.
8. R. Giles et al., Phys. Rev., *D30*, 2279, 1984.
9. J. L. Cortes, X. Y. Pham and A. Tounsi, Phys. Rev., *D25*, 188, 1982.
10. R. Rückl, Weak Decays of Heavy Flavours, CERN-preprint, Oct. 1983.
11. M. K. Gaillard and L. Maiani, Cargèse Lectures, Vol. 61, eds. M. Lévy, J.-L. Basdevant, D. Speiser, J. Weyers, R. Gastmans, M. Jacob, Plenum Press, London, New York, 1980.
12. K. Kleinknecht and B. Renk, Phys. Lett., *130B*, 459, 1983; C. Klopfenstein et al., Phys. Lett., *130B*, 444, 1983; E. Fernandez et al., Phys. Rev. Lett., *51*, 1022, 1983; M. S. Alam et al., Phys. Rev. Lett., *51*, 1143, 1983; N. S. Lockyer et al., Phys. Rev. Lett., *51*, 1316, 1983; J. Green et al., Phys. Rev. Lett. *51*, 347, 1983.

13. H. Pietschmann and H. Rupertsberger, *Z. Phys.*, *27C*, 73, 1985.
14. R. Bertlmann and H. Pietschmann, *Phys. Rev.*, *D15*, 683, 1977.
15. M. Klein and H. Rupertsberger, *Phys. Rev. D*, *34*, 1389, 1986.
16. M. Klein, H. Pietschmann and H. Rupertsberger, *Phys. Lett.*, *153B*, 341, 1985.
17. S. Stone, 10th Int. Workshop on Weak Interactions, Savonlinna, Finland, 1985.
18. A. Soni, *Phys. Rev. Lett.*, *53*, 1407, 1984; G. Fogli, *Phys. Rev.*, *D28*, 1153, 1983.
19. S. E. Csorna et al., *Phys. Rev. Lett.*, *54*, 1894, 1985.
20. F. E. Close, G. J. Gounaris, J. E. Paschalis, *Phys. Lett.*, *149B*, 209, 1984; A. Ali, J. G. Körner, G. Kramer and J. Willrodt, *Z. Phys.*, *C1*, 269, 1979; B. Stech, Heidelberg preprint HD-THP-85-8.
21. F. Schöberl and H. Pietschmann, *Europhys. Lett.*, *2*, 583, 1986.

ZERO MODES AND REPARAMETRIZATION INVARIANCE IN THE ISING SPIN GLASS*

I. KONDOR**

*Department of Theoretical Physics
University of Manchester, M13 9PL, UK*

and

R. NÉMETH

*Department of Atomic Physics, Roland Eötvös University
1088 Budapest, Hungary*

A particular class of massless fluctuations around Parisi's mean field solution for an Ising spin glass is identified as inducing reparametrization of the order parameter function.

Parisi's replica symmetry breaking mean field theory (MFT) [1] has now been accepted as the correct solution of the problem of the long-range Ising spin glass [2]. It has been shown to be (marginally) stable [3] and was given a physical interpretation in terms of overlaps of pure states [4]. The physical meaning of the fluctuations around this mean field solution has, however, remained unclear so far. The presence of infinitely many zero modes in what is basically a discrete model is particularly intriguing. Though we shall not be able to clarify the physical meaning of these zero modes in the present note, we will show how one particular class of them is related to a continuous symmetry of the theory known as reparametrization or gauge invariance.

For simplicity, the discussion will be restricted to the truncated model [1] valid near the transition temperature, but we wish to point out that the relationship between the zero modes and reparametrization is quite general: as has been realized recently [5], the structure of the eigenmodes depends only on the ultrametricity [6] of phase space, hence our conclusions remain valid in MFT at any temperature in the condensed phase, moreover, they also hold beyond MFT, for the real, short range model, provided its dimension is high enough to support a Parisi-type order.

We begin the discussion by recalling a few facts about Parisi's solution and about the fluctuations around it. The truncated free energy as functional of the $n \times n$ (n is the replica number) order parameter (OP) matrix $q_{\alpha\beta}$ is [1]:

$$F = - \lim_{n \rightarrow 0} \frac{1}{2n} \left(\tau \text{Tr} q^2 + \frac{1}{3} \text{Tr} q^3 + \frac{1}{6} \sum_{\alpha \neq \beta} q_{\alpha\beta}^4 \right), \quad (1)$$

where $\tau = (T_c - T)/T_c$ is the reduced temperature.

* Dedicated to Prof. K. Nagy on his 60th birthday

** Permanent address: Institute for Theoretical Physics, Roland Eötvös University, Budapest, Hungary

If $\{p_k\}_{k=0}^{R+1}$ ($p_0 \equiv n$, $p_{R+1} \equiv 1$) are the sizes of the subdivisions of the OP matrix in the Parisi construction [1] and $\{q_k\}_{k=0}^R$ are the corresponding OP components the free energy works out to be:

$$F = -\frac{1}{2} \left\{ \tau \sum_{r=0}^R (p_r - p_{r+1}) q_r^2 + \frac{1}{3} \sum_{r=0}^R (p_r - p_{r+1}) q_r \left[\sum_{t=0}^{r-1} (p_t - p_{t+1}) q_t^2 + (p_r - 2p_{r+1}) q_r^2 + 2q_r^2 \sum_{t=r+1}^R (p_t - p_{t+1}) q_t \right] + \frac{1}{6} \sum_{r=0}^R (p_r - p_{r+1}) q_r^4 \right\} \quad (2)$$

Now the p_r 's and q_k 's are treated as variational parameters to be fixed by the stationarity conditions [1]

$$\frac{\partial F}{\partial q_k} = 0 \quad (3a)$$

and

$$\frac{\partial F}{\partial p_k} = 0 \quad (3b)$$

which yield the requisite number of equations for the $2R + 1$ unknowns. In the replica limit $n \rightarrow 0$ the p_k 's and q_k 's are analytically continued into the interval $(0, 1)$ and for $R \rightarrow \infty$ they go over into two continuous functions $p(x)$ and $q(x)$, respectively [1]. In the same limit the stationarity conditions (3a–b) go over into two integral equations which, however, turn out to be *equivalent*. This loss of information, resulting from the large R limit, is the origin of reparametrization invariance. The remaining single stationarity condition cannot determine the two functions $p(x)$ and $q(x)$ separately, it fixes the relation between them only. In the simple case of the truncated model considered here this relationship works out to be

$$q(x) = \begin{cases} p(x)/2, & x < x_1, \\ p(x_1)/2, & x > x_1, \end{cases} \quad (4)$$

where the breakpoint x_1 is determined by

$$\tau - q(x_1) + q^2(x_1) = 0. \quad (5)$$

Since in the course of the derivation leading to (4) one repeatedly divides by $\dot{p}(x)$, $\dot{q}(x)$, the monotonicity of $p(x)$ and $q(x)$ has to be assumed, but otherwise any continuous, monotonic increasing function $p(x)$ defined over $(0, 1)$ with $p(0) = 0$, $p(1) = 1$ and a $q(x)$ linked to it by (4) is a solution of the extremum problem. The solution is therefore highly degenerate, instead of a single point we are dealing with an invariant orbit in functional space.

This built-in redundancy of the Parisi scheme, called reparametrization invariance, is somewhat analogous to a local gauge transformation; its origin lies in the continuous limit $R \rightarrow \infty$, and is quite distinct from that of the other, evident symmetries (time reversal and permutation symmetry) of the model. No physical consequence of

the theory depends on the actual choice of $p(x)$, so usually the simplest gauge $p(x) = x$ is employed, though in some instances the use of a different gauge may be more convenient. (We note that Sompolinsky's [7] anomalous susceptibility $\Delta(x)$ is related to $p(x)$ through $\dot{\Delta}(x) = -p(x)\dot{q}(x)$). The interpretation of $q(x)$ in terms of the overlap probability distribution $P(q) = \frac{dx}{dq}$ [4] may lead one to believe that there is something special about the standard gauge $p(x) = x$. This is not the case: a little reflection shows that for a generic $p(x)$ the overlap distribution is $P(q) = \frac{dp}{dq}$.

Let us turn to the zero modes now. Diagonalizing the stability matrix $\frac{\partial^2 F}{\partial q_{\alpha\beta} \partial q_{\gamma\delta}}$ at the Parisi stationary point three families of eigenfunctions were found [3] (see also [5] for more details). There are zero modes in all three of them. One class of zero modes is related to the plateau in the OP function $q(x)$ and will be of no concern for us here. The other class of zero modes is peculiar to family three. In order to explain the structure of these modes we have to step back to the discrete ($R < \infty$) case. The third family eigenvectors (represented, as usual, by symmetric, $n \times n$ matrices) can then be characterized by three integers: $r=0, 1, \dots, R$ and $k_1, k_2=r+1, r+2, \dots, R+1$. Of these we are interested in the $k_1, k_2=r+1$ modes only. For finite R they are, in fact, unstable with small, negative $O(R^{-2})$ eigenvalues [8]. In the limit $R \rightarrow \infty$ these $(r, r+1, r+1)$ modes become zero modes and, as shown below, they are the ones that induce reparametrization of $q(x)$.

The eigenvector of an $(r, r+1, r+1)$ mode is shown in Fig. 1. It has non-zero elements only in one particular diagonal block of size $p_r \times p_r$ (Fig. 1a) which, in turn, is divided into sub-blocks of size $p_{r+1} \times p_{r+1}$. Of these the diagonal ones are all zeroes again, while the rest of the structure is as suggested by Fig. 1b: identical shading means identical elements. The sum of matrix elements in each row is zero:

$$\begin{aligned} p_{r+1}A + (p_r - 2p_{r+1})B &= 0, \\ 2p_{r+1}B + (p_r - 3p_{r+1})C &= 0. \end{aligned} \tag{6}$$

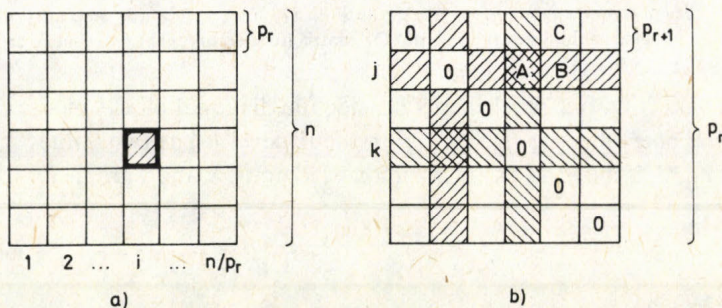


Fig. 1. The structure of $|i; j, k\rangle$, one of the $(r, r+1, r+1)$ modes. The eigenvector has non-vanishing elements inside a single diagonal block of size $p_r \times p_r$, (a), where it has the structure indicated in (b)

Conditions (6) ensure the orthogonality of this particular class of eigenvectors to all the others (i.e. to the first two families and to the third family vectors with a different r and/or k_1, k_2). The class $(r, r+1, r+1)$ is still degenerate: the distinguished $p_r \times p_r$ block can be moved to different positions in Fig. 1a and the two distinguished rows and columns of $p_{r+1} \times p_{r+1}$ blocks can be shifted to various places in Fig. 1b. The complete characterization of such an eigenvector therefore requires three numbers: $i=1, 2, \dots, n/p_r$ to specify the position of the $p_r \times p_r$ block and $j, k=1, 2, \dots, p_r/p_{r+1}$ to fix the distinguished rows inside. A vector so defined will be called $|i, j, k\rangle$ (the order of j, k does not matter).

Now for a particular choice of the series $\{p_k\}$ the solution of the discrete stationarity conditions (3a) yields the piecewise constant OP function shown in Fig. 2. Assume we want to introduce a new subdivision p' between p_r and p_{r+1} . This can be achieved by adding a properly chosen linear combination formed out of the $(r, r+1, r+1)$ modes to the original OP matrix. Before the analytic continuation all the p_r 's are divisors of the other p_k 's with $k < r$, by construction, so let us assume that the new subdivision p' is also divisible by p_{r+1} and itself a divisor of p_r . Let us form the following linear combination:

$$|r, p'\rangle = \frac{1}{2} \sum_{i=1}^{n/p_r} \sum_{l=1}^{p_r/p'} \sum_{j,k} |i, j, k\rangle, \quad (7)$$

where the summation over j, k is restricted to $(l-1)\frac{p'}{p_{r+1}} < j, k \leq l\frac{p'}{p_{r+1}}$. The result is a

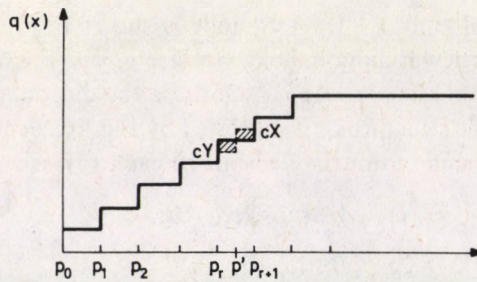


Fig. 2. Parisi's order parameter function $q(x)$ for finite R (heavy line), with a new step introduced at p' by adding a combination of zero modes to the original OP matrix (dashed line). The two shaded areas are equal

matrix with vanishing matrix elements outside the diagonal blocks of size $p_r \times p_r$ and also inside the ones of size $p_{r+1} \times p_{r+1}$; on the r^{th} level it has two values X and Y as indicated in Fig. 3. If the $|i, j, k\rangle$'s are normalized to unity X and Y are easily shown to be

$$X = -1/2 \frac{p' - p_r}{p_{r+1}} \frac{2}{(p_{r+1} - p_r)^{1/2} (3p_{r+1} - p_r)^{1/2}},$$

$$Y = -1/2 \frac{p' - p_{r+1}}{p_{r+1}} \frac{2}{(p_{r+1} - p_r)^{1/2} (3p_{r+1} - p_r)^{1/2}}. \quad (8)$$

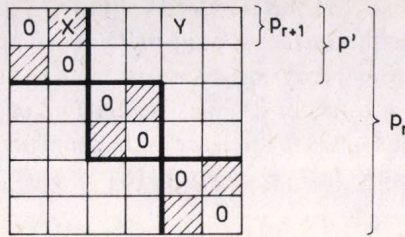


Fig. 3. The linear combination of the zero modes, Eq. (7), is composed of diagonal blocks with a structure shown here

If $|r; p'\rangle$ is added to a Parisi matrix $q_{\alpha\beta}$ with a small coefficient c it will not disturb it anywhere except on the r^{th} level, where it introduces the new subdivision p' (Fig. 2). Evidently, X and Y satisfy

$$(p' - p_{r+1})X + (p_r - p')Y = 0, \tag{9}$$

so the two shaded areas in Fig. 2 are equal and the integral of $q(x)$ (a gauge invariant quantity) is preserved.

In order to keep $q(x)$ monotonic the coefficient c must be negative (so that $cX > 0$) and since the steps between the q 's are of $O(1/R)$ while $X, Y \sim O(1/\sqrt{R})$, c must also be $O(1/\sqrt{R})$.

By an appropriate choice of c and/or p' we can, if we wish, make one of the new levels equal either to q_{r-1} or to q_{r+1} which corresponds to shifting p_r upwards or p_{r+1} downwards.

Repeated application of this kind of transformation on all levels r will carry any monotonic, piecewise constant function over into any other and in the continuous limit they will tend to different monotonic $q(x)$ functions but, for $R \rightarrow \infty$, there will be no change in the value of the free energy.

It must be clear that the point about the combination (7) is not that added to the original $q_{\alpha\beta}$ with a small coefficient c it produces a small, $O(c^2)$ variation in the free energy — any combination of any zero modes will do that. The point is that we can take infinitely many (7)-like steps in succession, and if we choose each step properly we can ensure that we remain on the invariant orbit, so that in the end we can produce a change of $O(1)$ in $q(x)$ without causing any variation (for $R \rightarrow \infty$) in F . (This is not true for the maximum of $q(x)$ whose value is fixed by the stationarity condition).

We conclude with a brief comment. In a sense, the transformation induced by adding the linear combination (7) to the stationary solution is analogous to an infinitesimal rotation in a Heisenberg model. In contrast to that, however, the meaning of reparametrization symmetry in the spin glass is totally obscure, in fact, the general belief is that it is a mere artefact of the replica formalism with no physical content at all. One expects therefore that the soft fluctuations that induce transitions between stationary points connected by this continuous symmetry also disappear from the

theory. As has been shown in [8], this is indeed so: the $(r, r + 1, r + 1)$ modes (along with some other, finite mass excitations) act as built-in Pauli–Villars regulators that subtract the Greens's functions of the theory and disappear from the physical results. We note, however, that these same subtractions, while essential for keeping the theory back from blowing up already on the Gaussian level, are at the same time responsible for the high infrared powers found in the free propagators [8].

References

1. G. Parisi, Phys. Rev. Lett., 43, 1754, 1979; J. Phys., A13, 1101; 1887; L115, 1980.
2. D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett., 35, 1972, 1975.
3. C. De Dominicis and I. Kondor, Phys. Rev., B27, 606, 1983.
4. G. Parisi, Phys. Rev. Lett., 50, 1946, 1983;
A. Houghton, S. Jain and A. P. Young, J. Phys., A16, L375, 1983.
5. C. De Dominicis and I. Kondor, in Lecture Notes in Physics 216, ed. L. Garrido, Springer Verlag, 1985, and to be published.
6. M. Mézard, G. Parisi, N. Sourlas, G. Toulouse and M. Virasoro, Phys. Rev. Lett., 52, 1156, 1984.
7. H. Sompolinsky, Phys. Rev. Lett., 47, 935, 1981.
8. C. De Dominicis and I. Kondor, J. Phys. Lett. (Paris), 45, L205, 1984.

INVERSION METHOD TO TEST RADIAL STRUCTURES IN CYLINDRICAL PLASMAS*

K. G. ANTAL, J. F. BITÓ and F. HALÁSZ

TUNGSRAM Co. Ltd.
1340 Budapest, Hungary

A new method was developed to investigate radial structures of cylindrically symmetrical plasmas, based on the evaluation of the integrated intensity distribution using modified Abel inversion. The method reduces the well known error sensitivity of measurements and the restrictions owing to analytically invertible functions are avoided. In addition to numerical tests, experiments were carried out with optically thin lines of low pressure Hg–Ar discharge.

1. Introduction

Up-to-date methods to investigate plasmas and the diagnostics based upon the same, can be classified into active and passive ones, depending on the grade of perturbation in the plasma, caused by the use of these methods [1]. In testing cylindrically symmetrical laboratory plasmas, mainly passive methods are applied, whereby variations caused in plasma parameters do not occur or are insignificant. In this range, spectroscopic methods are important [2] though these yield values characteristic in general but of the integrated intensity of radiation. Radial intensity distributions, i.e. the radial structure of each considered spectral line can be determined by measuring the integrated intensities of probe radiations or parallel sections of the radiation emitted by the plasma, by applying Abel inversion.

The present investigations are aimed on the one hand at deciding experimentally what a priori hypotheses concerning the microstructure are admissible in a simplified theoretical description of the positive column of discharge plasmas produced in laboratories, and, on the other hand, at the experimental verification of radial dependences obtained in complex models. In this paper restricted to the consideration of the positive column of low-pressure Hg–Ar discharges, a new special inversion will be used for radiating transitions of Hg $6^1D_2-6^1F_1$ (579 nm) and $7^3S_1-6^3P_1$ (436 nm), for which the plasma is optically thin. As published models that have so far been developed for the description of laboratory plasmas are in some cases inconsistent and using them significant features are neglected, a concise survey may be expedient.

* Dedicated to Prof. K. Nagy on his 60th birthday

2. Column models of low-pressure Hg-Ar plasmas

One of the most extensively investigated laboratory plasmas subject to both experimental [3–9] and theoretical [10–16] research is the positive column of the low pressure Hg-Ar discharge used mainly in light sources. The basic aim of the early theoretical models was to calculate the energy balance of the positive column, which can be easily accomplished by making simplified assumptions for the radial structure. Complex models giving “full” numerical descriptions without fitting parameters were set up [11, 16]. As these calculations are time-consuming, models using simplified assumptions to describe certain selected processes continue to be published [14, 15]. The analysis of microprocesses by reduced analytical formulas is again preferred to the solution of some 6 to 8 joint differential equations. However, the radial distribution of the excited Hg levels is a significant factor in the “big” models as well as in the simplified analytical ones. It is expedient to point out that according to the “most sophisticated” model [16] taking into account the majority of the physical processes, using up-to-date numerical methods, the radial distribution function of the 6^3P_2 level is saddle-shaped. At the same time, according to [15] this level may be expected to play an essential role in the stepwise ionisation of mercury. The a priori statement of the mentioned model is that the concentration of the given level is constant along the radius of the discharge and it slightly depends on the electron concentration. Only this simplification gives the possibility to obtain analytical solution.

Similar assumptions without satisfactory experimental basis frequently occur in the literature.

As an example for the discussion mentioned above, let us use the data originating from the numerical experiments based on the results of [16]. To ease the survey of the problem we also show the simplified Hg term schema (Fig. 1). The maximum set of levels that has ever been taken into account in the column-models of the available literature consists of the $6^1S_0(g)$, $6^3P_0(q)$, $6^3P_1(r)$, $6^3P_2(s)$, $6^1P_1(p)$ and the ionised (i) states.

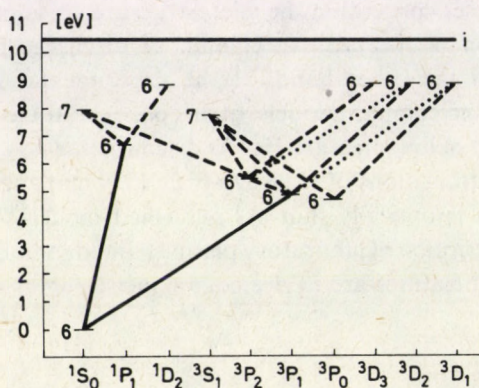


Fig. 1. Term schemes of lower levels of Hg, showing main optical transitions

In [16] the input parameters of the model are:

- the radius of the discharge tube 1.3 cm,
- the wall-temperature of the discharge tube $T=313$ K,
- the pressure of the argon buffer gas 2 mbar,
- the field strength of the column $E=0.78$ V/cm.

Unfortunately, the radial distributions calculated from the model cannot be directly measured by passive methods computed from the model (Fig. 2), because the

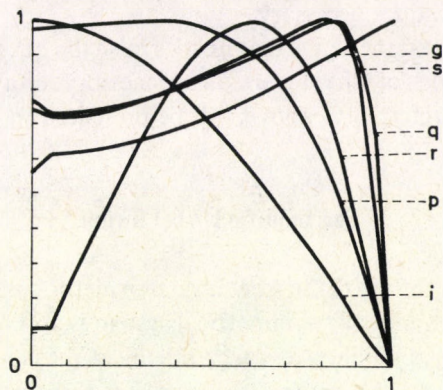


Fig. 2. Standardized radial structures of levels g , p , q , r , s , i , as calculated according to model [16]

Table of values

r/R	g	p	q	r	s	i
0	0.56	0.98	0.75	0.11	0.77	1
0.05	0.61	0.98	0.72	0.11	0.73	1
0.1	0.62	0.98	0.73	0.23	0.74	0.99
0.15	0.62	0.99	0.73	0.34	0.74	0.97
0.2	0.63	0.99	0.74	0.45	0.75	0.95
0.25	0.64	1	0.75	0.55	0.76	0.92
0.3	0.65	1	0.77	0.65	0.77	0.89
0.35	0.67	1	0.78	0.75	0.79	0.85
0.4	0.69	1	0.8	0.83	0.81	0.8
0.45	0.71	0.99	0.82	0.9	0.83	0.75
0.5	0.73	0.97	0.85	0.95	0.85	0.7
0.55	0.75	0.95	0.87	0.98	0.88	0.63
0.6	0.77	0.91	0.89	1	0.9	0.57
0.65	0.8	0.86	0.92	0.99	0.93	0.5
0.7	0.83	0.79	0.95	0.94	0.96	0.43
0.75	0.86	0.7	0.97	0.88	0.98	0.36
0.8	0.88	0.6	0.99	0.76	1	0.29
0.85	0.92	0.48	1	0.62	1	0.21
0.9	0.95	0.33	0.91	0.4	0.86	0.13
0.95	0.98	0.15	0.68	0.19	0.46	0.05
1	1	0	0	0	0	0

plasma cannot be regarded optically thin with respect to the resonance lines originating from the said levels and, at the same time, there are no radiation transitions between the metastable and the resonance states.

The occupation of higher levels has never been treated by any theoretical model up to now because of the lack of fairly reliable transition probability data. However, there is a possibility to analyse the radial structure by the use of a relatively simple technique. The occupation of the levels given in the model (Fig. 2) can be measured by the classical "passive" method, though the small differences occurring in that case make the use of lock-in techniques necessary.

Both the passive and the active diagnostics raise the question of the appropriate inversion method. One of the inversion processes suitably converted to the experimental methods referred to above will be presented in the next Section.

3. The modified Abel inversion

As is well known, the use of the Abel inversion method is expedient whenever the radial structure of a cylindrically symmetrical system is to be determined and at the same time there is no possibility to carry out local measurements. In the optical investigation of plasmas, the general formula for the measurable $I(x)$ intensity-coordinate function has the following structure:

$$I(x) = e^{-2 \int_x^R \frac{a(r)r dr}{\sqrt{r^2 - x^2}}} \left\{ I_0 + 2 \int_x^R \varepsilon(r) e^{\int_x^R \frac{a(r')r' dr'}{\sqrt{r'^2 - x^2}}} \frac{r}{\sqrt{r^2 - x^2}} dr \right\} \quad (1)$$

with the notations:

I_0 the intensity of the light source to be tested;

x the distance measured from the plasma axis;

$\varepsilon(r)$ the intensity radiated by the plasma in the direction of observation;

$a(r)$ the absorption of the plasma.

Investigations will be made for two extreme cases:

a) $a(r) = 0$; $I_0 = 0$;

$$I(x) = 2 \int_x^R \varepsilon(r) \frac{r}{\sqrt{r^2 - x^2}} dr; \quad (2)$$

b) $\varepsilon(r) = 0$,

$$I(x) = I_0 e^{-2 \int_x^R \frac{a(r)r}{\sqrt{r^2 - x^2}} dr}. \quad (3)$$

The first case pertains to the intensity measurement of the light emitted by the plasma in the wavelength range of negligible absorption. The second one is the

absorption measurement using an external light source, when an appropriate filtering of the radiation emitted by the plasma need to be provided (e.g. negligible own radiation compared to I_0 or application of the lock-in technics, etc.). The $\varepsilon(r)$ and $a(r)$ functions are to be determined and they will be proportional to the occupation of the upper and lower level of the transition, respectively.

It can be stated that by examining the logarithm of the transmission in case b) the same relation as in case a) can be established between the measured values ($I(x)$ or $\ln(I(x)/I_0)$) and the required quantities ($\varepsilon(r)$ or $a(r)$), so it is sufficient to consider formula (2) only. Formally, by applying the Abel inversion [17] relation (2) can be brought to the following form:

$$\varepsilon(r) = -\frac{1}{\pi} \int_r^R \frac{I'(x)}{\sqrt{x^2 - r^2}} dx. \quad (4)$$

However, in practical cases the $I'(x)$ function is generally not available and in fact, the $I(x)$ function is known in some discrete points only. (Of course, in most cases it is quite enough to determine $\varepsilon(r)$ in discrete points.)

In this way, the alternatives are as follows:

- i. numerical derivation to determine $I'(x)$ and, subsequently, a numerical approximation of integral (4) using e.g. the Simpson formula, etc.;
- ii. fitting an analytically Abel invertable (4) function series to the measured intensity $I(x)$;
- iii. based on theoretical considerations seeking $\varepsilon(r)$ in a given form and choosing the appropriate parameters to reach the best fitting of the $I(x)$ function computed from (2) to the measured values;
- iv. considering (2) in discrete points only, we can obtain a relation of the form

$$I_i = \sum_{j=1}^N A_{ij} \varepsilon_j, \quad (5)$$

where A_{ij} depends on the number of the discrete points and on the particular kind of numerical integration applied. This is a system of linear equations which can be solved either by the inversion of matrix A (if the measurements are made at identical points) or by other methods and in this way the values ε_j can be determined from the measured intensities I_i .

In alternatives i and iv the significant error due to the numerical derivation is even enhanced by the expression of $1/\sqrt{0}$ type in the integral. In case iv this is not evident but the term is comprised in matrix A and in its inverse, too.

In alternatives ii and iii this source of error can be avoided but at the cost of an explicit or implicit limitation of $\varepsilon(r)$ in conformity with the functions used in the fitting. This fact means a kind of a priori limitation of the form of $\varepsilon(r)$ which limitation of course also influences the results.

The method avoiding the difficulties inherent in the techniques described is based essentially on alternative iv. In this way the assumptions inevitable in cases ii and iii can be bypassed and, as a result, the inversion method becomes applicable to problems with unknown radial structures. One of the most powerful techniques for decreasing numerical errors is the method of smallest quadratic standard deviation according to which the radial distribution can be determined but in points less than the ones originally tested.

As the measured intensity values can vary by orders of magnitude, it is expedient to look for the minimum relative standard deviation. Hence the basic relation is:

$$\sum_{i=1}^N \left(\sum_{j=1}^M \frac{A_{ij}\varepsilon_j}{I_i} - 1 \right)^2 = \min, \quad (6)$$

where N is the number of points in which measurements were carried out; M is the number of the requested ε values; I_i is the intensity measured in the i th point; A_{ij} are the matrix elements depending on the integral approximation used, but not on ε_j .

Derivating relation (6) according to ε_l we can deduce the following system of linear equations:

$$\sum_{i=1}^N \frac{A_{il}}{I_i} \left(\frac{\sum_{j=1}^M A_{ij}\varepsilon_j}{I_i} - 1 \right) = 0, \quad l = 1 \dots M. \quad (7)$$

It is evident that if $x_i > r_j$, then in first approximation $A_{ij} = 0$, but depending on the integral approximation this is not necessarily so. A non-zero term may also occur when $x_i \sim r_j$. Of course the $N = M$ case corresponds to the iv-th one discussed above. As it is shown in Fig. 3 a decrease of M causes a significant reduction of the expected error of the computed radial distribution.

Let the analytically invertible function corresponding to the radial distribution have the form

$$\varepsilon(r) = 1 - (r/R)^2 \quad (8)$$

in the considered case, then

$$I(x) = (1 - (x/R)^2)^{3/2}. \quad (9)$$

For numerical investigations the measured $I(x)$ input intensity data were supposed to have an accuracy of at least 1 per cent while the errors pertaining to the discrete points were randomly generated. The number of the tested points was supposed to be $N = 27$. In Fig. 3 the absolute values of the relative deviations for the $M = 27, 21, 15$ values are shown. As can be readily seen, the 1 per cent error of the measurement can cause even 15 per cent in the case of classical Abel inversion while the present method yields less than 3 per cent.

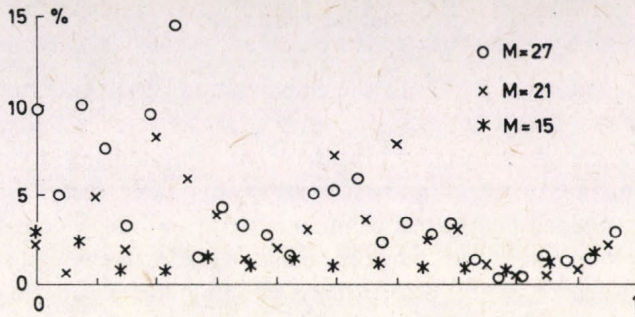


Fig. 3. Error analysis using analytically invertible test function (horizontal axis — radial positions, vertical axis — absolute values of relative errors)

Table of values

○		x		*	
v. t.	%	v. t.	%	v. t.	%
0	9	0	2	0	2
0.04	5	0.05	0	0.07	2
0.08	9	0.1	5	0.14	0
0.12	7	0.15	2	0.21	0
0.15	3	0.2	8	0.29	1
0.19	9	0.25	5	0.36	0
0.23	14	0.3	4	0.43	1
0.27	1	0.35	1	0.5	1
0.31	4	0.4	2	0.57	0
0.35	3	0.45	3	0.64	0
0.38	2	0.5	7	0.71	2
0.42	2	0.55	3	0.79	2
0.46	5	0.6	8	0.86	1
0.5	5	0.65	2	0.93	3
0.54	5	0.7	3		
0.58	3	0.75	1		
0.62	4	0.8	0		
0.65	3	0.85	0		
0.69	2	0.9	0		
0.73	1	0.95	2		
0.77	0				
0.81	1				
0.85	1				
0.88	3				
0.92	5				
0.96	1				

4. Radial intensity test

The experimental setup of both the emission and absorption measurement is presented in Fig. 4.

In tests of this kind utmost care must be given to minimizing any possible error source. Absorption measurements make it necessary to guarantee fairly good stability in time for the I_0 intensity of the test beam, otherwise an independent simultaneous measurement for "normalisation" purposes must be made, to eliminate the effects of varying optical features of the glass tube containing the plasma (e.g. curvature, transparency, etc.). This can be done by comparing the test results and those obtained with the same discharge tube tested for transmittivity when the discharge is switched off. It is also important to achieve fairly good reproducibility of positioning. (It is expedient to use step motors for control.) Precise slot-setting to choose the correct plasma section is less critical, though in order to obtain good resolution a minimum slot width is necessary, of course within the range determined by the detector-amplifier system.

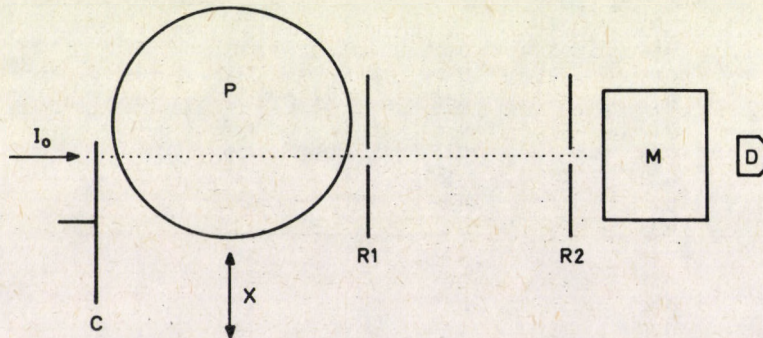


Fig. 4. Layout scheme for testing radial emission and absorption. I — probing radiation; C — beam chopper; P — plasma; X — moving mechanism; R1, R2 — slots; M — monochromator; D — detector

Requirements concerning the monochromators depend on the concrete transition to be investigated. In emission tests the separation from transitions starting from other upper levels should be accomplished by good wavelength resolution in cases when the intensity of these transition lines is not orders of magnitude lower than the intensity of the line to be measured. (E.g. the $6^3D_1-6^3P_1$ and $6^1D_2-6^3P_1$ transitions require a resolution better than 0.1 nm). In absorption tests this difficulty is not significant because the lower levels for close lines are in most cases common and in each test the sum of the absorptions associated with the lower level is measured.

It is the beamchopper-detector-amplifier system the common stability of which must be restricted. Normally, it is expedient to require 1 per cent minimum value for the stability and reproducibility of the measurement. A narrow band amplifier tuned to the

chopper's frequency can be applied in such measurements though better results can be achieved by the use of phase-synchronised (lock-in) amplifiers even if no time resolution is necessary.

Let us use the emission measurement data given in Fig. 5 to illustrate the method described in Section 3. The tests were carried out on the positive column of a Hg-Ar discharge of 26 mm diameter ($T=313$ K, $p=2$ mbar, resolution is 0.5 mm, the number of test points is 52).

The inversion was executed by a computer program specially developed for data processing. The integrals involved were approximated by the use of the Simpson formula. The calculated radial distributions for two selected lines are shown in Fig. 6.

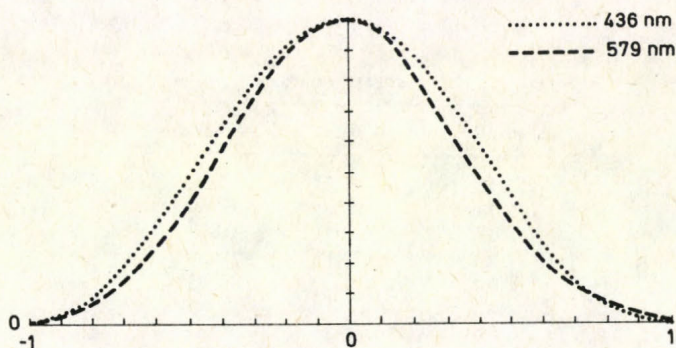


Fig. 5. Measured intensity distributions of transitions $6^1D_2 - 6^1P_1$ (579 nm) and $7^3S_1 - 6^3P_1$ (436 nm)

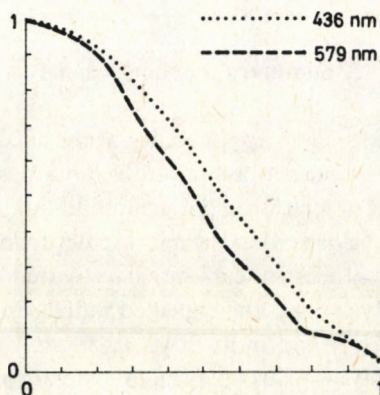


Fig. 6. Calculated radial distribution of intensity of transitions $6^1D_2 - 6^1P_1$ (579 nm) and $7^3S_1 - 6^3P_1$ (436 nm)

Values measured on the 436 nm line

	0.002	0.007	0.016	0.034	0.053	0.081	0.122	0.168	0.211	0.26	0.314	0.374	
	0.442	0.495	0.566	0.629	0.691	0.734	0.801	0.848	0.893	0.922	0.944	0.973	0.993
1	0.998	0.983	0.965	0.937	0.898	0.853	0.812	0.752	0.694	0.631	0.572	0.51	0.435
	0.388	0.316	0.26	0.202	0.164	0.119	0.088	0.062	0.038	0.022	0.011	0.007	0.007

Values measured on the 579 nm line

	0.007	0.011	0.018	0.039	0.052	0.075	0.091	0.118	0.164	0.186	0.255	0.305		
	0.373	0.42	0.477	0.559	0.623	0.689	0.752	0.798	0.873	0.895	0.945	0.97	0.991	1
	0.986	0.98	0.959	0.909	0.85	0.786	0.72	0.652	0.598	0.539	0.473	0.389	0.343	
	0.284	0.236	0.198	0.143	0.116	0.095	0.089	0.064	0.055	0.041	0.023	0.016	0.011	

Table of values

r/R	436 nm	579 nm
0	1	1
0.1	0.976	0.969
0.2	0.905	0.877
0.3	0.791	0.681
0.4	0.66	0.545
0.5	0.5	0.363
0.6	0.353	0.247
0.7	0.201	0.112
0.8	0.105	0.087
0.9	0.027	0.026
1	0.003	0.015

5. Summary and conclusions

The limitations of the area of applicability of the theoretical models described above together with their inherent assumptions and neglects require further refinement both in the fields of measurement techniques and theory. In the absence of some improvement in these fields neither the microstructure of the plasma, nor the finer details of the interaction processes can be investigated — the neglects and assumptions applied usually conceal the expected radial inhomogeneities concerning macroparameters both directly and indirectly.

In this paper a passive spectroscopical diagnostic method for investigating cylindrical plasmas is presented. Both the theoretical and experimental analyses mainly refer to the positive column of the well-known Hg-Ar laboratory plasmas investigated by several authors and widely applied for lighting.

Surveying the alternatives to verify complex and simplified theoretical models, we can state that optical methods, due to their good reproducibility, passive character and that they do not impose special restrictions on the system measured, can be regarded expedient for such purposes. The optical methods applicable in this field make the use of Abel inversion necessary to analyze experimental results.

The essence of the method suggested for Abel inversion is the application of the minimum root-mean-square treatment avoiding numerical derivations by which the numerical errors occurring in computation can be reduced seriously. The applicability of this procedure is demonstrated in Fig. 3, showing an example of the error analysis by an analytically invertible function. It is striking that the 1 per cent random error considered is amplified to some 15 per cent if the classical Abel inversion is used and this error is reduced to 3 per cent, with satisfactory radial resolution.

Numerical experiments evidence that 1 per cent testing accuracy is required to determine the radial structure reliably. To give an idea of the severity of this exigence the factors affecting the test accuracy were surveyed.

Experience hitherto gained with the tests and also theoretical considerations suggest that it is expedient to integrate the micromoving, data collecting and control systems in one microcomputer system in order to increase the number of test points.

The $I'(x)$ function can be directly measured by providing for the high uniformity of the motion and completing the amplifier with a differentiating circuit of appropriate time constant. In so avoiding any numerical derivation, the Abel inversion can be executed by simple integrations on the basis of (4).

Though the primary aim of this paper is to describe a new testing method, some remarks concerning the test results presented in Figs 5 and 6 may be useful to illustrate the applicability.

In Fig. 5 an intensity distribution measured in radiating transitions that start from two distinct upper levels, and in Fig. 6 the radial distribution obtained after the evaluation of the test are shown. Albeit the calculated radial structure shown in Fig. 2 does not appear in any of these transitions, the difference in distributions in Fig. 6 is nevertheless striking. This is an evidence that the obvious assumption in LTE models of the identity, and in some cases, the constancy of radial distributions of the excited levels is a marked oversimplification.

References

1. J. Bitó, I. Bolla and K. Antal, *Műszaki Tudomány*, 48, 303, 1974 (in Hungarian).
2. J. Bitó, *Magyar Fizikai Folyóirat*, XVI. No. 6, 541, 1968 (in Hungarian).
3. C. Kenty, *J. Appl. Phys.*, 21, 1309, 1950.
4. B. T. Barnes, *J. Appl. Phys.*, 31, 852, 1960.
5. W. Verweij, *Philips Res. Rep. Suppl.*, 2, 1, 1961.
6. M. Koedam and A. A. Kruithof, *Physica*, 28, 80, 1962.
7. M. Koedam, A. A. Kruithof and J. Riemans, *Physica*, 29, 565, 1963.

8. K. G. Antal and E. Gáti, *TUNGSRAM Technische Mitteilungen*, 29, 1226, 1976.
9. K. G. Antal and E. Gáti, *TUNGSRAM Technische Mitteilungen*, 31, 1303, 1976.
10. J. F. Waymouth and F. Bitter, *J. Appl. Phys.*, 27, 122, 1956.
11. M. A. Cayless, *Br. J. Appl. Phys.*, 14, 863, 1963.
12. T. G. Verbeek and P. C. Drop, *J. Phys.*, D7, 1677, 1974.
13. J. W. Denneman, J. J. de Groot, A. G. Jack and F. A. S. Ligthart, *J. Ill. Eng. Soc.*, 10, 2, 1980.
14. C. M. Ferreira and J. Loureiro, *J. Phys. D. Appl. Phys.*, 17, 1175, 1984.
15. W. L. Lama, C. F. Gallo, T. J. Hammond and P. J. Walsh, *Appl. Opt.*, 21, 180, 1982.
16. G. Molnárka, *Teubner-Texte zur Mathematik 61*, 181, 1984.
17. W. Lochte-Holtgreven, *Plasma Diagnostics*, North Holland Publishing Company, Amsterdam, 1968.

THE POSSIBILITY OF AVOIDING THE AXION IN THE PECCEI–QUINN MECHANISM*

T. NAGY

*Institute for Theoretical Physics, Roland Eötvös University
1088 Budapest, Hungary*

The axion connected with the Peccei–Quinn solution of the strong CP puzzle can be eliminated by the 't Hooft mechanism. In this case massless quarks must exist at the tree level. It is argued that the violation of the 't Hooft symmetry by instanton effects can make these quarks massive and in some models this effect might be large enough to generate the masses of the lightest quarks.

There have been several attempts to explain in a natural way the smallness of the strong CP violation, implied by the θ -term of the QCD Lagrangian. The most elegant mechanism was proposed by Peccei and Quinn [1]. The idea is that if the total Lagrangian has a global chiral U(1) symmetry, rotating the θ angle, and if at least one of the quarks gets its entire mass from spontaneous symmetry breaking, the effective θ parameter

$$\bar{\theta} = \theta - \arg \det m_L$$

obtained when one brings the quark mass matrix (with the left-handed part m_L) to diagonal form will automatically be equal to zero, as a consequence of the appropriate θ -dependence of the scalar expectation value. However, this mechanism is plagued by the so-called axion problem [2]: the spontaneous breakdown of the U(1) symmetry gives rise to an extra zero mass Goldstone boson at the tree level. Since the symmetry is explicitly violated by instanton effects, this particle, the axion, is actually a pseudo-Goldstone boson with a small mass. No such particle has been observed and new ideas have been suggested to save the Peccei–Quinn solution of the strong CP violation.

Given this state of affairs, it is quite straightforward to invoke the 't Hooft mechanism [3] as a possible solution. The mechanism works as follows: let X be a gauge generator and Y a generator of a global symmetry and let both of them be spontaneously broken while a combination $Z = aX + bY$ remains unbroken. Then there is only one corresponding Goldstone boson and this can be transformed away as in the usual Higgs mechanism. If we take here Y as the generator of the Peccei–Quinn U(1), no extra Goldstone boson is associated with its spontaneous breaking, there is no axion.

* Dedicated to Prof. K. Nagy on his 60th birthday

Such a solution of the axion problem was discussed within a concrete model by Deshpande and Mannheim [4]. However, also this kind of solution has its own difficulties. It is easy to show [5] that in such models at least one massless quark must exist: the anomaly $\text{Tr}_L Q \{F_\alpha, F_\beta\} - \text{Tr}_R Q \{F_\alpha, F_\beta\}$ (F_α : $\text{SU}(3)_c$ generator) must be zero for a $Q = \bar{X}$ gauge generator, by renormalizability, while it is different from zero for $Q = Y$, otherwise $U(1)_Y$ would not rotate θ ; thus Z is anomalous too, and only massless coloured fermions can contribute to its anomaly.

Now the point I want to make is this: as it was mentioned above, $U(1)_Y$ is not a real symmetry, it is violated by instanton effects. The same is true for Z . The effects, making the axion a pseudo-Goldstone boson give masses to the quarks, massless at the tree level. There is a possibility that the lightest quarks get their masses by this mechanism.

To illustrate all these ideas, let us consider a simple model which is essentially a standard model with two $U(1)$ factors and one fermion generation, showing a global chiral $U(1)_{PQ}$ symmetry. That is, we have

$$G_{\text{local}} = \text{SU}(3)_c^{F_\alpha} \times \text{SU}(2)_L^{T_\alpha} \times U(1)^{Y_1} \times U(1)^{Y_2}$$

with the corresponding gauge bosons $G_{\alpha\mu}$, $A_{\alpha\mu}$, $B_{1\mu}$ and $B_{2\mu}$ (the superscripts denote the group generators), and

$$G_{\text{global}} = U(1)_{PQ}^Y.$$

The quark multiplet is:

$$q_L = \begin{pmatrix} P \\ N \end{pmatrix} \begin{matrix} Y_r = \varepsilon_r, Y = 1 \\ L \end{matrix}$$

$$q_R: P_R^{Y_r = \varepsilon_r, Y = -1}, \quad N_R^{Y_r = \varepsilon_r, Y = -1}.$$

We introduce two $\text{SU}(2)_L$ Higgs doublets and one singlet:

$$\Phi_i = (\Phi_i^a)_{a,i=1,2}^{Y_r = \delta_i^a, Y = \delta^i}, \quad \Phi^{Y_r = \delta_r, Y = \delta}.$$

The singlet Φ does not play any role in the mechanism to be discussed, it is needed only to break the symmetry down to $\text{SU}(3)_c \times U(1)_Q$ and get an imitation of the "real" situation.

The Lagrangian of the model can be written in the form:

$$\begin{aligned} \mathcal{L} = & -\frac{1}{4} F_{\alpha\mu\nu} F_\alpha^{\mu\nu} + i\bar{q}_L \gamma^\mu \nabla_\mu q_L + i\bar{q}_R \gamma^\mu \nabla_\mu q_R + \mathcal{L}_Y^q + \\ & + \mathcal{L}^l + (\nabla_\mu \Phi_i^a)^* (\nabla^\mu \Phi_i^a) + (\nabla_\mu \Phi)^* \nabla^\mu \Phi - P(\Phi), \end{aligned} \quad (1)$$

where the first term contains the self-interactions of all the gauge bosons, ∇_r denotes the covariant derivative and $P(\Phi)$ is the scalar potential. The Yukawa interaction of the quarks has the general form:

$$\begin{aligned}
 -\mathcal{L}_Y^q = & \bar{P}_R [P_L(\Gamma_k^P \Phi_k^{1*} + \tilde{\Gamma}_k^P \Phi_k^2) + N_L(\Gamma_k^P \Phi_k^{2*} - \tilde{\Gamma}_k^P \Phi_k^1)] + \\
 & + \bar{N}_R [P_L(\Gamma_k^N \Phi_k^{1*} + \tilde{\Gamma}_k^N \Phi_k^2) + N_L(\Gamma_k^N \Phi_k^{2*} - \tilde{\Gamma}_k^N \Phi_k^1)] + \text{HC}. \quad (2)
 \end{aligned}$$

The leptonic part \mathcal{L}^l will not be specified.

The U(1) symmetries of \mathcal{L}_Y^q give:

$$\begin{aligned}
 \Gamma_k^P \neq 0: & \quad \varepsilon_r - \varepsilon_r^P - \delta_r^k = 0, \quad 2 - \delta^k = 0, \\
 \Gamma_k^N \neq 0: & \quad \varepsilon_r - \varepsilon_r^N - \delta_r^k = 0, \quad 2 - \delta^k = 0, \\
 \tilde{\Gamma}_k^P \neq 0: & \quad \varepsilon_r - \varepsilon_r^P + \delta_r^k = 0, \quad 2 + \delta^k = 0, \\
 \tilde{\Gamma}_k^N \neq 0: & \quad \varepsilon_r - \varepsilon_r^N + \delta_r^k = 0, \quad 2 + \delta^k = 0.
 \end{aligned} \quad (3)$$

The conditions for the absence of Adler-Bell-Jackiw anomalies yield further restrictions on the U(1) quantum numbers. The condition connected with the $\text{Tr } Y_r \{F_\alpha, F_\beta\}$ combination reads

$$2\varepsilon_r = \varepsilon_r^P + \varepsilon_r^N.$$

All the other conditions involve the U(1) quantum numbers of the leptons and can be satisfied if we treat the lepton multiplet in complete symmetry with the quarks.

