CWI Syllabi

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

The CWI is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

CWI Syllabus 6

Proceedings Seminar 1982-1983 Mathematical structures in field theories

P.J.M. Bongaarts, J.N. Buur, E.A. de Kerf,
R. Martini, H.G.J. Pijls, J.W. de Roever

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

1980 Mathematics Subject Classification:
81EXX, 81E10, 53BXX, 55RXX, 55N30.
ISBN 90 6196 290 0

TABLE OF CONTENTS

PREFACE

These proceedings cover the lectures given in the seminar "Mathematical Structures in Field Theories", held during the academic year 1982-1983. The first part concerns papers, related to "Quantum Field Theory", and the second part is devoted to mathematical methods in connection with Yang-Mills theories.

Chapter 1 by P.J.M. Bongaarts deals with the method of Feynman's path integral and its application to the anharmonic oscillator as a simplified model for perturbative quantum field theory.

E.A. de Kerf treats in Chapter 2 topological solitons in several dimensions, Derricks theorem and in particular, a gauge model in (1+2) dimension. The first part of these proceedings is concluded with a contribution by J. Buur and E.A. de Kerf, where a geometrical description of fields and Lagrangians is presented in order to define the generating functional for Green's functions.

Chapter 4 and 5 by H.G.J. Pijls and J.W. de Roever are devoted to Ward's construction of Yang-Mills potentials.

Chapter 6 by R. Martini is concerned with representations of holomorphic solutions of the massless field equations; these solutions are viewed as elements of a cohomology group.

Because of the importance of cohomology for field equations Chapter 7 by J.W. de Roever, gives a survey of cohomology theory, in particular, cohomology of sheafs.

The same author treats in the final Chapter the Penrose transformation, which is a very suitable tool for solving linear as well as nonlinear field equations. The author points out the relation between this transformation and the Ward construction.

Finally, we thank the Centre for Mathematics and Computer Science for the excellent technical production of this book.

The editors,

E.M. de Jager
H.G.J. Pijls

FEYNMAN'S PATH INTEGRAL. THE ANHARMONIC OSCILLATOR AS A SIMPLIFIED MODEL FOR PERTURBATIVE QUANTUM FIELD THEORY

by

P.J.M. BONGAARTS
Lorentz Institute, University of Leiden

SUMMARY

Feynman's path integral is at present widely used in quantum field theory, in particular as a means for generating in a systematic way the formal perturbation expansion of matrix elements of the S-operator.

As a mathematical model a quantum field is equivalent to an infinite system of oscillators. Instead of venturing directly into the full complexities of quantum field theory, one may therefore first study the essential features of the path integral method in the simple context of a single quantum oscillator.

With this in mind we shall in these lectures discuss the Feynman path integral for one-dimensional non-relativistic quantum mechanics, first quite generally, and then specialized to the particular case of an oscillator, with emphasis on the derivation of the formal series for the transition amplitudes under anharmonic perturbation.

ACKNOWLEDGEMENT

The research connected with these lectures is part of the program of the "Stichting voor Fundamenteel Onderzoek der Materie(FOM)", which is financially supported by the "Organisatie voor Zuiver Wetenschappelijk Onderzoek (ZWO)".