Our aim is to generate a situation where the scalar vacuum expectation values $\tilde{\Phi}_i, \tilde{\Phi}$ yield two conserved generators, a flavor generator Q (electric charge) and a combination

$$Z = X + Y,$$

where X is a flavor generator. These v.e.v.'s give mass to one of the quarks. Then the strong interactions generate a Z -forbidden contribution δP to the scalar potential, changing the v.e.v.'s by $\delta\tilde{\Phi}_i, \delta\tilde{\Phi}$, and the other quark becomes massive too. This can be achieved with the following scheme:

$$Q = T_3 + Y_1, \quad \varepsilon_1 = \frac{1}{6}, \quad \varepsilon_1^P = \frac{2}{3}, \quad \varepsilon_1^N = -\frac{1}{3}, \quad \delta_1^k = \frac{1}{2}, \quad \delta_1 = 0$$

(i.e. Y_1 is the "usual" U(1) generator). Then only Γ_k^N and $\tilde{\Gamma}_k^P$ are allowed; to have both kinds of couplings we choose

$$\delta^1 = 2, \quad \delta^2 = -2, \quad \delta_2^1 = \varepsilon_2 - \varepsilon_2^N = \varepsilon_2^P - \varepsilon_2 = \delta_2^2,$$

leaving Γ_1^N and $\tilde{\Gamma}_2^P$ different from zero. The combination Z , orthogonal to Q , can be written as

$$Z = \alpha(T_3 - Y_1) + Y_2 + Y.$$

There are two Higgs components which are neutral with respect to both Q and Z : Φ and Φ_1^2 , or Φ_2^2 . Making the choice Φ , Φ_1^2 (i.e. $\delta_2 + \delta = 0$, $2 - \alpha + \delta_2^{\frac{1}{2}} = 0$), the v.e.v.'s are:

$$\tilde{\Phi}_1 = \begin{pmatrix} 0 \\ \Lambda \end{pmatrix}, \quad \tilde{\Phi}_2 = 0, \quad \tilde{\Phi} = \Sigma.$$

The most general scalar potential which exhibits all the symmetries is of the following form:

$$\begin{aligned} P(\Phi) = & a_1 \Phi_1^* \Phi_1 + a_2 \Phi_2^* \Phi_2 + a_3 \Phi^* \Phi + b_1 (\Phi_1^* \Phi_1)^2 + b_2 (\Phi_2^* \Phi_2)^2 + \\ & + b_3 (\Phi^* \Phi)^2 + c_1 (\Phi_1^* \Phi_1) (\Phi_2^* \Phi_2) + c_2 (\Phi_1^* \Phi_1) (\Phi^* \Phi) + \\ & + c_3 (\Phi_2^* \Phi_2) (\Phi^* \Phi) + c_4 (\Phi_1^* \Phi_2) (\Phi_2^* \Phi_1), \end{aligned} \quad (4)$$

giving the equations

$$\begin{aligned} a_1 + 2b_1 |\Lambda|^2 + c_2 |\Sigma|^2 &= 0, \\ a_3 + 2b_3 |\Sigma|^2 + c_2 |\Lambda|^2 &= 0 \end{aligned} \quad (5)$$

for the v.e.v.'s. If we write

$$\begin{aligned} \Phi_1^2 &= \frac{e^{i\alpha}}{\sqrt{2}} (\sqrt{2} |\Lambda| + \varphi_1 + i\varphi_2), \\ \Phi &= \frac{e^{i\beta}}{\sqrt{2}} (\sqrt{2} |\Sigma| + \psi_1 + i\psi_2), \end{aligned}$$

the quadratic part of $P(\Phi)$ takes the form

$$P(\Phi)^{\text{quadr}} = \mu_1^2 |\Phi_1^1|^2 + \mu_2^2 |\Phi_2^2|^2 + \mu_{11} \varphi_1^2 + 2\mu_{12} \varphi_1 \psi_1 + \mu_{22} \psi_1^2 \quad (6)$$

with

$$\begin{aligned} \mu_1^2 &= a_2 + c_1 |\Lambda|^2 + c_3 |\Sigma|^2, \\ \mu_2^2 &= \mu_1^2 + c_3 |\Lambda|^2, \\ \mu_{11} &= 2b_1 |\Lambda|^2, \quad \mu_{12} = c_2 |\Lambda| |\Sigma|, \quad \mu_{22} = 2b_3 |\Sigma|^2. \end{aligned}$$

This shows that we have four Goldstone bosons: φ_2 , ψ_2 and the two real components of Φ_1^1 . These bosons correspond to the broken gauge symmetries and can be transformed away — there is no axion. However, as it was anticipated there is a massless quark:

$$m_P = 0, \quad m_N = \Gamma_1^N \Lambda^*. \quad (7)$$

As we said earlier, the instanton effects give a Z violating correction to the scalar potential. By calculating this contribution in the simplest approximation, described in [1], one gets

$$-\delta P = A \Phi_1^* \Phi_2 e^{-i\theta} + A^* \Phi_1 \Phi_2^* e^{i\theta} \quad (8)$$

with

$$A = a^2 \tilde{\Gamma}_2^P \Gamma_1^N,$$

where $a^2 > 0$ is the v.e.v. of a quartic expression of the quark fields. It is clear that δP is forbidden by $U(1)_Z$.

The modification $\delta\tilde{\Phi}_\alpha$ of the scalar v.e.v.'s which we assume to be small can be determined from the equations

$$\left. \frac{\partial^2 P}{\partial \Phi_\alpha \partial \Phi_\beta} \right|_{\tilde{\Phi}} \delta\tilde{\Phi}_\beta + \left. \frac{\partial \delta P}{\partial \Phi_\alpha} \right|_{\tilde{\Phi}} = 0.$$

One obtains

$$\delta\tilde{\Phi}_1^2 = \delta\tilde{\Phi} = 0, \quad \delta\tilde{\Phi}_2^2 = \frac{A^*}{\mu_2^2} \Lambda e^{i\theta} \tag{9}$$

which gives for the P -mass:

$$m_P = \frac{a^2 |\tilde{\Gamma}_2^P|^2}{\mu_2^2} \Gamma_1^{N*} \Lambda e^{i\theta} = \frac{a^2 |\tilde{\Gamma}_2^P|^2}{\mu_2^2} m_N^* e^{i\theta}. \tag{10}$$

Note that m_P and m_N , as given here, are the left-handed parts of the masses and that $\arg \det m_L = \arg m_P m_N = \theta$.

This simple example shows how quark masses are generated at different levels in axion free models. Of course, the question is whether the higher order masses are large enough to play a role in a phenomenological analysis. In the above expression we expect a to be of the order of a typical strong scale, μ_2 is a Higgs mass, thus the ratio m_P/m_N is probably very small unless $\tilde{\Gamma}_2^P$ is allowed to assume a sufficiently large value. Leaving aside the question of the reliability of the approximation in which δP was determined, it is clear that in a realistic situation we want the heavy quarks to be massive at the tree level. Thus a realistic analysis must involve several quark generations. If there are models which can accommodate quantitatively acceptable masses, their form is probably severely restricted. Such extended models will be investigated elsewhere.

References

1. R. D. Peccei and H. R. Quinn, Phys. Rev. Letters, 38, 1440, 1977; Phys. Rev., D16, 1791, 1977.
2. S. Weinberg, Phys. Rev. Letters, 40, 223, 1978; F. Wilczek, Phys. Rev. Letters, 40, 279, 1978.
3. G. 't Hooft, Nucl. Phys., B35, 167, 1971.
4. N. G. Deshpande and P. D. Mannheim, Phys. Letters, 94B, 355, 1980.
5. S. M. Barr and X. C. Gao, Phys. Rev., D25, 3423, 1982.

GLOBAL-SCALE CHANGES OF THE ENVIRONMENT: OBSERVATIONS FROM SPACE*

K. YA. KONDRATYEV

*Institute for Lake Research of the USSR
Academy of Sciences, Leningrad, USSR*

In connection with the development of the International Geosphere-Biosphere Programme, key aspects of the environment have been discussed that must specify the major objectives of the IGBP as a programme on studies of the interactive evolution of the geosphere and biosphere in conditions of an increased anthropogenic impact. Taking into account extensive international programmes (it concerns, in particular, the World Climate Programme as well as the Man and Biosphere Programme), the problem of global biospherical cycles of carbon, sulfur, nitrogen, phosphorus, and some other elements must be considered a key aspect of the IGBP. The processes determining the formation of biogeochemical cycles are responsible for anthropogenically induced changes in the atmospheric composition which, in their turn, can cause considerable changes in the climate and a depletion of the ozone layer which protects the biosphere. Ideas have been expressed of a global observational system to monitor the development of processes in the geosphere and biosphere.

The evolution of the geosphere and biosphere is characterized by an intensifying interaction of their components, which is, first of all, determined by an increasing anthropogenic impact on the environment. This impact is mainly manifested through the transformation of nature by man and growing anthropogenic loads on the environment which affect practically all the components of the geosphere and biosphere (in some cases on the global scale) and require a system approach to studies of the processes in question [1—35].

Uncovering the laws of interaction between the components of the geosphere and biosphere requires interdisciplinary studies of unprecedented complexity, putting the respective problems forward as those of primary importance in natural sciences.

Complex studies of the geosphere and biosphere have long since been needed. During the recent decades numerous large-scale programmes have been accomplished: the International Geophysical Year (IGY), the subsequent International Quiet Sun Year (IQSY), the World Weather Watch (WWW) as well as Global Atmospheric Research Programme (GARP). Recently, an accomplishment has been started of the large-scale Middle Atmosphere Programme (MAP), International Lithosphere Programme, World Climate Programme (with the World Climate Research Pro-

* Dedicated to Prof. K. Nagy on his 60th birthday

gramme as a component), a specialized International Satellite Cloud Climatology Project (ISCCP) as well as a large-scale Soviet programme "Sections" aimed at studies of the role of oceans in the short-term climate changes. Two new international programmes have the same objective: World Ocean Circulation Experiment (WOCE) and Tropical Ocean and Global Atmosphere (TOGA) programmes.

The following international projects on studies of biogeochemical cycles of various components are being accomplished under the aegis of SCOPE and UNEP: the American Programme of the Global Troposphere Chemistry, the International Hydrobiological Programme, and the Man and Biosphere Programme (within the framework of UNESCO, etc.).

Each of the enumerated (and some other) programmes is characterized by a complex approach to solving the respective problems. Analysis of the global ecological problem, on the whole, shows, however, an urgent need of coordinated efforts to achieve the goals of various programmes and substantiating a "super-programme" covering the key aspects of the geosphere-biosphere studies.

These circumstances have stimulated the development of a longterm (the full launching of the observation system is planned for the 90's) International Geosphere-Biosphere Programme (IGBP) aimed at studies of global changes, especially from the viewpoint of anthropogenic impact on the biogeochemical cycles of carbon, nitrogen, sulfur, phosphorus, water, and the dynamics of such factors of life support as solar radiation, the quality of air and natural waters, soil fertility, as well as interrelationships between biospheric and geospheric events [29]. Under the auspices of NASA (USA) an independent (but closely related to the IGBP) programme "Global Change: Impact on Habitability" is being developed [6, 7].

This situation requires an analysis of the key problems of the geosphere-biosphere studies. The following relevant circumstances are of primary importance. First, in time scales of decades and longer it is most clearly seen that the geospheric components constitute a coupled system, the interaction of individual components, being characterized by synergism manifestations (by the feedbacks that sometimes lead to a mutual intensification of various processes). This circumstance determines an exceptional importance of a system approach as well as interdisciplinary studies of physical, chemical and biological processes based on the realization of a joint programme.

Second, an increasing anthropogenic impact on the geosphere and biosphere necessitates studies of the global-scale processes and changes. From the viewpoint of planning the observational system this testifies to an important role of observations from space, with the preserved importance of conventional observation systems both now and in the future [10]. For the problems, on the whole, palaeo- and comparatively planetary analogies become particularly urgent [9, 11]. The latter means that geospheric studies must be closely connected with investigations of the origin and evolution of the solar system. This conclusion is important because the geosphere-biosphere system is not closed in a sense that it is affected by various extra-terrestrial

factors (solar activity, galactic cosmic rays, etc.). Thus, there is no doubt that the atmosphere, ocean, cryosphere, lithosphere, and extra-terrestrial factors are key objects in studies of the geosphere and biosphere.

What worries us now is that some anthropogenic impacts on the global geosphere have already become noticeable (it is manifested, for instance, through increased concentrations of CO_2 and nitrogen oxides, which illustrates an effect on the global cycles of carbon and nitrogen) and can grow during a period of several decades, whereas, our ability for countermeasures has the same (or even longer) characteristic time scale, which demonstrates an urgency of the discussed problems.

Naturally, in studying an exceptionally complicated interdisciplinary problem of primary importance is the determination of its key aspects, among them being the problem of climate and its changes. Climate is a result of interaction between the components of the climatic system (atmosphere–hydrosphere–lithosphere–cryosphere–biosphere), with numerous feedbacks. At present there is no problem in natural science globally more complex and requiring an account of complicated anthropogenic impacts.

In connection with the climate problem the role of the geosphere–biosphere interaction requires serious attention. This role is clearly seen, in particular, when the climatic effect of the anthropogenic CO_2 concentration increase is considered. During the recent years this problem has brought forth controversial conclusions about possible climate changes [4, 17, 19]. This has been caused by two circumstances: (i) lack of reliable enough models of the global carbon cycle (so far, the contribution by marine and continental biota is not clear); (ii) inadequacy of the climate theory from the viewpoint of interactive account of processes in the geosphere and biosphere. Though the climate theory can serve as an illustration of the importance of considering the effect of extra-terrestrial factors, the physical mechanisms determining this effect are still unknown.

The climate problem is also very important from the practical point of view. It is well known that human activity depends (sometimes critically) on climatic conditions (it particularly concerns agriculture). Recently, a new aspect of the problem has been outlined. It is the possible impact of a nuclear war on the climate and biosphere [8, 18]. The respective estimates show an unavoidable global ecological catastrophe due to a strong climate cooling ("nuclear winter"), which must result from a decreased solar radiation reaching the Earth's surface because of strong attenuation by urban and forest post-nuclear fires. Of still more importance is a climatic instability under conditions of such strong disturbances of the climatic system.

Bearing in mind the importance of the climate problem, a perspective interdisciplinary part within the IGBP must be formulated, taking into account the WCRP and other programmes (WOCE, TOGA, ISCCP, etc.) which are being now accomplished. Considering specific features of such programmes, it is clear that studies of interrelated problems of the physics and chemistry of the atmosphere as well as biogeophysical cycles of such components as carbon, nitrogen, phosphorus, sulfur, and

solar-atmospheric interrelationships within the IGBP claim most serious attention since they are not properly foreseen in the on-going programmes.

The problems of atmospheric physics and chemistry are very important from the point of view of assessing the prospects for the biosphere's evolution (its productivity) in conditions of growing anthropogenic impacts. Among other problems the emphasis must be placed on the problem of affecting the stratospheric ozone layer which protects the biosphere from the harmful effects of the UV solar radiation.

Though the conclusions about anthropogenic effects of ozone have been repeatedly changed [26, 27], no doubt, there is a great concern for the effect on total ozone content and vertical profile of ozone concentration of ejected chlorofluorocarbons (CFC), carbon dioxide, nitrogen oxides, methane, and other gases. It is now clear that the effects of these components are closely interrelated, and therefore estimates for individual (isolated) components cannot be considered correct.

There is evidence (from observational data) for increased concentrations of CFC, methylchloroform, carbon dioxide, nitrogen monoxide and methane. However, lack of reliable data on the sources and sinks of the enumerated gases does not permit one to obtain realistic forecasts of their trends in the future (it is particularly true for methane).

The solution of the ozone problem, which is an important component of MAP, requires a realization of an extensive programme of laboratory, field, and theoretical studies of interesting physical, chemical, and photochemical processes in the troposphere and stratosphere, as well as global-scale monitoring of changes in the chemical composition of the atmosphere (for a great number of components).

As for the chemistry of the troposphere, most urgent are studies of (i) water cycle; (ii) processes determining the dynamics of oxidants, especially O_3 , SO_2 and NO_2 ; (iii) mechanisms for dry and wet deposition of pollutants on the surface of land and water; (iv) processes determining the content in the lower atmosphere of aerosol and such gases as CO_2 , N_2O , CH_4 , NH_3 , H_2O ; (v) the effect of optically active atmospheric components on the radiative regime and climate. As in all other cases, an interaction between chemical processes in the troposphere and stratosphere is of principal importance.

To understand the biogeochemical cycles it is necessary to study such controlling factors as (i) natural and anthropogenic inputs of minor components to the troposphere; (ii) their further transport; (iii) chemical transformations (including gas-to-particle conversions); (iv) scavenging from the troposphere of minor gaseous and aerosol components (here the role of the cloud cover is significant). Of primary importance are studies of such biospheric components as forests in the tropics and midlatitudes, steppes and savannah, tundra, agricultural areas, upwellings, and open ocean.

Climatic changes are determined by an important role of minor optically active atmospheric components (including ozone) in the formation of the atmospheric greenhouse effect whose variations are an energy cause of anthropogenic impacts on climate.

An increased ozone content in the troposphere observed during the recent years illustrates this fact. Estimates showed that the contribution of this increase to an intensification of the greenhouse effect is equivalent to the effect of an increased CO_2 concentration.

In the field of numerical modelling the problem of development of interactive models with account of dynamical and photochemical processes is most urgent, to reproduce the global budgets of O_3 , CO and other critically important components.

As before, approximate 1-D and 2-D models must play an important role in studies of the regional transport, geochemical balances, and a search for new important chemical reactions. Longterm objectives must foresee studies of the role of complex heterogeneous processes in the formation of global cycles of various components, with emphasis on poorly studied processes in remote (background) regions of the globe.

In many respects the chemical composition of the atmosphere has been formed and is being changed as a result of interaction with the biosphere. The biosphere is a major source of various organic compounds and of such volatile halocarbon as methylchloroform. Correlations between respective processes are not always direct and unanimous. For instance, chemical reactions with participation of the oxides of nitrogen in the troposphere can lead to both an increase and decrease of the ozone content, which, in its turn, affects the vegetation cover. Variations in concentrations of N_2O , CH_3Cl and CH_4 in the troposphere affect the ozone content in the stratosphere, and this affects the biosphere, too.

A very important problem is to study the physical, chemical and biological processes responsible for release and assimilation of various components by the biosphere (within different ecosystems) and, on the other hand, a reaction of the biosphere to a variability of physical characteristics and chemical composition of the atmosphere. Here of the first priority are such gases whose content is being changed due to human activity: CO_2 , CH_4 , N_2O , NO_x , and SO_x , as well as carbon monoxide playing a substantial role in the chemistry of tropospheric hydroxyl.

There are grounds to believe that chemical and photochemical processes in the stratosphere and mesosphere, varying under the influence of extra-terrestrial factors (solar protons events, galactic cosmic rays) are responsible for the mechanisms that determine the effect of extra-terrestrial factors on the climate.

A key factor of the formation of the chemical composition of the atmosphere is its interaction with ocean which is a sink for CO_2 and nitrogen oxides, but a source (in the regions of coastal lowlands and productive zones) of CH_4 , H_2S and $(\text{CH}_3)_2\text{S}$, as well as a source of halogenes, nitrogen oxide and water vapor for the atmosphere. An urgent problem is a study of the ocean-atmosphere gas exchange as a result of corresponding chemical, physical and biological processes. The cardinal role of the ocean in the formation of the sulfur cycle claims special attention, as well as an interaction of processes in the ocean and on land through rivers, estuaries and coastal waters. Here estimates of the input to the ocean of fresh water, carbon, phosphorus and suspended matter are needed.

Atmospheric aerosol plays an important role as an optically active atmospheric component affecting the formation of the climate [34]. Volcanic stratospheric aerosol that causes strong and prolonged disturbances in the radiative regime and respective climate changes, is a most substantial manifestation of the effect of natural aerosol on radiation and global climate (dust storms are an important regional factor). The presence of anthropogenic aerosol in the troposphere is manifested as a trend of growing atmospheric turbidity and through the formation of atmospheric haze in high latitudes of the Northern hemisphere. In all cases of great importance is a mechanism of gas-to-particle conversion (in the case of volcanic stratospheric aerosol this mechanism prevails), which determines the interrelationship between biogeochemical cycles of sulfur and nitrogen and the processes of atmospheric aerosol formation. Of great concern is an almost unknown effect of aerosol on the processes of formation and dissipation of the cloud cover. As has been mentioned above, the smoke aerosol resulting from post-nuclear fires is critically important.

One more first-priority component of the IGBP is the problem of biological productivity on land and in the ocean. Conditions for biological productivity on land are determined by resources of water, biogenic components and the light, and for the ocean only the two last factors are important. It follows that to assess conditions on land, it is important to obtain information about the global spatial and temporal distribution and rainfall intensity. Fluxes of biogenic components driven by rivers to coastal zones of the ocean and due to mixing, coming eventually to deep oceanic layers, as well as the transport of matter from land to the ocean through the atmosphere play a decisive role as factors of oceanic bioproductivity.

An assessment of the role of the latter factor is hindered by the unreliability of present climate models from the viewpoint of simulating the global distribution of precipitation and the need of taking into account an increasing anthropogenic input of optically active minor gaseous components, which leads to a change of the temperature regime, affects the water cycle, causes acid rains and changes the concentration of toxic gases, soil fertility, and the content of nutrients in rivers and shelf zone. Variations in the intensity of the UV solar radiation at the surface level caused by atmospheric pollution can play a substantial role in affecting the bioproductivity. Having determined the natural variability of bioproductivity, one faces an important problem of analyzing the anthropogenic effect on the biosphere. Solution of this problem will require investigations in various directions.

One of them is the global moisture cycle, whose formation is determined by an interaction of the thermal regime and atmospheric circulation, the processes of evaporation (including evapotranspiration), transfer and phase transformation of water in the atmosphere, the atmospheric circulation being coupled with the radiative regime affected by anthropogenic variations in the atmospheric composition (including chemically and biologically active components). Major difficulties in solving the problem of water cycle are as follows: insufficient reliability of the estimates of evapotranspiration on land, evaporation and precipitation on the oceans, inadequate

parametrization of the processes of formation of clouds and rainfall. The role of the vegetation cover in the transport of water from soil to the atmosphere must further be studied. Data on global cloud cover climatology, distribution of precipitation over the oceans, and evaporation from the ocean surface are still fragmentary.

The leading role of the ocean in the formation of the global water cycle and climate determines the key importance of the problem of atmosphere-ocean interaction and, in this connection, the programme "Sections" [22]. An interaction between the atmosphere and land surface has been studied rather poorly (this has determined the convocation of the 1982 International Conference on Parametrization of Land Surface Processes and a recent development of an International Satellite Land Surface Climatology Project — ISLSCP).

Snow and ice cover and its variability is a fundamental climate forming factor. A high albedo of snow and ice determines a strong (and interactive) dependence of the climate on variations of the extent and properties of the snow and ice cover. Sea ice plays an important role as an insulating layer (it determines a large contribution by heat exchange through polynyas and leads in high latitudes) and as an inertial component that transforms the annual course of temperature by shifting extrema due to latent heat released during freezing in the fall and heat expenditure on melting in the spring. Being a sensitive component of the climate system, the ice cover is a useful indicator of climate changes.

Of principal importance is that the continental biomass has been estimated to an accuracy of a factor of 2, and changes in the time scale of decades are known still worse. Such a situation is hindered by a large-scale effect of human activity on natural ecosystems during the last 100 years (deforestation, wide use of monocultures in agriculture, etc.). Much more accurate estimations are needed of total biomass for individual ecosystems and for land surface, on the whole, which will permit a more reliable analysis of the effect of variations in land use, climate, and an input of nutrients to the global primary bioproductivity.

The following studies are of the first priority: (i) the effect on the input of biogenic components (nitrogen, phosphorus, and sulfur, in particular), variations in the concentration of oxidants, acid rain and heavy metals deposition on reservoirs of carbon; (ii) reaction of biota to environmental changes; (iii) factors that regulate the carbon ratio between living and dead components of ecosystems, and their dependences on anthropogenic impacts.

Solution of the problems requires a development of a global observational system with the use of both conventional and space-borne observation means, with an undoubtful priority of the latter. One must, first of all, analyze an adequacy of available observation means from the viewpoint of the IGBP problems, since it is clear that some of the present needs in information on the geosphere and biosphere can already be satisfied (either completely or partially) with the use of meteorological and Earth's resources satellites as well as manned orbital stations. Also, it is clear that all the requirements to observational data can only be met by improving, expanding, and

supplementing the available observation means. First of all, it necessitates planning an adequate and optimal global system of observations [10].

The importance of the problem of optimal planning of the system of remote sensing from outer space is determined by a number of circumstances, including the high cost of the experiments, limitations for the weight of instruments, power supply and the volume of the information stored on-board. Besides, a number of scientific aspects of the problem must be borne in mind. The satellite-derived data interpretation is connected with solving mathematically incorrect inverse problems: using the characteristics of either reflected or absorbed radiation, one must retrieve the parameters that characterize the state of the environment. The incorrect character of most of the inverse problems of remote sensing determines the critical importance of the choice of measurement conditions for the accuracy of the obtained results [10]. Therefore, a non-adequate choice of spectral channels or geometry of observations can considerably depreciate the observational results. On the other hand, a neglect of the information content of remote sensing data can result in redundant information.

A similarity of requirements to the systems of remote sensing of the objects, interesting for various branches of applications, has brought forth the problem of planning the multi-purpose systems of remote sensing from space. Below will be discussed two major strategies of the approach to the problem of optimal planning of such systems [35].

The economic strategy of the optimal planning of the systems is based on maximization of the integral economic efficiency factor for a multi-purpose observation system V under the condition that the efficiencies for individual problems v_i ($i = 1, \dots, N$) are regulated by special limitations of the type of inequalities from above and from below. One of the most efficient numerical algorithms for the solution of this problem is Balash's method used in calculations for planning the multi-purpose systems of remote sensing of environmental objects from space. All the problems are divided into four large groups: oceanology, hydrology, geology, forestry, and agriculture. Each of these groups is subdivided into sub-groups. The number of partial problems in each of the basic groups varies from 9 to 14 [10]. As an illustration, we shall consider the results of the choice of an optimal set of spectral channels for multi-purpose survey in the interests of oceanology (Table I).

The number of spectral intervals (channels) required for the solution of 12 problems ($N = 12$) in the group "oceanology" (Table I) totals 37 ($M = 37$), the number of particular channels m_i for the i -th specific problem ranging between 18 and 33 ($18 \leq m_i \leq 33$). It means that the volume of information for each specific problem is sufficiently large. It constitutes not less than half the whole volume of information for the entire group of problems "oceanology".

Calculations of the economic efficiency factors of the measurement channels K_j ($j = 1, \dots, N$) have shown that some channels are potentially unprofitable (negative values of K_j) from the viewpoint of multi-purpose application of the observational system. These are channels $0.40\text{--}0.45 \mu\text{m}$ and $0.69\text{--}0.73 \mu\text{m}$. Measurements in the

Table I
Specific problems within the group "Oceanology"

Number	Problem
1	Sea surface state
2	Water turbidity
3	Sea ice
4	General near-shore marine survey
5	Mapping the coastal currents and tides
6	Global-scale mapping of currents, sea surface survey
7	Mapping the coastline and shallow waters
8	Batimetry and topography of ice
9	Distribution and migration of aquatic organisms
10	Contamination of coastal waters
11	Impact of contamination on marine medium
12	Studies of bars, reefs, etc.

following spectral channels are economically justified best: 0.53–0.55; 0.66–0.63; 0.59–0.61; 0.61–0.64; 0.55–0.57 μm . Here maximum values of $\{K_j\}$ are reached. Thus, even simple calculations make it possible to select from a large initial list of spectral channels those most substantial from the viewpoint of economic criteria. Similarly, one can consider other groups of problems and characteristics of the observational system (spatial and temporal resolution, periodicity, geometry, etc.).

The strategy of multi-purpose planning of the space-borne remote sensing system is based on the use of statistical techniques for processing observational data [35]. Mathematical formalization of requirements can easily be made in this case [10]. As an example, we shall consider the problem of the choice of spectral intervals.

Combinations of the channels recommended for solution of specific problems form a multitude on the spectral axis. Having enumerated all the channels, we obtain a totality of elementary channels forming the basis for choosing any interval from the initial totality. We form a vector with the number of components n equal to that of the elementary channels. Let the indication vector x^j , the i -th component of which is either zero or unity (depending on whether the i -th elementary interval meets the requirements of the j -th specific problem or not) correspond to the j -th specific problem. With a set of requirement vectors $X = \{x^j\}$ the problem can be formulated of finding the vector x^* minimizing (maximizing) some purpose functional V determined by a set X . Further, the problem can be solved in two ways. With the totality of vectors $X = \{x^j\}$ considered as a statistical sample, the problem of multi-purpose planning can be formulated as a problem of stochastic optimization, and to solve it, the techniques of principal components and factor analysis can be applied.

The main idea of the factor analysis is that correlations between the components of the X vector are determined by a small number of independent latent variables y_l ($l=1, \dots, m$), with $m \ll n$. In the respective representation

$$x^j = \sum_{l=1}^m f_{jl} y_l + \delta_j \quad (1)$$

the random vector $\delta = (\delta_1, \dots, \delta_n)$ characterizes the level of errors. The elements of vector-columns that form the matrix of factor loads $F = \{f_{il}\}$ give an idea of the components of the initial vector X , substantial from the viewpoint of the information content of performed measurements. In the spectral problem the analysis is aimed at the selection of most informative channels. In practice, it is worthwhile to attempt to select single-bounded multitudes on the spectral axis. First, the columns F are found by the technique of principal components, but not always can their components meet the indicated conditions. To achieve the goal, one can use the fact that relationship (1) and statistical properties remain invariant with respect to the conversion $\tilde{y} = U \cdot y$, $\tilde{F} = F \cdot U^T$ with the orthogonal matrix U . The technique of varimax rotations [10] enables one to find a conversion for which the extremum of the functional

$$V(\tilde{F}) = \sum_{i=1}^m \left\{ n \sum_{j=1}^n (\tilde{f}_{ji}^2 / \tilde{f}_j^2)^2 - \left[\sum_{j=1}^n (\tilde{f}_{ji}^2 / \tilde{f}_j^2) \right]^2 \right\} \quad \left(\tilde{f}_j^2 = \sum_{i=1}^m \tilde{f}_{ji}^2 \right)$$

is realized which characterizes the moderate degree of inhomogeneity of the vector-columns of the factor loads. The parameter $V(\tilde{F})$ is called the variation matrix.

The choice of spectral intervals. Using the described procedure one can choose single-bounded spectral regions for multi-purpose survey carried out in the interests of the following four groups of problems: I — oceanology; II — hydrology; III — geology; IV — forestry and agriculture. The number of specific requirements in the groups totalled: I — 102; II — 33; III — 36; IV — 32. In numerical solution one succeeds in selecting 7–9 factors for each of the groups of problems. The first four channels provide the 70% coverage of the mean square deviation in the space of requirements. In all the cases, two IR channels are chosen: 0.73–0.8 and 0.9–1.1 μm and two intervals in the visible: 0.55–0.58 and 0.69 μm . The remaining channels vary depending on the group of problems. The next step is combining the enumerated groups of problems into complexes. With exclusion of the O_2 , H_2O , O_3 and NO_2 absorption bands, Table II gives the limits of optimal spectral intervals for the “oceanology” problems as well as the complexes of “land” problems and a combination of all the four groups. To achieve the 70% coverage of the initial MSDs in the space of requirements, five channels are necessary (three in the visible, two in the near IR).

Optimal conditions for the survey from space. Let us discuss the choice of parameters of the survey that characterize the geometry of observations as well as the spatial and temporal resolution of the obtained information. We shall consider the following parameters: sun elevation, nadir angle of viewing, spatial resolution, coverage of the survey, and maximum periodicity. The requirements produced by

Table II
Optimal choice of spectral channels for remote sensing

Group of problems	Spectral intervals [μm]				
	1	2	3	4	5
I	0.83–0.87	0.94–1.1	0.76–0.78	0.66–0.69	0.54–0.57
II–IV	0.94–1.1	1.2–1.25	0.51–0.57	0.66–0.69	0.40–0.49
I–IV	0.83–0.87	0.84–1.1	1.2–1.25	0.52–0.57	0.66–0.69

Table III
Optimization of conditions for survey from space

Parameter of the measuring system	Group of problems			
	I	II	III	IV
Sun elevation [deg.]	15–30	15–30	15–30	15–60
Nadir angle of viewing [deg.]	30	30	30	30
Spatial resolution [m]	50–300	5–10	10–30	10–30
Coverage [km]	400–1000	200–400	400–1000	50–200
Maximum periodicity [days]	1	1	3–20	2–20

specific problems with respect to these parameters totalled 73 [10]. For each of the five parameters three types of requirements were chosen: high, moderate and low, which has markedly simplified the interpretation of the results of application of the factor analysis algorithm (Table III).

Some similarity of requirements to the parameters of survey is observed in oceanology and hydrology. High requirements are made to spatial resolution in hydrology. Similarity of requirements is observed in geology, forestry and agriculture, with the difference only in the coverage of survey.

A comparison of the results obtained on the problem of optimization of the system of the spectral channels for remote sensing shows a certain consistency. The channels 0.53–0.57, 0.66, and 0.68 μm are most efficient for oceanology from the viewpoint of both strategies. Other channels located in the near IR (0.83–0.87; 0.94–1.1 μm) are only important from the viewpoint of multi-purpose application of the obtained information. Their economic efficiency is twice as low as compared to the channels of the optical range.

These results show that the developed techniques make it possible, on the strict mathematical basis, to substantiate the optimal composition of the observational system and conditions for observations. The multi-purpose applicability of the observational systems ensures the widest use of the obtained results. The proposed techniques for planning an observational system provide a certain economic effect of the space-borne remote sensing systems both for the economy on the whole and for its certain branches.

To summarize, note that with the efforts taken within some specialized programmes, studies of the interaction between the geosphere–biosphere components, with emphasis on physical, chemical and biological processes on land and in the ocean must be a central problem of the IGBP.

Since the key aspect of the IGBP is biogeochemical cycles and their changes as a result of anthropogenic impacts, man must be not only a cause of geosphere–biosphere transformations but also an important object of studies (from the viewpoint of varying conditions in the geosphere and biosphere that affect him). Such being the case, a major goal of the IGBP is an analysis of the present state of interaction between man and the environment as well as respective prognostic estimates.

References

1. E. P. Borisenkov, *Climate and its Changes*, Znanie Publ., Moscow, 1976.
2. M. I. Budyko, *Evolution of the Biosphere*, Gidrometeoizdat, Leningrad, 1981.
3. P. J. Crutzen and U. Schmailzl, *Planet. and Space Sci.*, 31, 1009, 1983.
4. H. Flohn, *Z. f. Meteorol.*, 35, 1, 1985.
5. Global biology research program. — "NASA Techn. Memo. 85629", Washington, D. C., 1983.
6. Global Change: Impacts on Habitability. A Scientific Basis for Assessment. A Report of the Executive Committee of a Workshop held at Woods Hole, Mass., June 21–26, 1982. — NASA, JPL, D-85, Washington, D. C., 1982.
7. Global Change: A Biochemical Perspective. — JPL Publ. 83–51, NASA, Washington, D. C., 1983.
8. Yu. A. Izrael, *Ecology and the Environmental Control*, Gidrometeoizdat, Leningrad, 1984.
9. K. Ya. Kondratyev and G. B. Hunt, *Weather and Climate on Planets*, Pergamon Press, Oxford, 1982.
10. K. Ya. Kondratyev and O. M. Pokrovsky, *Izv. AN SSSR, Ser. Geography*, 83, 1978.
11. K. Ya. Kondratyev, *Radiative Factors of the Present-Day Global Climate Changes*, Gidrometeoizdat, Leningrad, 1980.
12. K. Ya. Kondratyev, *Stratosphere and Climate*. VINITI, Moscow, 6, 1981.
13. K. Ya. Kondratyev, *World Climate Research Programme: State, Perspectives and the Role of Observations From Space*. VINITI, Moscow, 8, 1982.
14. K. Ya. Kondratyev, *Satellite Climatology*, Gidrometeoizdat, Leningrad, 1983.
15. K. Ya. Kondratyev, A. A. Grigoryev, O. M. Pokrovsky and E. V. Shalina, *Remote Sounding of Atmospheric Aerosol From Space*, Gidrometeoizdat, Leningrad, 1983.
16. K. Ya. Kondratyev and V. I. Binenko, *Impact of Cloudiness on Radiation and Climate*, Gidrometeoizdat, Leningrad, 1984.
17. K. Ya. Kondratyev and N. I. Moskalenko, *The Greenhouse Effect of the Atmosphere and Climate*, VINITI, Moscow, 13, 1984.
18. K. Ya. Kondratyev, S. N. Baibakov and G. A. Kondratyev, *Science in the USSR*, 2, 2, 1985.
19. K. Ya. Kondratyev, *Abh. des Meteorol. Dienstes der DDR*, 43, 5, 1985.
20. *Land-related global habitability science issues*. "NASA Techn. Memo. 85841", Washington, D. C., 1983.
21. G. I. Marchuk, K. Ya. Kondratyev and V. P. Dymnikov, *Some Problems of the Climate Theory*, VINITI, Moscow, 7, 1981.
22. G. I. Marchuk, V. P. Dymnikov, V. B. Zalesny, V. N. Lykosov and V. Ya. Gokhan, *Mathematical Modeling of the General Circulation in the Atmosphere and Ocean*, Gidrometeoizdat, Leningrad, 1984.
23. G. I. Marchuk, *Mathematical Modelling in the Environment Problem*, Nauka Publ., Moscow, 1982.
24. R. J. McNeal, J. P. Mugler, Jr., R. C. Harriss and J. H. Hoell, Jr. *EOS*, 64, 561, 1983.
25. V. A. Mohnen, *Problems of Tropospheric Chemistry for the 1990's*. Preprint, 1983.
26. *News Highlights, Upper Atmos. Progr. Bull.*, 1–3, 1984.
27. M. P. Nicoli and G. Visconti, *Pure and Appl. Geophys.*, 120, 626, 1982.
28. R. G. Prinn, P. G. Simmonds, R. A. Rasmussen, et al., *J. Geophys. Res.*, C88, 8353, 1983.

29. Toward an International Geosphere-Biosphere Program. A Study of Global Change. Report of National Research Council Workshop. Woods Hole, Mass., July 25-29, 1983. National Academy Press, Washington, D. C., 1983.
30. A. F. Treshnikov and Yu. V. Nikolaev, *Problemy Arktiki i Antarktiki*, N54, 25, 1978.
31. V. E. Zuev and I. E. Naats, *Inverse Problems of Laser Sounding of the Atmosphere*, Nauka Publ., Novosibirsk, 1982.
32. A. F. Treshnikov and K. Ya. Kondratyev, *Izv. of the All-Union Geographical Soc.*, 2, 89, 1985.
33. K. Ya. Kondratyev, *Z. f. Meteorol.*, 35, N6, 1985.
34. K. Ya. Kondratyev, *Global Climate*, VINITI, Moscow, 17, 1987.
35. O. M. Pokrovsky, *An Optimization of the Meteorological Sounding of the Atmosphere from Satellites*, Gidrometeoizdat, Leningrad, 1984.

YANG-FELDMAN FORMALISM OF DIPOLE FIELDS*

K. L. NAGY

*Institute for Theoretical Physics, Roland Eötvös University
1088 Budapest, Hungary*

The special form of the Yang–Feldman equations for dipole fields is examined. Comments are given concerning their applications to a *c*-number source and gauge theories.

1. In connection with Heisenberg's unified field theory of elementary particles [1] among other most important features it turned out that we have to consider dipole ghost states [2, 3], state vector spaces with indefinite metric, or more closely we have to work out the quantum theory of the dipole (multipole) fields. A paper appeared soon [4] on the Yang–Feldman formalism of the Froissart field [5] in a non-completely mature form. Dipole fields, however, violate the physical unitarity [6], which hampered further studies, in those times. From this early stage of developments an excellent compilation can be found in Nakanishi's paper [7].

It seems, however, that the correct description of the covariant Lagrangian formalism of gauge fields (including gravitation) cannot be carried out without multipole fields [7, 8]. In this case appropriate subsidiary conditions guarantee the physical unitarity. At the same time those fields may provide a natural explanation for confining potentials [10, 11, 12].

In this paper we examine in the usual framework of an indefinite metric theory the Yang–Feldman equations of dipole fields (c.f. mainly [8, 13, 14]), and make some observations and comments. We do not use anywhere integrals of the type

$$\int A(x-x')\psi(x')dx',$$

where ψ is a (possible free) field. The form of our equations versus [8, 13, 14] seems to us more inspiring, especially in applications and understanding.

2. The Lagrangian

$$L = \partial^\mu A \partial_\mu B - m^2 AB + \frac{\lambda}{2} A^2 + Aj_1 + Bj_2 \quad (1)$$

* Dedicated to Prof. K. Nagy on his 60th birthday

yields the field equations

$$(\square + m^2)A = j_2, \quad (2)$$

$$(\square + m^2)B = \lambda A + j_1. \quad (3)$$

Assuming that the sources j_1, j_2 behave properly (e.g. do not contain a discrete spectrum at $p^2 = m^2$) the Yang-Feldman equation corresponding to Eq. (2) is

$$A(x) = A_i(x) + \int \Delta_i(x-x')j_2(x')dx', \quad (4)$$

$$(\square + m^2)A_i(x) = \delta(x).$$

The suffix i denotes asymptotic indices "in" or "out", respectively. The solution of (3) for $m \neq 0$ can be found as follows:

Since

$$\square x^\mu \partial_\mu \varphi \equiv (2 + x^\mu \partial_\mu) \square \varphi, \quad (5)$$

from (2) we have

$$\square x^\mu \partial_\mu A = (2 + x^\mu \partial_\mu)(j_2 - m^2 A).$$

Looking at this form, it is easy to recognise that the field defined by

$$C_\alpha \equiv -\frac{\lambda}{2m^2}(\alpha + x^\mu \partial_\mu)A \quad (6)$$

satisfies

$$(\square + m^2)C_\alpha = \lambda A + j_\alpha,$$

$$j_\alpha = -\frac{\lambda}{2m^2}(2 + \alpha + x^\mu \partial_\mu)j_2.$$

α is an arbitrary number. Therefore

$$(\square + m^2)(B - C_\alpha) = J_\alpha \equiv j_1 - j_\alpha. \quad (7)$$

The general solution of (7) is

$$B(x) = B_i^0(x) - \frac{\lambda}{2m^2}(\alpha + x^\mu \partial_\mu)A(x) + \int \Delta_i(x-x')J_\alpha(x')dx', \quad (8)$$

where $B_i^0(x)$ is a free field.

$$B_{\text{out}}^0(x) = B_{\text{in}}^0(x) + \int \Delta(x-x')J_\alpha(x')dx'. \quad (9)$$

Substituting J_α and A , after a partial integration we have

$$\begin{aligned} B(x) = & B_i^0(x) - \frac{\lambda}{2m^2}(\alpha + x^\mu \partial_\mu)A_i(x) + \int \Delta_i(x-x')j_1(x')dx' - \\ & - \frac{\lambda}{2m^2} \int [(x^\mu - x\mu') \partial_\mu + 2] \Delta_i(x-x')j_2(x')dx'. \end{aligned} \quad (10)$$

It is interesting to notice that from Eq. (5)

$$(\square + m^2) [(x^\mu - x'^\mu)j_\mu + 2]A_i(x - x') = -2m^2\Delta_i(x - x'),$$

as it has to be. Using the identity

$$2m^2 \frac{\partial \Delta}{\partial m^2} = (2 + x^\mu \partial_\mu) \Delta, \tag{11}$$

Eq. (9) reads

$$B(x) = B_i^0(x) - \frac{\lambda}{2m^2} (\alpha + x^\mu \partial_\mu) A_i(x) + \int \Delta_i(x - x') j_1(x') dx' - \lambda \int \frac{\partial \Delta_i(x - x')}{\partial m^2} j_2(x') dx'. \tag{12}$$

Of course

$$(\square + m^2) \frac{\partial \Delta_i}{\partial m^2} = -\Delta_i. \tag{13}$$

Denoting

$$B_i(x) = B_i^0(x) - \frac{\lambda}{2m^2} (\alpha + x^\mu \partial_\mu) A_i(x), \tag{14}$$

one gets

$$B_{out}(x) = B_{in}(x) + \int \Delta(x - x') j_1(x') dx' - \lambda \int \frac{\partial \Delta(x - x')}{\partial m^2} j_2(x') dx'. \tag{15}$$

A_i, B_i satisfy

$$\begin{aligned} (\square + m^2) A_i &= 0, \\ (\square + m^2) B_i &= \lambda A_i. \end{aligned}$$

Demanding that for $j_i = 0$ the canonical commutation relations should give

$$[B(x), B(x')] = -i\lambda \frac{\partial}{\partial m^2} \Delta(x - x'), \tag{16}$$

we must have $\alpha = 1$. The other commutators are

$$[A(x), A(x')] = 0, \quad [A(x), B(x')] = i\Delta(x - x').$$

By means of the creation and annihilation parts of A_i and B_i^0 the indefinite metric Fock spaces $\{|i\rangle\}$ can be constructed in the usual way.

Equations (4) and (15) are valid also for $m = 0$;

$$\begin{aligned} -\frac{\partial}{\partial m^2} \Delta(x) \Big|_{m^2=0} &= E(x) = \frac{1}{2} \Delta^{-1}(x_0 \partial_0 - 1) D(x) = \\ &= -\frac{\varepsilon(x_0)}{8\pi} \theta(x^2). \end{aligned}$$

Instead of Eq. (14), from B_i one can extract a non-covariant simple pole free field $B_i^0(x)$ as:

$$B_i^0(x) = B_i(x) - \lambda \frac{1}{2} \Delta^{-1} \left(x_0 \partial_0 - \frac{1}{2} \right) A_i(x). \quad (17)$$

Perhaps as an example it is interesting to note that for a c -number source $j_1 = 0$,

$$j_2 = g \delta^{(3)}(\mathbf{r}) e^{-\alpha|t|} \cos \omega t, \quad \omega \neq 0 \quad (18)$$

in the asymptotic region $t \gg 0$

$$\begin{aligned} \langle \text{in}, 0 | A_{\text{out}}(x) | \text{in}, 0 \rangle &= \frac{g}{4\pi r} \cos \omega(r-t), \\ \langle \text{in}, 0 | B_{\text{out}}(x) | \text{in}, 0 \rangle &= \frac{g\lambda}{8\pi} \Delta^{-1} (r\partial_t - 1) \frac{\cos \omega(r-t)}{r} = \\ &= -\frac{g\lambda}{8\pi\omega^2 r} (\omega r \sin \omega(r-t) + \cos \omega(r-t)), \\ \langle \text{in}, 0 | B_{\text{out}}^0(x) | \text{in}, 0 \rangle &= \frac{g\lambda}{8\pi} \Delta^{-1} \left((r-t)\partial_t - \frac{1}{2} \right) \frac{\cos \omega(r-t)}{r} = \\ &= -\frac{g\lambda}{8\pi^2 \omega^2 r} \left(\omega(r-t) \sin \omega(r-t) + \frac{3}{2} \cos \omega(r-t) \right). \end{aligned}$$

3. Dipole fields, in general, violate the physical unitarity. In connection with the above equations, we mention a case where physical unitarity might survive. From (9) $B_{\text{in}}^0 = B_{\text{out}}^0$ if

$$j_1 = j_x = -\frac{\lambda}{2m^2} (3 + x^\mu \partial_\mu) j_2. \quad (19)$$

Let us suppose that a physical field φ is coupled to the above dipole field in the form e.g. j_1 of (19) and $j_2 = g\varphi^2$. Let the physical subspace be spanned by states $\{|\text{in}\rangle_{\text{phys}}\}$ containing only physical φ "in" particles (i.e. of the form $\varphi_{\text{in}}^+ \dots \varphi_{\text{in}}^+ |0\rangle$, with positive norm). Assuming asymptotic completeness $|\text{in}\rangle_{\text{phys}}$ can be expanded symbolically as

$$\varphi_{\text{in}}^+ \dots \varphi_{\text{in}}^+ |0\rangle = \sum c_{n,m,l} (\varphi_{\text{out}}^+)^n (A_{\text{out}}^+)^m (B_{\text{out}}^0)^l |0\rangle.$$

The states on the right-hand side are linearly independent, l must be zero since $B_{\text{in}}^0 = B_{\text{out}}^0$ and looking at the left-hand side, the state by definition must not contain B_{in}^+ . States with $l=0$, $m \neq 0$ have zero norm. Therefore, symbolically

$$1 = \sum c_{n,0,0}^* c_{n,0,0}.$$

Such a theory is, perhaps, non-trivial, since in the above example with j_2 of (18), and

$$\langle \text{in}, 0 | B_{\text{out}}^0(x) | \text{in}, 0 \rangle = 0, \quad \langle \text{in}, 0 | A_{\text{out}}(x) | \text{in}, 0 \rangle = \frac{g}{4\pi r} \cos \omega(r-t)$$

however for $B_{out}(x)$ we have

$$\langle in, 0 | B_{out}(x) | in, 0 \rangle = - \frac{\lambda g}{8\pi\omega^2 r} \left(\omega t \sin \omega(r-t) - \frac{1}{2} \cos \omega(r-t) \right).$$

In more realistic theories the vector version of (1) has to be considered too. There, however, terms proportional to λ and m^2 violate the gauge invariance. Those terms may appear as a consequence of a spontaneous symmetry breaking. Taking e.g. the Higgs mechanism, one complex scalar field Φ_1 coupled to A_μ in the form

$$\nabla_\mu \Phi_1^* \nabla^\mu \Phi_1, \quad \nabla_\mu \Phi_1 = (\partial_\mu - ie_1 A_\mu) \Phi_1,$$

reproduces $\lambda \neq 0$, while with another field Φ_2

$$\nabla_\mu \Phi_1^* \nabla^\mu \Phi_2 + h.c., \quad \nabla_\mu \Phi_2 = (\partial_\mu - ie_2 B_\mu) \Phi_2,$$

$m \neq 0$ can also be established. Thus the Higgs becomes a complex scalar dipole field.

In this article we strived to demonstrate that multipole fields still deserve some interest nowadays.

References

1. W. Heisenberg, Nucl. Phys., 4, 532, 1957; Rev. Mod. Phys., 29, 269, 1957; Introduction to the Unified Field Theory of Elementary Particles. Wiley and Sons, London, 1966.
2. G. Källén and W. Pauli, Dan. Mat. Fys. Medd., 30, No. 7. 1955.
3. W. Heisenberg, Nucl. Phys., 4, 532, 1957.
4. L. Horváth, Dissertation., unpublished, 1962.
5. M. Froissart, Nuovo Cim. Suppl., 14, 197, 1959.
6. N. Nakanishi, Prog. Theor. Phys., 38, 881, 1967; N. Nakanishi, Phys. Rev., D3, 1343, 1971; K. L. Nagy, Acta Phys. Hung., 28, 245, 1970.
7. N. Nakanishi, Prog. Theor. Phys. Suppl., No. 51, 1, 1972.
8. T. Kugo and I. Ojima, Prog. Theor. Phys. Suppl., No. 66, 1, 1979.
9. N. Nakanishi, Publ. Math. Sci., Kyoto Univ. Vol. 19, No. 3, 1983.
10. J. E. Kiskis, Phys. Rev., D12, 3583, 1975.
11. K. L. Nagy, Acta Phys. Hung., 42, 377, 1977.
12. H. Narnhofer and W. Thirring, Phys. Letters, 76B, 428, 1978.
13. N. Nakanishi, Prog. Theor. Phys., 54, 890, 1975.
14. N. Nakanishi and I. Ojima, Prog. Theor. Phys., 59, 242, 1978.

DENSITY PEAKS AND LARGE SCALE VELOCITIES*

ALEXANDER S. SZALAY** and LARS GERHARD JENSEN

*Theoretical Astrophysics Group, Fermi National Accelerator Laboratory
Batavia, Illinois, 60510, USA*

The velocity dispersion of density extrema is calculated for a general fluctuation spectrum and is shown to be smaller than for randomly selected points for most spectra by a factor $2/(n+3)$, where $|\delta_k|^2 = k^n$. This is valid for the peculiar velocity distribution of clusters of galaxies as well. Our location near the Virgo cluster changes the interpretation of the large scale velocity field measurements. The results imply that "biased" galaxy formation with cold dark matter has even more serious difficulties than previously suggested by Vittorio and Silk [1].

Recent observations on the large scale velocities of the galaxy distribution (Hart and Davies [2], de Vaucouleurs and Peters [3]) were used to constrain theories of galaxy formation (Clutton-Brock and Peebles [4], Kaiser [5], Vittorio and Silk [1]). The latter papers calculated the rms velocity of a randomly placed spherical shell of radius R for fluctuations given by their power spectrum, assuming random phases. With various normalization criteria like $\xi(5h^{-1}Mpc)$, $J_3(20h^{-1}Mpc)$ or $1+Z_{nl}$ expectations for $\langle v^2 \rangle$ were obtained. By comparing these to the observations the different theoretical models were accepted or rejected at respective significance levels.

We would like to point out a possible systematic bias in this method. The observations are hardly randomly selected or independent, they are all influenced by the Virgo cluster, the largest mass concentration in our neighbourhood. Attempting to model this configuration, we calculate the rms velocity of a spherical shell centered on an extremum of the smoothed density field, and show that this velocity is lower in all practical cases than for a random point. A more detailed calculation where the shell is not centered on the extremum will be published elsewhere.

Let us define the n 'th moment of the smoothed density fluctuation spectrum as

$$s_n^2 = \int d^3k k^{2n} |\delta_k|^2 W(kR), \quad (1)$$

where $W(kR)$ is the square of the window function, the Fourier transform of the galaxy selection function and R its radius. The dispersion of the velocity v_r and the smoothed overdensity D are given by

$$\langle D^2 \rangle = s_0^2; \quad \langle v_r^2 \rangle = (Haf)^2 s_{-1}^2, \quad (2)$$

* Dedicated to Prof. K. Nagy on his 60th birthday

** On leave from Eötvös University, Budapest, Hungary

with $H = \dot{a}/a$, the Hubble constant and $Hf = (\delta/\delta)$. Here we denote the smooth overdensity centered on \mathbf{x} as $D(\mathbf{x})$, the components of its gradient and second derivatives as D_i , D_{ij} and the components of its velocity as $v_i = (Haf)U_i$. The correlation coefficients between those quantities are simple due to the orthogonality of the odd and even ranks of the derivatives. The velocities are only coupled to the gradients independent of D and D_{ij} , hence the velocity of an extremum will not depend on whether it is a maximum, a minimum or a saddle point. The correlation between D_i and U_i are the following:

$$\begin{aligned} \langle DD_i \rangle &= \langle DU_i \rangle = \langle D_{ij}D_i \rangle = \langle D_{ij}U_i \rangle = 0, \\ \langle D_iD_i \rangle &= s_1^2/3; \quad \langle D_iU_i \rangle = s_0^2/3; \quad \langle U_iU_i \rangle = s_{-1}^2/3. \end{aligned} \quad (3)$$

Only the $i=j$ terms are coupled, and their correlation matrix is

$$\mathbf{C} = \begin{pmatrix} s_1^2/3 & s_0^2/3 \\ s_0^2/3 & s_{-1}^2/3 \end{pmatrix} \quad \text{and} \quad \mathbf{C}^{-1} = \begin{pmatrix} 3s_{-1}^2/Q & 3s_0^2/Q \\ -3s_0^2/Q & 3s_1^2/Q \end{pmatrix}, \quad (4)$$

where $Q = s_1^2s_{-1}^2 - s_0^4$.

Now, using \mathbf{C}^{-1} we can easily construct the joint probability distribution $P(U_i, D_i)$ (for a discussion, see Bardeen et al [6]). Constraining the value of D_i to 0 for the selection of extrema the velocity dispersion becomes

$$\langle v_x^2 \rangle = 3(Haf)^2 \langle U_i^2 \rangle = (Haf)^2 Q^2 s_1^{-2}, \quad (5)$$

where the subscript x refers to the extremum. This is to be compared to the velocity v_r of a random shell:

$$\frac{\langle v_x^2 \rangle}{\langle v_r^2 \rangle} = 1 - \frac{s_0^4}{s_1^2 s_{-1}^2}. \quad (6)$$

Approximating the power spectrum with $|\delta_k|^2 = Ak^n$ and the window function with $\exp(-k^2 R^2)$, following Kaiser [5], we obtain $s_m^2 = A\Gamma((2m+n+3)/2)R^{-(2m+n+3)}$, so

$$\frac{\langle v_x^2 \rangle}{\langle v_r^2 \rangle} = \frac{2}{n+3}. \quad (7)$$

This turns out to be a fair approximation even for cold dark matter, if R is sufficiently large (Table I). For the $n=1$ Zel'dovich spectrum $2/(n+3)$ is equal to 0.5, hence the rms velocity of the extrema is $1/\sqrt{2}$ lower than that of a random point.

If we assume, that the measured data reflect such a configuration, then the velocities should be multiplied by $\sqrt{2}$ before comparing them to calculations made for random points or the model velocities should be reduced. This effect will change the statistical significance attached to any particular model substantially, it makes the agreement with cold dark matter models worse, providing a very strong upper limit on the enhancement factor of the correlation function in "biased" galaxy formation models.

Table I

The ratio of the velocity of extrema (v_e) to the velocity of a random shell (v_r) as a function of the shell radius R for cold dark matter, $n=1$, $h=0.5$, $\Omega=1$. It is approaching the asymptotic analytic value $2/(n+3)=1/2$ rapidly. For a neutrino-dominated universe the value is very close to the asymptotic one

$R[Mpc]$	$\langle v_e^2 \rangle / \langle v_r^2 \rangle$
10	0.591
20	0.544
40	0.519
80	0.508

References

1. N. Vittorio and J. Silk, *Ap. J.*, 293, L1, 1985.
2. L. Hart and R. D. Davies, *Nature*, 197, 191, 1982.
3. G. H. de Vaucouleurs and W. L. Peters, *Ap. J.*, 187, 1, 1984.
4. M. Clutton-Brock and P. J. E. Peebles, *Astron. J.*, 86, 1115, 1981.
5. N. Kaiser, *Ap. J.*, 273, L17, 1983.
6. J. M. Bardeen, J. R. Bond, N. Kaiser and A. S. Szalay, *Ap. J.*, 304, 15, 1986.

THE GLASS TRANSITION FEEDBACK: REVIEW AND COMMENTS*

T. GESZTI

*Department of Atomic Physics, Roland Eötvös University
1088 Budapest, Hungary*

and

*Research Institute for Technical Physics of the Hungarian Academy of Sciences
1325 Budapest, Hungary*

Different implementations of the mode-coupling feedback scenario of the fluid-glass transition are compared. It is pointed out that the relaxation of density fluctuations is at least partially controlled by shear motions coupled to one-particle diffusion. Neutron inelastic scattering studies are suggested to be done on heat-treated samples, at low temperatures.

1. Introduction

To explain how a supercooled fluid gets solid without developing crystalline order, i.e. how it transforms into a glass, is one of the outstanding unsolved problems in the physics of condensed matter. In a way, what happens is natural: at temperature $T = 0$ a substance held together by attractive forces, if not superfluid, is in a state similar to classical mechanical equilibrium, which is — apart from pathological model systems — locally stable. Therefore the system resists shear, i.e. it is solid, having infinite viscosity η . The obvious phenomenology of the fluid-glass transition is then $\eta(T) \rightarrow \infty$ for $T \rightarrow 0$. At the somewhat arbitrary finite value $\eta = 10^{12}$ Pa · s (10^{13} Poise) the substance appears already perfectly solid; besides, as estimated from the Stokes-Einstein relation with slip boundary conditions

$$D = \frac{k_B T}{4\pi\eta a}, \quad (1)$$

where D is the coefficient of self-diffusion and a is the diameter of a rigidly diffusing unit (atom or molecule), at such a high viscosity it takes hours for an atom to diffuse an atomic diameter away from its starting position.** Therefore it is both macroscopically and microscopically justified to call the corresponding temperature T_g where

$$\eta(T_g) = 10^{12} \text{ Pa} \cdot \text{s} \quad (2)$$

* Dedicated to Prof. K. Nagy on his 60th birthday

** The corresponding value of the diffusion coefficient, $D \sim 10^{-19}$ m²/s is of the usual order of magnitude measured in metallic glasses below their glass transition temperatures, which are much lower than the temperatures at which solid-state diffusion is usually measured.

the temperature of glass transition. If a fluid is cooled below its T_g fast enough to give no time for crystallization, then it remains amorphous but structurally arrested, i.e. glassy.

At closer inspection of data about how η grows with cooling, $\eta(T)$, which is measurable only above T_g , seems to diverge at some temperature T_0 usually some 10% below T_g , according to the empirical law

$$\eta = \eta_0 e^{\frac{A}{T - T_0}} \quad (3)$$

called Vogel–Fulcher law [1, 2], or

$$\eta = B(T - T_0)^{-1.8} \quad (4)$$

which is derived from more recent theories to be discussed below [3, 4]. In the pre-vitrification region well above T_g the simpler Batchinski–Hildebrand law [5, 6]

$$\eta \propto (T - T^*)^{-1} \quad (5)$$

with some $T^* > T_g$ also gives a good fit to data.

The divergence of η at a finite temperature, which seems to be the rule with only a few exceptions (SiO_2 , GeO_2), suggests the existence of some kind of fluid–glass phase transition, the nature of which is still far from being clear.

This is not meant to say that no serious theoretical ideas have been advanced to explain the fluid–glass transition, or, as it is usually called, the glass transition. Adam and Gibbs in their suggestive paper [7] discuss vitrification in terms of a gradual loss of excitable configurational degrees of freedom, as measured by the configurational entropy. This work, however, leaves the relevant microscopic degrees of freedom largely unspecified. A related and microscopically much more specific theory has been suggested by Cohen and Turnbull [8] and worked out in detail by Cohen and Grest [9] in which atomic mobility is associated with inhomogeneously distributed free volume, enabling isolated atoms to hopping motion. This theory suffers from the need to specify how a tensorial shear flow can be composed of vectorial hopping currents [10], however, the inhomogeneous scenario it suggests may be true.

Recently an alternative, homogeneous scenario was suggested [11] and subsequently worked out in considerable detail [3, 4, 12, 13], according to which the viscosity is driven to infinity by a positive feedback in the coupling of collective modes of the fluid, growing in strength with decreasing temperature. In the remaining part of this paper the original suggestion [11] is compared with the later developments, particular attention being paid to conflicting predictions and possibilities of obtaining experimental information about the microscopic mechanism.

2. Mode coupling route to the glass transition

Shear motion is liable to emit sound waves, as exemplified by the file or the violin. Planes of atoms in a fluid are rough on the atomic scale, therefore their shear flow would emit most intensively waves of wavelengths close to atomic first-neighbour distances. Such short waves in a liquid are usually overdamped: they relax and do not oscillate.

The emission of sound waves carries away some of the energy of the shear flow, therefore it contributes to viscosity. Since this mechanism depends on the coupling between one mode of collective motion of the fluid (shear flow) and another (sound waves), the corresponding contribution to transport coefficients (in the present case: viscosity) is said to originate from mode-mode coupling, or more briefly: from mode coupling [14].

The coupling of sound waves to shear motion gets stronger with decreasing temperature, since atomic planes cling more together. More important: the longer an overdamped sound wave lives, i.e. the slower its relaxation, the more energy it dissipates, i.e. the more it contributes to viscosity.

That suggests the existence of a *viscosity feedback* loop [11]: growing viscosity makes the relaxation of overdamped waves slower, which makes their contribution to viscosity larger. At a critical strength of mode coupling, corresponding to a critical temperature T_0 , this feedback may drive the viscosity to infinity. This is a possible scenario for the glass transition.

Short-wavelength overdamped sound waves in a liquid can be divided into longitudinal and transverse modes, and it is a quite non-trivial question [15] which of them becomes actually slowly relaxing at the glass transition. The answer [11] is: the longitudinal one, because the autocorrelation function of density fluctuations in a fluid of point-like particles at positions $\mathbf{r}_i(t)$,

$$F(\mathbf{q}, t) = N^{-1} \langle \rho(\mathbf{q}, t) \rho(-\mathbf{q}, 0) \rangle = N^{-1} \left\langle \sum_{i,j} e^{i\mathbf{q} \cdot (\mathbf{r}_j(t) - \mathbf{r}_i(0))} \right\rangle \quad (6)$$

is *non-relaxing* in a glass in which particles are confined to oscillate around fixed equilibrium positions:

$$F(\mathbf{q}, t) \rightarrow f(\mathbf{q}) \neq 0 \quad (t \rightarrow \infty, \text{glass}).^* \quad (7)$$

If the viscosity is finite but high, the rigid confinement of atoms is lifted and $F(\mathbf{q}, t)$ relaxes to zero but *slowly*. No similarly systematic reason for slow relaxation exists in the case of the transverse mode.

The time Fourier transform of $F(\mathbf{q}, t)$ is the dynamical structure factor $S(\mathbf{q}, \omega)$ which has a narrow quasi-elastic peak around $\omega = 0$ if $F(\mathbf{q}, t)$ has a slowly relaxing long-

* The non-zero limit $f(\mathbf{q})$ is analogous to the Edwards-Anderson order parameter of spin glasses [16] and to the recoilless fraction characterizing the Mössbauer effect [17].

time tail. Since $S(q, \omega)$ is proportional to the doubly differential cross-section of coherent inelastic neutron scattering, some hope arises that the quasi-elastic peak may be observed by recent high-resolution inelastic scattering techniques [18]. Below we discuss those aspects of the theory which seem to have immediate relevance to the experiments.

3. Slow relaxation: shear viscosity

We begin the quantitative description by an improved argument for the physical mechanism of slow relaxation described in [11]. $F(q, t)$ can be written in the form

$$F(q, t) = \langle s_i(q, t) f_1(q, t) \rangle, \quad (8)$$

where

$$s_i(q, t) = \sum_j e^{i\mathbf{q} \cdot (\mathbf{r}_j(t) - \mathbf{r}_i(t))} \quad (9)$$

is expected to oscillate fast around its non-zero mean value: the static structure factor

$$S(q) = N^{-1} \left\langle \sum_{i,j} e^{i\mathbf{q} \cdot (\mathbf{r}_j - \mathbf{r}_i)} \right\rangle, \quad (10)$$

whereas

$$f_i(q, t) = e^{i\mathbf{q} \cdot (\mathbf{r}_i(t) - \mathbf{r}_i(0))}, \quad (11)$$

the thermal average of which is called the "incoherent scattering function" and denoted by $F_s(q, t)$, relaxes to zero by single-particle diffusion which is very slow if the system has high viscosity. Therefore the glass-transition region, due to the wide separation of time scales of $s(q, t)$ and $f_s(q, t)$, is a clear case for the Vineyard convolution approximation [19]

$$F(q, t) \approx S(q) F_s(q, t). \quad (12)$$

The above argument emphasizes that independently of approximation (12), it is slow translational one-particle diffusion that controls the relaxation of $F(q, t)$. Diffusion is, however, impossible if the viscosity is infinitely high, as described by the Stokes-Einstein relation, Eq. (1). Using that equation in the context of a quasi-hydrodynamical description of diffusion, one obtains for long times

$$F(q, t) \approx f(q) e^{-\delta t}, \quad (13)$$

where $f(q)$ is the quantity appearing in Eq. (7) ($f(q) < S(q)$ because of faster transient non-diffusive relaxation of $F_s(q, t)$), whereas

$$\delta = k_B T q^2 / (4\pi a \eta). \quad (14)$$

Equation (7) is recovered in the limit $\eta \rightarrow \infty$.

The mode coupling expansion [11] gives the viscosity in the form

$$\eta = \eta_{\text{coll}} + \int_0^{\infty} dt \int d^3k w(k) F^2(k, t), \quad (15)$$

where η_{coll} is the contribution of non-mode-coupling (collisional) processes, whereas $w(k)$ is a mode-coupling vertex function which increases if the peaks of the static structure factor $S(q)$ get sharper, which happens with decreasing T [20]. Substituting (13) and (14) into (15), we obtain an equation of the form

$$\eta = \eta_{\text{coll}}(T) + \alpha(T)\eta, \quad (16)$$

which determines η self-consistently. If $\alpha(T)$ remains analytic through the glass transition and can be approximated there by a linear function then the Batchinski-Hildebrand law, Eq. (5) is obtained.

4. Slow relaxation: frequency-dependent longitudinal viscosity

Leutheusser [3] and Bengtzelius et al [4] modified the above picture in two important respects, retained by subsequent workers [12, 13].

(i) They pointed out that by definition the slow relaxation of $F(q, t)$ is described by the longitudinal viscosity, a transport coefficient in principle independent of the shear viscosity η . Doing a lowest-order mode-coupling expansion analogous to Eq. (15) for the longitudinal viscosity itself, a closed feedback loop is obtained for the slowing-down of the self-coupled field of density fluctuations, without explicit participation of shear motions.

(ii) They recognized that transport coefficients containing a mode-coupling contribution are frequency-dependent, thus e.g. Eq. (15) should be written as

$$\eta(\omega) = \eta_{\text{coll}} + \int_0^{\infty} dt e^{i\omega t} \int d^3k w(k) F^2(k, t), \quad (17)$$

and this frequency dependence is very relevant in the feedback mechanism, since $F^2(k, t)$ in the integrand is just as slowly relaxing as the primary quantity $F(q, t)$ whose relaxation is governed by the viscosities.*

Inspection of the mode coupling vertex functions like $w(k)$ reveals that the strongest coupling is between density fluctuations of wave vectors \mathbf{q} with $|\mathbf{q}| = q_0 \approx 2\pi/a$, corresponding to the main peak of the static structure factor $S(q)$. Taking these fluctuations approximately as a closed self-coupled system and introducing

$$\Phi(t) = F(q_0, t)/S(q_0), \quad (18)$$

* The frequency dependence of η_{coll} is negligible in the context of slow relaxation.

on the basis of (i) and (ii) Leutheusser [3] and Bengtzelius et al [4] derived the model equation

$$\ddot{\Phi}(t) + \Omega^2 \Phi(t) + \gamma \dot{\Phi}(t) + 4\Omega^2 \lambda \int_0^t \Phi^2(t-t') \dot{\Phi}(t') dt' = 0, \quad (19)$$

where γ characterizes the collisional contribution to longitudinal viscosity, Ω^2 is determined by $\ddot{\Phi}(t=0)$, finally λ , the control parameter of the problem, contains an angular integral of the vertex function characterizing the strength of coupling of three longitudinal modes of wave number q_0 among themselves. Explicit calculations show that λ is the bigger the sharper are the features of $S(q)$, therefore it grows with decreasing temperature.

Equation (19) with the initial conditions

$$\Phi(0) = 1; \quad \dot{\Phi}(0) = 0, \quad (20)$$

the first of which follows from the definition (18) and the second from time reversal invariance, offers a very attractive theoretical model for the glass transition. First of all, for $\lambda > \lambda_c = 1$ its solution, in accordance with the postulated Eq. (7), for long times approaches the finite limit

$$f = \frac{1}{2} + \frac{1}{2} (1 - 1/\lambda)^{1/2}. \quad (21)$$

The critical value $\lambda_c = 1$, if λ is explicitly evaluated for a hard-sphere model [4], corresponds to a critical packing fraction 0.516, whereas computer simulations on the same model [21, 22] indicate a fluid–solid transition between 0.52 and 0.56, which — in view of the rough approximations used in deriving Eq. (19) — gives confidence in the correctness of the scheme.

For $\lambda < 1$, $\Phi(t)$ relaxes to zero, however very slowly if $\varepsilon \equiv 1 - \lambda \ll 1$. In this critical region Eq. (19) has a solution of the form

$$\Phi(t) \approx \frac{1}{2} \exp [2\varepsilon^{1/2} g(\varepsilon^\nu t)] \quad (22)$$

possessing the property of dynamical scaling.

On two competing time scales* divided by $\varepsilon^\nu t = 1$, the solution has different characteristic limiting behaviours:

(i) For $1 \ll t \ll \varepsilon^{-\nu}$ one has a scaling solution

$$\Phi(t) \approx \frac{1}{2} + at^{-\alpha}, \quad (23)$$

* Like in [4] time is measured here in units of t_{micr} , a characteristic time of microscopic collisional processes: the Enskog collision time for the hard-sphere fluid, scaled to 10^{-13} s for argon at 40 K.

where the exponent α is fixed by the corresponding limiting form of Eq. (19) to be $\alpha = 0.395$. The requirement that the short-time limit should be insensitive to the criticality parameter ε and still compatible with Eq. (22), gives $\nu = (2\alpha)^{-1} = 1.265$, setting the time scale of the critical dynamics.

(ii) For very long times, $t \gg \varepsilon^{-\nu}$, the relaxation is exponential:

$$\Phi(t) \propto e^{-\delta t}. \quad (24)$$

This is compatible with Eq. (22) if $\delta \propto \varepsilon^{-\mu}$ with $\mu = \nu + \frac{1}{2} = 1.765$. This slowest exponential relaxation dominates the mode-coupling contribution to both shear and longitudinal viscosities, which therefore diverge for $\varepsilon \rightarrow 0$ with the same exponent μ . If $\varepsilon = 1 - \lambda$, characterising the strength of mode coupling, is assumed to be a smooth function of temperature, then Eq. (4) of the Introduction is obtained, which gives a faithful representation of experimental data.

5. Discussion

The slow relaxation of density fluctuations in a glass-forming fluid can be measured in principle by means of high-resolution coherent neutron inelastic scattering, since the imaginary part $\tilde{\Phi}''(\omega)$ of the Laplace transform of $\Phi(t)$,

$$\tilde{\Phi}(\omega) = i \int_0^{\infty} dt e^{i\omega t} \Phi(t), \quad (25)$$

determines through Eq. (18) the dynamical structure factor at wave-vector $q \approx q_0$:

$$S(q_0, \omega) = \tilde{\Phi}''(\omega) S(q_0), \quad (26)$$

which is the quantity rather directly measured in such experiments. Such a study, utilizing the neutron spin echo [18] facility in Grenoble, is now being planned [23].

The features predicted by the Leutheusser [3] and Bengtzelius et al [4] theories are: a very narrow Lorentzian central quasielastic peak for $\omega t_{\text{micr}} < \varepsilon^{\nu}$, corresponding to the longest-time relaxation, Eq. (24), and a broader, relatively high-frequency scaling region with $S(q_0, \omega) \propto |\omega|^{-(1-\alpha)}$, for $\varepsilon^{\nu} < \omega t_{\text{micr}} < 1$. What are the chances of a successful detection?

The most specific prediction of the theory [3, 4] is the existence of the scaling region. It cannot be observed, however, unless we are sufficiently close to the critical point that $\varepsilon^{\nu} \ll 1$, or with $\nu \approx 1.3$, unless $\varepsilon \ll 1$. For the real control parameter, temperature, ε measures the relative distance from T_0 , the temperature where $\eta(T)$ diverges.

That seems to be a real difficulty. The vitrifying fluid is frozen in at a temperature T_g about 10% above T_0 , therefore structurally it cannot get sufficiently close to T_0 to make the scaling region observable.

We suggest a way out of the difficulty. The strength of mode coupling, as discussed in Sections 3 and 4, depends on the sharpness of the peaks of the static structure factor $S(q)$. On lowering the temperature, these peaks get sharper for two different reasons: structural relaxation, and decreasing amplitudes of vibrational motions [20]. Both can be enhanced to get closer to the critical point:

(i) Structural relaxation can be advanced by relaxing heat treatments at temperatures slightly below T_g . To avoid crystallization during heat treatment, at least three-component glasses (e.g. 50% SiO_2 , 20% MgO , 30% Al_2O_3 [24, 25] of which all components are good coherent scatterers) should be used. Natural boron, popular in glass-manufacturing, should be avoided because of its very high neutron capturing cross-section.

(ii) The vibrational width of the peaks of $S(q)$ can be reduced directly by doing the neutron scattering measurements at temperatures considerably below T_0 . Although a structure of higher temperature is frozen in, cooling will increase the strength of mode coupling and get the system closer to criticality.

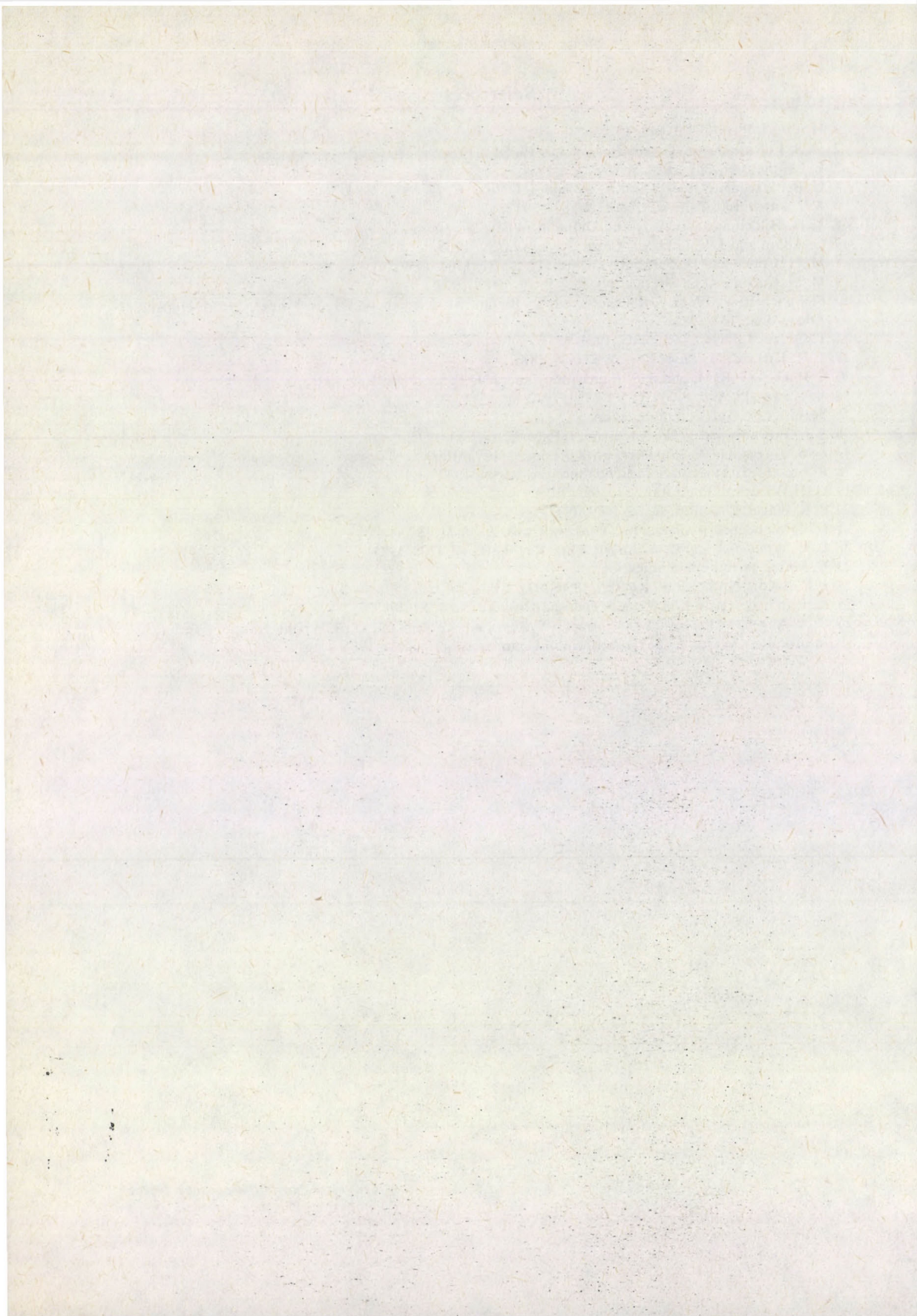
Finally, let us comment on the observability of the Lorentzian central peak. If the microscopic collision frequency is about 10^{-13} s and we get well into the critical region, the central peak should still reach out to GHz frequencies, easily detectable by high-resolution neutron inelastic scattering. Here, however, the theoretical prediction seems more doubtful than in the scaling region. The arguments given in Section 3 and [11] indicate that on the largest time scale the atomic motions may be controlled by the coupling to shear modes, which is missing in the Leutheusser [3] and Bengtzelius et al [4] theories. If this is true then the width of the central peak is determined by diffusion obeying the Stokes–Einstein law, Eq. (1), and a rough estimate for this width would be the “one-atomic-distance displacement per several hours” figure mentioned in the Introduction, viz. 10^{-4} Hz. Theoretical work to clarify this apparent contradiction is now in progress.

It seems more difficult to decide whether the whole homogeneous scenario of vitrification retains any predictive power in the close vicinity of T_0 , i.e. below the empirical glass transition temperature T_g . As a matter of fact, there is considerable evidence [26] that below T_g the temperature dependence of the shear viscosity is well described by a simple Arrhenius-type function, indicating the dominance of thermally activated defect motions and the decreasing importance of collective mode-coupling effects. The proper theoretical treatment of these data is another important task for future research.

The author is indebted to Dr. F. Mezei for inviting him to a Seminar for which the above review has been prepared, and to Dr. G. Bouquet for discussions giving valuable insights into real glasses.

References

1. H. Vogel, *Phys. Z.*, **22**, 645, 1921.
2. G. S. Fulcher, *J. Am. Ceram. Soc.*, **8**, 339, 1925.
3. E. Leutheusser, *Phys. Rev. A*, **29**, 2765, 1984.
4. U. Bengtzelius, W. Götze and A. Sjölander, *J. Phys. C*, **17**, 5915, 1984.
5. A. J. Batchinski, *Z. Phys. Chem.*, **84**, 643, 1913.
6. V. H. Hildebrand, *Viscosity and Diffusion*, Wiley, New York, 1977.
7. G. Adam and J. H. Gibbs, *J. Chem. Phys.*, **43**, 139, 1965.
8. M. H. Cohen and D. Turnbull, *J. Chem. Phys.*, **31**, 1164, 1959.
9. M. H. Cohen and G. Grest, *Phys. Rev. B*, **20**, 1077, 1979.
10. F. Spaepen, in: *Physics of Defects*, ed. R. Balian, M. Kléman and J. P. Poirier, North-Holland, Amsterdam, 133, 1981.
11. T. Geszti, *J. Phys. C*, **16**, 5805, 1983.
12. T. R. Kirkpatrick, *Phys. Rev. A*, **31**, 939, 1985.
13. S. P. Das, G. F. Mazenko, S. Ramaswamy and J. J. Toner, *Phys. Rev. Lett.*, **54**, 118, 1985.
14. K. Kawasaki, *Ann. Phys. (N.Y.)*, **61**, 1, 1970.
15. W. Götze, private communication.
16. S. F. Edwards and P. W. Anderson, *J. Phys. F*, **12**, 965, 1975.
17. K. S. Singwi and A. Sjölander, *Phys. Rev.*, **120**, 1093, 1960.
18. F. Mezei, ed., *Neutron Spin Echo*, Springer, Berlin, 1980.
19. G. H. Vineyard, *Phys. Rev.*, **110**, 999, 1958.
20. S. R. Nagel, *Phys. Rev. B*, **16**, 1694, 1977.
21. L. V. Woodcock, *J. Chem. Soc. Faraday Trans. 2*, **73**, 11, 1977.
22. L. V. Woodcock and C. A. Angell, *Phys. Rev. Lett.*, **47**, 1129, 1981.
23. F. Mezei, private communication.
24. J. F. Bacon, NASA Contract Rep. 1856, 1971.
25. K. I. Inagaki, Aichi-Ken Koguo Shidosho Hokoku No 9, 52, 1973.
26. C. A. Angell, in: *Relaxations in Complex Systems*, ed. K. Ngai and G. B. Wright, National Technical Information Service, U. S. Department of Commerce, Springfield, 1985.



FRAGMENT YIELDS IN THE MICROCANONICAL MODEL OF NUCLEAR DISASSEMBLY*

G. FAI**

*Department of Physics, Kent State University
Kent, OH 44242, USA*

Total relative fragment yields are calculated in the framework of the microcanonical model of nuclear disassembly. Results are compared to available data at intermediate energies (400 and 800 MeV/nucleon) and at beam energies at 32, 92 and 137 MeV/nucleon. The agreement of the calculated results with the experimental cross sections indicates that the microcanonical model of nuclear disassembly provides a satisfactory background calculation for the description of nuclear collisions.

1. Introduction

The primary goal for colliding nuclei at high energy is to learn about the properties of hot and dense hadronic matter. For example, it would be of great interest to observe consequences of phase transitions in hadronic matter (from nucleons to quarks and gluons at high enough energies [1] as well as the liquid-vapor transition at lower energies [2]). Phase transition-like phenomena can be embedded in the general context of statistical physics. From this perspective the experimental and theoretical efforts in the field of nuclear collisions can be viewed as a search for the equation of state of hadronic matter. Unfortunately, however, it is not possible to directly measure the thermodynamical properties of hot and dense hadronic matter. The most we can do is to measure the physical fragments in the final state of a nuclear collision. It is clear, therefore, that we have to rely on theoretical models when inferring the values of the quantities pertaining to the earlier stages of the time evolution of the system. Since the nuclear collisions under investigation are expected to be very complicated, the development of the necessary theoretical models is a difficult task. The first requirement for any theoretical model is of course that it reasonably reproduce available experimental information on fragment yields [3, 4].

By virtue of the complexity of the process, statistical considerations can be used when extrapolating back in time from the final state to the hot and dense state of hadronic matter. Statistical ideas were first applied to hadronic reactions by Fermi [5], who discussed pion production in high-energy proton-proton collisions. For relativistic nuclear collisions Mekjian [6] and others [7, 8] have developed statistical

* Dedicated to Prof. K. Nagy on his 60th birthday

** On leave from the R. Eötvös University, Budapest, Hungary

models for composite fragment production based on the grand canonical phase space. In the grand canonical ensemble conservation laws are satisfied on the mean, but not in each individual collision event. It should be noted that since in these collisions the system has "only" a few hundreds of degrees of freedom (as opposed to on the order of 10^{23} degrees of freedom for macroscopic systems), fluctuations may be large and the grand canonical approximation may not be adequate.

It was therefore necessary to develop a statistical model [9] based on the available microcanonical phase space. The model found a number of applications in the past couple of years. On the theory side it was successfully applied e.g. to estimate the effect of event to event fluctuations on exclusive observables [10], and to give a quantitative measure of the collective outward motion of fragments [11]. The investigation of equation of state and entropy questions is in progress [12, 13]. On the experimental front the model has proven to be a useful tool e.g. in the analysis of the bias introduced into the data by the very complex Plastic Ball/Wall detection system [14] and in the design of new equipment and experiments at Berkeley and Brookhaven [15, 16].

In addition to the apparent usefulness of the model I want to mention a more fundamental motivation for carrying out studies of nuclear collisions in the microcanonical framework: since the system has only a few hundred degrees of freedom (or sometimes much less) and the distributions are not extremely sharp, we clearly push our statistical ideas to the limit in these applications. We thereby get into a limiting situation (e.g. with respect to the rather delicate concept of the entropy), which is recognized as a potential source of new knowledge and insight.

Here I want to report on a systematic comparison of the results of the microcanonical model of nuclear disassembly [9] with experimental information on the fragment yields [3, 4]. It will be demonstrated that the model reproduces the fragment yield ratios reasonably well. This then lends credence to other results obtained with the use of the microcanonical model of nuclear disassembly. The rest of the paper is organized as follows: in Section 2 I briefly describe the model. Section 3 contains the comparison of the results to experimental data mentioned above. Finally, in Section 4 the results are discussed and some ideas are presented about how to improve the agreement further.

2. The model

The collision geometry plays an important role in nuclear reactions. Therefore, a statistical model needs to be augmented with a geometrical picture to obtain results that are relevant to the experimental situation. In the model geometrical ideas are used together with a minimum dynamical prescription to divide the collision system into a few subsystems. The subsystems defined in this manner are assumed to be independent. The independently disassembling subsystems will be called sources.

For the partitioning of the collision system into sources the model first invokes simple geometrical concepts that have proven their value for nuclear collisions at relativistic energies. The system is split into "participants" and "spectators" on the basis of the standard straight-trajectory clean-cut prescription associated with the nuclear fireball model [7]. Thus one participant source and up to two spectator sources are introduced. Next, the sharing of energy and momentum among the sources is decided. To partition the momentum among the sources the overall c.m. frame is utilized. In the c.m. frame of the collision

$$A_0 + B_0 \rightarrow A + B + C \quad (1)$$

the total momentum is zero:

$$\mathbf{P}_{A_0} + \mathbf{P}_{B_0} = 0. \quad (2)$$

Therefore the longitudinal momenta of the three sources are conveniently parametrized in the form

$$\mathbf{P}_A^1 = (1-y) \frac{A}{A_0} \mathbf{P}_{A_0}, \quad \mathbf{P}_B^1 = (1-y) \frac{B}{B_0} \mathbf{P}_{B_0}, \quad \mathbf{P}_C^1 = -\mathbf{P}_A^1 - \mathbf{P}_B^1. \quad (3)$$

Similarly, the transverse momenta are written as

$$\mathbf{P}'_A = z \frac{A}{A_0} \mathbf{P}_{A_0}, \quad \mathbf{P}'_B = z \frac{B}{B_0} \mathbf{P}_{B_0}, \quad \mathbf{P}'_C = -\mathbf{P}'_A - \mathbf{P}'_B. \quad (4)$$

The parameters y and z vary in the interval $[0, 1]$. The excitation energy generated (equal to the loss of translational energy implied by (3) and (4)) is parametrized with the help of a third parameter (denoted x). To do this the invariant source masses (including excitation) are written as

$$\begin{aligned} M_A c^2 &= A(mc^2 + xq) + V_A, \\ M_B c^2 &= B(mc^2 + xq) + V_B, \\ M_C c^2 &= C \left\{ mc^2 + \left[1 + \frac{A+B}{C} (1-x) \right] q \right\} + V_C. \end{aligned} \quad (5)$$

Here q is the excitation energy per nucleon. It has been assumed that the two spectators are excited in proportion to their mass ratio. The quantities V_A , V_B and V_C are ground state mass excesses and m denotes the nucleon mass. Demanding energy conservation then determines q in terms of x and thus gives the excitation energy per nucleon in the respective sources.

In general, the parameters x , y and z depend on the impact parameter. For central collisions (which provide the set of data the results of the present paper will be compared to) the values

$$x = x_0 = 0.05, \quad y = y_0 = 0.25, \quad z = z_0 = 0 \quad (6)$$

have been used in the calculation. These values were obtained from a fit to the Plastic Ball/Wall data [14].

For the description of the disassembly of the sources a random generation process is implemented which populates the available multifragment channels according to their statistical weight in the microcanonical ensemble. Sufficiently excited sources are assumed to explode quickly into a number of pions, nucleons and composite nuclei that are generally excited and particle unstable. Metastable products of the explosion are assumed to deexcite by sequential light particle evaporation.

The statistical generation of an event is based on the factorization property of the exclusive probability distribution proven in [9]. Factorization makes it possible to first select one fragment on the basis of the corresponding one-fragment distribution. Once a fragment has been selected with its baryon number, charge and four momentum, the values of these conserved quantities characterizing the remaining part of the source are also known. A subsequent random selection of a second fragment can then be made on the basis of the available information. This procedure is iterated until no residual system remains. (As shown in [9] this will necessarily happen after a finite number of steps in the generation sequence.)

The microcanonical one-fragment distributions appearing in the generation sequence are approximated with their truncated grand canonical equivalents. This approximation is essential in turning the method into a practical tool. The validity and detailed implementation of the method are discussed in [9]. Here I will only give a brief summary of the random selection procedure.

Once a source has been characterized by its baryon number, charge and four momentum the grand canonical constraint equations (expressing conservation of total energy, baryon number and charge on the mean) are solved. The corresponding Lagrange multipliers are the temperature parameter and the chemical potentials. Thus at each stage of the generation sequence the solution of three coupled equations yields the temperature and the chemical potentials characterizing the appropriate source in the grand canonical approximation. Since the logarithm of the grand partition function can be expressed as a computable sum over different fragment species [9], all required probabilities can be reconstructed with the given values of the Lagrange multipliers. On the basis of these probabilities a random selection of the fragment type, its intrinsic excitation and its momentum is made. The stability of the selected fragment against light particle emission is examined. If the state is metastable, the sequential light particle evaporation chain of the fragment is followed with a combination of traditional and random selection methods [9] until the daughter proves stable against light particle evaporation.

The above procedure yields a complete set of fragments produced in the final state of each event. These fragments are listed with their intrinsic excitation and momentum in the laboratory. The conservation laws taken into account are satisfied by the list of fragments pertaining to any individual event. A large number of events is generated. The final product of the procedure (which may appear on a magnetic tape

for example) is not unlike the data taken in a given experiment except that the theoretically generated events contain full information (e.g. neutral fragments are also present). Any desired quantity of experimental interest (total fragment yields for example) can of course easily be calculated from the theoretically generated "data".

3. Comparison to experimental cross sections

In this Section I will compare the results of the model described above to total experimental fragment cross sections from central collisions [3, 4]. Experimentally, the centrality of the collisions is assured by measuring at large angles [3], or by measuring at intermediate rapidity and applying different cuts to the data [4, 17]. In the present calculation centrality is effected by taking zero impact parameter, $b = 0$ fm. Each of the following Figures displays experimental and theoretical total fragment cross sections in barns. The order in which the different fragment species appear in the Figures simply follows the order of the experimental presentation. In all cases the calculated results are normalized to the data at the point representing the total deuteron cross section. First the comparison to experimental results of Nagamiya et al [3] will be discussed.

In Fig. 1 results for 800 MeV/nucleon Ar + Pb and Ar + KCl collisions are plotted. I choose this set of data from [3] as the subject of the first comparison because

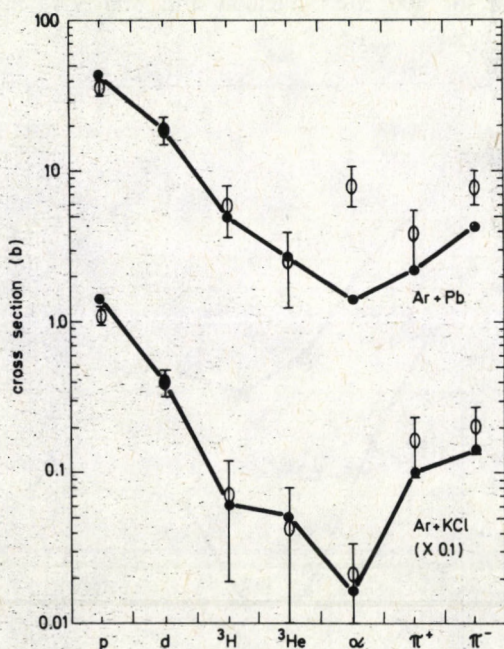


Fig. 1. Experimental (heavy dots) and theoretical (open symbols) total fragment cross sections (in barns) for Ar + Pb and Ar + KCl collisions at 800 MeV/nucleon beam energy. The Ar + KCl points have been scaled by a factor 0.1. The solid line connecting the experimental points (from [3]) are drawn to guide the eye. The error bars represent the statistical uncertainty associated with the sample of 200 theoretically generated events

it is here that the most detailed experimental information is available. The Ar + KCl results have been multiplied by 0.1 to make an efficient presentation (in one Figure) possible. Experimental points are indicated by heavy dots while the theoretical results are displayed as open symbols. The line connecting the experimental points is drawn (on this and later figures) to guide the eye. The theoretical points represent the ensemble average of 200 events. On this particular Figure I also show the statistical uncertainty associated with the sample of 200 events. Further calculations for other systems are also carried out on the basis of a sample of 200 events. The uncertainty is not plotted but comparable for all other theoretical calculations. It is seen that apart from an overestimate of the α -yield in Ar + Pb the agreement of the relative yields is quite satisfactory.

Fig. 2 displays the comparison between theory and experiment for 800 MeV/nucleon C+Pb and C+C collisions. Here (and in all remaining Figures) the same notation is used as in Fig. 1. In Fig. 3 Ne + Pb comparisons are shown at two different beam energies, 800 MeV/nucleon and 400 MeV/nucleon (*0.1), respectively. The trend of the calculation follows the data with the pion yield somewhat overestimated at both energies. Fig. 4 presents the Ne+Cu comparisons at 800 MeV/nucleon and 400 MeV/nucleon (*0.1) beam energies. The agreement is surprisingly good. In Fig. 5 calculated results for Ne+NaF are plotted together with the experimental data for 400 MeV/nucleon and 800 MeV/nucleon (*10) beam

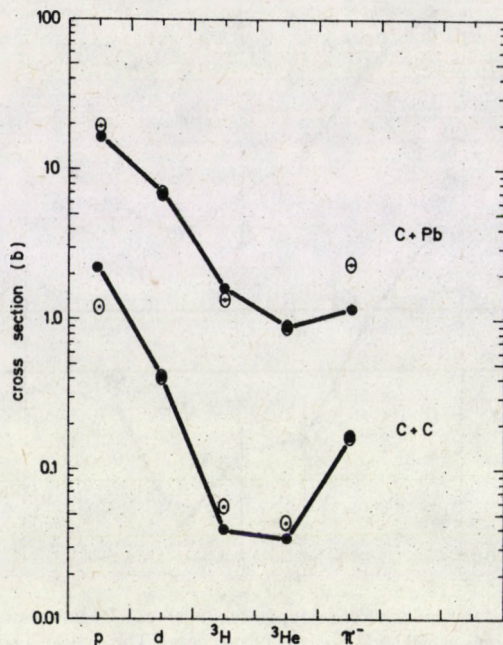


Fig. 2. The same as Fig. 1 for 800 MeV/nucleon C+Pb and C+C collisions. Data are from [3]

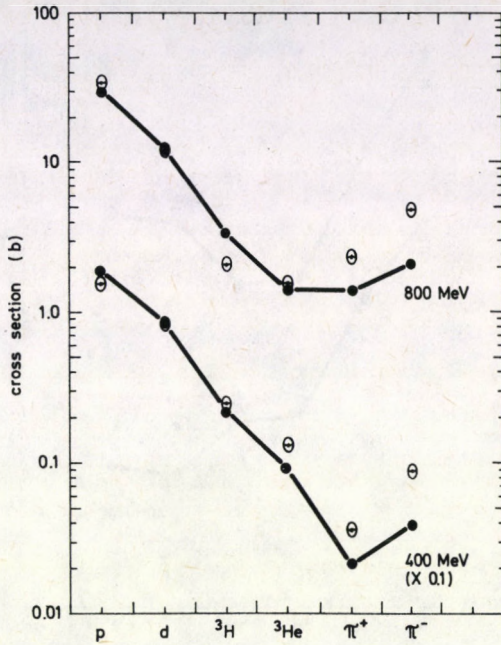


Fig. 3. The same as Fig. 1 for 800 MeV/nucleon and 400 MeV/nucleon Ne + Pb collisions. Data are from [3]

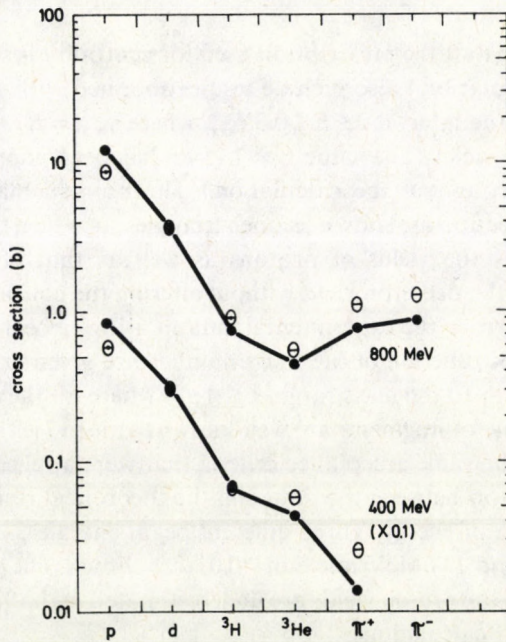


Fig. 4. The same as Fig. 1 for 800 MeV/nucleon and 400 MeV/nucleon Ne + Cu collisions. Data are from [3]

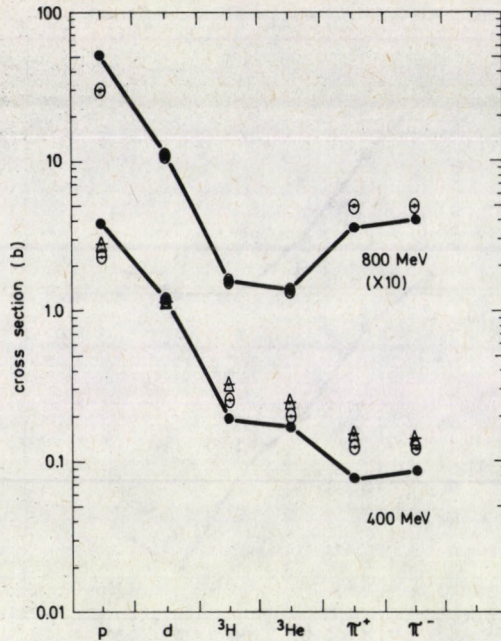


Fig. 5. The same as Fig. 1 for 800 MeV/nucleon and 400 MeV/nucleon Ne + NaF collisions. Open triangles indicate the theoretical results obtained with a randomly selected impact parameter in an interval representing central collisions. Data are from [3]

energies. Recognizing that the prescription used for centrality is somewhat arbitrary, for the 400 MeV calculation I also include results obtained with a randomly selected impact parameter in the interval $0 \leq b \leq 0.3b_{\max}$, where $b_{\max} = R_1 + R_2$, the sum of the radii of the colliding nuclei. (The value $r_0 = 1.15$ fm has been adopted for the nuclear radius parameter throughout the calculation.) The points obtained with random impact parameter selection are shown as open triangles. It is seen that such a modified prescription increases the yields of protons as well as that of heavier fragments somewhat relative to the deuteron yield without altering the conclusions qualitatively.

Let me now turn to the experimental data in a lower beam energy range [4]. Here relative yields as function of the mass number are given experimentally. I will restrict the comparison to the mass range $1 \leq A \leq 4$ where all the different thresholds and acceptance criteria for fragments are well-known to me [17]. The calculated results are filtered through the same acceptance criteria that were applied to the data. Fig. 6 displays the comparison between the data and the theoretical results for the Ar + Au collision as a function of energy. Three different beam energies, 42 MeV/nucleon, 92 MeV/nucleon (*0.1) and 137 MeV/nucleon (*0.01) are shown. The overall agreement is satisfactory, although the proton yields are underestimated in the 92 MeV/nucleon and the 137 MeV/nucleon calculations.

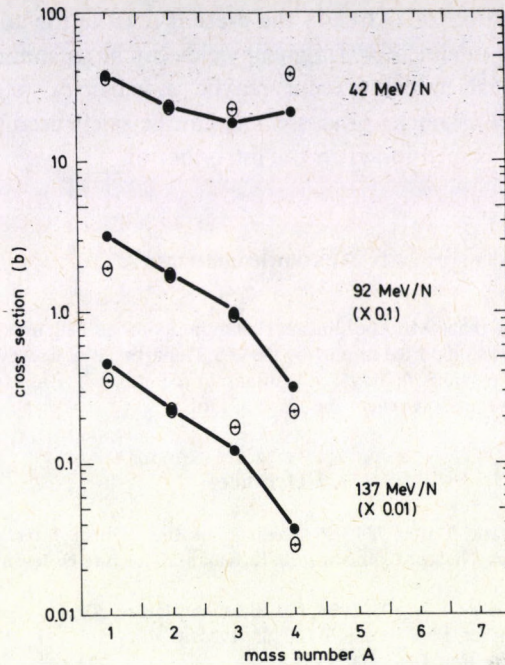


Fig. 6. Experimental and theoretical total fragment cross sections (in barns) as a function of the fragment mass number in Ar + Au collisions at three different beam energies. The experimental data are from [4]. The notation is the same as in Fig. 1

4. Discussion

I consider the agreement of the results of this simple model with the experimental data surprisingly good. There is a tendency to underestimate the proton yields and overestimate the pion yields in the calculation. Also, there is a warning sign concerning the α particle yield in at least one case (in Fig. 1). One possible reason for these disagreements is that Coulomb interactions between different fragments are neglected in the model. Another shortcoming of the calculation is related to the arbitrary definition of central collisions. Integrating up to a maximum impact parameter instead of taking strictly $b=0$ fm does influence the results somewhat as shown in Fig. 5. The energy and momentum transport parameters (6) also play a more important role in such a scenario. In particular, the impact parameter dependence of x , y and z provides some flexibility that could be used to fit the theoretical results to the data. (Note that the values of x , y and z at zero impact parameter are important for the present calculation since we consider both symmetric and asymmetric collisions.) Such a generalization on the other hand would introduce additional free parameters in the model. It is questionable whether such a simple model and the inclusive data considered warrant

further effort in this direction. Already the present level of the comparison seems to indicate that available total relative fragment yields can be reasonably well reproduced by a simple statistical model (except maybe α particle yields). Therefore the microcanonical model for nuclear disassembly can be safely used for the background and design calculations mentioned in the introduction.

5. Acknowledgements

The author wants to thank Mr. Abdunnasser F. Barghouty for his help in preparing the illustrations for this paper. This work was supported in part by the U.S. Department of Energy, Grant No. DE-FG02-86ER40251. However, any opinions, findings, conclusions, or recommendations expressed herein are those of the author and do not necessarily reflect the views of DoE.

References

1. See, for example, Quark Matter '84, Proceedings of the Fourth International Conference on Ultrarelativistic Nucleus-Nucleus Collisions, ed. K. Kajantie, Lecture Notes in Physics, Springer, New York, 1985.
2. A. S. Hirsch et al., Phys. Rev., *C29*, 508, 1984; P. J. Siemens, Nature, *305*, 410, 1983; G. Fai, L. P. Csernai, J. Randrup and H. Stöcker, Phys. Lett., *164B*, 265, 1985.
3. S. Nagamiya et al., Phys. Rev. *C24*, 971, 1981.
4. B. V. Jacak et al., Phys. Rev. Lett., *51*, 1846, 1983.
5. E. Fermi, Prog. Theor. Phys., *5*, 570, 1950.
6. A. Mekjian, Phys. Rev., *C17*, 1051, 1978.
7. J. Gosset, J. I. Kapusta and G. D. Westfall, Phys. Rev., *C18*, 844, 1978.
8. J. Randrup and S. E. Koonin, Nucl. Phys., *A356*, 223, 1981; G. Fai and J. Randrup, Nucl. Phys., *A381*, 557, 1982.
9. G. Fai and J. Randrup, Nucl. Phys., *A404*, 551, 1983.
10. L. P. Csernai, G. Fai and J. Randrup, Phys. Lett., *140B*, 149, 1984.
11. L. P. Csernai, J. Kapusta, G. Fai, D. Hahn, J. Randrup and H. Stöcker, Phys. Rev. *C35*, 1297, 1987.
12. G. Fai and E. Remler, in preparation.
13. L. P. Csernai, G. Fai and J. Kapusta, in preparation.
14. H.-A. Gustafsson et al., Phys. Rev. Lett., *52*, 1590, 1984.
15. D. Greiner, private communication.
16. H. Pugh, private communication.
17. G. D. Westfall, private communication.

A SYMBOLIC ALGORITHM FOR FINDING EXACTLY SOLUBLE STATISTICAL MECHANICAL MODELS*

P. RUJÁN**

*Institute for Solid State Research of the Nuclear Research Establishment Jülich
D-5170 Jülich, FRG****

In general it is a very difficult task to find statistical mechanical models which satisfy the Yang–Baxter equations and thus are completely integrable. We propose a new approach leading to a (overdetermined) set of *linear* equations. The formalism is applied to the Ising and the Ashkin–Teller models, which are both self-duals in two dimensions. Preliminary results for a symbolic algebra manipulation program is given, which would derive the relevant set of equations for an arbitrary internal spin symmetry group.

1. Introduction

“Everywhere is good but the best is home” says the proverb. I have not a few colleagues at the Institute for Theoretical Physics of the Eötvös University who at the moment feel good and remember with pleasure how was it at best. I think one of most important — and may be underestimated — contributions of Professor Károly Nagy was his encouragement of the relaxed but creative atmosphere in our Institute. He understood very early the potential of numerical calculations and helped us in having access to computing facilities. As a token of my appreciation I would like to present him this paper, an example of how computers do analytic work.

Although in recent years physicists became deeply involved in problems related to strongly non-analytic behaviour [1] it is still of high interest to find exactly soluble models which can help in establishing, for example, the critical behaviour of a whole universality class [2]. In two-dimensions the eight-vertex model solved first by Baxter [3] seems to be the primordial model incorporating almost all known fully integrable models [4]. There is also a class of “partially integrable” models [5], where only ground state information is available. By now there are excellent reviews available on the subject of completely integrable systems [6–9] and here we will only shortly summarize some facts used in our approach.

* Dedicated to Prof. K. Nagy on his 60th birthday

** On leave from: Institute for Theoretical Physics, Roland Eötvös University, Budapest

*** Address: Institut für Festkörperforschung der Kernforschungsanlage Jülich, D-5170 Jülich, BRD, Postfach 1913

Following the statistical mechanical line of thought we consider here two-dimensional models with nearest neighbour interactions between some $s_i = 0, 1, \dots, (q-1)$ q -state abelian variables (we call spins). Generalization to other variables, multispin interactions and higher dimensions will be discussed elsewhere. The partition function of a system consisting of M rows and N columns may be written in terms of the row-to-row transfer matrix $T_{\mathbf{s}, \mathbf{s}'}^{RR}$, "transferring" the interaction from one row to the next one as

$$Z_{M,N} = \sum_{\{\mathbf{s}^{(1)}\}} \sum_{\{\mathbf{s}^{(2)}\}} \dots \sum_{\{\mathbf{s}^{(M)}\}} \prod_{\langle \mathbf{s}^{(i)}, \mathbf{s}^{(i+1)} \rangle} e^{h(\mathbf{s}, \mathbf{s}')} = \text{Tr}(T^N), \quad (1)$$

$$T^{RR} = T_1^{1/2} T_2 T_1^{1/2}, \quad (2)$$

where $\mathbf{s}^{(i)}$ denotes the configurational space of a row ($\dim = q^N$), $\langle \mathbf{s}, \mathbf{s}' \rangle$ means indices involving two nearest neighbour rows. T_2 is a diagonal matrix (or "two-site" operator) whose elements contain the Boltzmann factor of the *intrarow* couplings while T_1 is the part ("one-site operator") which corresponds to *interrow* couplings. Because of the trace in (1) there are many possible forms for writing a transfer matrix and sometimes a symmetric form as (2) does not exist.

If we introduce the spin matrix algebra as [10]

$$\sigma_1 = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & & & \ddots & \\ & & & & 1 \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix}; \quad \sigma_3 = \begin{bmatrix} 1 & & & & \\ & \omega & & & \\ & & \omega^2 & & \\ & & & \ddots & \\ & & & & \omega^{q-1} \end{bmatrix} \quad (3)$$

$$\omega = e^{2\pi i/q}$$

satisfying the basic relations

$$\sigma_1^q = \sigma_3^q = \mathbf{1}; \quad \sigma_1 \sigma_3 = \omega \sigma_3 \sigma_1 \quad (4)$$

and

$$\sigma_3^1(m) = \mathbf{1} \otimes \dots \otimes \sigma_3^1 \otimes \dots \otimes \mathbf{1}, \quad (5)$$

where \otimes denotes a direct product and $\mathbf{1}$ a $q \times q$ unit-matrix, the transfer matrix (2) can be written in an operator form [11] as

$$T_1 = C^N \prod_{i=1}^N \exp \left[\sum_m K_m^*(y) (\sigma_1^m(i) + \text{h.c.}) \right], \quad (6)$$

$$T_2 = \prod_{i=1}^N \exp \left[\frac{1}{2} \sum_m K_m(x) (\sigma_3^m(i) \sigma_3^{+m}(i+1) + \text{h.c.}) \right]. \quad (7)$$

Here $K_m(x)$ and $K_m(y)$ ($m=1, \dots, \text{int} \left[\frac{q}{2} \right]$) are the Fourier components of the interaction along the x ("space") and y ("imaginary time") directions, respectively. K_m^* are the dual transforms which can be obtained as [11]

$$f_k^* = \sum_{m=0}^{q-1} f_m \omega^{km}; \quad f_m = \sum_{n=0}^{q-1} \exp(K_n \omega^{mn}). \quad (8)$$

The phase diagram of the Z_q model on a square lattice has been studied previously using different approximate methods [11–12]. From the point of view of complete integrability the case $q=2$ (Ising model) was solved first by Onsager [13], an event which marks the beginning of the modern theory of phase transition. The model is also exactly soluble on the Potts self-duality point [14] and for the antiferromagnetic Potts model on a square lattice [15] — these solutions were obtained using the mapping of the Potts model into a 6-vertex model [16]. A model which attracted much interest is the $q=4$ (Ashkin–Teller) model. Here the exactly soluble cases can also be obtained using a vertex representation [17] or the symmetries of the model [18]. A special solution was found for any q on the self-duality line [19], it is believed that this is a special bifurcation point of the phase diagram.

The difficulties involved in searching systematically for special soluble cases of such models are related to the fact that a generalized Bethe–Ansatz method is applicable only when there exist a family of commuting row-to-row transfer matrices [6]. The condition for the existence of a parameter v such that

$$[T(v), T(v')] = 0 \quad (9)$$

is called the Yang–Baxter generalised star–triangle transformation and corresponds in high energy physics language to the factorization of many particle scattering matrices into two-particle scattering processes. This is a set of highly overdetermined, non-linear equations which reduces at integrability due to some internal hidden symmetry [20]. In most cases it is an art to extract these solutions from the general equations (see [21] for a remarkable example) and as such, it remains the privilege of few.

The main goal of this paper is to reshape this problem in a form amenable to symbolic programming, that is, to transform art into manufacturing. The key concept of our approach will be the notion of an associated Hamiltonian.

2. The associated Hamiltonian

Eq. (9) involves the existence of an infinite set of conserved commuting charges [22] defined as

$$C_n = \frac{\partial^{n+1}}{\partial v^{n+1}} \ln T(v)|_{v=v_0}, \quad (10)$$

where

$$C_0 = H_L \quad (11)$$

is a "local" Hamilton operator. Note that T_1 and T_2 are exponentials of non-commuting operators and

$$T = e^{-vH_L} \quad (12)$$

would involve, in general, a highly non-local H_L operator. There is a limit, however, in which the form (12) can be obtained without recurring to (9). H corresponds to the extreme anisotropic limit $K_m(x) \rightarrow 0$; $K_m^*(y) \rightarrow 0$ but $K_m^*(y)/K_m(x) = \text{constant}$. This limit is known as the "Hamiltonian limit" [23] and interesting enough almost all local Hamiltonians obtained via Eqs (9–11) have a *structure* similar to the corresponding Hamiltonian limits. This is expected in view of universal critical behaviour and is an important ingredient of our approach. As discussed by Kasteleyn [6] the existence of an associated Hamiltonian is a necessary condition for a model to be completely integrable — we think it is also a sufficient one.

In our approach we will use extensively a second simplifying feature of the transfer matrix and its associated Hamiltonian, namely the duality transformation. As later it will become clear, this is not crucial but will help making our calculations more transparent. Consider, for instance, the Hamiltonian limit of the transfer matrix (2), (6), (7) which reads

$$H = - \left\{ \sum_i \sum_m K_m(x) [\sigma_3^m(i) \sigma_3^{+m}(i+1) + \text{h.c.}] + K_m^*(y) [\sigma_1^m(i) + \text{h.c.}] \right\}. \quad (13)$$

Now one introduces the bond operators [$i' = (i, i+1)$]

$$\begin{aligned} \mu_3(i') &= \sigma_3(i) \sigma_3^+(i+1), \\ \mu_1(i') &= \prod_{j \leq i} \sigma_1(j), \end{aligned} \quad (14)$$

which satisfy the same algebra as the original $\{\sigma\}$ operators. In terms of the new operators the sums in (6–7) and (13) are simply interchanged (some attention must be paid to the boundary conditions) but the part of the spectrum invariant under global spin Z_q and lattice dihedral symmetry is not changed [24]. This generates the usual self-duality relations for the Z_q models [11, 12].

Self-duality was used also to derive a simple condition for integrability [25–26]. It was shown that it is enough to require $[C_0, C_1]=0$ in order to construct an infinite set of charges. In this sense we believe that it is actually enough to require $[T, C_0]=0$, which involves less computational effort.

Our goal is to find the most general local Hamiltonian H whose couplings should be determined from the equation

$$[T, H]=0. \quad (15)$$

By locality we mean that H should be a general two-site operator of form

$$H = \sum_i h(i, i+1). \quad (16)$$

Let us define an intermediate Ansatz Hamiltonian as

$$H^A = T_1^{1/2} H T_1^{-1/2}. \quad (17)$$

First one makes the observation that since T_1 is a product of one-site operators H^A will remain also a local operator as long as H is one. Interesting enough if there exists a state $|\Phi\rangle$ such that $H^A|\Phi\rangle = \lambda|\Phi\rangle$, then the ground state of H is $T_1\left(\frac{K^*(y)}{2}\right)|\Phi\rangle$ which has a direct product form and corresponds to an order line solution [5]. Now in terms of the couplings (α, β, γ) entering in H^A Eq. (15) is satisfied if

$$T_1(-K^*(y))H^A(\alpha, \beta, \gamma)T_1(K^*(y)) = T_2(K(x))H^A(\alpha, \beta, \gamma)T_2(-K(x)). \quad (18)$$

Technically it is an easy matter to calculate a term like $T_1^{-1}H^AT_1$ but much more difficult to obtain $T_2H^AT_2^{-1}$. It is useful then to require that $H^A(\alpha, \beta, \gamma)$ is self-dual, so that one has

$$\begin{aligned} T_1(K^*(y))^D &= T_2(K^*(y)), \\ T_2(K(x))^D &= T_1(K(x)), \\ H^A(\alpha, \beta, \gamma)^D &= H^A(\beta, \alpha, -\gamma). \end{aligned} \quad (19)$$

With these notations the general Eq. (18) becomes

$$T_1(-K^*(y))H^A(\alpha, \beta, \gamma)T_1(K^*(y)) = [T_1(K(x))H^A(\beta, \alpha, -\gamma)T_1(-K(x))]^D \quad (20)$$

and the problem is reduced to moving a product of one-site operators over a sum of two-site operators. Note, however, that although the operator on the left-hand side is local, the one whose dual appears on the right side may contain in principle operators whose duals involve non-local strings (14). Therefore, we have first to satisfy the condition that the operator $[T_1(K(x))H^A(\beta, \alpha, -\gamma)T_1(-K(x))]^D$ is also dual. The resulting equations for the couplings (α, β, γ) are called compatibility conditions. If the compatibility conditions are met, Eq. (20) involves equating coefficients of two local operators with similar structure. We call these equations matching conditions. Once the

compatibility and the matching conditions are satisfied one may work with $H^A(\alpha, \beta, \gamma)$ (which commutes with the non-symmetric transfer matrix $T_1 T_2$) or one can calculate the local operator H from Eq. (17).

We now illustrate the method in two simple cases.

3. $q=2$. The Ising model

When $q=2$ the operators σ_1 and σ_3 become the usual Pauli matrices and

$$T_1 = \exp\left(K_y^* \sum_j \sigma_1(j)\right); \quad T_2 = \exp\left(K_x \sum_j \sigma_3(j)\sigma_3(j+1)\right). \quad (21)$$

The form of H^A suggested by the Hamiltonian limit will be

$$H^A = \sum_j h(j, j+1); \quad h(1, 2) = \frac{\alpha}{2}(\sigma_1(1) + \sigma_1(2)) + \beta\sigma_3(1)\sigma_3(2). \quad (22)$$

We add to this form a term obtained by combining the two basic expressions in (22), namely

$$h^A(1, 2) = h(1, 2) + \gamma(\sigma_1(1) + \sigma_1(2))\sigma_3(1)\sigma_3(2). \quad (23)$$

It is easy to convince oneself that this is the most general local and self-dual operator which satisfies Eq. (19). The only calculation one has now to do is to evaluate an expression of type

$$g(1, 2) = e^{-a(\sigma_1(1) + \sigma_1(2))} h^A(\alpha, \beta, \gamma) e^{a(\sigma_1(1) + \sigma_1(2))} \quad (24)$$

whose result can be used on both sides of Eq. (20). Using that $\sigma_1^2 = 1$ and $e^{a\sigma_1} = \cosh a + \sigma_1 \sinh a \equiv c + \sigma_1 \cdot s$ one has

$$\begin{aligned} g(1, 2) &= (c - \sigma_1(1)s)(c - \sigma_1(2)s) \cdot \left[\frac{\alpha}{2}(\sigma_1(1) + \sigma_1(2)) + \beta\sigma_3(1)\sigma_3(2) + \right. \\ &\quad \left. + \gamma(\sigma_1(1) + \sigma_1(2)) \right] (c + \sigma_1(1)s)(c + \sigma_1(2)s) = \\ &= \frac{\alpha}{2}[\sigma_1(1) + \sigma_1(2)] + [\beta c^2 - 2\gamma s c] \sigma_3(1)\sigma_3(2) + \\ &\quad + [-\beta s c + (c^2 + s^2)] (\sigma_1(1) + \sigma_1(2)) \sigma_3(1)\sigma_3(2) + \\ &\quad + [\beta s^2 - 2\gamma s \cdot c] \sigma_1(1)\sigma_1(2)\sigma_3(1)\sigma_3(2). \end{aligned} \quad (25)$$

Applying this formula to the left side of Eq. (20) one has to identify $\bar{s} = \sinh 2K_y^*$, $\bar{c} = \cosh 2K_y^*$, while when using it under the duality sign on the right side $\alpha \rightarrow \beta$, $\beta \rightarrow \alpha$, $\gamma \rightarrow -\gamma$, $s = \sinh 2K_x$, $c = \cosh 2K_x$, $\bar{s} \rightarrow -s$.

The compatibility conditions are therefore

$$\beta \bar{s}^2 - 2\gamma \bar{s} \cdot \bar{c} = 0 \quad \text{on the left side,} \quad (26a)$$

$$\alpha s^2 - 2\gamma s \cdot c = 0 \quad \text{on the right side,} \quad (26b)$$

leading to the determination of $\gamma(\alpha, \beta)$ (the trivial solutions $\sinh 2K_y^* = 0$ and $\sinh 2K_x = 0$ are disregarded for the moment)

$$\beta = 2\gamma \bar{c} / \bar{s}, \quad (27a)$$

$$\alpha = 2\gamma c / s. \quad (27b)$$

On the right side one can now perform one more duality transformation and one obtains finally the *matching* conditions

$$\beta = \beta \bar{c}^2 - 2\gamma \bar{s} \cdot \bar{c} \equiv \beta \quad \text{from (27),} \quad (28a)$$

$$\alpha = \alpha c^2 - 2\gamma s \cdot c \equiv \alpha \quad \text{from (27),} \quad (28b)$$

$$-\alpha s \cdot c + \gamma(c^2 + s^2) = -\beta \bar{s} \cdot \bar{c} + \gamma(\bar{c}^2 + \bar{s}^2), \quad (28c)$$

leading to

$$\gamma = \gamma$$

if (27) is satisfied. As we will also see in the following Section the matching conditions are automatically satisfied if the equations are *compatible*. Actually, the matching conditions give information only on how to pair the possible solutions of (27). Eq. (28c) is fulfilled only by pairing the solution $K_x = 0$ (*decoupling in 1D chains*) to $K_y^* = 0$ ($K_y = \infty$, the system becomes a *single 1D chain*), and (27a) to (27b), respectively. Note that for ($K_x = 0, K_y^* = 0$) $T \sim 1$ and commutes of course with any H . Expanding in first order $T \simeq 1 + K_x \cdot H$ as in the Hamiltonian limit. It gives correct commutation relations to $O(K_x^2)$.

The final results are

$$H^A = \gamma \sum_j \{ -(2 \cotanh 2K_x) \sigma_1(j) - (2 \cotanh 2K_y^*) \sigma_3(j) \sigma_3(j+1) + (\sigma_1(j) + \sigma_1(j+1)) \sigma_3(j) \sigma_3(j+1) \} \quad (29)$$

and

$$H = \sum_j \{ (1+c) \sigma_1(j) \sigma_1(j+1) + (1-c) \sigma_2(j) \sigma_2(j+1) + 2h \sigma_3(j) \}, \quad (30)$$

$c = 2 \cosh 2K_y^*$; $h = \tanh 2K_x \cotanh 2K_y^*$, a result obtained originally by Suzuki [27]. It is interesting to notice that the commutation relation can be satisfied with real K_x, K_y only *outside* the circle

$$c^2 + h^2 = 1, \quad (31)$$

which is the equation of an order line [28].

Furthermore, H is dual to a Hamiltonian operator commuting with the diagonal-to-diagonal transfer matrix of a *triangular* Ising model [29]. A big advantage of the diagonal-to-diagonal transfer matrix is that K_x and K_y enter symmetrically and the corresponding Hamiltonian is (K_x, K_y, K_z are now the couplings along the three axis of the triangular lattice)

$$H = \sum_j [\mu_1(j) + k_1 \mu_1(j) \mu_1(j+1) + k_2 \mu_3(j-1) \mu_1(j) \mu_3(j+1)], \quad (32)$$

$$k_1 (\cosh K_z)^2 = \sinh 2K_x \sinh 2K_y \cosh 2K_z + \cosh 2K_x \cosh 2K_y \sinh 2K_z,$$

$$k_2 (\cosh K_z)^2 = -(\sinh K_z)^2.$$

$K_z = 0$ corresponds to the boundary (31). The form (32) is especially useful since k_1 and k_2 appear as in the parametrization of the star-triangle transformation [6] and make the choice of v (Eq. (10)) obvious. It will be therefore of some interest to extend the present method to diagonal-to-diagonal transfer matrices.

4. $q = 4$. The Ashkin-Teller model

The next example we consider is the Ashkin-Teller model, or Z_4 model [30]. Since the spin variables take now on the values 0, 1, 2, and 3 they can be represented by two Ising variables per lattice site [31]. The energy function of the two-dimensional model on a square lattice has then the parametrization

$$\begin{aligned} -H = \sum_{x,y} & K_1(x) s_{x,y} s_{x+1,y} + K_2(x) t_{x,y} t_{x+1,y} + K_4(x) s_{x,y} t_{x,y} s_{x+1,y} t_{x+1,y} + \\ & + K_1(y) s_{x,y} s_{x,y+1} + K_2(y) t_{x,y} t_{x,y+1} + K_4(y) s_{x,y} t_{x,y} s_{x,y+1} t_{x,y+1}, \end{aligned} \quad (33)$$

where $s_{x,y}, t_{x,y} = \pm 1$. The one-dimensional transfer matrix can be written as

$$a = e^{K_1 + K_2 + K_4} e^{A + K_1^* \sigma^x + K_2^* \tau^x + K_4^* \sigma^x \tau^x}, \quad (34)$$

where

$$A = \frac{1}{4} \ln \Sigma_1 \Sigma_2 \Sigma_3 \Sigma_4,$$

$$K_1^* = \frac{1}{4} \ln \frac{\Sigma_1 \Sigma_2}{\Sigma_3 \Sigma_4}; \quad K_2^* = \frac{1}{4} \ln \frac{\Sigma_1 \Sigma_3}{\Sigma_2 \Sigma_4}; \quad K_4^* = \frac{1}{4} \ln \frac{\Sigma_1 \Sigma_4}{\Sigma_2 \Sigma_3};$$

$$\Sigma_1 = 1 + \gamma_1 + \gamma_2 + \gamma_3;$$

$$\Sigma_2 = 1 - \gamma_1 - \gamma_2 + \gamma_3; \quad \gamma_1 = e^{-2(K_2 + K_4)},$$

$$\Sigma_3 = 1 + \gamma_1 - \gamma_2 - \gamma_3; \quad \gamma_2 = e^{-2(K_1 + K_4)},$$

$$\Sigma_4 = 1 - \gamma_1 + \gamma_2 - \gamma_3; \quad \gamma_3 = e^{-2(K_1 + K_2)}. \quad (35)$$

Since the T_1 transfer matrix is the direct product of one-dimensional transfer matrices its form will be (after omitting the constant factor)

$$T_1(\mathbf{K}^*(y)) \equiv \exp \sum_j (K_1^*(y) \sigma_j^x + K_2^*(y) \tau_j^x + K_4^*(y) \sigma_j^x \tau_j^x), \quad (36a)$$

$$T_2(\mathbf{K}(x)) \equiv \exp \sum_j (K_1(x) \sigma_j^z \sigma_{j+1}^z + K_2(x) \tau_j^z \tau_{j+1}^z + K_4(x) \sigma_j^z \sigma_{j+1}^z \tau_j^z \tau_{j+1}^z), \quad (36b)$$

where σ_j^x , σ_j^z and τ_j^x , τ_j^z are Pauli operators acting on the s_j and the t_j spins, respectively.

Starting from the Hamiltonian limit form and adding all possible self-dual local terms one finally obtains the Ansatz Hamiltonian $H^A(\alpha, \beta, \gamma)$ (17) as

$$\begin{aligned} H^A(\alpha, \beta, \gamma) &= \sum_j h_{j,j+1}^A(\alpha, \beta, \gamma), \\ h_{1,2}^A &= \alpha_1 \sigma_1^x + \beta_1 \sigma_1^z \sigma_2^z + \gamma_1 (\sigma_1^x + \sigma_2^x) \sigma_1^z \sigma_2^z + \\ &+ \alpha_2 \tau_1^x + \beta_2 \tau_1^z \tau_2^z + \gamma_2 (\tau_1^x + \tau_2^x) \tau_1^z \tau_2^z + \\ &+ \alpha_3 \sigma_1^x \tau_1^x + \beta_3 \sigma_1^z \sigma_2^z \tau_1^z \tau_2^z + \gamma_3 (\sigma_1^x \tau_1^x + \sigma_2^x \tau_2^x) \sigma_1^z \sigma_2^z \tau_1^z \tau_2^z + \\ &+ \alpha_4 (\sigma_1^x + \sigma_2^x) \tau_1^z \tau_2^z + \beta_4 (\tau_1^z + \tau_2^z) \sigma_1^z \sigma_2^z + \\ &+ \alpha_5 (\sigma_1^x \tau_1^x + \sigma_2^x \tau_2^x) \sigma_1^z \sigma_2^z + \beta_5 (\sigma_1^z + \sigma_2^z) \sigma_1^z \sigma_2^z \tau_1^z \tau_2^z + \\ &+ \alpha_6 (\tau_1^x \sigma_1^x + \sigma_2^x \tau_2^x) \tau_1^z \tau_2^z + \beta_6 (\tau_1^z + \tau_2^z) \sigma_1^z \sigma_2^z \tau_1^z \tau_2^z. \end{aligned} \quad (37)$$

Introduce the notations

$$\mathbf{K}^* = (K_1^*(y), K_2^*(y), K_4^*(y)),$$

$$\mathbf{K} = (K_1(x), K_2(x), K_4(x)),$$

$$\alpha = (\alpha_1 \dots \alpha_6),$$

$$\beta = (\beta_1 \dots \beta_6),$$

$$\gamma = (\gamma_1, \gamma_2).$$

The duality transformation then acts as

$$T_1(\mathbf{K})^D = T_1^2(\mathbf{K}), \quad (38)$$

$$H^A(\alpha, \beta, \gamma, \gamma_3)^D = H^A(\beta, \alpha, -\gamma, \gamma_3).$$

Now we apply the method of Section 2 to obtain the *compatibility* and the *matching conditions*.

The "only" calculation one has to do is again similar to (24):

$$g(1, 2) = \bar{a}_1^1(\mathbf{a}) \bar{a}_2^1(\mathbf{a}) h_{1,2}^A(\alpha, \beta, \gamma, \gamma_3) a_2(\mathbf{a}) a_1(\mathbf{a}). \quad (39)$$

This is a rather lengthy but straightforward calculation leading to the following compatibility conditions:

for the left-side of Eq. (18):

$$\gamma_1 = \frac{\bar{v}_1}{2} \beta_1 \quad \text{or} \quad \bar{v}_4^2 = 1 \quad \text{or} \quad \bar{v}_1 = 0, \quad (40a)$$

$$\alpha_5 = \frac{\bar{v}_4}{2} \beta_1 \quad \text{or} \quad \bar{v}_1^2 = 1 \quad \text{or} \quad \bar{v}_4 = 0, \quad (40b)$$

$$\beta_4 = \frac{v_1 v_4}{2} \beta_1 \quad \text{or} \quad \bar{v}_4 = 0 \quad \text{or} \quad \bar{v}_1 = 0, \quad (40c)$$

$$\gamma_2 = \frac{\bar{v}_2}{2} \beta_2 \quad \text{or} \quad \bar{v}_4^2 = 1 \quad \text{or} \quad \bar{v}_2 = 0, \quad (40d)$$

$$\alpha_6 = \frac{\bar{v}_4}{2} \beta_2 \quad \text{or} \quad \bar{v}_2^2 = 1 \quad \text{or} \quad \bar{v}_4 = 0, \quad (40e)$$

$$\alpha_4 = \frac{\bar{v}_2 \bar{v}_4}{2} \beta_2 \quad \text{or} \quad \bar{v}_2 = 0 \quad \text{or} \quad \bar{v}_4 = 0, \quad (40f)$$

$$\beta_5 = \frac{\bar{v}_1}{2} \beta_3 \quad \text{or} \quad \bar{v}_2^2 = 1 \quad \text{or} \quad \bar{v}_1 = 0, \quad (40g)$$

$$\beta_6 = \frac{\bar{v}_2}{2} \beta_3 \quad \text{or} \quad \bar{v}_1^2 = 1 \quad \text{or} \quad \bar{v}_2 = 0, \quad (40h)$$

$$\gamma_3 = \frac{\bar{v}_1 \bar{v}_2}{2} \beta_3 \quad \text{or} \quad \bar{v}_1 = 0 \quad \text{or} \quad \bar{v}_2 = 0, \quad (40i)$$

where $\bar{v}_i = \tanh 2K_i^*(y)$ and or = and/or.

For the right side of Eq. (18) one has similar conditions, except for the substitutions

$$\alpha \rightarrow \beta; \quad \gamma \rightarrow -\gamma; \quad \gamma_3 \rightarrow \gamma_3; \quad \bar{v} \rightarrow -v,$$

($v_i = \tanh 2K_i(x)$).

Next one derives the *matching* conditions. After long calculations one obtains the same result as in the Ising model, namely that once the compatibility equations are satisfied, the matching conditions become identities.

After solving the *linear* equations for the parameters ($\alpha, \beta, \gamma, \gamma_3$) only three *independent* equations remain (we consider here only the first row of Eqs (40) — the decoupling solutions given by row 2 or 3 are not discussed).

$$v_1^2 = \bar{v}_2^2, \quad (41a)$$

$$v_4^2 = \bar{v}_4^2, \quad (41b)$$

$$v_1 \bar{v}_1 = v_2 \bar{v}_2 \cdot \frac{\bar{v}_4}{v_4}. \quad (41c)$$

In terms of the original couplings the four possible solutions can be represented as

$$K_1^*(y) = \eta_1 K_2(x), \quad (42a)$$

$$K_4^*(y) = \eta_2 K_4(x), \quad (42b)$$

$$K_2^*(y) = \eta_3 K_1(x), \quad (42c)$$

where $\eta_i = \pm 1$ and $\eta_1 \eta_2 \eta_3 = 1$.

These are the solutions found also in [18] for the case $K_1 = K_2$ (symmetric Ashkin-Teller model). I have also calculated the form of the Hamiltonian (16) which is a sum of 15 different local operators.

5. A Reduce algorithm—preliminary results

From the two examples presented above it is clear that the method involves a rather lengthy but straightforward computing step once the form of H^A has been properly established. This step corresponds to Eq. (24) for $q=2$ and to Eq. (39) for $q=4$. We propose to use a symbolic algebraic language — as for example Reduce [32] — to carry out this step. In a later stage the program might be further developed to solve also the system of ensuing linear equations.

In order to give the following statements a proper weight, the author's inexperience with Reduce should be first emphasized. I took the most orthodox user's approach and after reading [32] — but not the further references indicated therein — I set up a simple program which deals with the (σ_1, σ_3) algebra. I defined sig 1, sig 3 non-commuting operators and $(\text{sig } 1)^2 = (\text{sig } 2)^2 = 1$. The first problem arises when one would like to define the operator algebra as

$$\sigma_1 \sigma_3 = -\sigma_3 \sigma_1, \quad (43)$$

rather than

$$\sigma_i \sigma_j = i \varepsilon_{ijk} \sigma_k. \quad (44)$$

It turns out that Reduce is not very effective in making substitutions involving three operators, as for example,

$$\sigma_1 \sigma_3 \sigma_1 = -\sigma_3, \quad (45a)$$

$$\sigma_3 \sigma_1 \sigma_3 = -\sigma_1. \quad (45b)$$

It seems far better to define a "canonical operator form" (which I choose to be $\sigma_1 \sigma_3$) on which no substitution attempt should be made. This approach works nicely and as shown in Protocol 1 (Appendix 1) a full multiplication Table can be generated for a

basis of the two-site operator algebra. Defining as generators of the two-site operator basis

$$\hat{s}_1 = \sigma_1(1) + \sigma_1(2), \quad (46a)$$

$$\hat{s}_2 = \sigma_1(1) \cdot \sigma_1(2), \quad (46b)$$

$$\hat{s}_3 = \sigma_3(1) \cdot \sigma_3(2), \quad (46c)$$

and closing this algebra by adding

$$\hat{s}_4 = \hat{s}_1 \cdot \hat{s}_3, \quad (46d)$$

$$\hat{s}_5 = \hat{s}_2 \cdot \hat{s}_3, \quad (46e)$$

one gets the multiplication table:

		second factor				
		s_1	s_2	s_3	s_4	s_5
first factor	s_1	$2(1 + s_2)$	s_1	s_4	$2(s_3 + s_5)$	s_4
	s_2	s_1	1	s_5	s_4	s_3
	s_3	$-s_4$	s_5	1	$-s_1$	s_2
	s_4	$-2(s_3 + s_5)$	s_4	s_1	$-2(1 + s_2)$	s_1
	s_5	$-s_4$	s_3	s_2	$-s_1$	1

Protocol 2 (Appendix 2) shows the derivation of Eq. (25).

In conclusion we remark that although in principle very simple, a calculation of this style is not easy to implement. However, we hope that by a better understanding of the way Reduce operates or by using a more "user friendly" symbolic manipulation language, a general algorithm dealing also with other types of algebras will become soon operational.

Appendix 1

>>>>>

THIS IS THE FIRST PROTOCOL

<<<<<<<

COMMENT THIS IS PART OF THE PROJECT COMM(H,T)=0 AND APPLIES TO THE ISING MODEL. IN THIS FIRST PART A DIRECT APPROACH IS USED BASED ON THE CANONICAL FORM SIG I (K) * SIG J (N) IS FIRST ACCORDING TO INDEX I<J AND THEN ACCORDING TO INDEX K<N, I,J,K,N=1,2. A MULTIPLICATION TABLE IS GENERATED FOR A FULL BASIS OF TWO SITE OPERATOR H(1,2);

COMMENT DEFINITION OF A FULL BASIS FOR THE ALGEBRA;

```
OPERATOR S;
NONCOM S(K);
LET S(1)=SIG1(1),
    S(2)=SIG1(2),
    S(3)=SIG1(1)*SIG1(2),
    S(4)=SIG3(1)*SIG3(2),
    S(5)=S(1)*S(4),
    S(6)=S(2)*S(4),
    S(7)=S(3)*S(4);
```

COMMENT DEFINITION OF THE COMMUTATION RULES;

```
OPERATOR SIG1, SIG3;
NONCOM SIG1, SIG3;
FOR ALL K, J SUCH THAT J>K LET SIG1(J)*SIG1(K)=SIG1(K)*SIG1(J),
                                SIG3(J)*SIG3(K)=SIG3(K)*SIG3(J),
                                SIG3(J)*SIG1(K)=SIG1(K)*SIG3(J),
                                SIG3(K)*SIG1(J)=SIG1(J)*SIG3(K);
```

```
FOR ALL K LET
                                SIG1(K)*SIG1(K)=1,
                                SIG3(K)*SIG3(K)=1,
                                SIG3(K)*SIG1(K)=-SIG1(K)*SIG3(K);
```

COMMENT CHECKING THE ALGEBRA AND CREATING THE MULTIPLICATION TABLE;

COMMENT NOTE THE FOLLOWING PECULIARITY (!!??);

```
AA:= SIG1(1)*SIG1(1)*SIG1(2)*SIG3(1)*SIG3(2);
AA := SIG1(2)*SIG3(1)*SIG3(2)
```

```
AB:= SIG1(2)*SIG1(1)*SIG1(2)*SIG3(1)*SIG3(2);
AB := SIG1(1)*SIG3(1)*SIG3(2)
```

```
AA+AB;
SIG1(2)*SIG3(1)*SIG3(2) + SIG1(1)*SIG3(1)*SIG3(2)
```

COMMENT BUT!!!!!!;

```
(SIG1(1)+SIG1(2))*SIG1(1)*SIG1(2)*SIG3(1)*SIG3(2);
- SIG1(2)*SIG3(1)*SIG3(2) + SIG1(1)*SIG3(1)*SIG3(2)
```

```
ARRAY TAB(7,7);
FOR K:= 1:7 DO <<TAB(1,K):=S(1)*S(K)>>;
FOR K:= 1:7 DO <<WRITE TAB(1,K)>>;
```



```

1
SIG1(1)*SIG1(2)
SIG1(2)
SIG1(1)*SIG3(1)*SIG3(2)
SIG3(1)*SIG3(2)
SIG1(1)*SIG1(2)*SIG3(1)*SIG3(2)
SIG1(2)*SIG3(1)*SIG3(2)

FOR K:= 1:7 DO <<TAB(2,K):=S(2)*S(K)>>;
FOR K:= 1:7 DO <<WRITE TAB(2,K)>>;

```

```

SIG1(1)*SIG1(2)
1
SIG1(1)
SIG1(2)*SIG3(1)*SIG3(2)
SIG1(1)*SIG1(2)*SIG3(1)*SIG3(2)
SIG3(1)*SIG3(2)
SIG1(1)*SIG3(1)*SIG3(2)

FOR K:= 1:7 DO <<TAB(3,K):=S(3)*S(K)>>;
FOR K:= 1:7 DO <<WRITE TAB(3,K)>>;

```

```

SIG1(2)
SIG1(1)
1
SIG1(1)*SIG1(2)*SIG3(1)*SIG3(2)
SIG1(2)*SIG3(1)*SIG3(2)
SIG1(1)*SIG3(1)*SIG3(2)
SIG3(1)*SIG3(2)

FOR K:= 1:7 DO <<TAB(4,K):=S(4)*S(K)>>;
FOR K:= 1:7 DO <<WRITE TAB(4,K)>>;

```

```

- SIG1(1)*SIG3(1)*SIG3(2)
- SIG1(2)*SIG3(1)*SIG3(2)
SIG1(1)*SIG1(2)*SIG3(1)*SIG3(2)
1
- SIG1(1)
- SIG1(2)
SIG1(1)*SIG1(2)

FOR K:= 1:7 DO <<TAB(5,K):=S(5)*S(K)>>;
FOR K:= 1:7 DO <<WRITE TAB(5,K)>>;

```



```

- SIG3(1)*SIG3(2)
- SIG1(1)*SIG1(2)*SIG3(1)*SIG3(2)
SIG1(2)*SIG3(1)*SIG3(2)
SIG1(1)
(-1)
- SIG1(1)*SIG1(2)
SIG1(2)

```

```

FOR K:= 1:7 DO <<TAB(6,K):=S(6)*S(K)>>;
FOR K:= 1:7 DO <<WRITE TAB(6,K)>>;

```

```

- SIG1(1)*SIG1(2)*SIG3(1)*SIG3(2)
- SIG3(1)*SIG3(2)

```

```

SIG1(1)*SIG3(1)*SIG3(2)
SIG1(2)
- SIG1(1)*SIG1(2)
(-1)
SIG1(1)

```

```

FOR K:= 1:7 DO <<TAB(7,K):=S(7)*S(K)>>;
FOR K:= 1:7 DO <<WRITE TAB(7,K)>>;

```

```

- SIG1(2)*SIG3(1)*SIG3(2)
- SIG1(1)*SIG3(1)*SIG3(2)
SIG3(1)*SIG3(2)
SIG1(1)*SIG1(2)
- SIG1(2)
- SIG1(1)
1

```

```

OPERATOR R;
NONCOM R(X);
LET

```

```

    SIG3(1)*SIG3(2)=R(4),
    SIG1(2)=R(2),
    SIG1(1)=R(1);

```

COMMENT ACCORDING TO THE CANONICAL FORM ONE SHOULD HAVE
 ALWAYS PRODUCTS AS $R(I)*R(4)$, $X=1,2,3$. HOWEVER, I
 GET ALWAYS THE INVERSED ORDER (!??)
 SO THE FOLLOWING STATEMENTS ARE INVERSED ACCORDINGLY;


```
LET R(4)*R(3)=R(7),  
    R(4)*R(2)=R(6),  
    R(4)*R(1)=R(5),  
    R(2)*R(1)=R(3);
```

```
FOR K:= 1:7 DO <<WRITE TAB(1,K)>>;
```

```
1  
R(3)  
R(2)  
R(5)  
R(4)  
R(7)  
R(6)
```

```
FOR K:= 1:7 DO <<WRITE TAB(2,K)>>;
```

```
R(3)  
1  
R(1)  
R(6)  
R(7)  
R(4)  
R(5)
```

```
FOR K:= 1:7 DO <<WRITE TAB(3,K)>>;
```

```
R(2)  
R(1)
```

```
1  
R(7)  
R(6)  
R(5)  
R(4)
```

```
FOR K:= 1:7 DO <<WRITE TAB(4,K)>>;
```

```
- R(5)  
- R(6)  
R(7)  
1  
- R(1)  
- R(2)  
R(3)
```


FOR K:= 1:7 DO <<WRITE TAB(5,K)>>;

- R(4)
 - R(7)
 R(6)
 R(1)
 (-1)
 - R(3)
 R(2)

FOR K:= 1:7 DO <<WRITE TAB(6,K)>>;

- R(7)
 - R(4)
 R(5)
 R(2)
 - R(3)
 (-1)
 R(1)

FOR K:= 1:7 DO <<WRITE TAB(7,K)>>;

- R(6)
 - R(5)
 R(4)
 R(3)
 - R(2)
 - R(1)
 1

COMMENT THIS WAY ONE CAN DEFINE A FULL MULTIPLICATION TABLE
 WHICH CAN BE USED TO SIMPLIFY THE FINAL RESULT AS
 SHOWN IN THE NEXT PROTOCOL;

Appendix 2

>>>>>>> SECOND PROTOCOL: THE ACTUAL CALCULATION <<<<<<<<

COMMENT THIS PROGRAM IS PART OF THE PROJECT COMM(H,T)=0.
 IN THE SECOND PART WE USE THE GROUP TABLE FOR THE ALGEBRA
 OF THE TWO-SITE OPERATORS GENERATED IN THE PREVIOUS STEP.
 NOTE THAT TAKING INTO ACCOUNT THE $H(1,2)=H(2,1)$ SYMMETRY
 HAS REDUCED THE BASIS FROM 7 TO 5 ELEMENTS.;

OPERATOR S1,S2,S3,S4,S5;
 NONCOM S1,S2,S3,S4,S5;
 KORDER S1,S2,S3,S4,S5;
 FACTOR S1,S2,S3,S4,S5,B,C;

COMMENT THE DEFINITION OF THE REDUCTION TABLE;

LET

$S1(X)*S1(X)=2*(1+S2(X))$,
 $S1(X)*S2(X)=S1(X)$,
 $S1(X)*S3(X)=S4(X)$,
 $S1(X)*S4(X)=2*(S3(X)+S5(X))$,
 $S1(X)*S5(X)=S4(X)$,

$S2(X)*S1(X)=S1(X)$,
 $S2(X)*S2(X)=1$,
 $S2(X)*S3(X)=S5(X)$,
 $S2(X)*S4(X)=S4(X)$,
 $S2(X)*S5(X)=S3(X)$,

$S3(X)*S1(X)=-S4(X)$,
 $S3(X)*S2(X)=S5(X)$,
 $S3(X)*S3(X)=1$,
 $S3(X)*S4(X)=-S1(X)$,
 $S3(X)*S5(X)=S2(X)$,

$S4(X)*S1(X)=-2*(S3(X)+S5(X))$,
 $S4(X)*S2(X)=S4(X)$,
 $S4(X)*S3(X)=S1(X)$,
 $S4(X)*S4(X)=-2*(1+S2(X))$,
 $S4(X)*S5(X)=S1(X)$,

$S5(X)*S1(X)=-S4(X)$,
 $S5(X)*S2(X)=S3(X)$,
 $S5(X)*S3(X)=S2(X)$,
 $S5(X)*S4(X)=-S1(X)$,
 $S5(X)*S5(X)=1$;

COMMENT WE DEFINE NOW THE FOURIER SERIES FOR THE
 EXPONENTIAL FUNCTION;

OPERATOR FEXP;
 FOR ALL A LET

$FEXP(A/2)=(COSH(A)+1+S2(X)*(COSH(A)-1)+S1(X)*SINH(A))/2$,
 $FEXP(-A/2)=(COSH(A)+1+S2(X)*(COSH(A)-1)-S1(X)*SINH(A))/2$;

COMMENT START THE CALCULATION;

A1:=FEXP(-F/2);

A1 := (- SINH(F)*S1(X) + S2(X)*(COSH(F) - 1) + COSH(F) + 1)
/2

A2:=B*S3(X)+C*S4(X);

A2 := B*S3(X) + C*S4(X)

A3:=A1*A2;

A3 := (S3(X)*B*(COSH(F) + 1) - (2*C*SINH(F))*S3(X) - (B*SINH
(F))*S4(X) + 2*C*COSH(F)*S4(X) + S5(X)*B*(COSH(F)
- 1) - (2*C*SINH(F))*S5(X))/2

A4:=A3*FEXP(F/2);

A4 := (S3(X)*B*(COSH(F)² + SINH(F)² + 1) - (4*C*COSH(F)*SINH
(F))*S3(X) - (2*B*COSH(F)*SINH(F))*S4(X) + 2*S4(X)
C(COSH(F)² + SINH(F)²) + S5(X)*B*(COSH(F)² + SINH(F)
)² - 1) - (4*C*COSH(F)*SINH(F))*S5(X))/2

COMMENT THIS RESULT IS IDENTICAL TO EQ. (25).;

Acknowledgement

I am grateful to J. M. Maillard and T. T. Truong for helpful comments.

I am also indebted to Mária Zimányi, who was so kind to improve my Reduce program and to run it on different machines. This helped also to identify a LISP-implementation error in our system.

References

1. For an example see G. Györgyi and P. Ruján, *J. Phys.*, C17, 4207, 1984 and references therein.
2. For example the exact solution of the hard hexagon model (R. J. Baxter, *J. Phys.*, A13, L61, 1980, and *J. Stat. Phys.*, 26, 427, 1981) confirmed earlier conjectures regarding the critical exponents of the 3-state Potts model by M. P. M. den Nijs, *J. Phys.*, A12, 1857, 1979.
3. R. J. Baxter, *Ann. Phys. (N. Y.)*, 76, 1, 1973; *ibid.* 76, 25, 1973, *ibid.* 76, 48, 1973.
4. G. E. Andrews, R. J. Baxter and P. J. Forrester, *J. Stat. Phys.*, 35, 193, 1984.

5. P. Ruján, in *Lecture Notes in Physics*, 226, p. 286, ed. N. Sanchez, Springer, New York, 1985.
6. R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic, London, 1982; P. W. Kasteleyn, in *Fundamental Problems in Statistical Mechanics*, Vol. 3, p. 103, ed. E. D. G. Cohen, North-Holland, Amsterdam, 1975.
7. L. D. Faddeev, *Sov. Sci. Rev. Section C (Math. Phys.)*, Vol. 1, ed. P. Novikov, 1981; L. A. Takhtadzhand and L. D. Faddeev, *Russian Math. Surveys*, 34, 11, 1979.
8. H. B. Thacker, *Rev. Mod. Phys.*, 53, 253, 1981.
9. S. K. Pokrovsky and Yu. A. Bashilov, *Comm. Math. Phys.*, 84, 103, 1982.
10. L. Mittag and M. J. Stephen, *J. Math. Phys.*, 12, 441, 1971.
11. S. Elitzur, R. B. Pearson and J. Shigemitsu, *Phys. Rev. D*19, 3698, 1979.
12. J. L. Cardy, *J. Phys.*, A13, 1507, 1980; P. Ruján, G. O. Williams, H. L. Frisch and G. Forgács, *Phys. Rev.*, B23, 1362, 1981.
13. L. Onsager, *Phys. Rev.*, 65, 117, 1944.
14. R. J. Baxter, *J. Phys. C*6, L445, 1973.
15. R. J. Baxter, *Proc. Roy. Soc. London*, A383, 43, 1982.
16. H. N. V. Temperley and E. H. Lieb, *Proc. Roy. Soc. London*, A322, 251, 1971; R. J. Baxter, S. B. Kelland and F. Y. Wu, *J. Phys. A*9, 397, 1976.
17. F. Wegner, *J. Phys.*, C5, L131, 1972.
18. J. M. Maillard, P. Ruján and T. T. Truong, *J. Phys.*, A18, 3399, 1985.
19. V. A. Fateev and A. B. Zamolodchikov, *Phys. Lett.*, A92, 37, 1982.
20. A. A. Belavin, *Nucl. Phys.*, B180(FS2), 189, 1981.
21. R. J. Baxter, *Comm. Math. Phys.*, 88, 185, 1983.
22. M. Lüscher, *Nucl. Phys.*, B117, 475, 1976.
23. E. Fradkin and L. Susskind, *Phys. Rev.*, D17, 2637, 1978.
24. M. J. Stephen and L. Mittag, *J. Math. Phys.*, 13, 1944, 1972; R. Savit, *Rev. Mod. Phys.*, 52, 453, 1980.
25. L. Dolan and M. Grady, *Phys. Rev.*, D25, 1587, 1982.
26. G. V. Gehlen and V. Rittenberg, to be published.
27. M. Suzuki, *Progr. Theor. Phys.*, 46, 1377, 1971.
28. P. Ruján, *J. Stat. Phys.*, 29, 231, 1982; J. Kurmann, H. Thomas and G. Müller, *Physica*, 112A, 235, 1982.
29. see Ref. [24a].
30. J. Ashkin and E. Teller, *Phys. Rev.*, 64, 178, 1943.
31. C. Fan, *Phys. Lett.*, A39, 136, 1972.
32. "Reduce User's Manual" by A. C. Hearn, Version 3.1, Rand Publications, 1984 and Seven Reduce Interactive Lessons, by D. R. Stoutmayer.

CANONICAL QUANTIZATION OF THE RELATIVISTIC THEORY OF THE DIRAC MONOPOLE*

A. FRENKEL

Central Research Institute for Physics
1525 Budapest, Hungary

After a short review of the non-relativistic theory, Dirac's procedure of canonical quantization of gauge theories is applied to the relativistic theory of the magnetic charge.

1. The non-relativistic theory

1.1 Classical theory

Let us consider a spinless "electron" of charge e and mass M in the magnetic field

$$\mathbf{H} = g \frac{\mathbf{r}}{r^3} \quad (1)$$

of an infinitely heavy spinless magnetic monopole of magnetic charge g . Into the equation of motion of the electron

$$M\ddot{\mathbf{r}} = \frac{e}{c} [\dot{\mathbf{r}} \times \mathbf{H}] \quad (2)$$

no potential enters, but if one wishes to derive this equation in the framework of the Lagrangian or of the Hamiltonian formalism, one needs a vector potential \mathbf{A} . Since over any closed surface S enclosing the monopole the integral $\int (d\mathbf{S} \cdot \mathbf{H}) = 4\pi g \neq 0$, the relation $\mathbf{H} = \text{rot } \mathbf{A}$ cannot hold everywhere on S . This relation must be violated at one point (at least) of any such surface, that is along a line (the "Dirac string") going from the monopole to infinity. If the string lies on the negative Z axis, it is easy to find \mathbf{A} [1]. In cylindrical and in spherical coordinates (see Fig. 1) only its φ -component is different from zero:

$$A_\varphi = \frac{g}{\rho} \left(1 - \frac{z}{\sqrt{\rho^2 + z^2}} \right) = \frac{g}{r} \text{tg } \frac{\theta}{2}. \quad (3)$$

* Dedicated to Prof. K. Nagy on his 60th birthday

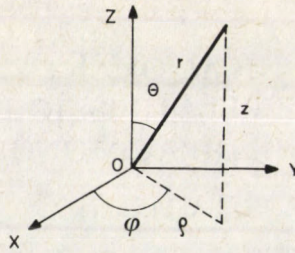


Fig. 1. Notations for coordinates

Near the string (that is for $z < 0$, $\rho \approx 0$)

$$A_\varphi \approx \frac{2g}{\rho}, \quad (4)$$

on the string itself $A_\varphi = \infty$, and $\text{rot } \mathbf{A}$ does not exist there. Everywhere else $\text{rot } \mathbf{A} = g\mathbf{r}/r^3 = \mathbf{H}$.

The Lagrangian for the electron in an external magnetic field reads

$$\mathcal{L} = \frac{1}{2} M \dot{\mathbf{r}}^2 + \frac{e}{c} \dot{\mathbf{r}} \mathbf{A}(\mathbf{r}). \quad (5)$$

In our case the variational principle $\delta I = 0$, $I = \int dt \mathcal{L}$ gives indeed (2) at any point \mathbf{r} except on the string. There \mathbf{A} is infinite in such a way that δI is ill-defined. Thus if the electron arrives to the string, the Lagrangian theory does not say how it will go further. However, the investigation of Dirac [2] showed that this imperfection of the classical Lagrangian (and Hamiltonian) theories is not inherited by the quantum theory. There the behaviour of the wave functions of the electron near and on the string is well defined. Therefore Dirac proposed to apply the classical variational principle only to those virtual trajectories of the electron which avoid the string, and to go over first to the classical and then to the quantized Hamiltonian theory. This prescription, usually referred to as the "veto" of Dirac, sometimes has been erroneously carried over to the quantum theory where it allegedly should mean that all the wave functions of the electron must vanish on the string. As we shall see, this is not so. On the contrary, there should be — and there are — wave functions which do not vanish on the string. Without them the string would be observable, and the "monopole" would not be a pointlike, but a stringlike object.

The passage to the classical Hamiltonian theory is straightforward. We write

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}} = M \dot{\mathbf{r}}, \quad (6)$$

$$(\mathbf{p} \cdot \dot{\mathbf{r}}) - \mathcal{L} = \frac{1}{2M} \mathbf{K}^2 \equiv \mathcal{H}(\mathbf{r}, \mathbf{p}), \quad (7)$$

where

$$\mathbf{K} = \mathbf{p} - \frac{e}{c} \mathbf{A} \quad (8)$$

is the kinetic momentum of the electron. The Hamiltonian is infinite on the string. The Hamiltonian equations of motion coincide with (2) everywhere, except on the string. Following Dirac [2], we do not try to explain this insufficiency of the classical theory away, but go over to the quantum theory via the canonical quantization procedure.

1.2. Quantum theory

We replace the Poisson Brackets (PB's) $\{r_i, p_k\} = \delta_{ik}$ by the commutators

$$[r_i, p_k] \equiv r_i p_k - p_k r_i = i\hbar \delta_{ik} \quad (9)$$

and accordingly we write $\mathbf{p} = -i\hbar \nabla$ in \mathbf{K} . The solutions of the Schrödinger equation with the Hamiltonian (7) were investigated in the classic paper of Dirac [1] in 1931. The complete set of the eigenfunctions of \mathcal{H} for the case when the dimensionless parameter

$$\nu \equiv 2 \frac{eg}{\hbar c} \quad (10)$$

is an integer, $\nu = n$, was given in the same year by Tamm [3], who worked then with Dirac. For the case $\nu \neq n$ the complete set of eigenfunctions was found by Hurst [4] in 1968. The very existence of these complete sets means that for any real value of ν the Schrödinger equation has a mathematically sound solution. The physical interpretation of these solutions, in particular the question whether the string (that is the singularity of \mathbf{A} along the string) is observable or not gave rise to passionate discussions even in the late 60-ies. As far as the non-relativistic theory is concerned, the correctness of Dirac's original statement [1] is nowadays firmly established and generally accepted: the necessary and sufficient condition for the string being unobservable is that ν be integer. Below we shortly recapitulate the main features of the non-relativistic quantum theory of the monopole, because we shall have to rely on them when discussing the relativistic theory.

As is well known, in the presence of a vector potential \mathbf{A} the kinetic momentum \mathbf{K} is an observable, the canonical momentum \mathbf{p} is not. In our case K_φ contains A_φ , and A_φ is infinite on the string. However, K_φ contains also the differential operator $\nabla_\varphi = \partial_\varphi / \rho$. Thus near the string we have (see (4) and (10))

$$K_\varphi \approx -\frac{\hbar}{\rho} (i\partial_\varphi + \nu). \quad (11)$$

If K_φ acts on a wave function $\psi(\rho, \varphi, z)$, $K_\varphi \psi$ will be infinite on the string except in two cases: either ψ should vanish sufficiently rapidly when $\rho \rightarrow 0$, or if $\psi \neq 0$ on the string,

then its phase must be $v\varphi$. Indeed, then the zero coming from the relation

$$(i\partial_\varphi + v)e^{iv\varphi} = (-v + v)e^{iv\varphi} = 0 \quad (12)$$

may overrun the infinity coming from $1/\rho$. In the latter case v must be an integer, otherwise the wave function will not be single valued: when carried along a closed loop round the Z axis, it would get an extra factor $\exp(i2\pi v) \neq 1$.

The ψ 's in the complete system of \mathcal{H} do have these features. When $v = n$, there are eigenfunctions which do not vanish on the string and they all have the phase factor $\exp(in\varphi)$ [1, 3]. There are also others with phases $m\varphi$, $m \neq n = v$, and they all do vanish on the string. On the other hand, for the case $v \neq n$ all the eigenfunctions of \mathcal{H} vanish on the string [4], as they should.

Let us point out that in the case $v = n$ the wave functions with phase $n\varphi$ are single valued at any point outside of the string (the $-Z$ axis in our case), but on the string itself they are not. It is easy to see this: $\exp(in\varphi) = +1$ if we approach the string along the positive X axis ($\varphi = 0$), but $\exp(in\varphi) = -1$ if we approach along the negative X axis ($\varphi = \pi$). As stressed in [5], this multivaluedness of the wave functions on the string does not invalidate the probabilistic interpretation, because the scalar product (ψ_1, ψ_2) of any two energy eigenfunctions is single valued. If both have the "string-type singularity" $\exp(in\varphi)$, then it drops out from the scalar product. If only one of them has it, the other is zero on the string. In either case the scalar product is single valued. Since the argument obviously holds for any linear combination of the wave functions, all the probability amplitudes are single valued.

It remains to be seen whether the string is observable or not. If it is not, a change in the position of the string should correspond to a gauge transformation.

The vector potential for the magnetic field (1) with the string S lying on an arbitrarily chosen line $\mathbf{u}(\sigma)$ has been given by Jordan [6] in 1938:

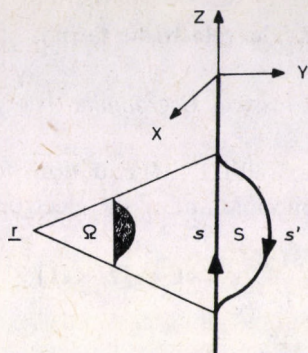
$$\mathbf{A}^S(\mathbf{r}) = g \int_0^\infty d\sigma \frac{\left[\frac{\partial \mathbf{u}(\sigma)}{\partial \sigma} \times (\mathbf{u}(\sigma) - \mathbf{r}) \right]}{|\mathbf{u}(\sigma) - \mathbf{r}|^3}. \quad (13)$$

As shown in [5] a change in the position of the string leads to a gauge transformation of \mathbf{A} :

$$\mathbf{A}^S(\mathbf{r}) \rightarrow \mathbf{A}^{S'}(\mathbf{r}) = \mathbf{A}^S(\mathbf{r}) + g \nabla \Omega(S, S'; \mathbf{r}), \quad (14)$$

where Ω is the solid angle under which the oriented surface $S(S, S')$ subtended by the strings is seen from the point \mathbf{r} (Fig. 2). Notice that the gauge term $\nabla \Omega$ is singular along the edge of the surface S . Indeed, $\Omega \approx 2\varphi'(\mathbf{r})$ when \mathbf{r} tends to a point on the edge S' , and $\Omega \approx -2\varphi(\mathbf{r})$ when $\mathbf{r} \rightarrow S$. Here $\varphi(\mathbf{r})$ stands for the azimuthal angle of a local coordinate system with its origin at the point to which \mathbf{r} tends, and with its $-Z$ axis tangent to the string at that point. Thus for $\mathbf{r} \rightarrow S$

$$g \nabla_\varphi \Omega \approx -\frac{2g}{\rho}, \quad (15)$$

Fig. 2. The solid angle $\Omega(s, s'; r)$

and comparing this with (4) we realize that the gauge term in (14) cancels the singularity of \mathbf{A}^S along the old string S and introduces the required singularity of $\mathbf{A}^{S'}$ along S' .

In quantum mechanics (14) must be completed with a transformation of the wave functions:

$$\psi^S(\mathbf{r}) \rightarrow \psi^{S'}(\mathbf{r}) = e^{i \frac{eg}{\hbar c} \Omega(S, S'; \mathbf{r})} \psi^S(\mathbf{r}) = e^{i \frac{v}{2} \Omega} \psi^S(\mathbf{r}), \quad (16)$$

and the wave functions must be single valued everywhere, except on their respective strings where they must have the string-type singularity unless they vanish there. Let us first choose a point \mathbf{r} which is not on the strings, and let us follow the change of the factors in (16) along a closed path which avoids the strings but intersects the surface S subtended by them. At the end $\psi^{S'}$ and ψ^S should come back to their initial values, while Ω , as well known, changes by 4π . Therefore we get the condition

$$e^{i \frac{v}{2} 4\pi} = 1, \quad (17)$$

that is v must be integer. Let now \mathbf{r} approach a string. Taking into account the behaviour of Ω near the strings explained above, we see that $\psi^{S'}$ will indeed have the required $\exp(in\varphi')$ phase factor on S' , while the old string-type singularity $\exp(in\varphi)$ on S will be cancelled in $\psi^{S'}$. It is now easy to realize that any scalar product $(\psi_1^{S'}, \psi_2^{S'})$ is single valued and equals the corresponding untransformed product (ψ_1^S, ψ_2^S) . This concludes the outline of the proof that $v = n$ is the necessary and sufficient condition for the string being unobservable.

When $v \neq n$, the transformation (16) does not lead to single valued new wave functions. Accordingly (16) is not a gauge transformation, and the string is observable. It represents then a solenoid with vanishing radius and with magnetic flux $4\pi g$ inside the solenoid (that is along the string). The condition $v = n$ is then the condition for the vanishing of the Aharonov–Bohm interference picture [7]. All other observable effects caused by the string also do vanish then, only the monopole sitting at the end of the string is observable [5].

We see that Dirac was close to the discovery of the Aharonov–Bohm effect in 1931. He just did not ask himself whether there is a possible physical interpretation of the case $v \neq n$. . .

2. The relativistic theory

2.1. Classical Lagrangian theory

The Lorentz and the Maxwell ($L-M$) equations for a system composed of a spinless electron, a spinless monopole and of the electromagnetic field are

$$\frac{M_e \ddot{\mathbf{z}}_e}{\sqrt{1-\dot{\mathbf{z}}_e^2}} = e\mathbf{E} + e[\dot{\mathbf{z}}_e \times \mathbf{H}], \quad (18)$$

$$\frac{M_g \ddot{\mathbf{z}}_g}{\sqrt{1-\dot{\mathbf{z}}_g^2}} = g\mathbf{H} - g[\dot{\mathbf{z}}_g \times \mathbf{E}], \quad (19)$$

$$\operatorname{div} \mathbf{E} = 4\pi e \delta^{(3)}(\mathbf{x} - \mathbf{z}_e), \quad (20)$$

$$\operatorname{rot} \mathbf{H} - \frac{\partial \mathbf{E}}{\partial t} = 4\pi e \dot{\mathbf{z}}_e \delta^{(3)}(\mathbf{x} - \mathbf{z}_e), \quad (21)$$

$$\operatorname{div} \mathbf{H} = 4\pi g \delta^{(3)}(\mathbf{x} - \mathbf{z}_g), \quad (22)$$

$$\operatorname{rot} \mathbf{E} + \frac{\partial \mathbf{H}}{\partial t} = -4\pi g \dot{\mathbf{z}}_g \delta^{(3)}(\mathbf{x} - \mathbf{z}_g), \quad (23)$$

Here \mathbf{z}_e , \mathbf{z}_g stand for the coordinates of the electron and of the monopole; to simplify the writing, $c=1$.

The Lagrangian leading to these equations contains A_μ not only in the interaction term $j_\mu A_\mu$ but also in $F_{\mu\nu} F_{\mu\nu}$. Similarly to the non-relativistic case, if monopoles exist the usual definition $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$ cannot be correct, because it leads to $\partial_\mu \tilde{F}_{\mu\nu} \equiv 0$ in contradiction with (22–23). In his fundamental paper on the relativistic monopole [2] Dirac showed how to define $F_{\mu\nu}$ in such a way that (22–23) be satisfied identically. One has to write

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu + 4\pi \tilde{G}_{\mu\nu}, \quad (24)$$

where $\tilde{G}_{\mu\nu}$ is the dual of the "string tensor"

$$G_{\mu\nu}(\mathbf{x}, t) \equiv g \int_0^\infty d\sigma \left(\dot{y}_\mu \frac{\partial y_\nu}{\partial \sigma} - \frac{\partial y_\mu}{\partial \sigma} \dot{y}_\nu \right) \delta^{(3)}(\mathbf{x} - \mathbf{y}(\sigma, t)). \quad (25)$$

We see that $G_{\mu\nu}(\mathbf{x}, t)$ is different from zero only along a string $\mathbf{y}(\sigma, t)$. (The time component $y_0 \equiv t$.)

It is important to realize that the electric and magnetic fields \mathbf{E} , \mathbf{H} should be identified not with $\partial_\mu A_\nu - \partial_\nu A_\mu$, but with $F_{\mu\nu}$. Now, the \mathbf{E} 's and \mathbf{H} 's which solve the $L-M$ equations are infinite only at the points \mathbf{z}_e , \mathbf{z}_g , but not on the string. The latter does not show up at all in the $L-M$ equations. On the other hand, $G_{\mu\nu}$ and $\tilde{G}_{\mu\nu}$ are singular on the string by definition. Therefore in (24) the singularity of $\tilde{G}_{\mu\nu}$ must be compensated

by $\partial_\mu A_\nu - \partial_\nu A_\mu$. So we foresee that in the solutions of the $L - M$ equations A_μ must be singular along the string, and in the quantized field theory the wave functions will have a string-type singularity.

As stressed in [2], the string coordinates can be considered either as dynamical variables, or as given external variables. We shall follow Dirac in treating them as dynamical variables. The present author found [8] that it is then convenient to introduce the relative string coordinates \mathbf{u} through the relation

$$\mathbf{y}(\sigma, t) \equiv \mathbf{z}_g(t) + \mathbf{u}(\sigma, t). \quad (26)$$

This makes it possible to disentangle \mathbf{z}_g from the string variable in an easier way than in [2]. The results are of course the same.

The Lagrangian of the system is*

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3, \quad (27)$$

$$\mathcal{L}_1 = -M_e \sqrt{1 - \dot{\mathbf{z}}_e^2} - M_g \sqrt{1 - \dot{\mathbf{z}}_g^2}, \quad (28)$$

$$\mathcal{L}_2 = -\frac{1}{16\pi} \int d\mathbf{x} F_{\mu\nu}(\mathbf{x}) F_{\mu\nu}(\mathbf{x}), \quad (29)$$

$$\mathcal{L}_3 = -e[A_0(\mathbf{z}_e) - (\dot{\mathbf{z}}_e \cdot \mathbf{A}(\mathbf{z}_e))]. \quad (30)$$

As shown in [2], the variation of \mathbf{z}_e , $A_\mu(\mathbf{x})$, \mathbf{z}_g and $\mathbf{y}(\sigma)$ (\mathbf{z}_g and $\mathbf{u}(\sigma)$ in [8]) leads to the correct $L - M$ equations except for terms proportional to $\delta^{(3)}(\mathbf{z}_e(t) - \mathbf{y}(\sigma, t))$. This again means that when calculating the variation of $\int dt \mathcal{L}$ one has to disregard those virtual paths of the electron which intersect the string:

$$\mathbf{z}_e \neq \mathbf{y}(\sigma). \quad (31)$$

As a matter of fact, the $L - M$ equations are obtained when \mathbf{z}_e , \mathbf{z}_g and A_μ are varied. The variation of \mathbf{u} does not lead to an equation of motion for the string, it merely reproduces one of the already obtained equations. This means that the time evolution of the string is not determined by the variational principle. We also know that even without monopoles the time evolution of A_μ is not uniquely determined; there is a gauge freedom in it, supplemented now by a freedom needed to cancel the singularity along the freely moving string so that $F_{\mu\nu}$ be regular and uniquely defined on the string.

Let us point out that \mathcal{L}_2 seems to be a Lagrangian of the free electromagnetic field, but it is not. Indeed, $F_{\mu\nu}$ now contains not only A_μ , but also $\mathbf{y} = \mathbf{z}_g + \mathbf{u}$. Thus \mathcal{L}_2 describes the interaction of the electromagnetic field with the monopole. That is why \mathcal{L}_3 contains only the usual electron-electromagnetic field interaction.

* The time argument of the dynamical variables will be often omitted.

2.2. Classical Hamiltonian theory

The next step towards canonical quantization is the construction of the classical Hamiltonian formalism. The canonical momenta are

$$\mathbf{p}_e \equiv \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{z}}_e} = \frac{M_e \dot{\mathbf{z}}_e}{\sqrt{1 - \dot{\mathbf{z}}_e^2}} + e\mathbf{A}(\mathbf{z}_e). \quad (32)$$

$$\mathbf{p}_g \equiv \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{z}}_g} = \frac{M_g \dot{\mathbf{z}}_g}{\sqrt{1 - \dot{\mathbf{z}}_g^2}} + g \int_0^\infty d\sigma \left[\mathbf{E}(\mathbf{y}(\sigma)) \times \frac{\partial \mathbf{u}(\sigma)}{\partial \sigma} \right], \quad (33)$$

$$\Pi_0(\mathbf{x}) \equiv \frac{\delta \mathcal{L}}{\delta \dot{A}^0(\mathbf{x})} = -\frac{1}{4\pi} F_{00}(\mathbf{x}) = 0, \quad (34)$$

$$E^i(\mathbf{x}) \equiv 4\pi \frac{\delta \mathcal{L}}{\delta \dot{A}^i(\mathbf{x})} = -\left(\dot{A}^i(\mathbf{x}) + \frac{\partial A^0(\mathbf{x})}{\partial x_i} + 4\pi g \tilde{G}^{0i}(\mathbf{x}) \right), \quad (35)$$

$$\mathbf{p}_u(\sigma) \equiv \frac{\delta \mathcal{L}}{\delta \dot{\mathbf{u}}(\sigma)} = g \left[\mathbf{E}(\mathbf{y}(\sigma)) \times \frac{\partial \mathbf{u}(\sigma)}{\partial \sigma} \right]; \quad (\sigma > 0). \quad (36)$$

The velocities \dot{A}_0 and $\dot{\mathbf{u}}$ cannot be expressed with the help of these relations through coordinates and momenta, because they do not even enter them. On the other hand, we have constraints of the form $\xi_a(q, p) = 0$. One of them is $\Pi_0(\mathbf{x}) = 0$, the other, (36), can be conveniently written in the form

$$\chi(\sigma) \equiv \mathbf{p}_u(\sigma) - g \left[\mathbf{E}(\mathbf{y}(\sigma)) \times \frac{\partial \mathbf{u}(\sigma)}{\partial \sigma} \right] = 0. \quad (37)$$

Lagrangians with such properties are called "singular" or "degenerate". The Hamiltonian theory for them was worked out by Dirac in the 50-ies [9]. The lack of this Hamiltonian theory with constraints is felt on his paper on the relativistic monopole, the formalism of which is indeed "extremely complicated" [10]. The systematic application of the procedure of canonical quantization to the degenerate Lagrangian (27), carried out in detail in an unpublished preprint [8] of the present author and outlined below, makes the formalism simpler and corroborates Dirac's main result: the condition for the string being unobservable is $v = n$ in the relativistic quantum theory of the monopole, too.

The constraint $\Pi_0(\mathbf{x}) = 0$ contradicts the canonical PB

$$\{A_0(\mathbf{x}), \Pi_0(\mathbf{x}')\} = \delta^{(3)}(\mathbf{x} - \mathbf{x}'). \quad (38)$$

Therefore, as explained in [9], the constraint functions $\xi_a(q, p)$ should be equated to zero only after the PB's have been opened. This rule is expressed by saying that the constraints are equal to zero not "strongly", but "weakly": $\Pi_0(\mathbf{x}) \stackrel{W}{=} 0$, $\chi(\sigma) \stackrel{W}{=} 0$.

Now, it turns out that when the constraints hold (that is weakly) the expression $\sum_r p_r \dot{q}_r - \mathcal{L}(q, \dot{q})$ depends only on q and p , so that it is a Hamiltonian:

$$\sum_r p_r \dot{q}_r - \mathcal{L}(q, \dot{q}) \stackrel{W}{=} \mathcal{H}(q, p). \quad (39)$$

In our case one finds

$$\mathcal{H} = \mathcal{H}_e + \mathcal{H}_g + \mathcal{H}_\gamma \quad (40)$$

with

$$\mathcal{H}_e = \sqrt{M_e^2 + \mathbf{K}_e^2} + eA_0(\mathbf{z}_e), \quad (41)$$

$$\mathcal{H}_g = \sqrt{M_g^2 + \mathbf{K}_g^2}, \quad (42)$$

$$\mathcal{H}_\gamma = \frac{1}{8\pi} \int d\mathbf{x} (\mathbf{E}^2 + \mathbf{H}^2 - 2A_0 \operatorname{div} \mathbf{E}), \quad (43)$$

where $\mathbf{K}_e, \mathbf{K}_g$ are the kinetic momenta of the particles:

$$\mathbf{K}_e \equiv \mathbf{p}_e - e\mathbf{A}(\mathbf{z}_e) = \frac{M_e \dot{\mathbf{z}}_e}{\sqrt{1 - \dot{\mathbf{z}}_e^2}}, \quad (44)$$

$$\mathbf{K}_g \equiv \mathbf{p}_g - g \int_0^\infty d\sigma \left[\mathbf{E}(\mathbf{y}(\sigma)) \times \frac{\partial \mathbf{u}(\sigma)}{\partial \sigma} \right] = \frac{M_g \dot{\mathbf{z}}_g}{\sqrt{1 - \dot{\mathbf{z}}_g^2}}. \quad (45)$$

The constraints (34) and (37) should hold (weakly) at any moment of time. Therefore their PB's with \mathcal{H} should vanish at least weakly. This consistency requirement may lead to secondary constraints. In our case this happens for Π_0 . Indeed, one finds

$$4\pi \{ \Pi_0(\mathbf{x}), \mathcal{H} \} = \operatorname{div} \mathbf{E}(\mathbf{x}) - 4\pi e \delta^{(3)}(\mathbf{x} - \mathbf{z}_e), \quad (46)$$

so that one should impose the secondary constraint

$$\eta(\mathbf{x}) \equiv \operatorname{div} \mathbf{E}(\mathbf{x}) - 4\pi e \delta^{(3)}(\mathbf{x} - \mathbf{z}_e) \stackrel{W}{=} 0. \quad (47)$$

This is nothing else but Maxwell's equation (20). It does not contain velocities, this is why it can turn up as a constraint. Notice, however, that the equation (22) which is also free of velocities is not a constraint but an identity (that is a strong equality) due to the new definition of $F_{\mu\nu}$.

No further constraints arise from the consistency conditions of χ and η , and it turns out that the PB of any two constraints vanishes at least weakly*. Such constraints

* when the "veto" (31) holds. As explained above, this is always understood in the classical Lagrangian and Hamiltonian theories, but not in the quantum theory.

are "first class constraints", and as shown in [9] they must be added to \mathcal{H} with arbitrary coefficients. One obtains then the "extended Hamiltonian" \mathcal{H}_E

$$\mathcal{H}_E = \mathcal{H} + \int dx v(\mathbf{x}) \Pi_0(\mathbf{x}) + w(\mathbf{x}) \eta(\mathbf{x}) + \int_0^\infty d\sigma (v(\sigma) \cdot \chi(\sigma)). \quad (48)$$

The arbitrary functions v , w and ν carry the largest possible gauge freedom inherent in the theory. The variables \mathbf{z}_e , \mathbf{z}_g and $F_{\mu\nu}$ are unaffected by this freedom, because their PB's vanish with all the constraints at least weakly. Thus their equations of motion are in fact determined by \mathcal{H} , and they turn out to be identical with the $L - M$ equations.

The PB's of A_μ and of \mathbf{u} with the constraints do not all vanish. Their time evolution is therefore dependent on the arbitrary functions, and this gauge dependence is generated by the first class constraints. Dirac postulates [9] that such variables with unpredictable time evolution cannot correspond to observables.

The constraints Π_0 and η generate the gauge dependence of A_μ familiar from electrodynamics. They do not influence the motion of the string. The constraints χ contain both string and field variables, so that they generate interdependent arbitrariness in the motion of \mathbf{u} and of \mathbf{A} (but not of A_0 , since χ does not contain Π_0). The relation between \mathbf{u} and \mathbf{A} is easily found. From (24) and (25) we have

$$\mathbf{H}(\mathbf{x}, t) = \text{rot } \mathbf{A}(\mathbf{x}, t) + 4\pi g \int_0^\infty d\sigma \frac{\partial \mathbf{u}(\sigma)}{\partial \sigma} \delta^{(3)}(\mathbf{x} - \mathbf{y}(\sigma, t)), \quad (49)$$

and since \mathbf{H} is gauge invariant, $\mathbf{H}(\mathbf{x}, t)$ does not depend on the choice of two different functions $\nu^{(1)}(\sigma)$, $\nu^{(2)}(\sigma)$:

$$\mathbf{H}^{(1)}(\mathbf{x}) = \mathbf{H}^{(2)}(\mathbf{x}) \quad (50)$$

at any moment of time, so that

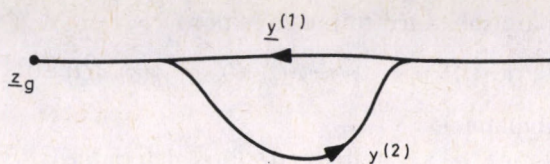
$$\text{rot } (\mathbf{A}^{(2)}(\mathbf{x}) - \mathbf{A}^{(1)}(\mathbf{x})) = 4\pi g \oint_{S(1,2)} d\sigma \frac{\partial \mathbf{u}(\sigma)}{\partial \sigma} \delta^{(3)}(\mathbf{x} - \mathbf{y}(\sigma)). \quad (51)$$

The integration in (51) goes along the closed path $S(1, 2)$ formed by those portions of the strings $\mathbf{y}^{(1)}$ and $\mathbf{y}^{(2)}$ where they do not coincide due to the difference of the ν 's (Fig. 3). For simplicity we chose $\nu^{(1)}$ and $\nu^{(2)}$ different from each other only on a finite interval of σ .

The solution of Eq. (51) is

$$\mathbf{A}^{(2)}(\mathbf{x}) = \mathbf{A}^{(1)}(\mathbf{x}) + g \nabla \Omega(\mathbf{y}^{(1)}, \mathbf{y}^{(2)}; \mathbf{x}), \quad (52)$$

where Ω is the solid angle under which the surface subtended by the closed string $S(1, 2)$ is seen from the point \mathbf{x} . The similarity with (14) is obvious (notice that $\mathbf{y}(\sigma) = \mathbf{u}(\sigma)$ if the

Fig. 3. The closed path $S(1,2)$

origin is at \mathbf{z}_g). One sees that in the relativistic theory, too, \mathbf{A} behaves near the string as the Dirac-Jordan potential. However, one has to keep in mind that while in the non-relativistic theory presented in Part 1 \mathbf{A} was treated as an external, time-independent potential and it was a functional of the chosen string \mathbf{u} from the outset, now \mathbf{A} and \mathbf{u} are independent dynamical variables and \mathbf{A} becomes a functional of \mathbf{u} only in the solutions of the Lagrangian or Hamiltonian equations of motion.

2.3. Quantum theory

When all the constraints are first class, the canonical quantization consists of two steps [9]: one has to substitute the commutators for the corresponding PB's of the canonical coordinates and momenta, and one has to impose the constraints as subsidiary conditions on the wave functions. In short, $\{q_i, p_k\} = \delta_{ik} \rightarrow [q_i, p_k] = i\hbar\delta_{ik}$, and $\xi_a(q, p) \stackrel{W}{=} 0 \rightarrow \xi_a(q, p)\psi(q) = 0$.

We shall choose a representation where the coordinates $\mathbf{z}_e, \mathbf{z}_g, A_0(\mathbf{x})$ and $\mathbf{u}(\sigma)$ are diagonal, so that their respective momenta are represented by

$$\mathbf{p}_e = -i\hbar \frac{\partial}{\partial \mathbf{z}_e}, \quad \mathbf{p}_g = -i\hbar \frac{\partial}{\partial \mathbf{z}_g}, \quad (53)$$

$$\mathbf{p}_u(\sigma) = -i\hbar \frac{\delta}{\delta \mathbf{u}(\sigma)}, \quad \Pi_0(\mathbf{x}) = -i\hbar \frac{\delta}{\delta A_0(\mathbf{x})}. \quad (54)$$

As far as the conjugate variables \mathbf{A} and \mathbf{E} are concerned, it turns out convenient to choose \mathbf{E} diagonal. Then \mathbf{A} is represented by (see (35))

$$\mathbf{A}(\mathbf{x}) = 4\pi i\hbar \frac{\delta}{\delta \mathbf{E}(\mathbf{x})}. \quad (55)$$

The reason for this choice is that the constraints η and χ do not depend on \mathbf{A} but on \mathbf{E} , and in our representation they will be diagonal in \mathbf{E} .

The ψ 's will now depend on the coordinate Q

$$Q = (\mathbf{z}_e, \mathbf{z}_g, A_0(\mathbf{x}), \mathbf{E}(\mathbf{x}), \mathbf{u}(\sigma)), \quad (56)$$

and those ψ 's which are submitted to the conditions

$$\Pi_0(\mathbf{x})\psi(Q)=0, \quad \eta(\mathbf{x})\psi(Q)=0, \quad \chi(\sigma)\psi(Q)=0 \quad (57)$$

form the physical submanifold.

The time evolution of a variable $f(Q, P)$ is determined by the well-known equation

$$\dot{f} = \frac{1}{i\hbar} [f, \mathcal{H}_E] = \frac{1}{i\hbar} [f, \mathcal{H} + \sum_a v_a \xi_a]. \quad (58)$$

It is clear that the condition for f being gauge independent is:

$$[f, \xi_a]\psi = 0 \quad \text{for all the } \xi_a \text{'s,} \quad (59)$$

because then the arbitrary coefficients v_a drop out from (58).

A straightforward but lengthy calculation shows that $\mathbf{z}_e, \mathbf{z}_g, \mathbf{E}$ and \mathbf{H} are gauge independent, A_μ and \mathbf{u} are not. However, this is not the end of the story. One has still to show that the gauge transformations of the ψ 's which go together with those of A_μ and \mathbf{u} (for comparison see (14) and (16)) are such that the wave functions remain single valued as required by the probabilistic interpretation. We shall show below that this happens only if $v = n$.

It should be stressed from the outset that while in the nonrelativistic theory the derivations can be carried out with mathematical rigor, in the relativistic theory one is dealing with operators acting on a manifold of wave functions (more exactly, the ψ 's are functions of $\mathbf{z}_e, \mathbf{z}_g$ and functionals of $A_0(\mathbf{x}), \mathbf{E}(\mathbf{x})$ and $\mathbf{u}(\sigma)$) the mathematical structure of which is poorly known. Such tools as the complete set of energy eigenfunctions are lacking here. Accordingly one can show only that when $v = n$ the wave functions have reasonable properties analogous to those of the wave functions of the non-relativistic theory.

This is achieved by making use of the subsidiary conditions (57). The condition $\chi\psi = 0$ written out in detail reads

$$\left(-i\hbar \frac{\delta}{\delta u_r} - \frac{g}{c} \varepsilon^{rls} E^l(\mathbf{y}(\sigma)) \frac{\partial u^s}{\partial \sigma} \right) \psi(Q) = 0; \quad r, l, s = 1, 2, 3. \quad (60)$$

(From now on we reintroduce the velocity of light c .) This equation can be rewritten in the form

$$\delta\psi(Q) = \frac{ig}{\hbar c} (\mathbf{E}(\mathbf{y}(\sigma)) \cdot \delta\mathbf{S}(\sigma))\psi(Q), \quad (61)$$

where

$$\delta\mathbf{S}(\sigma) = \left[\frac{\partial \mathbf{u}(\sigma)}{\partial \sigma} \times \delta \mathbf{u}(\sigma) \right] \quad (62)$$

is the infinitesimal surface swept out by the string under an infinitesimal variation $\delta \mathbf{u}(\sigma)$. A finite variation of the string from $\mathbf{u}^{(0)}$ to \mathbf{u} (with all the other coordinates in Q fixed) leads to

$$\psi(Q^{(0)}) \rightarrow \psi(Q) = \exp \left\{ \frac{ig}{\hbar c} \int_{S(y^{(0)}, y)} (\delta \mathbf{S} \cdot \mathbf{E}) \right\} \psi(Q^{(0)}) \quad (63)$$

where $S(y^{(0)}, y)$ is the surface swept out by the string (Fig. 4).

It is well known that the wave functions must be regular at all points except possibly at those where the Hamiltonian is singular [11]. Our \mathcal{H}_E is singular (infinite) "on the string", that is at such points Q^s for which $\mathbf{z}_e = \mathbf{y}(\sigma)$ at some value of σ . To see this, let $\psi(Q) \neq 0$ at some point Q . Then to fulfil the condition $\eta\psi = 0$ we must require

$$\eta(\mathbf{x}) \equiv \operatorname{div} \mathbf{E}(\mathbf{x}) - 4\pi e \delta^{(3)}(\mathbf{x} - \mathbf{z}_e) = 0, \quad (64)$$

and the solution of this equation is

$$\mathbf{E}(\mathbf{x}) = \mathbf{e} \frac{\mathbf{x} - \mathbf{z}_e}{|\mathbf{x} - \mathbf{z}_e|^3} + \mathbf{e}(\mathbf{x}); \operatorname{div} \mathbf{e}(\mathbf{x}) = 0, \quad (65)$$

where \mathbf{e} is regular at $\mathbf{x} = \mathbf{z}_e$ and is irrelevant for us. Now \mathcal{H}_E contains $\mathbf{E}(\mathbf{y}(\sigma))$ both in \mathbf{K}_g and in χ , and (65) shows that $\mathbf{E}(\mathbf{y}(\sigma)) = \infty$ if $\mathbf{y}(\sigma) = \mathbf{z}_e$.

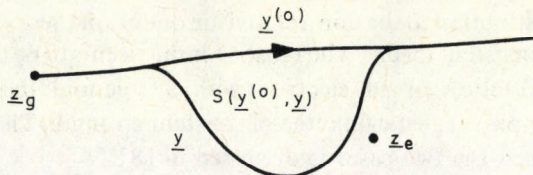


Fig. 4. The surface $S(y^{(0)}, y)$ swept out by the string

To avoid these singular points let us consider a wave function which is different from zero at some point $Q^{(0)}$ which is not on the string. Moreover, let us carry out the variation of the string in such a way that the point \mathbf{z}_e be always avoided (Fig. 4). Then in (63) both $\psi(Q^{(0)})$ and $\psi(Q)$ must be single valued, and in the exponent \mathbf{E} is finite everywhere on the surface swept out by the string. Let us now come back with the string to its original position in such a way that \mathbf{z}_e be in the interior of the closed surface swept out by the string. The result obviously is

$$\psi(Q^{(0)}) = \exp \left[i \frac{g}{\hbar c} \oint_s (\delta \mathbf{S} \cdot \mathbf{E}) \right] \psi(Q^{(0)}). \quad (66)$$

Since $\psi(Q^{(0)})$ must be single valued and $\mathbf{E}(\mathbf{x})$ is given by (65), we must have

$$\frac{g}{\hbar c} \oint_s (\delta \mathbf{S} \cdot \mathbf{E}) = \frac{g}{\hbar c} \cdot 4\pi e = 2\pi n \quad (67)$$

and this is just Dirac's condition $v = n$.

Let us now vary the string in such a way that one of its points tend to \mathbf{z}_e . From (63) we see that $\psi(Q)$ depends on the varied portion of the string only through the exponent. Therefore $|\psi(Q)| = |\psi(Q^{(0)})|$ during the whole variation, and since we had $\psi(Q^{(0)}) \neq 0$, we shall have a wave function which do not vanish on the string when $Q \rightarrow Q^s$. But then \mathbf{E} in the exponent is given by (65), and taking into account that $\nu = n$, the exponent becomes

$$i \frac{n}{2} \int_{S(\mathbf{y}^{(0)}, \mathbf{y})} \left[\delta \mathbf{S} \cdot \frac{\mathbf{x} - \mathbf{z}_e}{|\mathbf{x} - \mathbf{z}_e|^3} \right] = i \frac{n}{2} \Omega(\mathbf{y}^{(0)}, \mathbf{y}; \mathbf{z}_e). \quad (68)$$

(We neglected the term $\mathbf{e}(\mathbf{x})$ in (65), since it is regular at $\mathbf{x} = \mathbf{z}_e$.) As we have seen repeatedly, when $\mathbf{z}_e \rightarrow \mathbf{y}(\sigma)$ (or $\mathbf{y}(\sigma) \rightarrow \mathbf{z}_e$), $\Omega \rightarrow 2\varphi_{\mathbf{y}}(\mathbf{z}_e)$. Therefore when $Q \rightarrow Q^s$ we get

$$\psi(Q) \rightarrow \psi(Q^s) = e^{i n \varphi_{\mathbf{y}}(\mathbf{z}_e)} \psi(Q^{(0)}),$$

our wave function develops a string-type singularity as expected.

These considerations show that in the case $\nu = n$ a change in the position of the string is a gauge transformation, and the string is unobservable. It should be pointed out that in contradistinction to the non-relativistic quantum theory, now the case $\nu \neq n$ does not lead to a consistent theory. The reason for this seems to be that in the theory of the relativistic interaction of an electron with a solenoid the latter cannot be represented by a simple magnetic flux line of constant strength. The formal reason for the difference between the two cases is discussed in [8].

At present various kinds of monopoles play an important role in the non-abelian gauge theories of elementary particle physics. It is the author's hope that this work will be of some help to those who wish to go back to the origins of these modern monopole theories.

References

1. P. A. M. Dirac, Proc. Roy. Soc., *A113*, 60, 1931.
2. P. A. M. Dirac, Phys. Rev., *74*, 817, 1948.
3. I. E. Tamm, Ztschr. Phys., *71*, 1411, 1931.
4. C. A. Hurst, Ann. Phys. N. Y., *50*, 51, 1968.
5. A. Frenkel and P. Hráskó, Ann. Phys., N. Y., *105*, 288, 1977.
6. P. Jordan, Ann. Physik, 5 Folge, Band *32*, 66, 1938.
7. Y. Aharonov and D. Bohm, Phys. Rev., *115*, 485, 1959.
8. A. Frenkel, Reformulation of the Relativistic Theory of the Dirac Monopole, Preprint KFKI-1977-95, Budapest (unpublished).
9. P. A. M. Dirac, Lectures on Quantum Mechanics, Belfer Graduate School of Science, Yeshiva University, New York, 1964.
10. T. T. Wu and C. N. Yang, Phys. Rev., *D13*, 437, 1976.
11. J. von Neumann, Göttinger Nachrichten, Math. Phys. Kl., p. 1, 1927.

UNUSUAL MAPS AND THEIR USE TO APPROACH USUAL ONES*

Z. KAUFMANN

*Institute for Solid State Physics, Roland Eötvös University
1088 Budapest, Hungary*

P. SZÉPFALUSY

*Institute for Solid State Physics, Roland Eötvös University
1088 Budapest, Hungary*

and

*Central Research Institute for Physics
1525 Budapest, Hungary*

and

T. TÉL

*Institute for Theoretical Physics, Roland Eötvös University
1088 Budapest, Hungary*

One-dimensional maps coupled to discrete valued variables are introduced. They are designed to describe the motion in Lorenz and Hénon type systems on branched manifolds arising by expanding the maps in powers of the inverse dissipation strength, or by coarse graining. The maps are studied in detail along the crisis line where they exhibit complex behaviour with periodic and chaotic attractors. The convergence to the Hénon map is investigated numerically and found to be satisfactory for not too weak dissipations.

1. Introduction and summary

Our view concerning mechanical and other types of motion has drastically changed owing to recent developments in the theory of dynamical systems [1–3]. It is nowadays clear that the long-time behaviour of deterministic nonlinear systems with at least one and a half degrees of freedom generically exhibits unpredictable, chaotic motion in a certain region. Since this type of motion has stochastic features its complete description requires statistical methods [1–3].

Following an idea of Poincaré [4] it is often convenient to consider a discrete dynamics, a mapping generated by either the intersection points of the continuous trajectories with a certain surface of the phase space (Poincaré map) or by taking subsequent snapshots of the motion with a given periodicity (stroboscopic maps). The form of the mapping follows uniquely from the continuous motion, the inverse is,

* Dedicated to Prof. K. Nagy on his 60th birthday

however, not true, the same map can belong to several systems. General properties of chaotic motions are most commonly studied by investigating maps exhibiting chaotic behaviour [1–3, 5, 6].

Besides general common features, there are certain differences in the chaotic behaviour of conservative (Hamiltonian) and dissipative systems. In the former case, the fact whether a trajectory is chaotic depends strongly on its initial conditions. Chaotic trajectories then wander in a region of phase space the volume of which is nonzero [1, 2, 7]. In dissipative systems practically all trajectories are attracted towards a zero volume object, the attractor, of the phase space. Chaotic attractors, are strange sets [1–3, 8] characterized by noninteger, or fractal dimensions [9]. In this paper we shall be interested in dissipative systems which show up, besides mechanical phenomena, in the theory of nonequilibrium systems exhibiting instabilities leading to a new, turbulent macroscopic state.

In the simplest cases the associated Poincaré or stroboscopic map describes a dynamics of two variables, and the chaotic attractor has a dimensionality between 1 and 2. We derive the most essential features of the map in two important classes of systems, namely in systems where trajectories do not pass close to a singular point, and in systems where they do pass close to a hyperbolic point. The first class contains among many well-known examples the Rössler model [10] and the periodically kicked harmonic oscillator, where even an exact derivation of the map is possible. The maps generated in this class possess analytic forms and a smoothly position dependent area contracting ratio, the Jacobian of the map, which can be considered to be constant near the attractors. This type of maps will be called Hénon type one since Hénon's famous map [11, 12] belongs to this family. A standard example of the second class is the Lorenz model [13], but the Rikitake dynamo [14] and another model of Rössler [15] are also of this type. (For a discussion of general properties of the flow see [16]). The corresponding maps, called Lorenz type maps, are characterized by a singular form and a strongly position dependent Jacobian which vanishes or diverges along a certain line of the plane. The singularities are described by power laws the exponent of which is given by the ratios of the eigenvalues of the linearized equation of motion around the hyperbolic fixed point [17–21].

The asymptotic behaviour in dissipative chaotic systems is described by means of a stationary distribution concentrated on the chaotic attractor. Unfortunately, little is known about the existence and the properties of such distributions for maps of the plane. The situation simplifies considerably in the limiting case of an extremely small Jacobian, which, however, occurs quite often as a consequence of a strong dissipation in the continuous system or of rather long periods of snapshots or of turnover times.

In Hénon type cases the map typically reduces in this limit to a map of the interval defined by a continuous single humped function. Parameter settings when the maximum is mapped in two steps into an unstable fixed point, are of importance, since the chaotic character of the motion may then be shown with mathematical rigor. An essential condition for the existence of a unique stable stationary distribution in this

situation is the negativity of the so-called Schwarzian derivative of the map [5, 6]. In common cases this condition is fulfilled automatically.

If at the above-mentioned parameter setting the map generates chaotic trajectories the case of fully developed chaos [22] is realized. This is, at the same time, a crisis configuration [23] since an unstable orbit collides with the chaotic attractor. Crisis is, however, not restricted to chaotic attractors. It is rather a configuration on a capture/escape boundary [24].

Reduced Lorenz type maps obtained in the limit of vanishing Jacobian are more complicated. Firstly, the dynamics which is still a map of the interval, turns out to have a one-step memory since the state of the system depends also on the sign of the preimage of the variable [21]. Secondly, if a certain internal symmetry is maintained, the memory can be transformed out but the resulting 1D map is not obviously a single-humped one, and its Schwarzian derivative can be also positive. Our numerical simulations show that crisis situations in these cases do not necessarily imply chaos: chaotic regions are interrupted by periodic windows when changing a parameter of the map by keeping its maximum mapped in two steps into an unstable fixed point. In a special case when the effect of a positive Schwarzian derivative has been amplified by making the map discontinuous we were able to find even an asymptotic analytic formula specifying the position of such periodic windows.

Furthermore, we investigated maps with strong but finite dissipation. A perturbative expansion is worked out for the dynamics based on the method of [25] designed originally to determine the shape of chaotic attractors. In a first order calculation in Lorenz type maps we obtain a dynamics with a two-step memory expressed through the presence of the sign of two subsequent preimages as additional discrete variables. In symmetric systems this map can be reduced to a dynamics with a single additional discrete variable. We show that such a dynamics characterizes also the first order approximation of Hénon type maps. In higher order calculations the dynamics will have more and more discrete variables and the attractor will be approximated by a larger and larger number of branches. Finally, this attractor is hard to be distinguished from the chaotic attractor appearing in the map of the plane but using a finite resolution. In this sense unusual maps of an interval with several additional discrete variables, which can also be considered as 1D maps with several discontinuities, may be useful approximations of mappings of the plane. In higher orders, however, to carry out the calculation requires rapidly increasing numerical efforts. Instead, we have followed here an approach in the same spirit but technically easier to handle. Namely, we have considered a discretized approximation of the Hénon map in which one of the recursions is replaced by a step function with several steps, leaving the other recursion unchanged. The condition for crisis and the position and structure of a few periodic cycles have been compared in approximate and exact maps and a satisfactory agreement has been found at step numbers as low as 20–40, illustrating that the discretized version can be a reasonable approximate dynamics.

The paper is organized as follows. Sections 2 and 3 contain the derivation of Hénon and Lorenz type maps, respectively. Section 4 is devoted to the study of reduced symmetric Lorenz type maps in the limit of extremely strong dissipation. Emphasis is laid on a discontinuous version in which periodic windows of the map in crises are analytically specified. The characterization of certain satellite windows is given in the Appendix. In Section 5 the perturbative method valid in the case of strong but finite dissipation is applied for determining approximate dynamics. Finally in Section 6 we study a discretized Hénon map and compare certain properties of it with those of the exact Hénon map.

2. Hénon type maps

The relation between a flow and the associated discrete dynamics can be conveniently studied in periodically kicked systems. For kick lengths negligible on the time scale of the macroscopic motion, the succession of kicks is described by a periodic Dirac delta term in the equation of motion. The effect of a kick is then a jump in the momentum. If the evolution of the system is known between subsequent kicks, an exact form follows for the stroboscopic map.

The map

We consider here the linearly damped one-dimensional harmonic oscillator under the influence of periodic Dirac delta kicks the amplitude of which is position dependent in a nonlinear way. Let T denote the period of the kicks and let $\tilde{f}(x)$ represent the velocity jump caused by a kick acting at an actual position coordinate x . The form of $\tilde{f}(x)$ is arbitrary. By means of the well-known solution of a damped oscillator of eigenfrequency ω_0 and friction coefficient 2γ one obtains for the stroboscopic map (see e.g. [1])

$$\begin{aligned}x' &= xE \left(C + \frac{\gamma}{\omega} S \right) + v \frac{E}{\omega} S, \\v' &= vE \left(C - \frac{\gamma}{\omega} S \right) - xE \frac{\omega_0^2}{\omega} S + \tilde{f}(x'),\end{aligned}\quad (2.1)$$

where x , v denote position and velocity after the n th kick and the following abbreviations have been used

$$\begin{aligned}\omega &= (\omega_0^2 - \gamma^2)^{1/2}, \\C &= \cos \omega T, \quad S = \sin \omega T, \quad E = \exp(-\gamma T).\end{aligned}\quad (2.2)$$

For a general $\tilde{f}(x)$ (2.1) is a nonlinear map of the x , v plane. Its form simplifies considerably by introducing another position type coordinate y defined by

$$y = x \frac{1}{E} \left(C - \frac{\gamma}{\omega} S \right) - v \frac{S}{E\omega} + \tilde{f}(x) \frac{S}{E\omega}.\quad (2.3)$$

This leads to

$$\begin{aligned}x' &= 2ECx + \frac{ES}{\omega} f(x) - E^2y, \\y' &= x.\end{aligned}\tag{2.4}$$

Note that the Jacobian of this map is E^2 , in accordance with the fact that the amplitude of the oscillator decreases in an interval of length T by a factor $\exp(-\gamma T)$. After introducing

$$f(x) \equiv 2ECx + \frac{ES}{\omega} \bar{f}(x), \quad b \equiv -E^2,\tag{2.5}$$

(2.4) can be rewritten in the form

$$\begin{aligned}x' &= f(x) + by, \\y' &= x.\end{aligned}\tag{2.6}$$

We call the map (2.6) of Hénon type since for

$$f(x) = 1 - ax^2\tag{2.7}$$

(in dimensionless units) Hénon's map [11, 12] is recovered, which is known to be the most general quadratic map of the plane with a constant Jacobian. For $f(x) = 1 - a|x|$ (2.6) is the Lozi map [26]. As illustrated by these examples, Hénon type maps often describe chaotic (at least in a numerical sense) behaviour and possess for $|b| < 1$ chaotic attractor in the x, y plane.

The fact that there are no singular points in the phase space of the kicked damped harmonic oscillator turns out to be a crucial property. In all systems the trajectories of which do not pass close to a singular point around the attractor the associated map of the plane, describing the motion on or near the attractor, is expected to be Hénon's map since it must then be analytic in x and y . As an example we mention the Rössler model at its standard parameter values [10, 27], the Poincaré map of which is really of the form of (2.6), (2.7).

The limit of extremely strong dissipation

Strong dissipation means that the Jacobian of the map is tending toward zero. In the limit $b \rightarrow 0$ (2.6) becomes a one-dimensional map

$$x' = f(x),\tag{2.8}$$

since y is then a dummy variable. Such maps of the interval have been extensively studied in the literature from both of the point of view of the bifurcation sequence [28, 29, 6], of universal features [30, 31] and of the properties of the chaotic state [5, 6, 22]. We shall use this type of maps as a point of reference when investigating more general cases.

3. Lorenz type maps

In cases when trajectories on or near the attractor pass close to saddle points the corresponding Poincaré map drastically differs from (2.6). The reason is that a saddle point has at least one invariant hypersurface on which it is attracting. Therefore, trajectories approaching this hypersurface may stay for arbitrarily long time in the vicinity of the saddle and cause singularities in the form of the Poincaré map (Shilnikov's method [32]). As a consequence, the Jacobian of the map will be strongly position-dependent. The standard example of such systems is the Lorenz model [13] where the origin is a hyperbolic point belonging to the Lorenz attractor. General systems with this type of singularity on the attractor, and having no other singularities, we shall call Lorenz type and the corresponding Poincaré map Lorenz type map. Different approximate forms for this map have been deduced by using basically Shilnikov's method [17–21], among which that of [21] seems to be the most general one. It is worth mentioning that systems with a saddle type focus point have also been studied [33–36] and possess singular maps but different from Lorenz type.

The map

For sake of completeness we outline the derivation of the map and refer for the details to [21].

A three-dimensional dynamical system is considered with variables X_i , $i = 1, 2, 3$ the time evolution of which is governed by autonomous ordinary differential equations. Let the origin of the phase space X_1, X_2, X_3 be a hyperbolic point with a two-dimensional stable manifold and a one-dimensional unstable manifold, $W^u(0)$ like in the Lorenz model. For simplicity, the variables X_i are chosen to be the normal modes of the linearized equations around the origin with eigenvalues λ_i , so that X_1 belongs to the unstable mode, i.e. $\lambda_1 > 0$ but $\lambda_2 < \lambda_3 < 0$.

The Poincaré surface is chosen as the $X_3 = z = \text{const}$ plane where z is adjusted in such a way that the unstable manifold of the hyperbolic point should intersect the plane. In a certain reference frame on this plane the coordinates are denoted by x, y . We use the convention that only intersections from above belong to the map. The points D^+ and D^- will be of special importance, where $D^+(D^-)$ represents the first intersection point between the $X_3 = z$ plane and that branch of the unstable manifold $W^u(0)$ which emanates into the positive (negative) X_1 direction (Fig. 1).

Since the plane $X_3 = z$ is generally outside the region where the motion can be well approximated by the linearized equations around the hyperbolic fixed point, we introduce an auxiliary surface defined by $X_3 = Z$, where Z is a sufficiently small constant. The reference frame X, Y on this surface is chosen in such a way that the origin $X = Y = 0$ is the intersection point of the plane and the X_3 axis, and the $X(Y)$ axis is parallel with the $X_1(X_2)$ axis.

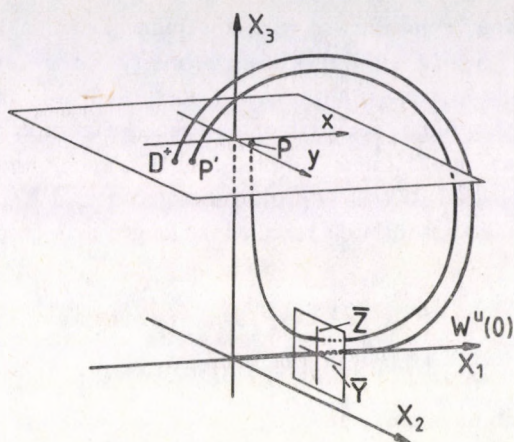


Fig. 1. The unstable manifold $W^u(0)$ and a trajectory passing close to the hyperbolic point in Lorenz type systems

Trajectories passing close to the hyperbolic point must start from the neighbourhood of the stable manifold. The trajectory emanated from $P=(x, y)$ crosses the $X_3=Z$ plane at a certain point (X, Y) . More generally, the flow generates a map $X=X(x, y)$, $Y=Y(x, y)$ between the two surfaces, where $X(x, y)$ and $Y(x, y)$ are smooth functions of their variables since no singular point lies between the planes $X_3=z$ and $X_3=Z$. An appropriate choice for Z always guarantees that both coordinates X, Y of the intersection with the auxiliary plane will be small. The subsequent motion of the point $(X(x, y), Y(x, y), Z)$ is thus described by the solution of the linearized equations around the origin.

As, after having left the hyperbolic point, the trajectory does not pass near any singular point, the deviation between the next intersection $P'=(x', y')$ with the Poincaré surface and $D^+(D^-)$, if $X(x, y)>0$ (if $X(x, y)<0$), is an analytic function of the coordinates \bar{Y} and \bar{Z} (Fig. 1). For small values of X and Y it is sufficient to keep the first terms of the Taylor expansion only, apart from exceptional cases when their coefficients vanish. Finally, one finds [21] as a typical form of the map near $X(x, y)=0$

$$\begin{aligned}x' &= (u + a_{11}|X(x, y)|^\beta) \operatorname{sgn}(X(x, y)) + a_{12}Y(x, y)|X(x, y)|^\delta, \\y' &= (v + a_{21}|X(x, y)|^\beta) \operatorname{sgn}(X(x, y)) + a_{22}Y(x, y)|X(x, y)|^\delta,\end{aligned}\quad (3.1)$$

where

$$\beta = |\lambda_3|/\lambda_1, \quad \delta = |\lambda_2|/\lambda_1, \quad (3.2)$$

$\operatorname{sgn}(X)$ denotes the sign of X , the coefficients a_{ij} are constants, and u, v are the coordinates of the point D^+ . For the sake of simplicity we assumed by writing down

(3.1) that the equations of motion are invariant under the transformation $X_1 \rightarrow -X_1$, $X_2 \rightarrow -X_2$, $X_3 \rightarrow X_3$, which is a well-known property of the Lorenz model.

The form of the map (3.1) simplifies considerably if we consider the recursions only in a small region around the origin of the Poincaré surface. The functions $X(x, y)$ and $Y(x, y)$ are then given as linear combinations of x and y , and we may choose the reference frame such that $X(x, y)$ is proportional to x and $Y(x, y)$ to y in the new variables. After having appropriately rescaled the length scales used in both directions, one arrives at

$$\begin{aligned}x' &= (-e + a|x|^\beta) \operatorname{sgn}(x) + cy|x|^\delta, \\y' &= (d + |x|^\beta) \operatorname{sgn}(x) + by|x|^\delta,\end{aligned}\quad (3.3)$$

with an x -dependent Jacobian

$$J(x) = (ab - c)\beta|x|^\eta, \quad (3.4)$$

where $\eta = \beta + \delta - 1$. In the following $a, b, e > 0$, $ab \geq c$ will be assumed.

The general form of the Poincaré map contains, of course, additional terms, analytic or less singular as those given already by Eq. (3.3). In order to illustrate the consequences of the singular feature of the map, however, it is sufficient to keep the most singular part. Therefore, we consider in the following the map obtained by extending the validity of equation (3.3) to the whole plane. More precisely, we regard the map (3.3) as a model which is designed to simulate some essential features of Lorenz type systems. In numerical simulations the map was found to possess a chaotic attractor at several values of the parameters [21] (as an example see Fig. 2).

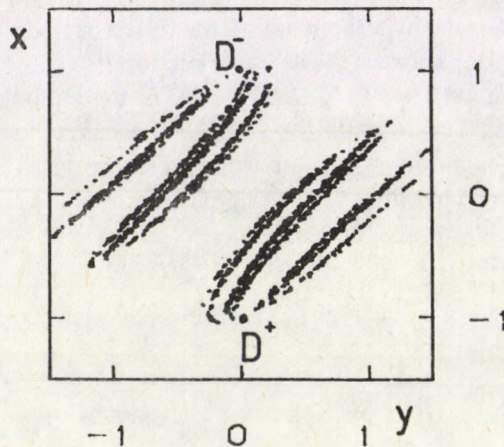


Fig. 2. The chaotic attractor of the map (3.3) obtained in a numerical simulation after 2000 steps. The parameters are $a = 1.5$, $b = 0.7$, $c = 0.25$, $d = 0$, $e = 1$, $\beta = 0.6$, $\delta = 0.2$

The limit of extremely strong dissipation

The limiting case of an identically vanishing Jacobian is realized if the parameters fulfil the relation $c = ab$. It follows then from (3.3) that any starting point jumps immediately on one of the straight lines $x = ay \mp (e + ad)$ where the attractor is situated. The sign $+(-)$ is to be taken if the point (x, y) is on the lower (upper) branch. Once, however, a point is on the lower (upper) branch, the x -coordinate of its preimage has to be positive (negative). Thus, the sign before the parenthesis is identical with that of the preimage of x .

Using this relation between x and y as well as (3.3) x' can be expressed in terms of x . We obtain

$$x' = \operatorname{sgn}(x) (-e + a f_{\sigma \operatorname{sgn}(x)}(|x|)), \quad (3.5)$$

where

$$f_{\pm}(|x|) = |x|^{\beta} + \frac{b}{a} (|x| \pm (e + ad)) |x|^{\delta} \quad (3.6)$$

and σ denotes the sign of the preimage of x . The y' coordinate is then determined by x' through

$$y' - \operatorname{sgn}(x)d = (x' + \operatorname{sgn}(x)e)/a. \quad (3.7)$$

The special two-step nature of the dynamics can be made clearer by rewriting (3.5) as

$$\begin{aligned} x' &= (-e + a|x|^{\beta})\operatorname{sgn}(x) + bx|x|^{\delta} + g\sigma|x|^{\delta}, \\ \sigma' &= \operatorname{sgn}(x), \end{aligned} \quad (3.8)$$

where $g = b(e + ad)$. Considering (3.8) as a map of the interval, it can be specified by the functions $f_{\pm}(x)$ given in (3.6) and by the rule that $f_{+}(f_{-})$ is to be taken if $\sigma \operatorname{sgn}(x)$ is positive (negative) (see [21] where also plots of f_{\pm} are given). Without any internal symmetry, a two-step dynamics cannot be simplified further.

Owing, however, to the symmetry property mentioned after (3.2), the form of the map may be reduced. By introducing $p = -\sigma x$ as a new variable branches of f_{\pm} pairwise coincide. This procedure is similar in spirit to Lanford's treatment of the Lorenz model [37]. The recursion for p is then obtained as

$$p' = h(p) \equiv e - a|p|^{\beta} - b|p|^{1+\delta} + g \operatorname{sgn}(p) |p|^{\delta}, \quad (3.9)$$

which is characterized by a single-valued continuous (but asymmetric) function, h .

The representation (3.9) is well suited for discussing the basic questions of the existence of a unique stationary probability distribution with a density since this problem has extensively been studied in the case of such continuous 1D maps.

Four cases should be distinguished: a) $\beta < \delta$, $\beta < 1$ (like in the standard Lorenz model); b) $\beta < \delta$, $\beta > 1$; c) $\beta > \delta$, $\delta > 1$; d) $\beta > \delta$, $\delta < 1$ (Fig. 3). In case a) the map has a cusp,

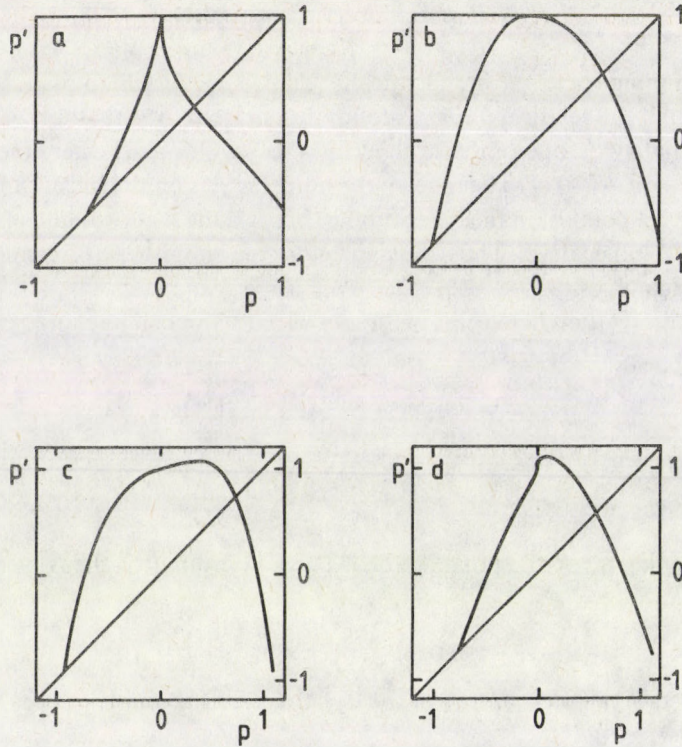


Fig. 3. Qualitatively different shapes of $p' = h(p)$ (3.9) obtained at different choices of the exponents β and δ . The parameters a, b are adjusted in such a way that the maximum is mapped in two steps into the negative fixed point; $e = 1, d = 0$ everywhere.

Case a): $\beta = 0.5, \delta = 1.5, a = 1.560, b = 0.6,$

Case b): $\beta = 2, \delta = 3, a = 1.839, b = 0.5,$

Case c): $\beta = 4, \delta = 1.2, a = 1.430, b = 0.5,$

Case d): $\beta = 1.5, \delta = 0.5, a = 1.491, b = 0.5.$

(Numerically all cases seem to be chaotic at these values)

while in cases b), c) and d) it has a smooth maximum (at certain special choices of the parameters two or three local maxima may be present). In the former case, if the map is everywhere expanding, well-known theorems apply and the existence of the unique probability density is ensured (see [5, 6] and references therein). In the latter cases the map can produce chaotic iterations for typical initial conditions only at particular control parameter values. The situation when the maximum point is mapped in two steps to an unstable fixed point, i.e., when fully developed chaos can exist, has been most extensively studied. Two basic conditions under which the existence of a unique absolutely continuous invariant measure has been proved in this situation are that the first derivative of the map is nonzero except at the maximum and that its Schwarzian derivative is negative [5, 6]. Both can be valid in case b), but in cases c) and d) the first and the second condition is violated near $x = 0$, respectively.

4. A discontinuous limiting case of Lorenz type maps

Our aim is here to investigate in some detail the characteristic features of maps belonging to case *d*. In the limit of strong dissipation — as it was pointed out in the previous section — the one-dimensional map $p' = h(p)$ as given by (3.9) has a positive Schwarzian derivate around $p=0$. In order to study the consequences of this property we turn to the extreme case of $\delta=0$ where the rapidly changing part of $h(p)$ around $p=0$ is replaced by a jump. After setting $e=1$ and choosing, as a typical value, $\beta=2$, the map still has three independent parameters: a defining the quadratic part, b giving the modulus of the slope at $p=0$, and g characterizing the jump there. For a sufficiently large value of b the map can be everywhere expanding. An interesting interplay between periodic and chaotic behaviour is, therefore, expected to be present for small b -s. The phenomena found at $b \leq 1$ are qualitatively similar to those at $b=0$, thus in the following we consider the map

$$p' = h(p) \equiv 1 - ap^2 + g \operatorname{sgn}(p). \quad (4.1)$$

We restrict our attention to the region $0.8 < a \leq 2$, $0 \leq g \leq 1$. The plot of $h(p)$ is displayed in Fig. 4. (4.1) is a straightforward extension of the parabola map $p' = 1 - ap^2$ which is recovered as a special case for $g=0$.

The situation when the maximum point of a single humped continuous map is mapped in two steps to an unstable fixed point is of importance since fully developed chaos [22] may then exist. This corresponds, at the same time, to a crisis situation [23] since an unstable orbit collides with the chaotic attractor. Such a crisis configuration may also be found in (4.1) with an appropriate choice of the parameters a and g but, as we shall see, it does not necessarily imply chaos.

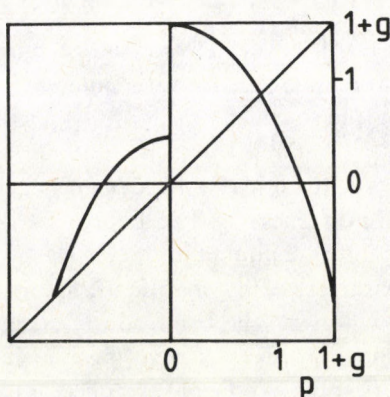


Fig. 4. Plot of the function $h(p)$ defining (4.1). A crisis situation is shown at $a=1.140$, $g=0.556$

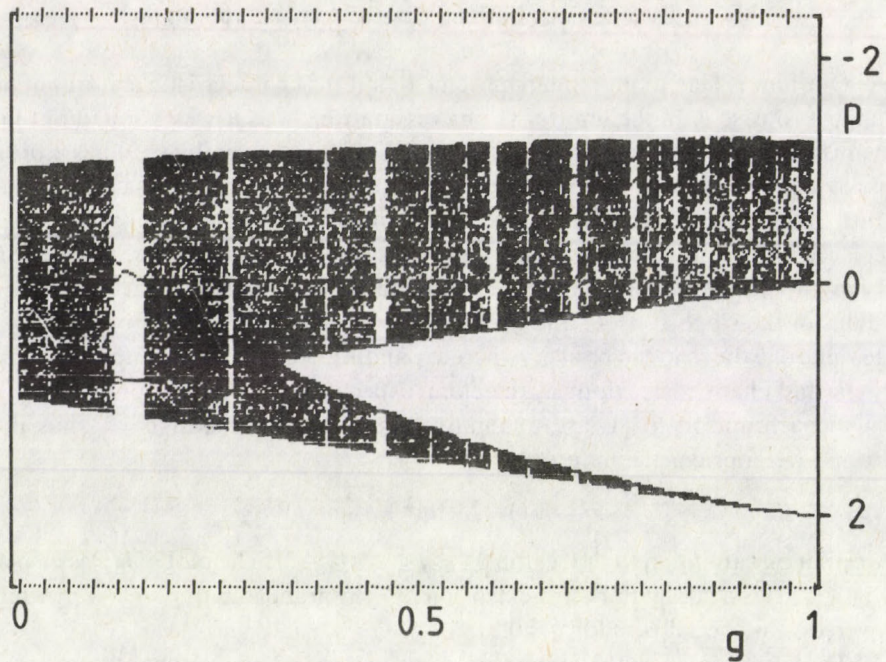


Fig. 5. Bifurcation diagram along the crisis line of (4.1) in the region $0 \leq g \leq 1$. A regular sequence of relatively broad periodic windows can be seen

The maximum point $h(+0) = 1 + g$ is mapped into the fixed point $p^* = -(1 + \sqrt{1 + 4a(1 - g)}) / (2a)$ if the relation

$$a^3(1 + g)^4 - 2a^2(1 + g)^3 + 2g = 0 \quad (4.2)$$

is fulfilled. The solution $a_c(g)$ of (4.2) defines a crisis line in the parameter plane a, g . For $a > a_c(g)$ trajectories are able to escape from the interval $(p^*, 1 + g)$. Only trajectories belonging to a Cantor set (the so-called repeller [38]) remain then bounded, provided no attracting cycle exists.

We have investigated numerically the dynamics generated by (4.1) along the crisis line $a = a_c(g)$ at different values of g . The 'bifurcation diagram' of Fig. 5 shows the results, where attractor points are plotted in the vertical direction. It is clear from this diagram that regions characterized by chaotic attractors (in a numerical sense) are interrupted by periodic windows. The condition of crisis is for maps like (4.1) not sufficient for chaotic behaviour. This is clear by noticing that the large negative contribution to the Lyapunov exponent of trajectories starting from the right neighbourhood of the origin is compensated by staying for a long time around the unstable fixed point, but this is no longer true for trajectories starting from the left neighbourhood of $p = 0$, which may, thus, produce a negative Lyapunov exponent at certain values of g , along the crisis line.

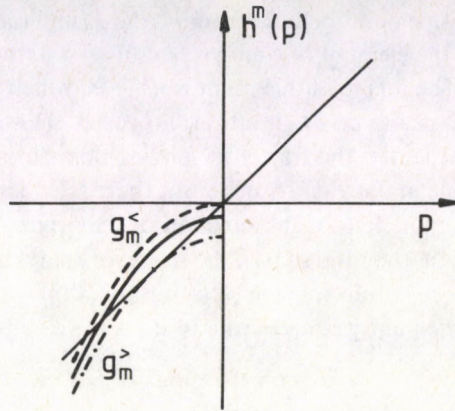


Fig. 6. The shape of the m th iterate of $h(p)$ in a left neighbourhood of the origin (bold line) for $g_m^< < g < g_m^>$

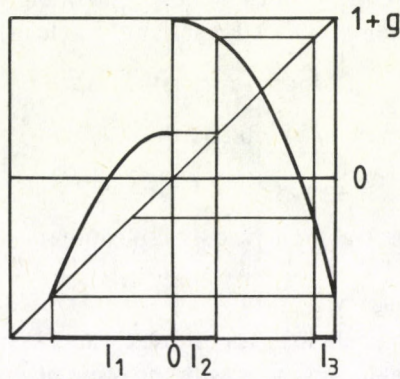


Fig. 7. The action of the map (4.1) on the interval I_1 in crisis configuration

There is a striking regularity in the sequence of periodic windows for $g \rightarrow 1$ (Fig. 5). In the following we study how to specify these windows. First, it is to be noted that the m th iterate of $h(p)$, the stable fixed point of which is an element of the period- m cycle of the original map, has the shape sketched in Fig. 6 in the vicinity of the origin. When increasing g the left branch of $h^m(p)$ is shifted downwards and at a certain value $g_m^<$ (dashed line) $h^m(-0)$ reaches the origin. Then a stable fixed point, and by that a stable limit cycle appears. By increasing g further, it exists till a certain $g_m^>$, the stable and unstable fixed points of h^m coincide. The m -cycle disappears at $g_m^>$ via a tangent bifurcation (dashed dotted line).

Next, we specify the position of the periodic windows. For $g > g_3^<$ the intervals I_1 , I_2 , I_3 shown in Fig. 7 are mapped as follows:

$$I_1 \rightarrow I_1 + I_2, \quad I_2 \rightarrow I_3, \quad I_3 \rightarrow I_1. \quad (4.3)$$

To every periodic cycle one point belongs from I_2 and I_3 and the rest of points from I_1 . Let us consider a cycle the element of which is $h(-0) = 1 - g$ (this is the case at $g = g_m$). For $g \rightarrow 1$ $1 - g \equiv \varepsilon \ll 1$. The image of this point is $h^2(-0)$, which is mapped into $h^3(-0) = h(1 + g) - h'(2 - \varepsilon)a\varepsilon^2$. In the crisis situation, however, $h(1 + g) = p^*$, thus $h^3(-0) = p^* - h'(2 - \varepsilon)a\varepsilon^2$. Consequently, the trajectory passes close to the unstable fixed point p^* . It will stay for a while in the vicinity of p^* but there is also another region where it will be captured, namely the left neighbourhood of the origin (where the map is flat).

Let us estimate first, the time steps k the trajectory stays around p^* . Since $f'(p^*) = 2$, $h(p) - h(p^*) \cong 2(p - p^*)$ and for the k th iterate $h^k(p) - h^k(p^*) \cong 2^k(p - p^*)$. This difference will be of order unity for $p = h^3(-0)$ if

$$k = \text{const} - 2 \log_2 \varepsilon. \quad (4.4)$$

The trajectory must return to $p = -0$. Therefore, after leaving the vicinity of p^* it must reach a point \bar{p} for which $h^l(\bar{p}) = -0$. Since the map is $h(p) = \varepsilon - ap^2$ for $p < 0$, $h^{l-1}(\bar{p}) = -(\varepsilon/a)^{1/2}$ holds. On the other hand, by iterating forward $p = \bar{p}$ one finds $h^l(\bar{p}) = -(a|\bar{p}|)^\omega a^{-1}$, where $\omega = 2^{l-1}$. From here it follows that

$$l = \log_2(-\log_2 \varepsilon) + \text{const}. \quad (4.5)$$

For the total length of a cycle m we obtain, thus,

$$m = -2 \log_2 \varepsilon + \log_2(-\log_2 \varepsilon) + \text{const}. \quad (4.6)$$

for $\varepsilon \rightarrow 0$. The inverse of this relation yields the position of the m -cycle at

$$g_m^< = 1 - \varepsilon_m = 1 - cm^{1/2} 2^{-m/2} \quad (4.7)$$

which is valid for asymptotically long trajectories. Table I contains this approximate value for cycles of finite length compared with the result of a numerical solution of $g_m^<$.

These considerations show that there is an infinite sequence of periodic windows in the map (4.1) when changing the parameters in such a way that the condition for crisis (4.2) is maintained. Note that there is a geometric sequence of periodic windows also in the parabola map for $a \rightarrow 2$ [31], however, in a pre-crisis situation. The sequence described by (4.7) is of quite different nature. The presence of the prefactor $m^{1/2}$ is a consequence of the smooth local maximum of the map at $p = -0$ while the exponent $-m/2$ follows from the fact that the map is discontinuous at the origin. It is worth mentioning the result obtained for an arbitrary exponent $\beta > 1$, i.e. for the map $h(p) = 1 - a|p|^\beta + g \text{sgn}(p)$. The same argumentation then yields

$$g_m^< = 1 - cm^\rho (h'(p^*))^{-m/\beta} \quad (4.8)$$

with

$$\rho = \frac{\log_2 h'(p^*)}{\beta \log_2 \beta}. \quad (4.9)$$

Table I

Stable periodic orbits in (4.1) along its crisis line. m denotes the length of the periodic orbit, g_m is the control parameter value where it appears by increasing g . The third column contains the results of the formula (4.7) with the fitted value $c = 1.05$. Numbers in the last column yield a measure of the accuracy of (4.7)

m	$g^<$	$1 - 1.05(m2^{-m})^{1/2}$	$\frac{1.05(m2^{-m})^{1/2}}{1 - g^<}$
3	0.452875	0.357009	1.175
4	0.593984	0.475000	1.293
5	0.696769	0.584951	1.368
6	0.771194	0.678504	1.405
7	0.826018	0.754454	1.411
8	0.867098	0.814384	1.396
9	0.898288	0.860788	1.368
10	0.922182	0.896238	1.333
11	0.940582	0.923048	1.295
12	0.954777	0.943167	1.256
13	0.965722	0.958172	1.220
14	0.974137	0.969307	1.186
15	0.980580	0.977535	1.156
16	0.985489	0.983594	1.130
17	0.989207	0.988042	1.107
18	0.992008	0.991299	1.088
19	0.994105	0.993679	1.072
20	0.995668	0.995414	1.058
21	0.996827	0.996677	1.047
22	0.997682	0.997595	1.037
23	0.998311	0.998261	1.029
24	0.998772	0.998744	1.022
25	0.999109	0.999094	1.016
26	0.999354	0.999346	1.011
27	0.999533	0.999529	1.008
28	0.999663	0.999661	1.005
29	0.999757	0.999756	1.003
30	0.999824	0.999824	0.9982

Since, however, for $g \rightarrow 1$ the fixed point is $p^* = -a^{-\frac{1}{\beta-1}}$, $h'(p^*) = \beta$ and, thus

$$g_m^< = 1 - cm^{1/\beta} \beta^{-m/\beta} \quad (4.10)$$

follows for the position of the m -cycle.

Satellite series of windows approaching $g_m^<$ and $g_m^>$ can also be specified. A detailed description of them is given in the Appendix by means of itinerary sequences.

Finally, it is worth emphasizing that an interplay between chaotic and periodic regions along the crisis line is characteristic not only for the map (4.1) but also for certain subclasses of (3.9). Namely, the same phenomenon can be found in maps with $\delta \ll 1$ and $\beta > 1$ and also in maps where $h(p)$ possesses two points with a vanishing first derivative (case c).

5. Discrete dynamics on branched manifolds

Chaotic attractors of systems with finite dissipation consist of infinitely many branches arranged in a fractal structure. If, however, the dissipation is strong only a few branches can be observed owing to the limited accuracy of the measurement. It may then be convenient to consider these branches to be infinitely narrow. They are called branched manifolds [16]. The dynamics on such branched manifolds is, of course, noninvertible since several original branches are considered as identical. Our aim is here to specify branched manifolds of Lorenz and Hénon type maps and to determine the dynamics on them. We shall apply the perturbative method worked out in [25] which determines the shape of the attractor as a power series in the inverse dissipation strength. The dynamics will turn out to be a map of the interval which depends, however, also on certain discrete variables, in lowest orders on the sign of subsequent preimages of the continuous variable.

Lorenz type maps

We start to investigate this type of maps since here even in the limit of extremely strong dissipation a dynamics with a special two-step nature has been found (cf. (3.8)). The deviation from this limiting case is measured now by

$$\varepsilon = ab - c, \quad (5.1)$$

a factor of the Jacobian (3.4). From (3.3) it follows that

$$x' - ay' + \operatorname{sgn}(x)g/b = -\varepsilon y|x|^\delta. \quad (5.2)$$

In a first order calculation in ε , it is sufficient to use the 0th order result (3.5) and (3.7) on the right hand side. Thus, we obtain for the shape of the attractor

$$x = ay - \sigma g/b - \frac{\varepsilon}{a} (\sigma f_{\sigma\tau}^{-1}(y\sigma - d) + \tau g/b) (f_{\sigma\tau}^{-1}(y\sigma - d))^\delta, \quad (5.3)$$

where f_{\pm}^{-1} denotes the inverse of f_{\pm} defined in (3.6) and σ and τ stand for the sign of the first and second preimage of x , respectively. An elimination of y from the first equation of (3.3) by means of (5.3) yields the dynamics of the x -variable as

$$x' = (-e + a|x|^\beta) \operatorname{sgn}(x) + bx|x|^\delta + \sigma g|x|^\delta + \frac{\varepsilon}{a} |x|^\delta \left[-x - \sigma g/b + \left(\sigma f_{\sigma\tau}^{-1} \left(\frac{e + x\sigma}{a} \right) + \tau(e + ad) \right) \right] \left(f_{\sigma\tau}^{-1} \left(\frac{e + x\sigma}{a} \right) \right)^\delta, \quad (5.4)$$

$$\sigma' = \operatorname{sgn}(x),$$

$$\tau' = \sigma.$$

Note that both the shape and the dynamics depend on two discrete variables σ and τ , and that for extremely strong dissipation, $\varepsilon = 0$, the dependence on the second preimage disappears.

Since the map (3.3) is inversion symmetric, the memory of the dynamics (5.4) can be reduced by introducing again, like in Section 3, $p = -\sigma x$ as a new variable. In this representation (5.4) has the form

$$p' = e - a|p|^\beta - b|p|^{1+\delta} + g \operatorname{sgn}(p) |p|^\delta + \\ + \frac{\varepsilon}{a} |p|^\delta \left\{ |p| + \operatorname{sgn}(p) \left[-g/b + \left(f_{-\rho}^{-1} \left(\frac{e-p}{a} \right) - \rho g/b \right) \left(f_{-\rho}^{-1} \left(\frac{e-p}{a} \right) \right)^\delta \right] \right\}, \quad (5.5) \\ \rho' = \operatorname{sgn}(p).$$

The p -dynamics has only a one-step memory since it depends on (the sign of) the immediate preimage only. Anyhow, these results show that in the case of a strong but finite dissipation, when the branched manifold of the extremely dissipative case splits into two, the number of discrete variables specifying the dynamics (either in x or in p) on the manifolds increases by one.

Hénon type maps

We investigate here maps of the class of (2.6) defined, as typically, by a single humped $f(x)$ the maximum of which is chosen to be at $x=0$. As for the shape of the attractor, the special case of Hénon's map has been extensively studied in the literature [25, 39].

Since the Jacobian J is constant in this class, the quantity $b = -J$ can be considered as the small parameter. As we have seen, in the limit of extremely strong dissipation the shape of the attractor is $x = f(y)$ and the dynamics on it is given by $x' = f(x)$. In first order, y is to be expressed through $x = f(y)$ leading to the dynamics

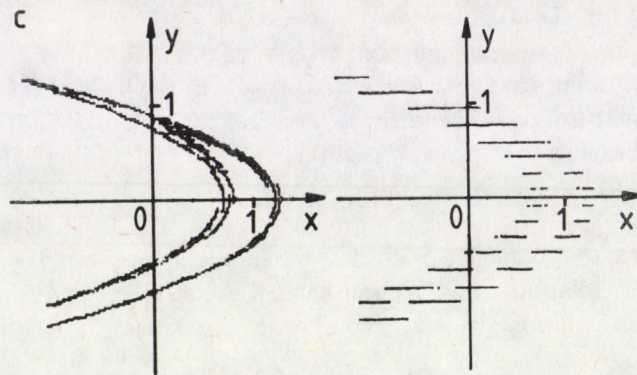
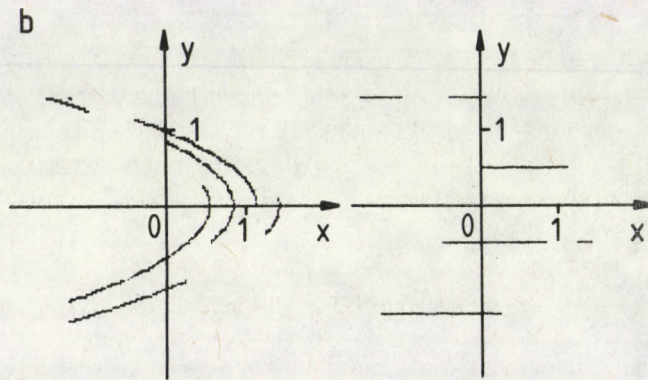
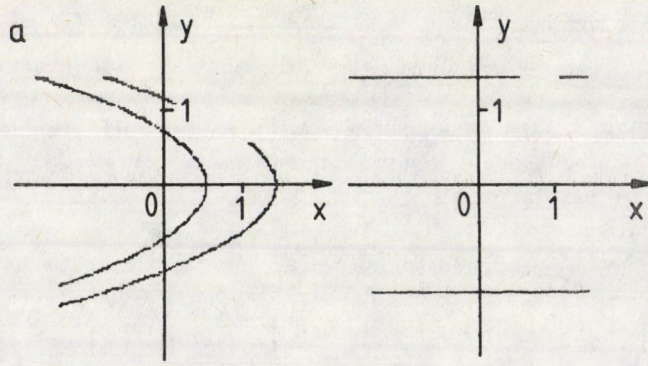
$$x' = f(x) + b f_\sigma^{-1}(x), \quad (5.6)$$

where f_σ^{-1} denotes that branch of f^{-1} where the sign of y is σ . Since $y' = x$, $\sigma' = \operatorname{sgn}(x)$. By means of these relations and (5.6) the shape of the attractor immediately follows.

A similar calculation yields in second order the dynamics

$$x' = f(x) + b f_\sigma^{-1}(x - b f_\tau^{-1}(f_\sigma^{-1}(x))), \quad (5.7)$$

where τ is the sign of the second preimage of x : $\tau' = \sigma$. This shows that inspite of the differences between Lorenz and Hénon type maps the dynamics on branched manifolds is, in both cases, a map of the interval with additional discrete variables; in lowest order the sign of preimages, the number of which increases with the number of branched manifolds. From this point of view there is only a minor difference between



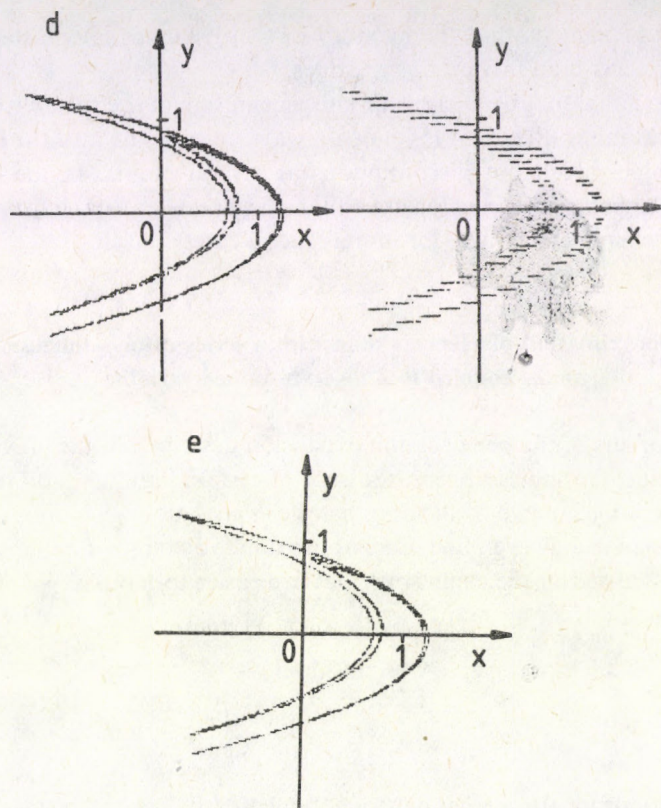


Fig. 8. The chaotic attractor obtained in numerical simulations of (6.1) (shown in the right column) and of (6.2) (left column) at increasing step numbers (m or N) in crisis configuration.

- a) $N = 2$ ($m = 0.35$), $a = a_c = 1.068$, $b = 0.3$,
- b) $N = 4$ ($m = 1.05$), $a = a_c = 1.164$, $b = 0.3$,
- c) $N = 16$ ($m = 5.85$), $a = a_c = 1.404$, $b = 0.3$,
- d) $N = 38$ ($m = 14.45$), $a = a_c = 1.413$, $b = 0.3$,
- e) The Hénon map at $a = a_c = 1.427$, $b = 0.3$.

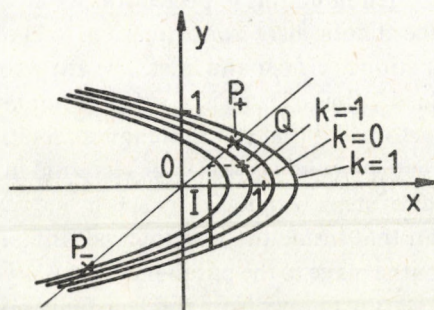


Fig. 9. The parabolas $x = 1 + b(k + 1/2)/m - ay^2$ for $k = -2, -1, 0, 1$, the preimage I of a point Q , and the fixed points P_+ , P_- of (6.2). The parameter values agree with those of Fig. 8b

these two classes, namely, that the memory of the dynamics is one step shorter in Hénon type systems than in Lorenz type ones.

It is to be noted that by an appropriate redefinition of the variable, multivalued one-dimensional maps (like (5.5)–(5.7)) can always be brought into the form of (2.8) where f is single-valued but discontinuous at several points. In the light of this comment (4.1) appears to be a prototype of such systems the qualitative features of which may be characteristic also for more general cases.

6. Approximation of Hénon's map with a series of one-dimensional maps coupled to a discrete-valued variable

In high orders of the perturbation expansion described in the previous Section several two-valued variables appear, the sign of certain preimages of a continuum-valued variable, which may be difficult to handle in analytic calculations or numerical simulations. One can, however, find other approximate dynamics for maps of the plane which are constructed in the same spirit but are easier to handle. We consider here approximations for Hénon's map. Firstly, let us take

$$x' = 1 - ax^2 + by, \quad (6.1a)$$

$$y' = S(x) \equiv \frac{\text{Int}(mx) + 1/2}{m}, \quad (6.1b)$$

where $\text{Int}(q)$ stands for the integer part of q . Note that the second recursion, $y' = x$, of (2.6) has been replaced by a piecewise constant form (6.1b) with segment length $1/m$. In the following another approximation will be used where the first equation of (2.6) is modified and the second one remains unchanged:

$$\begin{aligned} x' &= 1 - ax^2 + bS(y), \\ y' &= x. \end{aligned} \quad (6.2)$$

An approach toward the Hénon map is expected for $m \gg 1$.

Figure 8 displays the chaotic attractor of (6.1) and (6.2) obtained numerically by making the multi-step function $S(x)$ finer and finer. The attractors are plotted in a crisis configuration to be defined below. There is a striking similarity in the geometrical appearance of these attractors and that of the Hénon map [12] at already rather low step numbers, especially in the second version (6.2). Here and in the following it is more natural to use the number of steps N induced by $S(x)$ inside the maximal extension of the chaotic attractor than that inside the unit interval ($\text{Int}(m)$).

Equation (6.2) maps the plane to the parabolas $x = 1 + b(k + 1/2)/m - ay^2$, where k is an integer. An unusual feature follows from this fact. Only points on these parabolas possess preimages, and such a preimage is a straight line segment (see Fig. 9). Consequently, the chaotic attractor of (6.2) lies on a finite number of parabolas only,

and the dynamics on the attractor is essentially a map of the interval coupled to an extra multi-valued variable.

We have studied in detail the attractors of (6.2) in crisis configuration. Crisis in two-dimensional maps is defined by the existence of heteroclinic tangencies of stable and unstable manifolds [40, 23]. In Hénon's map the crisis when the unstable manifold W_+^u of the hyperbolic fixed point P_+ touches the stable manifold W_-^s of the hyperbolic fixed point P_- , situated in the negative quadrant, defines a critical line $a = a_c(b)$ in the parameter space. Beyond this line no one-piece chaotic attractor, which is the closure of W_+^u [40], may exist.

In the map (6.2) there are hyperbolic fixed points analogous with those of the Hénon map. Due to the special multi-step nature of (6.2) the stable manifold W_-^s is discontinuous and consists of straight line segments, which are the preimages of the fixed point P_- . The unstable manifold W_+^u is also discontinuous and consists of parabola segments. (If there exists a chaotic attractor which contains P_+ , it agrees in this case with W_+^u .)

In a certain region of the parameter space these manifolds have no common points (Fig. 10a).

Let w denote the highest index k at which the maximum point of the corresponding parabola still belongs to W_+^u . There are two topologically different possibilities for tangencies between W_+^u and W_-^s .

Case a): The w th parabola touches W_-^s ,

Case b): The endpoint of a segment of W_+^u (with $k > w$) collides with W_-^s .

It is clear from (6.2) that together with a tangency also intersections appear since the preimages of a point of tangency are segments of straight line. Figure 10b illustrates case a).

Beyond such a special configuration tangencies turn to intersections. Simultaneously, the region from which trajectories escape to infinity (a part of which is displayed in Fig. 10c as shaded area) overlaps with W_+^u . Therefore, a chaotic attractor containing P_+ cannot exist then (and gives way for a chaotic repeller [38]).

Thus, for maps like (6.2) the crisis line $a = a_c(b)$ is to be defined as the borderline of that region of the parameter space where W_+^u and W_-^s have no common points. The crisis configuration realized at $a = a_c(b)$ is then characterized by a simultaneous appearance of heteroclinic tangencies and intersections but without any escape.

The critical value $a_c(b)$ obtained at a fixed b for increasing step numbers N (which can be expressed in terms of other parameters as $N = 2m + 1 + (2w + 1)b$) have been found to approach rapidly the crisis value for the Hénon map at the same b (for data see caption of Fig. 8).

We have determined numerically also the bifurcation diagram along the crisis line at different fixed values of the step number N . Fig. 11 illustrates at $N = 8$ that there is an interplay between chaotic and periodic regimes. Thus, in this system crisis is again not a sufficient condition for chaos (not even in a numerical sense).

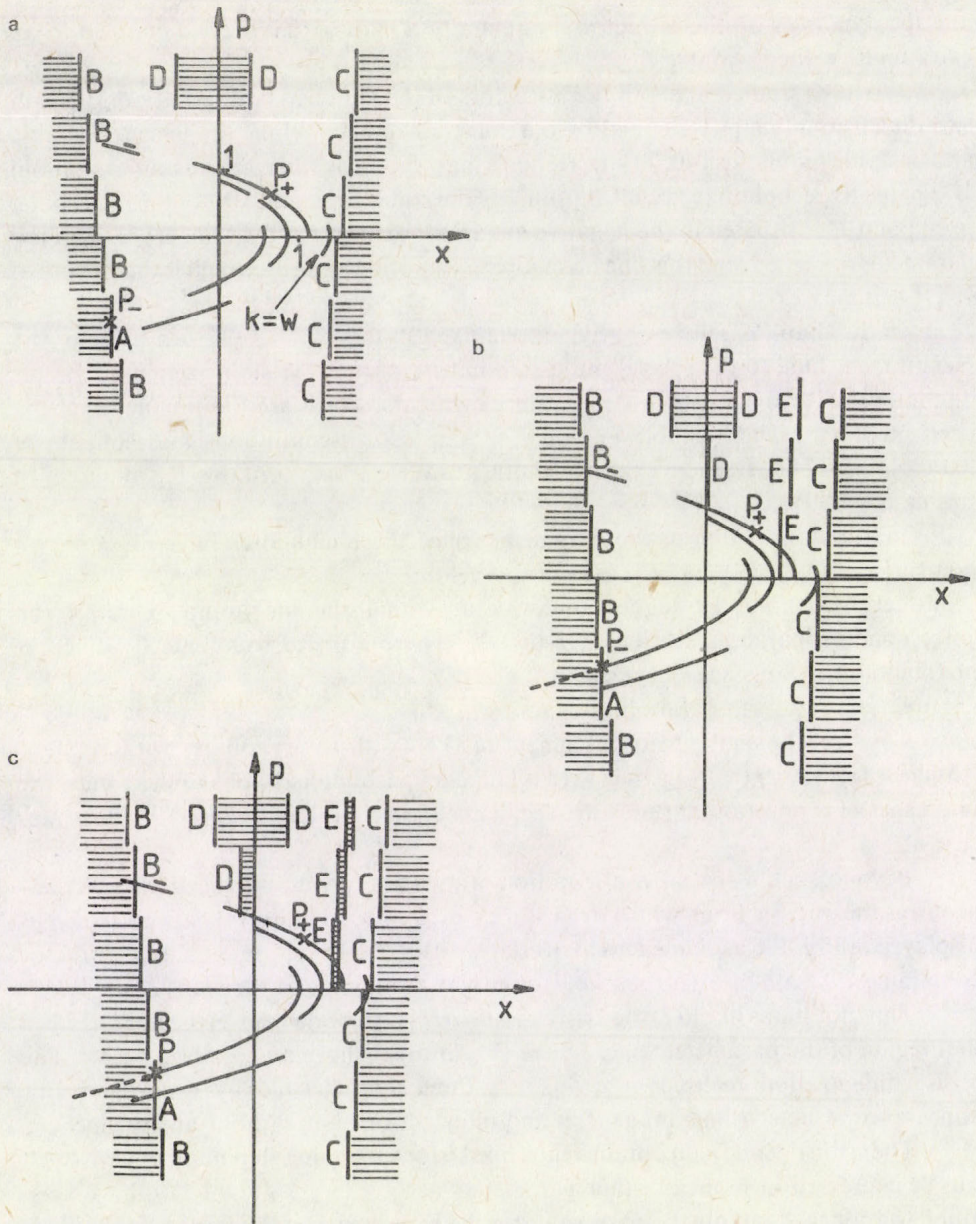


Fig. 10. The unstable manifold W^u (bold line) and a few branches of the stable manifold W^s (straight line segments) in (6.2) at different values of a ($b=0.3$). The segment A is the first preimage of P_- , the B -s are its second preimages. Subsequent further preimages are denoted by C , D and E . Trajectories starting from the dashed region are mapped in at most 5 iterations on the parabola arch drawn by a dashed line, where they go to infinity from.

- a) $a < a_c = 1.164$, no heteroclinic points,
- b) $a = a_c$, heteroclinic tangencies and intersections, but no escape from W^u ,
- c) $a > a_c$, escape occurs

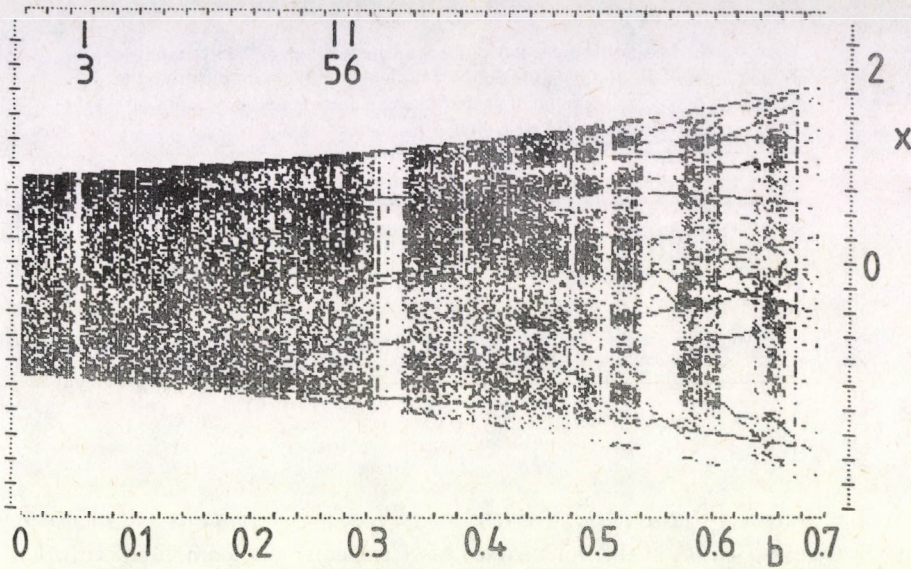


Fig. 11. Bifurcation diagram of (6.2) along its crisis line obtained at a fixed step number $N=8$. The slashes on the top denote the position of stable 3,5,6-cycles in the Hénon map

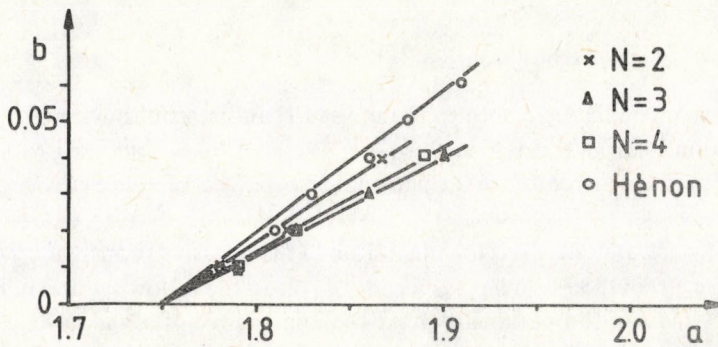


Fig. 12. Position of the stable 3-cycle in the parameter plane a, b for the Hénon map and for (6.2) with step numbers $N=2, 3, 4$

The position of the largest periodic windows has been specified and compared with that of Hénon's model appearing along its crisis line. It can be seen from Table II that, although not all periodic cycles of the Hénon map can be observed in (6.2) at arbitrary values of N , if a window disappears at a certain N it reappears at a somewhat larger one and the values specifying its position tend toward those of the Hénon map.

Table II

A few stable periodic orbits in (6.2) found at different values of the step number N , and those of Hénon's map. m denotes the length of the periodic orbit, the couples (a, b) specify that point of the parameter plane where the cycle appears along the crisis line when increasing b

m	$N=8$	$N=10$	$N=16$	$N=24$	$N=39$	Hénon
3	$a=1.891$ $b=0.052$	1.893 0.047	1.893 0.051	1.895 0.050	1.895 0.051	1.895 0.053
5	—	1.50 0.23	1.46 0.27	—	— 0.26	1.48 0.27
6	1.37 0.32	—	—	1.50 0.25	—	1.46 0.28

The line of a stable period-3 cycle in the parameter plane a, b has also been determined for $a \leq a_c(b)$ at different values of N and has been compared with that of Hénon's map [41]. Fig. 12 shows a qualitative agreement already at rather small step numbers.

These investigations support the view that the map (6.2) can be a reasonable approximation of Hénon's map for not too weak dissipations.

Appendix

In this Appendix we return to the map (4.1) and restrict ourselves again to the crisis situation (4.2) in the region $0 < g < 1$. We shall use itineraries to classify the windows of periodic attractors in the parameter space (see reference [6] Section II.1 for definitions).

The first iterate s of -0 is the maximum of the negative branch of h , i.e. $s = 1 - g$. The itinerary $I(s)$ of the point s plays a central role in the following. It will be denoted by M . It depends on the parameter g . At the appearance of a window the point -0 becomes the element of a stable periodic orbit and correspondingly M is finite: $M = DC$, where D is a finite sequence of R and L symbols. Inside the window the itinerary of s is infinite $M = (DL)^\infty$. One can use the repeated sequence $E = DL$ to identify the window.

Let us introduce the ordering between itineraries [6] here. Considering two different itineraries A, B let i be the first index for which $A_i \neq B_i$. Let n denote the number of R 's in $A_0 A_1 \dots A_{i-1}$. By definition $L < C < R$. It is said that $A < B$ if either

- a) $A_i < B_i$ and n is even, or
- b) $A_i > B_i$ and n is odd.

It is easy to see that the orderings $p \leq q$ and $I(p) \leq I(q)$ are equivalent.

The itinerary of the iterates of a point p can be obtained by shifting the itinerary of p $I(h^n(p)) = \{I_n, I_{n+1}, I_{n+2}, \dots\}$, which can be written with the help of the shift operator as

$$I(h^n(p)) = S^n I(p). \quad (\text{A.1})$$

For any itinerary we obtain $I(h^n(p)) < I(-0) \equiv LM$ if $h^n(p) < 0$, thus

$$S^{n+1} I(p) < M \quad \text{for} \quad I_n = L. \quad (\text{A.2})$$

For $p = s$ it gives

$$S^{n+1} M < M \quad \text{for} \quad M_n = L. \quad (\text{A.3})$$

The itinerary M plays a similar role in these conditions as the kneading sequence for unimodal maps [6].

In the following we restrict ourselves to the case when the separate intervals I_1, I_2, I_3 are mapped into each other according to (4.3) (see Fig. 7). It is fulfilled in the region $g_3 < g < 1$. It can be easily seen that R 's inside the itineraries may occur only in RR pairs followed by L . So itineraries of points in I_2 can be written as a product

$$I = N_{l_1} N_{l_2} \dots N_{l_n} C \quad \text{or} \quad I = N_{l_1} N_{l_2} \dots,$$

where

$$N_l = R^2 L^l, \quad l, l_i > 0.$$

Correspondingly, the itinerary M has this form. The situation when M is finite corresponds to the appearance of a window, the repeated sequence belonging to the stable periodic orbit in the window is

$$E = DL = N_{k_1} N_{k_2} \dots N_{k_n}, \quad k_i > 0. \quad (\text{A.4})$$

The allowed sequence $\{k_i\}_{i=1}^n$ of the subscripts is selected by the requirement (A.3) for $M = DC$.

On the basis of this description we get for the windows investigated in Section 4

$$E_m = N_{m-2} = R^2 L^{m-2}, \quad m > 2 \quad (\text{A.5})$$

and we find satellite series of windows

$$\begin{aligned} E_{m,-k} &= N_{m-3} N_{m-3+k}, & m > 3, \quad k > 0, \\ E_{m,k} &= (N_{m-2})^k L = (N_{m-2})^{k-1} N_{m-1}, & m > 2, \quad k > 1. \end{aligned}$$

The parameter values of $E_{m,-k}(E_{m,k})$ tend to the $g_m^<(g_m^>)$ value where the stable period E_m appears (disappears). Some of these windows were observed in bifurcation diagrams.

References

1. R. H. G. Helleman, in *Fundamental Problems in Statistical Mechanics*, Vol. 5, ed. E. G. D. Cohen, North Holland, Amsterdam, 1980.
2. A. J. Lichtenberg and M. A. Lieberman, *Regular and Stochastic Motion*, Springer, N. Y., 1983.
3. H. G. Schuster, *Deterministic Chaos*, VHC Publishers, Weinheim, 1984.
4. H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*, Gauthier Villars, Paris, 1892.
5. M. Misiurewicz, Maps of an interval, in: *Chaotic Behaviour of Deterministic Systems*, ed. G. Ioo, R. H. G. Helleman and R. Stora, North-Holland, Amsterdam, 1983.
6. P. Collet and J.-P. Eckmann, *Iterated Maps on the Interval as Dynamical Systems*, Birkhäuser, Boston, 1980.
7. B. V. Chirikov, *Phys. Reports*, 52, 263, 1979.
8. J.-P. Eckmann and D. Ruelle, *Rev. Mod. Phys.*, 57, 617, 1985.
9. B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, San Francisco, 1982.
10. O. E. Rössler, *Phys. Lett.*, 57A, 397, 1976.
11. M. Hénon and Y. Pomeau, *Lecture Notes in Math.*, 565, 29, 1976.
12. M. Hénon, *Commun. Math. Phys.*, 50, 69, 1976.
13. E. N. Lorenz, *J. Atmospheric Sci.*, 20, 130, 167, 1963.
14. T. Rikitake, *Proc. Cambridge Phil. Soc.*, 54, 89, 1958.
15. O. E. Rössler, *Z. Naturforsch.*, 31a, 1664, 1976.
16. R. F. Williams, *Lecture Notes in Math.*, 615, 94, 1977.
17. J. A. Yorke and E. D. Yorke, *J. Stat. Phys.*, 21, 263, 1979.
18. A. Arneodo, P. Coulet and C. Tresser, *Phys. Lett.*, 81A, 197, 1981.
19. C. Sparrow, *The Lorenz Equation: Bifurcations, Chaos and Strange Attractors*, *Appl. Math. Sci.*, 41, Springer, N. Y. 1982.
20. D. V. Lyubimov and M. A. Zaks, *Physica*, 9D, 52, 1983.
21. P. Szépfalusy and T. Tél, *Physica*, 16D, 252, 1985.
22. G. Györgyi and P. Szépfalusy, *J. Stat. Phys.*, 34, 451, 1984.
23. C. Grebogi, E. Ott and J. A. Yorke, *Physica*, 7D, 181, 1983.
24. L. Hitzl, *Physica*, 2D, 370, 1981.
25. R. Bridges and G. Rowlands, *Phys. Lett.*, 63A, 189, 1977.
26. R. Lozi, *J. de Physique*, Colloque 39, C5, 9, 1978.
27. J. P. Crutchfield, D. Farmer, N. Packard, R. Shaw, G. Jones, R. J. Donnely, *Phys. Lett.*, 76A, 1, 1980.
28. R. B. May, *Nature*, 261, 459, 1976.
29. S. Grossmann and S. Thomaé, *Z. Naturforsch.*, 32a, 1353, 1977.
30. M. J. Feigenbaum, *J. Stat. Phys.*, 19, 25, 1978; 21, 669, 1979.
31. T. Geisel and J. Nierwetberg, *Phys. Rev. Lett.*, 47, 975, 1981.
32. L. P. Shilnikov, *Math. USSR Sbornik*, 6, 427, 1968.
33. P. Gaspard, *Phys. Lett.*, 97A, 1, 1983.
34. P. Glendinning and C. Sparrow, *J. Stat. Phys.*, 35, 645, 1984.
35. P. Gaspard, R. Kapral and G. Nicolis, *J. Stat. Phys.*, 35, 697, 1984.
36. A. Arneodo, P. H. Coulet, E. A. Spiegel and C. Tresser, *Physica*, 14D, 327, 1985.
37. O. E. Lanford, *Limit Theorems in Statistical Mechanics*, *Troisième Cycle de la Physique*, E. N. Suisse Romande, Semestre d'Eté, 1978.
38. H. Kantz and P. Grassberger, *Physica*, 17D, 75, 1985.
39. H. Daido, *Prog. Theor. Phys.*, 63, 1190, 1980.
40. C. Simo, *J. Stat. Phys.*, 21, 465, 1979.
41. D. L. Hitzl and F. Zele, *Physica*, 14D, 305, 1985.

NUCLEON-NUCLEON INTERACTION WITH ULTRA SHORT RANGE ATTRACTION*

I. LOVAS and ERIKA LOVAS

*Central Research Institute for Physics
1525 Budapest, Hungary*

K. SAILER

*Kossuth Lajos University, Institute for Experimental Physics
4000 Debrecen, Hungary*

A nucleon-nucleon interaction with ultra short range attraction is introduced. The attractive regions are separated by a narrow, very high repulsive barrier. To find the parameters of this kind of potential, some properties of the two-nucleon system are calculated. The aim is to decrease the gap between the traditional nuclear physics and the quark model.

1. Introduction

In the last three decades, especially after the discovery of the quark model of hadrons [1] it became absolutely clear that the hadrons can not be considered as structureless, elementary particles. Nevertheless, in the framework of nuclear physics we try to formulate the nuclear many body problem in terms of the non-relativistic quantum mechanics of pointlike particles. This kind of approach is reasonable until the role of the quark degrees of freedom can be neglected. In the majority of nuclear phenomena this is the case indeed. Nevertheless, it seems to be desirable to modify the tools of the "classical" nuclear physics in order to decrease the gap between the conventional approach and the quark model [2]. Among the notions of nuclear physics the most fundamental is the nucleon-nucleon interaction potential. We know since the pioneering work of Jastrow [3] that at small values of the relative coordinate the potential shows a strong repulsion. This repulsion was represented by an infinitely high core having a radius of 0.4 fm.

Later on this hard core was substituted by a finite, soft core. The repulsive core was interpreted as the result of the exchange of vector mesons mediating a repulsive interaction. Since the masses of the vector mesons are rather high the range of the repulsive effect is small. Nowadays we are tempted to associate the repulsion with the finite radius of the quark bag. The repulsion, however, turns out to be too strong if we

* Dedicated to Prof. K. Nagy on his 60th birthday

identify the bag radius, or the measured radius of the nucleon ($R=0.8$ fm) with the radius of the repulsive core.

In this paper we try to introduce such a nucleon–nucleon potential which is able to account for the observed properties of the two-nucleon system, on the other hand, it is compatible with the large value of the nucleon radius.

The essential points of this approach are the following. The wave function of the deuteron in the quark model [4], [5] has two terms:

$$\Psi = (1 - \alpha^2)^{1/2} \psi(3q, 3q) + \alpha \psi(6q),$$

where the first term describes the relative motion of two colour-singlet three-quark clusters. This corresponds to the traditional deuteron wavefunction which describes the relative motion of two nucleons. The second term describes the motion of a six-quark system produced by the “temporary fusion” of the two three-quark clusters. This later component, having 2–3% contribution to the norm of the wave function, corresponds to such a configuration which is unstable in itself.

In the framework of the nonrelativistic quantum mechanics the nucleons are considered as pointlike objects therefore it is desirable to provide such a description in which the formation of the six-quark configuration at least can be imitated. The simplest way to do this is the following. We assume that the potential is attractive at short distances and it is also attractive at larger distance, however, these regions are separated by a rather high repulsive barrier. (See Fig. 1).

If the attraction is not too strong, then no bound state concentrated in the ultra short ranged attractive potential well is formed and then this type of potential is able to imitate the temporary formation of the six-quark configuration in the central region of the system.

In such a way the traditional potential resembles in a sense the picture implied by the quark model:

- the radius of the repulsive potential can be associated with the radius of the quark bag;
- the contribution to the norm of the wave function coming from the central region can be associated with the relative probability of the six-quark configuration.

The successful application of these ideas is possible only if the parameters of this type of potential can be fitted to the physical observables.

The physical observables in this context are the following:

- the two-nucleon data;
- the three-nucleon data and
- the properties of the nuclear matter, extrapolated from the properties of finite nuclei.

These data among the rich variety of nuclear data are distinguished by the fact that they are measured accurately and at the same time they can be calculated in a rather reliable manner starting from the two-nucleon potential.

In this paper we try to solve only the first part of the task. Namely, we will search for the parameters of the potential which is able to account for some of the properties of the two-nucleon system.

In particular we will calculate

- the binding energy of the deuteron ($E = -2.226$ MeV);
- the scattering length ($a = 5.424$ fm);
- the effective range ($r_0 = 1.759$ fm);
- the scattering phase shift as the function of the energy;
- the deuteron charge formfactor as the function of the momentum transfer.

The calculations, which are meant only for orientation, are performed for a simplified potential constructed from square well potentials [6].

2. Two-nucleon properties

The states of the two-nucleon system can be characterised by the total spin $S = S_1 + S_2$, the total isospin $I = I_1 + I_2$, and the total angular momentum J . If one disregards the tensor force, then the states can be characterised by the orbital momentum L of the relative motion instead of the total angular momentum J . Here we will consider only the $L = 0$ states. The spin-triplet, isospin-singlet state ($S = 1, I = 0$) is very important since the single existing bound state of the two-nucleon system, namely the ground state of the deuteron is carrying these quantum numbers.

The solution of the task sketched in the Introduction requires the study of the solutions of the Schrödinger equation given by

$$\frac{d^2 u(r)}{dr^2} + \frac{m}{\hbar^2} (E - V(r)) u(r) = 0.$$

From the solution of this equation we can calculate the physical observables of the two-nucleon system and we can compare them with the measured data. We will minimise the deviation by varying the parameters of the potential given by (see Fig. 1)

$$V(r) = \begin{cases} -V_1, & 0 \leq r \leq R_1, \\ V_2, & R_1 < r \leq R_2, \\ -V_3, & R_2 < r \leq R_3, \\ 0, & R_3 < r. \end{cases}$$

If the energy is negative, that is we are looking for bound state solution then the radial wave function $u(r)$ must vanish in the infinity

$$u(r) \rightarrow 0, \quad r \rightarrow \infty.$$

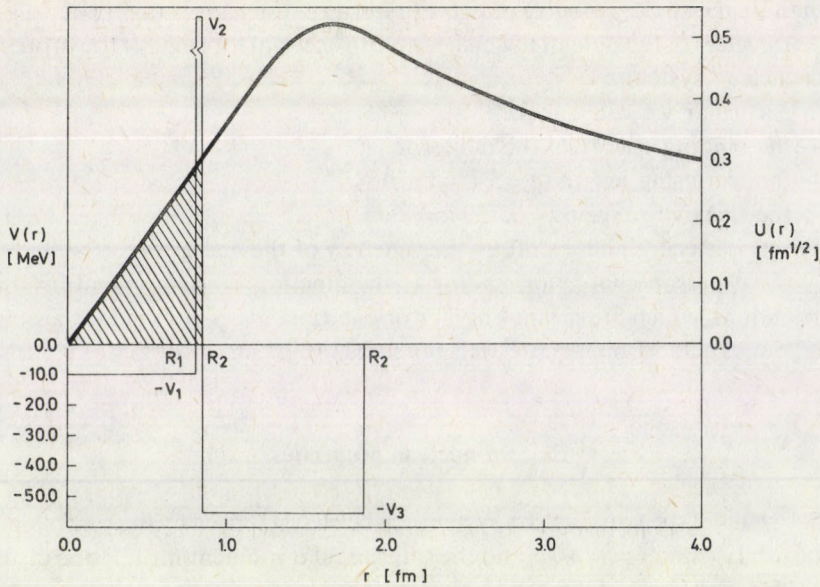


Fig. 1. The potential $V(r)$ as a function of the distance r between the two nucleons. The radial wave function $u(r)$ of the ground state of the deuteron

The solution in this case can be written in the following form:

$$u(r) = \begin{cases} a_1 \sin k_1 r, & 0 \leq r \leq R_1, \\ a_2 e^{k_2 r} + a_4 e^{-k_2 r}, & R_1 < r \leq R_2, \\ a_3 \sin k_3 r + a_5 \cos k_3 r, & R_2 < r \leq R_3, \\ a_6 e^{-k_4 r}, & R_3 < r, \end{cases}$$

where

$$k_1 = \left(\frac{m}{\hbar^2} (-|E| + V_1) \right)^{1/2}, \quad k_2 = \left(\frac{m}{\hbar^2} (|E| + V_2) \right)^{1/2},$$

$$k_3 = \left(\frac{m}{\hbar^2} (-|E| + V_3) \right)^{1/2}, \quad k_4 = \left(\frac{m}{\hbar^2} (|E|) \right)^{1/2},$$

and m is the nucleon mass. The solution and its first derivative must be continuous everywhere, among others in the points $r=R_1$, $r=R_2$ and $r=R_3$, too. From this requirement the amplitudes a_i can be determined. By varying a parameter, namely the value of V_3 , the binding energy of the deuteron can be fitted exactly. The charge formfactor of the deuteron $F(\mathbf{q})$ can be obtained from the solution $\psi(\mathbf{r})$ by Fourier transformation:

$$F(\mathbf{q}) = \int |\psi(\mathbf{r})|^2 e^{i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} = \frac{4\pi}{q} \int_0^\infty u^2(r) \frac{\sin qr}{r} dr.$$

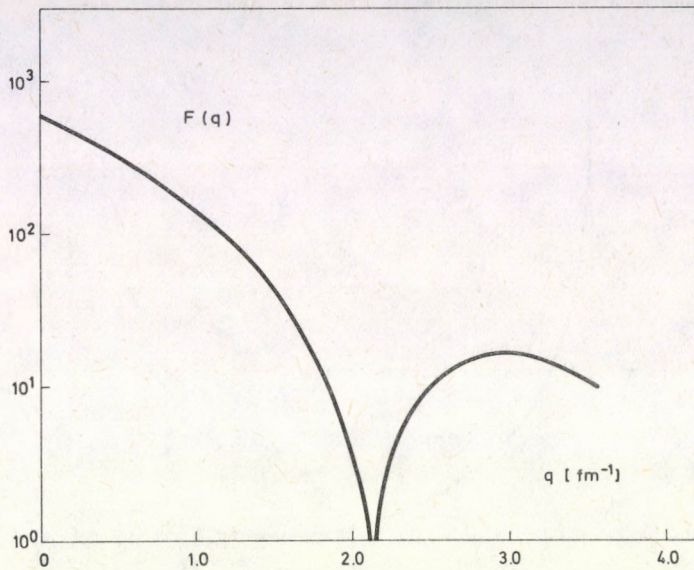


Fig. 2. The charge form factor of the deuteron as the function of the momentum transfer

The result can be seen in Fig. 2. By the help of the normalised wavefunction we can calculate the probability for finding the two nucleons closer than R_2

$$\alpha^2 = \int_0^{R_2} |\psi(\mathbf{r})|^2 d\mathbf{r}.$$

This is proportional to the shaded region of Fig. 1. The probability amplitude α defined in this way can be interpreted as the amplitude for finding the system in a six-quarks configuration.

If the energy is positive, that is we are interested in scattering states, then the radial wavefunction $u(r)$ tends to the wavefunction of a free system shifted by a phase

$$u(r) \rightarrow \sin(kr + \delta), \quad r \rightarrow \infty.$$

The scattering phase shift δ depends on the kinetic energy of the relative motion of the two nucleons. At low energies the scattering can be characterised by two parameters very well. These two parameters, the scattering length a and the effective range r_0 , are defined as follows:

$$k \operatorname{ctg} \delta(E(k)) = -\frac{1}{a} + \frac{1}{2} r_0 k^2.$$

In the spin-triplet, isospin-singlet state the measured values are given by $a = 5.424$ fm and $r_0 = 1.759$ fm, respectively, and they are fairly well reproduced. The phase shift vanishes there, where the effect of the attraction and that of the repulsion compensate

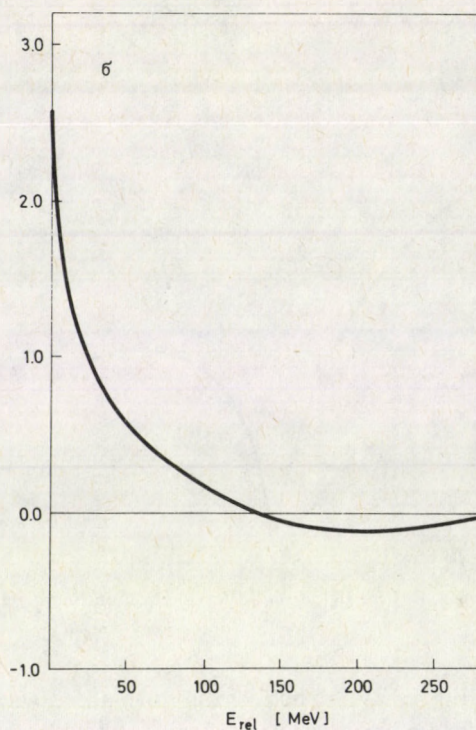


Fig. 3. The scattering phase shift as the function of the relative energy of the two nucleons

Table I

The physical quantities of the two-nucleon system, namely the scattering length a , the effective radius r_0 , the relative energy E_0 , where the phase shift vanishes, the probability α_2 of finding the two nucleons within the distance R_2 and finally the momentum transfer q_{\min} where the charge form factor $F(q)$ is minimal at $R_1=0.8$ fm, $R_2=0.81$ fm and $R_3=1.85$ fm as the function of V_1 , V_2 and V_3 . The binding energy of the deuteron is reproduced exactly $E = -2.226$

V_1	V_2	V_3	a	r_0	E_0	α^2	q_{\min}
8	3500	60.7	5.42	1.79	149	2.36	2.20
8	4000	60.2	5.43	1.80	138	2.21	2.13
8	4500	63.7	5.44	1.82	138	2.08	2.13
18	3500	59.8	5.42	1.78	160	2.49	2.20
18	4000	61.4	5.43	1.79	149	2.34	2.20
18	4500	62.9	5.44	1.81	138	2.19	2.13
MeV	MeV	MeV	fm	fm	MeV	%	fm ⁻¹

each other. The scattering phase shift vanishes at about $E_{\text{rel}} = 175$ MeV. The energy dependence of this phase shift is shown in Fig. 3. Typical results are presented in Table I.

If the strength of the ultra short range attraction reaches a critical value, then the radial wavefunction may have a zero in the range $r < R_2$. The idea to use such a wavefunction was proposed ten years ago [7] and it is supported by some recent calculations in the framework of the quark model [4]. As it can be judged from the preliminary results this type of potential seems to work as far as the two-nucleon system is concerned. It must be noted that the generalisation for the tensor force is necessary if we want to have a more realistic description. To draw a final conclusion the three-nucleon systems and the nuclear matter has to be examined.

References

1. M. Gell-Mann, Phys. Lett., 8, 214, 1964.
2. I. Lovas, Few-Body Nuclear Physics, IAEA, Vienna, 1978, p. 203.
3. R. Jastrow, Phys. Rev., 81, 165, 1951.
4. A. Faessler, F. Fernandez, G. Lübeck and K. Shimizu, Nucl. Phys., A402, 555, 1981.
5. Yu. A. Simonov, Phys. Lett., 107B, 1, 1981.
6. J. M. Blatt and V. F. Weisskopf, Theoretical Nuclear Physics, Springer, New York, 1979.
7. V. G. Neudatschin, I. T. Obukhovskiy, V. I. Kukul'in and N. F. Golovanova, Phys. Rev., C11, 128, 1975.

A QUANTITATIVE ANALYSIS OF THE COLLECTIVE OUTWARD MOTION OF NUCLEAR MATTER IN CENTRAL Ca + Ca AND Nb + Nb COLLISIONS*

L. P. CSERNAI**

*School of Physics and Astronomy, University of Minnesota
Minneapolis, MN 55455, USA*

and

G. FAI***

*Department of Physics, Kent State University
Kent, OH 44242, USA*

The collective outward motion of nuclear matter in the final stages of central collisions is analysed with the help of a statistical model calculation. Calculated relative fragment yields are fitted to the Plastic Ball data to infer the breakup properties of the collision system. The difference between the total available energy and the energy in thermal excitation is identified with the energy in collective motion. A quantitative measure of this collective flow is provided for systems where data are available.

1. Introduction

The success of phenomenological models based on the analogy between the nucleus and a droplet of liquid [1] in the explanation of ground and excited state nuclear properties led naturally to the attempt to describe energetic nuclear collisions in terms of fluid-dynamical concepts [2]. Fluid-dynamical models predict a collective sideways emission of the nuclear material in central collisions [3]. The blast wave model [4] also assumes an ordered outward motion of the fragments of a nuclear collision. It is therefore not surprising that a continuing effort to extract this collective flow from the data has become one of the central themes of the study of energetic nuclear collisions.

Angular distributions of protons from inclusive measurements [5] gave a hint of the expected collective behavior. However, since neutron angular distributions from central collisions do not show any forward suppression [6] and the proton angular distributions can essentially be explained as a Coulomb effect [7], we are led to

* Dedicated to Prof. K. Nagy on his 60th birthday

** On leave from the CRIP, Budapest, Hungary

*** On leave from the R. Eötvös University, Budapest, Hungary

conclude that inclusive proton data (and inclusive measurements in general) do not provide satisfactory evidence for the collective flow of nuclear matter.

Fortunately, we have witnessed the evolution of a new generation of experiments in the last couple of years. The GSI-LBL Plastic Ball/Wall Collaboration [8] played a leading role in the development of these new exclusive measurements detecting electronically a large fraction of the charged particles produced in each collision event. The analysis of the Plastic Ball/Wall data qualitatively shows the presence of a collective flow in nuclear collisions [9, 10]. A quantitative analysis can be performed utilizing the microcanonical model of nuclear disassembly [11]. This model provides a least biased statistical description of the final stages of a nuclear collision based on an approximate statistical treatment of the microcanonical phase space. It carefully takes into account available experimental information [12] and — in contrast to earlier statistical models — is well suited for the description of exclusive measurement since conservation laws are obeyed exactly on an event-by-event basis in the model. The importance of fragmentation in detecting the collective flow was pointed out earlier [13].

The aim of the present contribution is to extract a quantitative measure of the collective flow present in the collisions for which the exclusive Plastic Ball/Wall data have been analysed so far. The paper is organized as follows: in Section 2 we briefly describe the model. We emphasize the fact that our calculation is a straightforward application of the concepts of statistical physics. We also explain how we use the model to map the temperature–chemical potential space for nuclear disassembly. Section 3 contains the results presented as least-squared fits to the available experimental data. Section 4 is reserved for discussion and conclusions.

2. Statistical description of the collision

We intend to describe central collisions at beam energies of a few hundred MeV/nucleon. A large number of fragmentation channels is open for the disassembly of the hot nuclear matter formed in these reactions. Little is known about the detailed reaction mechanism except the fact that it is very complicated. We take advantage of the complexity of the process by making a statistical model of the final stages of the reaction. We will call this final process of fragment formation nuclear disassembly. The beam energies considered correspond to center of mass energies high enough for classical statistics to be applicable but low enough for the specific properties of light nuclei to play an important role.

Statistical models can be viewed as descriptions in which all the relevant S -matrix elements are assumed to be equal. The calculation of final state probabilities then amounts to an enumeration of available phase space. Models differ in the definition of the phase space for the system and in the statistical ensemble applied in the calculation. A calculation based on the grand canonical ensemble [14] carefully took

into account available experimental information and included all particle stable and unstable states with a width less than 1 MeV and with mass number $A \leq 16$ explicitly. This work was subsequently extended [11] to an approximate microcanonical treatment relevant for the exclusive data considered. To date the complete event generation model [11] remains the only statistical model on the market based on the microcanonical phase space. We therefore use this model in our analysis of the Plastic Ball/Wall data.

We start by associating with an ensemble of events a temperature T , a baryon chemical potential μ and an isospin chemical potential ν . The logarithm of the grand canonical partition function is expressed as a computable sum over different fragment species containing these three parameters. Reconstructing the partition function yields all required probabilities necessary to make a random choice of a fragment in the grand canonical treatment. In order to obtain an approximate microcanonical probability distribution the grand canonical distributions are truncated accordingly [11]. Once a fragment has been selected with its baryon number, charge and four-momentum, the values of these conserved quantities characterizing the remaining part of the source are also known. The grand canonical constraint equations pertaining to the residual source (and expressing conservation of total energy, baryon number and charge on the mean) are then solved for the three Lagrange multipliers T' , μ' and ν' . With these parameters the partition function characterizing the residual source can be calculated, making the subsequent random selection of a second fragment possible. This procedure is iterated until no residual system remains. In this way we obtain a description that is manifestly energy-, momentum- and number-conserving and is convenient to generate a large sample of events within a reasonably short CPU time.

In order to take into account the finite size of nuclear fragments the excluded volume approximation is invoked in the calculation. This approximation consists of replacing each integral over the position (of the center of mass) of the fragments with an average available volume expressed in terms of the standard nuclear volume of the source. Thus the parameter χ is introduced according to

$$\langle \int dr_i \rangle = \chi 4\pi r_0^3 A_0 / 3, \quad (1)$$

where A_0 is the mass number of the source and $r_0 = 1.15$ fm is the nuclear radius parameter. In other words, the volume blocked by one nucleon v_0 is assumed to be

$$v_0 = 4\pi r_0^3 \chi / 3. \quad (2)$$

It can be shown [15] that the thermodynamic quantities in the excluded volume approximation (denoted by the subscript xv) can be obtained from those of the point particle limit (subscript pt) with a shifted chemical potential μ^* , defined by the equation

$$\mu^* = \mu - P_{xv}(T, \mu, \nu)v_0, \quad (3)$$

where the pressure in the excluded volume approximation $P_{xv}(T, \mu, v)$ is equal to the point particle pressure with the shifted chemical potential:

$$P_{xv}(T, \mu, v) = P_{pt}(T, \mu^*, v). \quad (4)$$

The densities of the different extensive quantities scale with the factor

$$f = 1 + n_{pt}(T, \mu^*, v)v_0, \quad (5)$$

where n_{pt} is the baryon density in the point particle limit. More precisely, the baryon density and the energy density, for example, in the excluded volume approximation are given by

$$n_{xv}(T, \mu, v) = n_{pt}(T, \mu^*, v)/f \quad (6)$$

and

$$\rho_{xv}(T, \mu, v) = \rho_{pt}(T, \mu^*, v)/f. \quad (7)$$

This then makes it possible to construct a map in the (T, μ^*) plane. (Note that the isospin chemical potential v is determined self-consistently in each case.)

The temperature and chemical potential of the hot nuclear matter created in the collision is of course unknown. Starting from a given T and μ^* , however, the model produces a statistical sample of final states, that will be compared to the data (see Section 3). Utilizing additional information on the temperature T (such additional information may come e.g. from the observed kinetic energy spectra of protons [16, 17]) the thermal energy of the system can be extracted. Since the total energy is well known, subtraction yields the energy not available for thermal excitation which we identify with the energy in the collective flow.

3. Comparison to the data

We compared the ratio of "deuteron-like clusters" [16] to charged fragments and the sum of the squared deviations of fragment-to-proton ratios weighted by the fragment mass for all measured fragments to the experimental data for central collisions [16]. We calculated a large number of breakup states by choosing pairs of starting T and μ^* values without making any assumption about the breakup density or other quantities. In this way we mapped the thermodynamic parameter space and to each point we evaluated the corresponding microcanonical fragment distributions. Our ensembles consisted of 300 events in the case of Ca + Ca collisions, while samples of 150 events were used for the Nb + Nb reactions.

First we show the comparison of our results to the central Ca + Ca data at 1050 MeV/nucleon beam energy. We introduce the quantity

$$\begin{aligned} \Sigma = & 4 \times [(d/p)_{th} - (d/p)_{exp}]^2 + 9 \times [(t/p)_{th} - (t/p)_{exp}]^2 + \\ & + 9 \times [(^3He/p)_{th} - (^3He/p)_{exp}]^2 + 16 \times [(\alpha/p)_{th} - (\alpha/p)_{exp}]^2 \end{aligned} \quad (8)$$

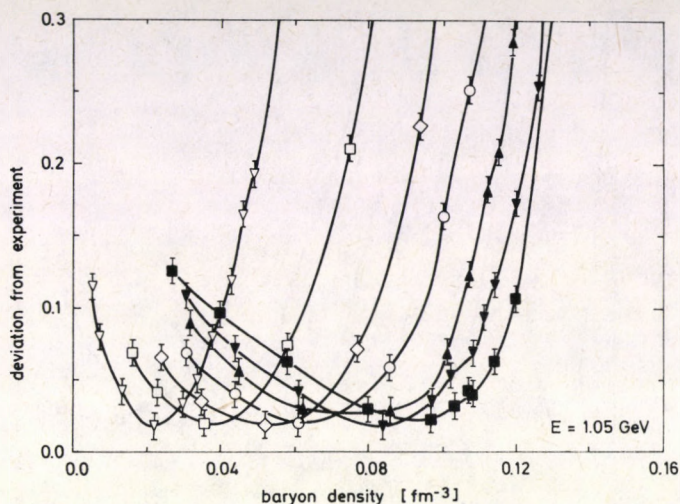


Fig. 1. Deviation from experiment (8) for central Ca + Ca collisions at 1050 MeV/nucleon beam energy as a function of the breakup density n at breakup temperatures $T = 30, 45, 60, \dots, 120$ MeV (open triangles, open squares, diamonds, circles, closed up triangles, closed down triangles and closed squares, respectively). Each point represents a breakup state (T, n) calculated in the model. The experimental data are from [16]

as a measure of the “deviation from experiment”, where (d/p) , (t/p) , $({}^3\text{He}/p)$ and (α/p) stand for the deuteron-to-proton, triton-to-proton, ${}^3\text{He}$ -to-proton and alpha-to-proton ratios respectively. The expression (8) was evaluated for every pair of the thermodynamical variables. Fig. 1* displays the deviation from experiment (8) versus the breakup density of the system at different breakup temperatures in the range $30 \text{ MeV} \leq T \leq 120 \text{ MeV}$. It is seen that the deviation from experiment exhibits a minimum at every temperature. However, these minima occur at different breakup densities. Thus neither the breakup temperature nor the breakup density can be determined accurately. On the other hand the apparent temperature (as determined from the slope of the kinetic energy spectra of fragments) is approximately 85 MeV in this case [16]. This piece of experimental information enables us to limit the breakup density into the region $n_{bu} = 0.07 \pm 0.02 \text{ fm}^{-3}$. In terms of the standard nuclear density $n_0 = 0.17 \text{ fm}^{-3}$ the breakup density is $n_{bu} = (0.41 \pm 0.12)n_0$.

In Fig. 2 we show the deviation from experiment Σ versus the internal energy per baryon at different breakup temperatures $T = 30 \text{ MeV}, 45 \text{ MeV}, 60 \text{ MeV}, 75 \text{ MeV}, 90 \text{ MeV}, 105 \text{ MeV}$ and 120 MeV for the same experiment. Again, using the value $T = 85 \text{ MeV}$ from the slope of the kinetic energy spectra of protons, we conclude that the internal energy per nucleon at breakup is approximately $\varepsilon = (110 \pm 10) \text{ MeV}$. (Note that the physical temperature is expected to be lower than the apparent temperature extracted from the slopes of kinetic energy spectra which may receive other

* The versions of the Figures presented here which appear in ref.-s [18] are based on a calculation containing an unfortunate numerical error. They should be replaced with the present Figures.

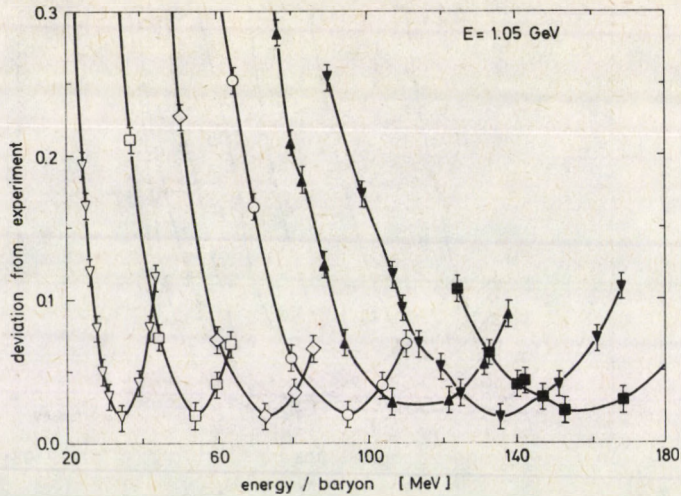


Fig. 2. Deviation from experiment (8) for central Ca + Ca collisions at 1050 MeV/nucleon beam energy versus the internal specific energy for different breakup temperatures. The temperature values and the notation used are the same as in Fig. 1. The experimental data are from [16]

contributions [2, 4]. The extracted internal energy per nucleon, therefore, represents an upper limit.) From other experiments [19] we estimate that pions carry about 30 MeV/baryon at this beam energy. The sum of these two contributions, 140 ± 15 MeV is substantially less than the available center of mass energy. The excess energy, which is about $40 \pm 10\%$ has to reside in degrees of freedom we have not accounted for thus far. If transparency is negligible the most natural assumption is that the excess energy is in the form of collective kinetic energy at the time of breakup. Therefore we identify this excess energy with the collective outward flow of nuclear matter. A similar analysis for the Ca + Ca system at the lower beam energy (400 MeV/nucleon), where we estimate that pions carry approximately 2 MeV/nucleon, shows that $29 \pm 10\%$ of the available energy is unaccounted for. We identify this with the flow energy at breakup in 400 MeV/nucleon central Ca + Ca collisions.

In Figs 3 and 4 we present the deviation from experiment (8) as a function of the internal energy per baryon for different breakup temperatures $30 \text{ MeV} \leq T \leq 120 \text{ MeV}$ for the Nb + Nb system at 400 MeV/nucleon and 650 MeV/nucleon beam energies, respectively. According to [10] the temperature extracted from the slope of the proton spectra in Nb + Nb at 400 MeV/nucleon is 65 MeV in the highest multiplicity bin (central collisions). Using this value as a (good) upper limit we read from Fig. 3 that the internal energy at breakup is 60 ± 10 MeV in this system. This estimate together with the estimate made earlier on the energy carried by pions leads to $35 \pm 10\%$ collective energy in this system. For the case of the 650 MeV/nucleon Nb + Nb collisions the temperature obtained from the slopes of the proton kinetic energy distributions is 70 MeV [16]. The pion contribution is taken from [19] to be 12 MeV. On the basis of these values and Fig. 4 we arrive at a flow energy of $42 \pm 10\%$.

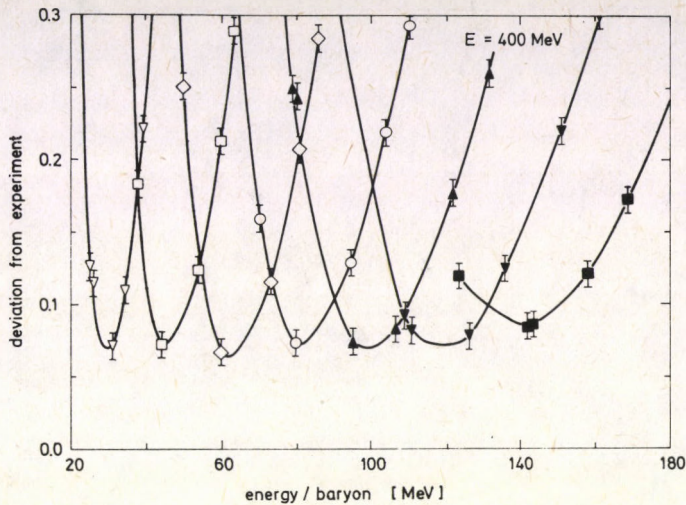


Fig. 3. Same as Fig. 2 for central Nb+Nb collisions at 400 MeV/nucleon beam energy

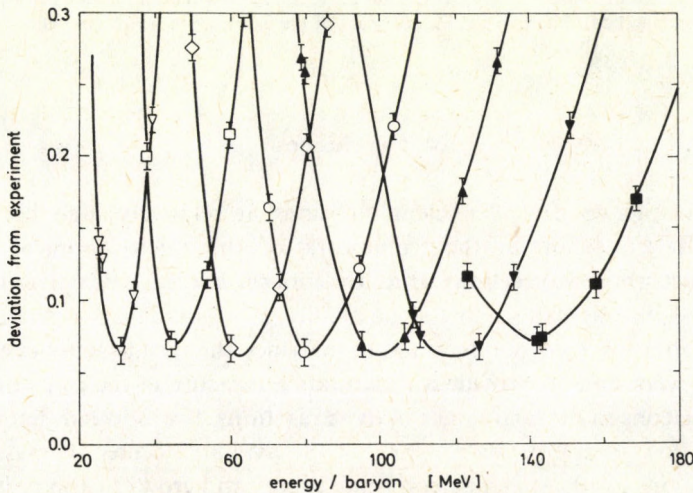


Fig. 4. Same as Fig. 2 for central Nb+Nb collisions at 650 MeV/nucleon beam energy

Fig. 5 displays the deviation from experiment Σ as defined by Eq. (8) as a function of the entropy per nucleon at different breakup temperatures in the range $30 \text{ MeV} \leq T \leq 120 \text{ MeV}$ (different symbols). It is seen from the Figure that all points fall approximately on the same curve irrespective of their temperature. This means that the entropy of the system at breakup can be determined more accurately than the other quantities. From Fig. 5 we conclude that the entropy at breakup is 3.7 ± 0.2 per nucleon for central 650 MeV/nucleon Nb+Nb collisions.

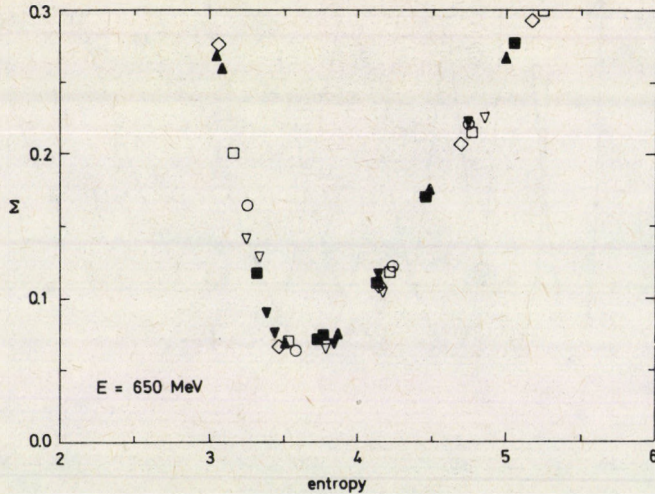


Fig. 5. Deviation from experiment Σ (8) for central Nb + Nb collisions at 650 MeV/nucleon beam energy as a function of the breakup entropy at different breakup temperatures. The temperature values and the notation used are the same as in all the other Figures. The experimental data are from [16]. Points corresponding to different temperatures fall approximately on the same curve. An approximate minimum of the deviation from experiment can be extracted independent of the assumed breakup temperature

4. Conclusions

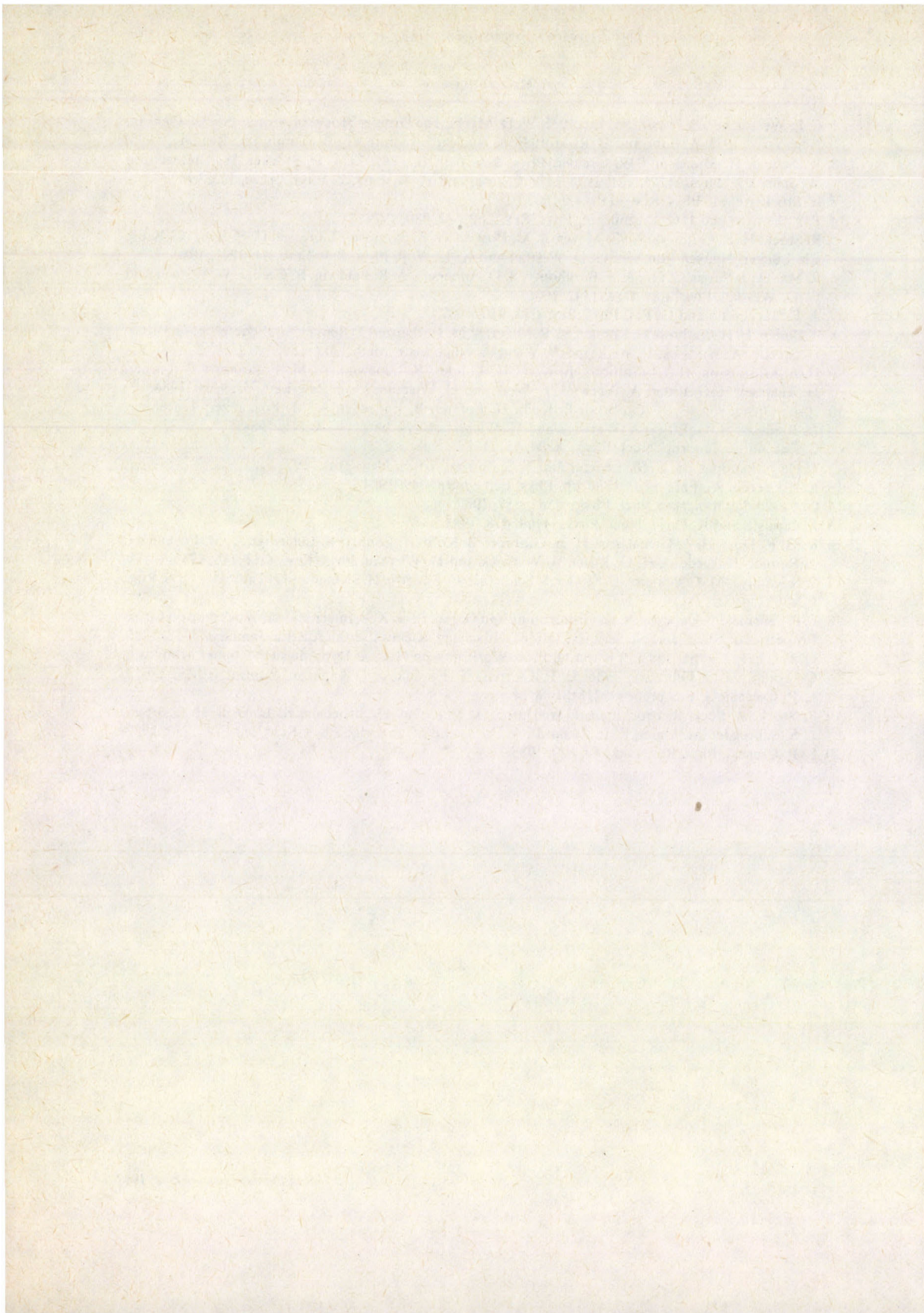
We investigated central nuclear collisions at relatively high beam energies ($E_{\text{lab}} \geq 400$ MeV/nucleon) in the framework of the microcanonical model for nuclear disassembly. Corrections due to surface effects, attractive interactions among composite fragments and the possibility of a liquid-vapor phase transition in nuclear matter [20] are not expected to influence the results significantly at these energies. We were able to provide a quantitative measure of the outward collective motion of nucleons in the final stages of these reactions. It was found that the amount of collective flow is approximately 30% of the available center of mass energy in Ca + Ca collisions at 400 MeV/nucleon beam energy and grows to approximately 40% at 1050 MeV/nucleon in the same system. For Nb + Nb the collective flow amounts to about 35% of the available center of mass energy at 400 MeV/nucleon beam energy and to about 42% at 650 MeV/nucleon.

5. Acknowledgements

This work was supported in part by the U.S. Department of Energy under Grant No. DE-FG02-86ER40251 and Contract No. DE-AC02-79ER10364 and by the Supercomputer Institute of the University of Minnesota. The authors are indebted to Prof. H. Stöcker for pointing out a numerical error in an earlier version of the calculation.

References

1. C. F. Weizsäcker, *Z. Physik*, *96*, 431, 1935; W. D. Myers, *The Droplet Model of Atomic Nuclei*, Plenum, New York, 1977; A. Bohr and B. R. Mottelson, *Nuclear Structure, II.*, Benjamin, Inc. Reading, 1975.
2. W. Scheid, H. Müller and W. Greiner, *Phys. Rev. Lett.*, *32*, 741, 1974; H. Stöcker, W. Greiner and W. Scheid, *Z. Phys.*, *A286*, 121, 1978; L. P. Csernai and H. W. Barz, *Z. Phys.*, *A296*, 173, 1980.
3. H. Stöcker et al., *Phys. Rev. Lett.*, *47*, 1807, 1981.
4. P. J. Siemens and J. O. Rasmussen, *Phys. Rev. Lett.*, *42*, 880, 1979.
5. R. Stock, H. H. Gutbrod, W. G. Meyer, A. M. Poskanzer, A. Sandoval, J. Gosset, C. H. King, G. King, Ch. Lucker, Nguyen Van Sen, G. D. Westfall and K. L. Wolf, *Phys. Rev. Lett.*, *44*, 1243, 1980.
6. R. Madey, J. Varga, G. Fai, A. F. Barghouty, B. D. Anderson, A. R. Baldwin, R. Cecil, J. W. Watson and G. D. Westfall, *Phys. Rev. C34*, 1342, 1986.
7. A. F. Barghouty and G. Fai, *Phys. Rev. C35*, 950, 1987.
8. A. Baden, H. H. Gutbrod, H. Lohner, M. R. Maier, A. M. Poskanzer, T. Renner, H. Riedesel, H.-G. Ritter, H. Spieler, A. Warwick, F. Weik and H. Wieman, *Nucl. Instr. Meth.*, *203*, 189, 1982.
9. H.-A. Gustafsson, H. H. Gutbrod, B. Kolb, H. Lohner, B. Ludewigt, A. M. Poskanzer, T. Renner, H. Riedesel, H.-G. Ritter, A. Warwick, F. Weik and H. Wieman, *Phys. Rev. Lett.*, *52*, 1590, 1984.
10. H.-A. Gustafsson, H. H. Gutbrod, B. Kolb, H. Lohner, B. Ludewigt, A. M. Poskanzer, T. Renner, H. Riedesel, H.-G. Ritter, A. Warwick and H. Wieman, *Phys. Lett.*, *142B*, 141, 1984.
11. G. Fai and J. Randrup, *Nucl. Phys.*, *A404*, 551, 1983.
12. *Table of Isotopes*, ed. C. M. Lederer and V. S. Shirley, Wiley, New York, 1978.
13. L. P. Csernai, G. Fai and J. Randrup, *Phys. Lett.*, *140B*, 149, 1984.
14. G. Fai and J. Randrup, *Nucl. Phys.*, *A381*, 557, 1982.
15. J. Kapusta and K. Olive, *Nucl. Phys.*, *A408*, 478, 1983.
16. K. G. R. Doss, H.-A. Gustafsson, H. H. Gutbrod, B. Kolb, H. Lohner, B. Ludewigt, A. M. Poskanzer, T. Renner, H. Riedesel, H.-G. Ritter, A. Warwick and H. Wieman, *Phys. Rev.*, *C32*, 116, 1985.
17. S. Nagamiya, M.-C. Lemaire, E. Moller, S. Schnetzer, G. Shapiro, H. Steiner and I. Tanihata, *Phys. Rev.*, *C24*, 971, 1981.
18. L. P. Csernai, J. I. Kapusta, N. K. Glendenning and G. Fai, *Proc. XIV. International Workshop on Gross Properties of Nuclei and Nuclear Excitations*, Hirschegg, Kleinwalsertal, Austria, January 13–18, 1986. G. Fai, L. P. Csernai and J. I. Kapusta, *Proc. Workshop on Nuclear Dynamics IV*, Copper Mountain, Colorado, 24–28 February, 1986. L. P. Csernai, G. Fai and J. I. Kapusta, preprint, UMSI 85/12, L. P. Csernai, G. Fai, preprint, UMSI 86/14.
19. R. Stock, R. Bock, R. Brockman, J. W. Harris, A. Sandoval, H. Stroebele, K. L. Wolf, H. G. Pugh, L. S. Schroeder, M. Maier, R. E. Renfordt, A. Dacal and M. E. Ortiz, *Phys. Rev. Lett.*, *49*, 1236, 1982.
20. L. P. Csernai, *Phys. Rev. Lett.*, *54*, 639, 1985.



A THREE-FLUID MODEL OF HEAVY ION COLLISIONS*

J. NÉMETH

*Theoretical Physical Department, Roland Eötvös University
1088 Budapest, Hungary*

A time dependent three fluid model is presented to describe intermediate energy heavy ion collisions. Both the mean field interaction and the direct nucleon–nucleon collisions are taken into account. The effect of the collisions term is considered in detail.

Introduction

In heavy ion collisions around the Fermi energy and just above both the mean field and the two-body collisions are expected to play an important role. In the last years attempts have been made to treat this energy domain by using the Vlasov equation with the collision term [1]. The distribution function of the system has been determined numerically. However, the solution of this equation requires enormous numerical efforts. A much simpler method to treat this systems is to apply nuclear hydrodynamics [2]. However the nuclear hydrodynamics is only applicable if one can assume that the changes in the system occur through thermal equilibrium or near-equilibrium states. To assume that the transition of the system from two cold fast moving nuclei (initial state) into a hot expanding nucleus (intermediate state) occurs through near-equilibrium states is very unrealistic. To avoid this assumption, a modified version of the three-fluid model introduced by Csernai et al [3] has been applied. This model seems to us reasonable if we consider the fact that at high bombarding energy the participant–spectator picture gives a fairly good understanding of heavy ion collisions. It distinguishes three zones in the interaction process: the target and the projectile spectator domains (fluid 1 and 2) and the participants (fluid 3), which belong to the region common to the projectile and the target. In the model we consider every particle, which collided at least once as belonging to this third domain. For each of these domains separately it is probably not too bad to assume that they are in local equilibrium, while the total system is far from it. We neglect the deviation for each component from its local thermal equilibrium value, which means we neglect the

* Dedicated to Prof. K. Nagy on his 60th birthday

usual viscosity terms. All our viscosity and dissipation is caused by the collision term is the Boltzmann equation. One advantage of this model is that all the entropy increase occurs during the collision of the two nuclei, while the expansion of the hot system is already isentropic.

In Part 1 we describe shortly our way of treatment of the one-fluid model, to introduce the concepts to be used later. In Part 2 a short description of the three-fluid model is given, while in Part 3 the effect of the collision term is considered in details.

1. The one-fluid hydrodynamics

The Boltzmann–Vlasov equation can be written as:

$$\frac{df}{Dt} = I^{(+)} - I^{(-)}, \quad (1)$$

where f is the one body distribution function ($f = f(\bar{r}, \bar{v}, t)$) and

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \bar{v} \frac{\partial}{\partial \bar{r}} + \frac{\bar{F}}{m} \frac{\partial}{\partial \bar{v}}. \quad (2)$$

\bar{F} is here the force coming from the mean field and m the nucleon mass. $I^{(+)}$ and $I^{(-)}$ are the gain and loss terms, respectively to be considered later in detail. The local density and the average velocity are defined by

$$\rho(\bar{r}, t) = m \int d^3v f(\bar{r}, \bar{v}, t), \quad (3)$$

$$\rho(\bar{r}, t) \bar{u}(\bar{r}, t) = m \int d^3v \bar{v} f(\bar{r}, \bar{v}, t). \quad (4)$$

The hydrodynamical equations are obtained by multiplying both sides of Eq. (1) by m , $m\bar{v}$ and $1/2m\bar{v}^2$ and integrating with respect to \bar{v} . Thus we obtain the continuity equation:

$$\frac{\partial \rho}{\partial t} + \text{div}(\rho \mathbf{u}) = 0, \quad (5)$$

the Euler equation:

$$\frac{d\bar{u}}{dt} = \frac{\partial \bar{u}}{\partial t} + (\bar{u} \nabla) \bar{u} = -\frac{1}{\rho} \nabla P + \frac{\bar{F}}{m} \quad (6)$$

(P is the pressure tensor), and the energy conservation equation, respectively.

Using a Skyrme interaction and applying the Thomas–Fermi approximation, it has been shown in [4] that Eq. (6) can be rewritten as

$$\frac{d\bar{u}}{dt} = -\frac{1}{m} \nabla \frac{\delta f}{\delta \rho}, \quad (6a)$$

where f is the free energy density per unit volume and $\frac{\delta}{\delta\rho}$ the functional derivative with respect to ρ . Assuming an irrotational fluid and introducing a velocity potential S defined by $\vec{u} = \frac{\hbar}{m} \nabla S$, the Euler equation can be transformed into the following equation:

$$-\hbar S = \frac{\delta f}{\delta\rho} + \frac{\hbar^2}{2m} (\nabla S)^2. \quad (7)$$

Eqs (5) and (7) can be regarded as the time dependent Thomas–Fermi equations. They can be reformulated in a homogeneous form by introducing a function ψ defined as

$$\psi = \sqrt{\rho} e^{iS}. \quad (8)$$

This transformation, due to Madelung, allows us to rewrite the TDTF equations into a single equation for the complex function ψ :

$$i\hbar \frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2\psi + V\psi, \quad (9)$$

where the potential energy can be expressed as

$$V = \frac{\delta f}{\delta\rho} + \frac{\hbar^2}{2m} \frac{\nabla\rho^{1/2}}{\rho^{1/2}}.$$

Eq. (9) has the same structure as a Schrödinger equation where ψ stands for the “wave function”. Its numerical solution is quite easy and fast, one can apply all the powerful numerical techniques developed for time dependent Hartree–Fock calculations.

The above described procedure can be extended to viscous fluids, too, by using the Navier–Stokes equation instead of the Euler one. This means that on the right hand side of Eq. (6) an additional term appears containing the viscous tensor σ . If there exists a function G for which

$$\nabla G = \frac{1}{\rho} \nabla\sigma, \quad (10)$$

then the TDTF equations can be again rewritten in the form (9), only the potential is modified by the viscous term

$$V_{\text{visc}} = V - G. \quad (11)$$

2. The three-fluid model

In our model we assume that the spectators moving with a large kinetic energy will remain at zero temperature, while the participants increase their temperature with time. Each fluid is described by a spherical Fermi sphere centred around its $\bar{u}_i(\bar{r}, t)$ mean velocity. At the early stage of the collision, fluid 3 (participants) is not present, it is created by the collisions between nucleons of fluid 1 and 2 (spectators). Nucleons are removed from a fluid of spectators if they collide with any of the other two fluids.

The model based on the above assumptions can be developed as follows.

The distribution function of the total system f can be split into three parts: $f = f_1 + f_2 + f_3$, where each part is associated with a given fluid. They will obey to the following Boltzmann–Vlasov equations [3]:

$$\frac{Df_1}{Dt} = I_{11} - I_{12}^{(-)} - I_{13}^{(-)}, \quad (12)$$

$$\frac{Df_2}{Dt} = I_{22} - I_{21}^{(-)} - I_{23}^{(-)}, \quad (13)$$

$$\frac{Df}{Dt} = I, \quad (14)$$

where in the last equation instead of f_3 we introduced the total distribution function f . In Eqs (12)–(13) $I_{ij}^{(+)}$ and $I_{ij}^{(-)}$ are the gain and loss terms associated with the collisions between fluid i and j , respectively, and $I_{ii} = I_{ii}^{(+)} - I_{ii}^{(-)}$.

Applying the usual multiplication–integration procedure discussed above, Eq. (14) leads to the usual hydrodynamical equations (however, quantities like pressure etc. will be complicated functions of $\rho_1, \rho_2, \rho, \bar{u}_1, \bar{u}_2, \bar{u}$ and T). On the other hand, equations (12) and (13) will lead to unusual continuity and velocity field equations:

$$\frac{\partial \rho_i}{\partial t} + \text{div } \rho_i \bar{u}_i = \frac{2}{\hbar} \rho_i W_i, \quad (15)$$

$$-\hbar S_i = \frac{\delta f}{\delta \rho} + \frac{\hbar^2}{2m} (\nabla S_i)^2 - G_i, \quad (16)$$

where the loss of the flux (W) and the viscosity term (G_i) are due to the two body collisions. (15) and (16) can be transformed into a complex Schrödinger-like equation

$$i\hbar \frac{\partial \psi_i}{\partial t} = -\frac{\hbar^2}{2m} \Delta \psi_i + (V_i - G_i) \psi_i + iW_i \psi_i. \quad (17)$$

The complex potential W_i has two parts in our model: one is temperature independent, coming from the collision of particles of fluid 1 and 2, the other is temperature dependent, coming from the collision of particles of the fluids i and 3. In the last term the Fermi integrals have to be calculated exactly (no T^2 expansion) because the effect of the surface is very important.

3. The collision integral

The total collision integral in (14) can be written as:

$$I = \int [f(v, r, t)f(v_1, rt) - f(v'rt)f(v_1'rt)] \sigma \delta^3(\mathbf{p} + \mathbf{p}_1 - \mathbf{p}' - \mathbf{p}'_2) \cdot \delta(E + E_1 - E' - E'_1) d^3 p_1 d^3 p' d^3 p'_1, \quad (18)$$

where the first term is the loss, the second the gain term, that is to say the first describes the fact that there is loss from the distribution function due to two-body collisions, while the second term gives the gain into it. σ is the collision cross section and $E_i = p_i^2/2m$. As we have seen in Part 1, the integrals of the loss and gain terms do not contribute into the mass-momentum-energy conservation terms, because they cancel each other. However in (11)–(12) only loss terms occur, so they will give an important contribution into (15)–(16). The result of it is that for fluid 1 and 2 we do not have mass conservation, fluid 1 and 2 slowly disappears, while fluid 3 increases. Both W_i and G_i consist of two parts, one coming from the collision of fluid 1 and 2, the second coming from the collisions of fluid i and 3. The first term is temperature independent, the second term is temperature dependent. The collision terms can be calculated exactly analytically, they are complicated functions of the variables. They become simple, however, for the collisions of the cold nuclei if their relative velocities are big compared to the sum of the Fermi velocities. In this case W_1 has the form

$$W_1 = \rho_2 \hbar \sigma |\mathbf{u}_1 - \mathbf{u}_2|, \quad (19)$$

which is the well known formula given by Bohr and Mottelson for the complex potential.

To get the complex Schrödinger type of Eq. (17) we applied again the method we described in Eq. (10)

$$\nabla G_i = \Gamma \frac{1}{m \rho_1}, \quad (20)$$

where Γ can be written as

$$\Gamma_1 = \int (I_{12} + I_{13}) \mathbf{p} d^3 p. \quad (21)$$

Eq. (17) shows clearly one of the main problems which arises in the application of the Boltzmann–Vlasov equation, namely the double counting. The collision integral gives contribution both to the real and the imaginary part of the potential, so it contributes to the mean field, too. Several attempts to solve this problem are under way [5].

As a first step we applied the model for exactly central collisions. The results show that during the collision the two nuclei first slow down quite significantly, due to

the Coulomb and the mean field. The absorptive potential starts to be important only somewhat later. Detailed calculations are in progress, however, already the first results show that the effect of the mean field can be important even at relatively high energies. It also seems to us that the assumption that fluid 1 and 2 have constant velocities in the considered energy domain is unjustified.

References

1. C. Gregoire, B. Remaud, F. Scheuter and F. Sebelle, Nucl. Phys., *A346*, 365, 1985; J. Aichelim and H. Stöcker, to be published.
2. For a review see: H. Stöcker and W. Greiner, to be published in Phys. Rep.
3. L. Csernai, I. Lovas, J. A. Maruhn, A. Rosenhauer, I. Zimányi and W. Greiner, Phys. Rev., *C26*, 149, 1982.
4. J. Németh, M. Barranco, C. Ngo, E. Tomasi, Z. Phys., *A320*, 691, 1985 and to be published in Z. Phys.
5. H. Feldmeier, private communication.

THE EFFECT OF CORRELATIONS ON THE ENTROPY AND HADROCHEMICAL COMPOSITION IN HEAVY ION REACTIONS*

H.-W. BARZ

*Central Nuclear Research Institute
Rossendorf, GDR-8051, Dresden, GDR*

T. S. BIRÓ**, B. LUKÁCS and J. ZIMÁNYI

*Central Research Institute for Physics
1525 Budapest, Hungary*

It is shown that spatial correlations between the constituent particles of a gas lead to roughly excluded volume type corrections in the equation of state. These corrections will appreciably change the hadrochemical composition of fireballs formed in heavy ion reactions.

1. Introduction

Even though energetic heavy ion collisions offer in laboratory the only possibility of achieving such exotic states as several times normal nuclear density with 100 MeV temperature, it is not quite clear how to extract the information about these states from what we finally observe. In particular, the detected particles have emerged from the breakup stage, which is cool and below normal nuclear density, so they cannot bring direct information on the matter at maximum density and temperature, which states may be of great importance because of such possibilities as e.g. deconfinement transition expected at $5-7 n_0$ and/or 100–200 MeV temperature ($n_0 = 0.17 \text{ fm}^{-3}$ being the normal nuclear density) [1, 2].

Siemens and Kapusta have called the attention to the specific entropy [3] which is, of course, a monotonously increasing quantity; furthermore, its increase is moderate after the maximum compression [4, 5]. Since the beam energy (in the CM system) is used up by compression and heating, if one measures the specific entropy of the final state the value of the maximum density can be reconstructed if the equation of state is known [3]; otherwise one can use the information to determine the equation of state.

Nevertheless [3] immediately demonstrated a serious problem called “entropy excess” or “entropy puzzle”. The suggested method for measuring the specific entropy s/n was to observe the final deuteron/proton ratio R_{dp} . It is shown in [3] that for a two-component Boltzmann gas the equilibrium value of this ratio depends only on the

* Dedicated to Prof. K. Nagy on his 60th birthday

** Present address: Institut für Theoretische Physik, Justus Liebig Universität, Heinrich-Buff-Ring 16., D-6300 Giessen, FRG

specific entropy; since they are not too low, deuteron yields are a convenient means of measuring it. Near to the breakup stage the matter is actually diluted, so the Boltzmann approximation is not expected to be too coarse. Now, the reconstructed specific entropy turned out to be too high (even below 1 GeV beam energy) such an extent that it was in contradiction even with the absurd supposition that there was no compression at all.

This "entropy excess" may be explained in two ways. Either a possibility is shown to generate extra entropy, or a correction must be found for the formula needed to reconstruct s/n from R_{dp} . This problem must be clarified because it is an obstacle to hadrochemical calculations. Obviously, until this problem is solved, such calculations yield either a wrong deuteron number or a wrong slope in the detected spectrum.

There are, of course various mechanisms for generating extra entropy. For example, if the equation of state is extremely soft, then more energy may be used for heating. However, if the reconstructed entropy is inconsistent with vanishing compression, then strong attractive forces are needed during the maximum density stage [6]. Today, with experience of higher beam energies, one can rule out the possibility of very exotic stages reached at a beam energy of 0.8 GeV. The second possibility mentioned in [6] was a phase transition. Although equilibrium phase transitions do not generate extra entropy [7], nonequilibrium transitions do [8, 9], and it is not too difficult to produce the required specific entropy excess $\delta(s/n) \simeq 2$ [10, 11]. However, since the energy balance considerations are still valid, below the threshold energy of nuclear transparency such mechanisms offer possibilities for more heating and less compression but cannot explain the total lack of compression or even a rarefaction at full overlap.

The alternative is a deuteron deficiency. One may consider the idea that the deuteron concentration remains below its equilibrium value. Hadrochemical calculations, however, show rapid and efficient chemical equilibration for the deuteron component [5]. In this case the only remaining possibility is to reconsider the Boltzmann approximation even for actually diluted matter. The only support for such an idea is that the deuteron is a very loose system. One may also observe that in the double Boltzmann model the deuteron degree of freedom is dynamically redundant [12]; in this context the inconsistency between predicted and measured R_{dp} values indicates that the deuteron component is really nonredundant. Then the Boltzmann model cannot be completely satisfying.

In a previous paper [13] we demonstrated that an ad hoc volume correction in the equation of state, with a reasonable deuteron volume, can eliminate the problem. Dimensional considerations suggest that this may be true for a quite wide class of equations of state [14], in which, of course, the deuteron component is not redundant in the dynamic sense. Nevertheless, one should show a reasonable way to introduce a deuteron volume into the equation of state since the usual way to obtain van der Waals corrections is to assume strong short-range repulsion [15], which is groundless between a nucleon and deuteron.

In this paper we show that considerations about the correlation volumes of fluctuations really indicate characteristic volumes in the equation of state; by manufacturing such an equation of state the situation discussed in [5] (a relativistic heavy ion collision in the hadrochemical model) is recalculated. The result is that at reasonable values of the volume parameter the entropy excess vanishes.

2. The model

In order to describe the heavy ion collisions we use the same hadrochemical model as in [5]; in fact, our goal is to improve the approximation used there taking finite constituent size effects into consideration. The details of this model can be found in [16], while [5] gives the necessary modifications for the relativistic case, so here we only recapitulate the main points.

The model describes the collision by means of three steps. The first one is ignition: in this stage the thermalisation is described by particle collisions, and the motions of the interpenetrating nuclei are simply extrapolated from the initial conditions. The second step — expansion — starts at total overlap of the colliding nuclei; the still cold nucleons are decoupled, and the remaining hot components are assumed to form a spherical fluid in complete thermal but not chemical equilibrium. The expansion of the sphere is governed by hydrodynamic equations. This step ends when the flow velocities at two neighbouring particles become too different to permit further collisions. Then, as the third step, there is a breakup of the fluid, the particles become decoupled, and one has to sum up the flow and thermal velocities to obtain the detected spectrum [15].

Since, as we shall see, volume effects influence only the equations of state and the chemical reactions, it will be sufficient to discuss the details for the expansion stage; the ignition stage contains the same chemical terms as the expansion one, and our breakup condition remains the same as in [5]. The expansion stage is governed by the equations

$$T^{ir}_{;r} = 0, \quad (2.1)$$

$$(n_a u^r)_{;r} = \psi_a. \quad (2.2)$$

Here T^{ik} is the energy-momentum tensor of the fluid, u^i is the flow four-velocity, while n_a stands for the particle number densities of the hadrochemical components, and ψ_a describes source terms from chemical transmutations. The semicolon denotes a covariant derivative (the coordinates are not necessarily Cartesian [17]) and the Einstein convention [17] is used for the indices, i.e. there is a summation for indices occurring twice, above and below. We take a perfect fluid energy-momentum tensor

$$T^{ik} = (\varepsilon + p)u^i u^k - p g^{ik}, \quad (2.3)$$

where g^{ik} is the metric tensor [17]; ε and p stand for energy density and pressure,

respectively. Then, if the equations

$$p = p(\varepsilon, n_\alpha), \quad (2.4)$$

$$\psi_\alpha = \psi_\alpha(\varepsilon, n_\beta) \quad (2.5)$$

are specified, all the necessary evolution equations are contained by the set of equations (2.1–5) for the independent hydrodynamic and thermodynamic variables u^i, ε and n [7].

Since the full dynamic problem is rather complicated, in [5] we assumed an approximate form for the spatial part of the velocity field

$$\mathbf{u} = \frac{\mathbf{V}}{\sqrt{1-v^2}} = \mathbf{r} \frac{\dot{R}(t)}{R(t)}, \quad (2.6)$$

where $R(t)$ is the radius of the firecloud; by other words that the radial component of the velocity is proportional to the radial coordinate. There are such analytic solutions in the classical limit [18], and this approximation still preserves the global conservation laws for the whole sphere [5].

3. The van der Waals gas

The simplest system which is not an ideal gas is the so called van der Waals gas; neglecting the long range attraction and choosing a specific heat as ideal as possible it satisfies the equations

$$\varepsilon = \frac{3}{2} nT, \quad (3.1)$$

$$p = \frac{nn_0}{n_0 - n} T. \quad (3.2)$$

Here T is the temperature, and n_0 is a constant. Usually these equations are derived from the approximation that repulsive forces are very strong between the particles within a distance R_0 [15]

$$\frac{4\pi}{3} R_0^3 n_0 \simeq 1. \quad (3.3)$$

Let us now start with a thermodynamic description of the system. A simple, one component gas is described by 3 extensive parameters, which may be chosen as the volume V , energy E and particle number N [19]. With these variables the appropriate potential is the entropy $S = S(V, E, N)$, which is also an extensive, i.e. it is a homogeneous linear function of its variables

$$S(\lambda V, \lambda E, \lambda N) = \lambda S(V, E, N). \quad (3.4)$$

Hence one obtains the Euler identity

$$S = X^R \frac{\partial S}{\partial X^R}, \quad (3.5)$$

where X^I stands for the independent extensives ($X^1 = V$, $X^2 = E$, $X^3 = N$). By introducing the shorthand notation

$$Y_I = \frac{\partial S}{\partial X^I} \quad (3.6)$$

and using the individual names

$$\begin{aligned} Y_1 &= \frac{p}{T}, \\ Y_2 &= \frac{1}{T}, \\ Y_3 &= -\frac{\mu}{T}, \end{aligned} \quad (3.7)$$

one can recognize the entropic version of the First Law of Thermodynamics in Eq. (3.5). By identifying T with the temperature and p with the mechanical pressure, as recommended [20], partial differential equations are obtained for the entropy density $s(\varepsilon, n)$ from Eqs (3.1–2), whose general solution is

$$s = n \left[\frac{3}{2} \ln \varepsilon - \frac{5}{2} \ln n + K + \ln \left(1 - \frac{n}{n_0} \right) \right]. \quad (3.8)$$

Here the constant K affects the zero point of the entropy; since for Boltzmann gases the entropy is not bounded from below, it cannot directly be fixed by means of the Third Law, however, if necessary, K can be calculated in a roundabout way through the Boltzmann limit of a Fermi gas.

It is interesting that for a given system the entropy density is almost uniquely fixed solely by the thermodynamic behaviour; even without any mechanical identification of the Y_I 's the only remaining transformation is [21]

$$s \rightarrow A^2 s + B^2 \varepsilon + Cn + D \quad (3.9)$$

common for all the systems which may become connected, where capital letters stand for real constants. Eq. (3.6) shows that this free transformation is a temperature rescaling plus zero point shifts for the entropic intensives Y_I .

4. Spatial correlations

Due to the entropy maximum principle, the matrix g_{IK}

$$g_{IK} = - \frac{\partial^2 s}{\partial \rho^I \partial \rho^K},$$

$$\rho^I = \frac{1}{V} X^I \quad (4.1)$$

is positive definite for stable states. Therefore it (or the analogous entity obtained by mutually substituting s by ε) generates a distance in the thermodynamic state space, and one can introduce Riemannian geometry too when generic coordinates are used [22, 23, 24]. This description is equivalent with the usual phenomenological formulation, moreover, it yields some prediction even for the fluctuation probabilities [25].

By means of the probability distribution of the fluctuation one can calculate the correlation volumes of the system, i.e. the characteristic volumes in which the fluctuations are spatially correlated. As it is shown in [25], such volumes V_c are given by the eigenvalue equation

$$(M^{IR} - V_c g^{IR})_{V_R} = 0, \quad (4.2)$$

where in extensive coordinates M^{IK} takes the form

$$M^{IK} = g^{IK}_{,RS} g^{RS}, \quad (4.3)$$

where the comma stands for partial derivative.

One can observe an intimate connection between the van der Waals parameter n_0 and a correlation volume. The detailed description is given in [26]; here we only recapitulate the main results. Consider first the equation of state (3.8). Then Eq. (4.2) possesses two solutions

$$V_{c1} = \frac{4}{3} \frac{1}{n} (1 + O(n/n_0)), \quad (4.4a)$$

$$V_{c2} = - \frac{4}{n_0} (1 + O(n/n_0)). \quad (4.4b)$$

(The negative sign in Eq. (4.4b) indicates anticorrelation.) Now, the first correlation volume persists even for an ideal Boltzmann gas, so it cannot be interpreted as a result of interactions; in fact, it simply expresses the existence of indivisible particles in the gas [25, 26]. On the other hand, the second one vanishes for an ideal gas, therefore it is the consequence of interactions. In view of this one can conclude that the van der Waals gas possesses a characteristic correlation volume of the order of $1/n_0$, which is by no means surprising because such an equation of state can be obtained by assuming volume exclusion.

Thus, V_{c2} is proportional to the excluded volume for van der Waals gases. This relationship suggests that V_{c2} should appear in the equation of state. However, it is not a trivial task to invert the connection between van der Waals corrections and V_{c2} because Eq. (4.2) is a partial differential equation of fourth order for $s(\varepsilon, n)$. Nevertheless, if

- 1) the correlation volume is constant, and there is anticorrelation;
- 2) the deviation from the ideal gas entropy can be expanded into a Taylor series at $n=0$;
- 3) there is an ideal gas limit for $n \rightarrow 0$;
- 4) the diluted high temperature states remain near to ideal behaviour, then, for first order in n/n_0 , one obtains

$$s = n \left\{ s_0 + \frac{3}{2} \ln \varepsilon - \frac{5}{2} \ln n - \frac{n}{n_0} + O((n/n_0)^2) \right\}. \quad (4.5)$$

By comparing this with the van der Waals entropy (3.8) they are identical up to first order terms in the corrections.

Therefore, in some sense, spatial anticorrelations result in a van der Waals correction in the approximations described above. This conclusion is further supported by dimensional considerations as well as by the observation that a nonvanishing V_{c2} necessarily leads to pressure corrections [26].

5. Nucleon-deuteron mixtures

Consider now some kind of nuclear matter in which the existence of nucleons and deuterons is postulated. Obviously, the deuteron, being a weakly bound system, is always affected by the neighbouring particles. Nevertheless, because here we will finally be interested in deuteron yields, any bound two-nucleon state may be considered as a deuteron. Then, the thermodynamic potential possesses the form

$$s = s(n_n, n_d, \varepsilon), \quad (5.1)$$

where n_n and n_d stand for the particle densities of nucleons and deuterons, respectively. It is reasonable to assume that at low densities the system goes to a two component ideal gas thus [15]

$$s = n \left(\frac{3}{2} \ln \varepsilon - \frac{3}{2} \ln n \right) - n_n \ln n_n - n_d \ln n_d + n\eta(n_n, n_d, \varepsilon),$$

$$n = n_n + n_d, \quad (5.2)$$

with an η vanishing in the limit $n_n \rightarrow 0$, $n_d \rightarrow 0$. In this system Eq. (4.2) would yield 3 different correlation volumes of which the first in the low density limit is the same as in Eq. (4.4a); the other two are nontrivial depending on the form of function η .

Now, it seems that the correlation of longest range established by interactions is the deuteron breakup: since the deuteron is weakly bound, it dissolves into two nucleons if a third one is too near. This distance seems to be in the same order of magnitude as the deuteron size defined from the wave function, which is definitely not smaller than the interparticle distances in a normal nuclear matter. In view of this there is no indication by virtue of which we could neglect this anticorrelation volume.

If the density and energy dependence of this breakup volume were known, one would again obtain a fourth order partial differential equation for $\eta(n_n, n_d, \varepsilon)$. Nevertheless, obviously the full information about the breakup volume is not within reach, so this way would not be constructive. Instead one can look for approximating model systems with similar anticorrelation volumes, and the one component results as well as dimensional considerations [14] again suggest a van der Waals gas. Therefore, as a generalization of Eqs (3.1–2), we assume

$$\varepsilon = \frac{3}{2} nT,$$

$$p = \frac{n}{1 - n_n V^n - n_d V^d} T, \quad (5.3)$$

where V^n and V^d are some characteristic volumes analogous to $1/n_0$; one may assume that $V^n \ll V^d$, but that is not necessary at this stage. The general solution for entropy density with ε and p given by Eq. (5.3) is as follows:

$$s = s_{id} + n\Phi(n_d/n_n) + n \ln(1 - n_n V^n - n_d V^d), \quad (5.4)$$

where s_{id} is given by Eq. (5.2) when $\eta = 0$, while Φ is a free function of one variable. Since Φ depends on density ratios, s/n has no unique limit for $n_n \rightarrow 0$, $n_d \rightarrow 0$ unless $\Phi = \text{const.}$ which, consequently, will be assumed here as a formulation of the existence of a Boltzmann limit. Then a constant Φ can be absorbed by the zero point constant not discussed here.

6. The evolution equations

Let us return to the evolution equations (2.1–6). The functions to be specified are the pressure and source functions. The pressure $p = p(n_n, n_d, \varepsilon)$ can be obtained via Eqs (3.7) from the entropy function, which is given by Eq. (5.4). The source terms ψ_x may be quite complicated, however, near to chemical equilibrium their forms are rather transparent. For the general structure see [16]; here we demonstrate the situation for ψ_d which is directly affected by volume effects.

There are two competing processes in ψ_d : deuteron production and breakup. There are equilibrium states, when $n_d = n_{deq}$, so

$$\psi_d(\varepsilon, n_n, n_{deq}(\varepsilon, n_n)) = 0. \quad (6.1)$$

The breakup is dominantly a result of two-body collisions, whereas deuteron production needs three nucleons for energy and momentum conservation, therefore

$$\begin{aligned}\psi_d &\simeq -\langle \sigma_{nd} v_{rel} \rangle n_n (n_d - \rho_d^{nn} n_n^2), \\ \rho_d^{nn} &= (n_d/n_n^2)_{eq}.\end{aligned}\quad (6.2)$$

Here σ_{nd} is the breakup cross section, and v_{rel} is the relative velocity between n and d ; ρ_d^{nn} is expected to depend on local thermodynamic quantities. The symbol $\langle \sigma_{nd} v_{rel} \rangle$ stands for the average value of the product of $\sigma_{nd} v_{rel}$.

Now, ρ_d^{nn} belongs to chemical equilibrium, thus it can be calculated from the entropy function. Namely, in equilibrium

$$\mu_d = 2\mu_n \quad (6.3)$$

(where μ stands for chemical potential), otherwise there would be entropy production in a transmutation [5]. But then, according to Eqs (3.6–7), (5.2) and (5.4), one obtains

$$\left(\frac{n_d}{n_n^2}\right)_{eq} = \rho_{d,id}^{nn} = \rho_{d,id}^{nn} \frac{1}{1 - n_n V^v - n_d V^d} e^{-\frac{(n_n + n_d)(V^d - 2V^n)}{1 - n_n V^n - n_d V^d}}, \quad (6.4)$$

where $\rho_{d,id}^{nn}$ belongs to the Boltzmann limit

$$\rho_{d,id}^{nn} = \frac{3}{2} \left(\frac{\pi \hbar^2}{mT}\right)^{3/2}. \quad (6.5)$$

Now let us consider the limiting case when the nucleons are pointlike and the deuterons are actually diluted, i.e.

$$\begin{aligned}n_n V^n &\rightarrow 0, \\ n_d V^d &\rightarrow 0.\end{aligned}\quad (6.6)$$

Then

$$\begin{aligned}p &\rightarrow (n_n + n_d)T = \frac{2}{3}\varepsilon, \\ \rho_d^{nn} &\rightarrow \rho_{d,id}^{nn} e^{-(n_n + n_d)(V^d - 2V^n)}.\end{aligned}\quad (6.7)$$

Thus, even in this limit there remains a correction in the equilibrium concentration, suppressing the deuterons. Another possible interpretation is that, according to Eq. (5.4), at a fixed n_d value the entropy density is lower than calculated from an ideal gas model.

The characteristic combination occurring in ρ_d^{nn} is the volume difference $V^d - 2V^n$ expressing the volume excess when a deuteron is produced. This quantity will be denoted here by δV . V^d can be estimated in two different ways. First, it can be calculated from the deuteron wave function as

$$V^d \simeq \frac{4\pi}{3} \langle r_d \rangle^3 \simeq 20 \text{ fm}^3 \quad (6.8)$$

where the value is obtained by using the wave function of [27], containing a repulsive core. Alternatively, the deuteron breakup volume can be estimated from the breakup cross section σ_{bu} as

$$V^d \sim V^{bu} \simeq \frac{4}{3\sqrt{\pi}} \sigma_{bu}^{3/2}. \quad (6.9)$$

With $\sigma_{bu} \sim 100$ mb [28] V_{bu} is again approximately 20 fm^3 . For V^n one may choose $V^n = (4\pi/3)r_0^3 = 5 \text{ fm}^3$. Thus, $\delta V \simeq 10 \text{ fm}^3$ seems to be a decent order of magnitude estimation, which will be used in the calculations.

According to [3, 13] and the results of this Section, the specific entropy is connected to the d/p ratio as

$$s_B/n_B = 3.95 - \ln R_{dp} - 1.25 R_{dp}/(1 + R_{dp}) - (n_n + n_d)(V^d - V^n). \quad (6.10)$$

7. Results and conclusions

The results of the calculations are summarized in three Figures. The calculation ends with a breakup of the firecloud at the moment when there is no further possibility to maintain thermal equilibration. This happens when the cooling becomes faster than the collisions between particles; the situation can be characterized by a dimensionless parameter a :

$$a = \tau_{\text{cooling}}/\tau_{\text{collision}} \quad (7.1)$$

(for the complete definition cf. [5]). At the breakup $a \simeq 1$; the actual value — according to the best possible proton spectrum — was chosen as 1.6. Figure 1 displays the hadrochemical processes for Ar + Ar collision at 0.8 GeV/A. One can observe that the $\Delta + \pi$ number is already saturated shortly after maximal compression (as the deuteron number in [5] without volume corrections), while now the deuteron number is monotonously increasing due to volume effects becoming more and more negligible with decreasing density. At breakup the total deuteron number is about 6, which is a roughly 25% decrease compared with $\delta V=0$ [5]. Figure 2 shows the predicted deuteron/proton ratios for Ne + NaF with and without volume corrections; observe the good agreement with medium energy experiments at $\delta V=10 \text{ fm}^3$. At 2.1 GeV beam energy no volume correction is needed for explaining the deuteron yield [5], this may be connected with a temperature dependence of the correlation volume. Finally, Fig. 3 compresses the results into a modified version of the figure of Siemens and Kapusta [3]. One can see that for $\delta V=10 \text{ fm}^3$ the specific entropy reconstructed from the observables is just within the expected range therefore no entropy excess is found.

A possible conclusion is that the low deuteron yield is a consequence of the great (geometrical or correlation) deuteron volume and can be explained by a deuteron

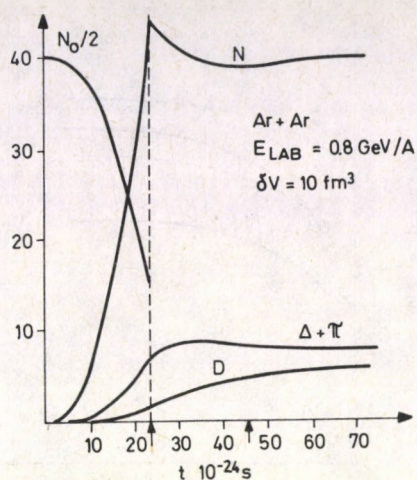


Fig. 1. Hydrochemical evolution in a 0.8 GeV beam energy Ar + Ar collision; the chosen deuteron volume excess is 10 fm^3 . Observe that the deuteron number is not saturated at breakup (indicated by an arrow), whereas the saturation actually occurs just after maximum compression when the volume excess is neglected [5]. The deuteron number decrease caused by volume effects is roughly 25%.

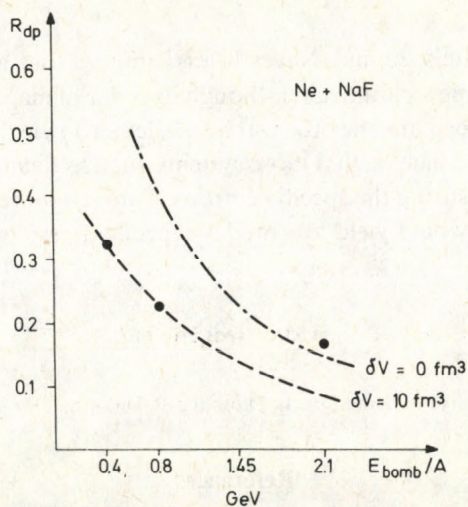


Fig. 2. Predicted deuteron/proton ratios in Ne + NaF collisions at different beam energies. The dots indicate the experimental values (cf. [29]). Observe the good agreement for $\delta V = 10 \text{ fm}^3$ at moderate energies

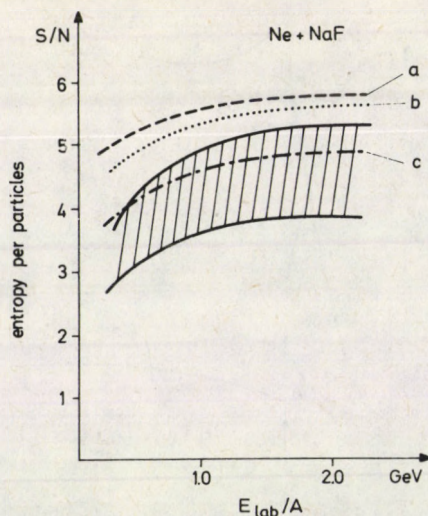


Fig. 3. Specific entropy S/N versus beam energy for Ne + NaF. Curves a, b and c are specific entropies calculated from the measured R_{dp} values: a for an ideal one component gas, b for an ideal two-component one, while c belongs to our two-component van der Waals equation of state (Eq. (3.8)) with $V^d - 2V^n = 10 \text{ fm}^3$. The shaded area is the prediction of Siemens and Kapusta [3] for hadronic matter. Since Curve c lies in the shaded area, the entropy excess vanishes for $\delta V = 10 \text{ fm}^3$

volume of approximately 20 fm^3 . Nevertheless, observe that at the present stage of understanding V^d is a new parameter, although its order of magnitude is not arbitrary. Therefore, in order to explain one observable (the deuteron yield) one new parameter is needed. One thus can conclude that loose systems, such as deuterons, are not the most suitable tools for measuring the specific entropy. Perhaps more compact small nuclei such as t , He^3 and α would yield less model-dependent results.

Acknowledgement

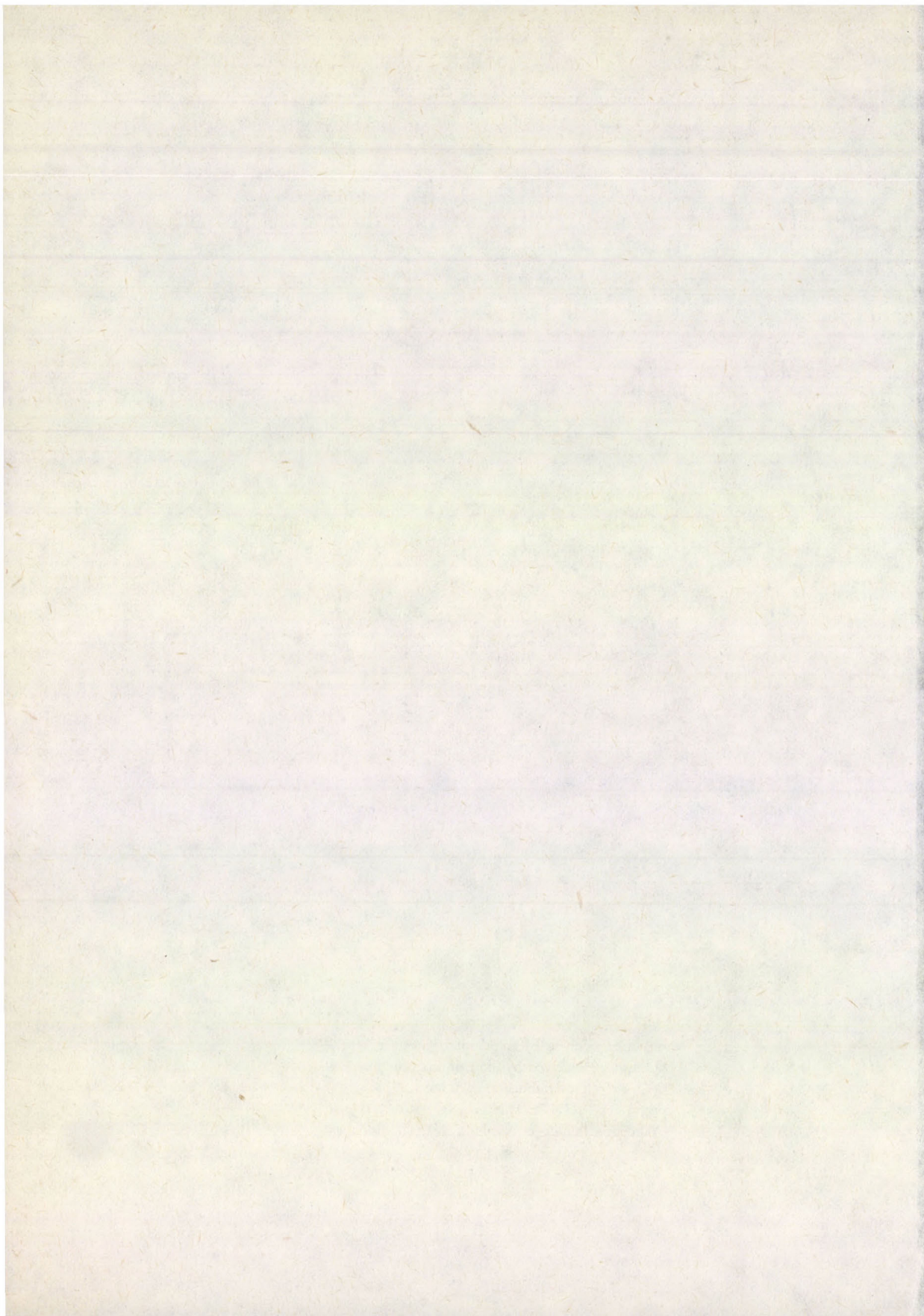
The authors would like to thank Drs. L. Diósi and P. Doleschall for illuminating discussions.

References

1. A. S. Chin, Phys. Lett., 78B, 552, 1978.
2. J. Kuti, B. Lukács, J. Polónyi and K. Szlachányi, Phys. Lett., B95, 75, 1980.
3. P. J. Siemens and J. I. Kapusta, Phys. Rev. Lett., 43, 1486, 1979.
4. L. P. Csernai, B. Lukács and J. Zimányi, Lett. Nuovo Cim., 27, 111, 1980.
5. T. S. Biró, H. W. Barz, B. Lukács and J. Zimányi, Phys. Rev., C27, 2695, 1983.
6. I. M. Mishustin, F. Myhrer and P. I. Siemens, Phys. Lett., 95B, 361, 1980.
7. B. Lukács, Acta Phys. Pol., B14, 33, 1983.
8. L. P. Csernai and B. Lukács, Phys. Lett., 132B, 295, 1983.

9. L. P. Csernai and B. Lukács, Proc. 6th Balaton Conf. on Nucl. Phys. (ed. J. Erő), 1983, p. 363.
10. H. W. Barz, B. Kämpfer, L. P. Csernai and B. Lukács, Phys. Lett., *143B*, 334, 1984.
11. L. P. Csernai, H. W. Barz, B. Kämpfer and B. Lukács, Phys. Rev., *C31*, 268, 1985.
12. B. Lukács and K. Martinás, Acta Phys. Slovaca, *36*, 81, 1986.
13. T. S. Biró, H. W. Barz, B. Lukács and J. Zimányi, KFKI-1981-90.
14. B. Lukács, Proc. 6th Balaton Conf. on Nucl. Phys. (ed. J. Erő), 1983, p. 367.
15. T. M. Reed and K. E. Gubbins, Applied Statistical Mechanics. Thermodynamic and Transport Properties of Fluids, McGraw-Hill, New York, 1974.
16. I. Montvay and J. Zimányi, Nucl. Phys., *316A*, 490, 1979.
17. L. P. Eisenhardt, Riemannian Geometry, Princeton Univ. Press, New York 1950.
18. J. Bondorf, S. Garpman and J. Zimányi, Nucl. Phys., *A296*, 320, 1978.
19. H. B. Callen, Thermodynamics, J. Wiley, New York, 1960.
20. L. Tisza, Ann. Phys., *13*, 1, 1961.
21. B. Lukács and K. Martinás, Phys. Lett., *114A*, 306, 1986.
22. F. Weinhold, J. Chem. Phys., *63*, 2479, 1975.
23. G. Ruppeiner, Phys. Rev., *A20*, 1608, 1979.
24. L. Diósi, G. Forgács, B. Lukács and H. L. Frisch, Phys. Rev., *A29*, 3343, 1984.
25. L. Diósi and B. Lukács, Phys. Lett., *112A*, 13, 1985.
26. L. Diósi and B. Lukács, J. Chem. Phys. *84*, 5081, 1986.
27. B. Gyarmati and J. Zimányi, Phys. Lett., *28B*, 363, 1969.
28. H. C. Catron, M. D. Goldberg, R. W. Hill, J. M. Le Blanc, J. P. Stoering, C. J. Taylor and M. A. Williamson, Phys. Rev., *123*, 218, 1961.
29. S. Nagamiya, M. C. Lemaire, E. Moeller, S. Schnetzer, G. Shapiro, H. Steiner and I. Tanihata, Lawrence Berkeley Laboratory Report LBL-12123, 1981.

Manuscript received by Akadémiai Kiadó: 6 June 1986
Manuscript received by the Printers: 16 June 1986
Date of publication: 20 May 1988
PRINTED IN HUNGARY
Akadémiai Kiadó és Nyomda Vállalat, Budapest



NOTES TO CONTRIBUTORS

I. PAPERS will be considered for publication in *Acta Physica Hungarica* only if they have not previously been published or submitted for publication elsewhere. They may be written in English, French, German or Russian.

Papers should be submitted to

Prof. I. Kovács, Editor
Department of Atomic Physics, Technical University
1521 Budapest, Budafoki út 8, Hungary

Papers may be either articles with abstracts or short communications. Both should be as concise as possible, articles in general not exceeding 25 typed pages, short communications 8 typed pages.

II. MANUSCRIPTS

1. Papers should be submitted in three copies.
2. The text of papers must be of high stylistic standard, requiring minor corrections only.
3. Manuscripts should be typed in double spacing on good quality paper, with generous margins.
4. The name of the author(s) and of the institutes where the work was carried out should appear on the first page of the manuscript.
5. Particular care should be taken with mathematical expressions. The following should be clearly distinguished, e.g. by underlining in different colours: special founts (italics, script, bold type, Greek, Gothic, etc.); capital and small letters; subscripts and superscripts, e.g. x^2 , x_3 ; small l and I ; zero and capital O ; in expressions written by hand: e and l , n and u , v and ν , etc.
A List of Symbols on a separate sheet should be attached to each paper.
6. References should be numbered serially and listed at the end of the paper in the following form: J. Ise and W. D. Fretter, *Phys. Rev.*, 76. 933, 1949.
For books, please give the initials and family name of the author(s), title, name of publisher, place and year of publication, e.g.: J. C. Slater, *Quantum Theory of Atomic Structures*, I. McGraw-Hill Book Company, Inc., New York, 1960.
References should be given in the text in the following forms: Heisenberg [5] or [5].
7. Captions to illustrations should be listed on a separate sheet, not inserted in the text.
8. In papers submitted to *Acta Physica* all measures should be expressed in SI units.

III. ILLUSTRATIONS AND TABLES

1. Each paper should be accompanied by three sets of illustrations, one of which must be ready for the blockmaker. The other sets attached to the copies of the manuscript may be rough drawings in pencil or photocopies.
2. Illustrations must not be inserted in the text.
3. All illustrations should be identified in blue pencil by the author's name, abbreviated title of the paper and figure number.
4. Tables should be typed on separate pages and have captions describing their content. Clear wording of column heads is advisable. Tables should be numbered in Roman numerals (I, II, III, etc.).

IV. RETURN OF MATERIAL

Owing to high postage costs, the Editorial Office cannot undertake to return *all* material not accepted for any reason for publication. Of papers to be revised (for not being in conformity with the above Notes or other reasons) only *one* copy will be returned. Material rejected for lack of space or on account of the Referees' opinion will not be returned to authors outside Europe.

Periodicals of the Hungarian Academy of Sciences are obtainable
at the following addresses:

AUSTRALIA

C.B.D. LIBRARY AND SUBSCRIPTION SERVICE
Box 4886, G.P.O., Sydney N.S.W. 2001
COSMOS BOOKSHOP, 145 Ackland Street
St. Kilda (Melbourne), Victoria 3182

AUSTRIA

GLOBUS, Höchstädtplatz 3, 1206 Wien XX

BELGIUM

OFFICE INTERNATIONAL DES PERIODIQUES
Avenue Louise, 485, 1050 Bruxelles
E. STORY-SCIENTIA P.V.B.A.
P. van Duyseplein 8, 9000 Gent

BULGARIA

HEMUS, Bulvar Ruszki 6, Sofia

CANADA

PANNONIA BOOKS, P.O. Box 1017
Postal Station "B", Toronto, Ont. M5T 2T8

CHINA

CNPICOR, Periodical Department, P.O. Box 50
Peking

CZECHOSLOVAKIA

MAD'ARSKA KULTURA, Národní třída 22
115 66 Praha
PNS DOVOZ TISKU, Vinohradská 46, Praha 2
PNS DOVOZ TLAČE, Bratislava 2

DENMARK

EJNAR MUNKSGAARD, 35, Nørre Søgade
1370 Copenhagen K

FEDERAL REPUBLIC OF GERMANY

KUNST UND WISSEN ERICH BIEBER
Postfach 46, 7000 Stuttgart 1

FINLAND

AKATEEMINEN KIRJAKAUPPA, P.O. Box 128
00101 Helsinki 10

FRANCE

DAWSON-FRANCE S.A., B.P. 40, 91121 Palaiseau
OFFICE INTERNATIONAL DE DOCUMENTATION ET
LIBRAIRIE, 48 rue Gay-Lussac
75240 Paris, Cedex 05

GERMAN DEMOCRATIC REPUBLIC

HAUS DER UNGARISCHEN KULTUR
Karl Liebknecht-Straße 9, DDR-102 Berlin

GREAT BRITAIN

BLACKWELL'S PERIODICALS DIVISION
Hythe Bridge Street, Oxford OX1 2ET
BUMPUS, HALDANE AND MAXWELL LTD.
Cowper Works, Olney, Bucks MK46 4BN
COLLET'S HOLDINGS LTD., Denington Estate,
Wellingborough, Northants NN8 2QT
WM DAWSON AND SONS LTD., Cannon House
Folkstone, Kent CT19 5EE
H. K. LEWIS AND CO., 136 Gower Street
London WC1E 6BS

GREECE

KOSTARAKIS BROTHERS INTERNATIONAL
BOOKSELLERS, 2 Hippokratous Street, Athens-143

HOLLAND

FAXON EUROPE, P.O. Box 167
1000 AD Amsterdam
MARTINUS NIJHOFF B. V.

Lange Voorhout 9-11, Den Haag
SWETS SUBSCRIPTION SERVICE
P.O. Box 830, 2160 Sz Lisse

INDIA

ALLIED PUBLISHING PVT. LTD.
750 Mount Road, Madras 600002
CENTRAL NEWS AGENCY PVT. LTD.
Connaught Circus, New Delhi 110001
INTERNATIONAL BOOK HOUSE PVT. LTD.
Madame Cama Road, Bombay 400039

ITALY

D. E. A., Via Lima 28, 00198 Roma
INTERSCIENTIA, Via Mazzè 28, 10149 Torino
LIBRERIA COMMISSIONARIA SANSONI
Via Lamarmora 45, 50121 Firenze
SANTO VANASIA, Via M. Macchi 58
20124 Milano

JAPAN

KINOKUNIYA COMPANY LTD.
Journal Department, P.O. Box 55
Chitose, Tokyo 156
MARUZEN COMPANY LTD., Book Department
P.O. Box 5050 Tokyo International, Tokyo 100-31
NAUKA LTD., Import Department
2-30-19 Minami Ikebukuro, Toshima-ku, Tokyo 171

KOREA

CHULPANMUL, Phenjan

NORWAY

TANUM-TIDSKRIFT-SENTRALEN A.S.
Karl Johansgata 43, 1000 Oslo

POLAND

WĘGIERSKI INSTYTUT KULTURY
Marszałkowska 80, 00-517 Warszawa
CKP I W, ul. Towarowa 28, 00-958 Warszawa

ROUMANIA

D. E. P., Bucuresti
ILEXIM, Calea Grivitei 64-66, Bucuresti

SOVIET UNION

SOYUZPECHAT — IMPORT, Moscow
and the post offices in each town
MEZH DUNARODNAYA KNIGA, Moscow G-200

SPAIN

DIAZ DE SANTOS Lagasca 95, Madrid 6

SWEDEN

ESSELTE TIDSKRIFTS-CENTRALEN
Box 62, 101 20 Stockholm

SWITZERLAND

KARGER LIBRI AG, Petersgraben 31, 4011 Basel

USA

EBSCO SUBSCRIPTION SERVICES
P.O. Box 1943, Birmingham, Alabama 35201
F. W. FAXON COMPANY, INC.
15 Southwest Park, Westwood Mass. 02090
MAJOR SCIENTIFIC SUBSCRIPTIONS
1851 Diplomat, P.O. Box 819074,
Pallas, Tx. 75381-9074
READ-MORE PUBLICATIONS, INC.
140 Cedar Street, New York, N. Y. 10006

YUGOSLAVIA

JUGOSLOVENSKA KNJIGA, Terazije 27, Beograd
FORUM, Vojvode Mišića 1, 21000 Novi Sad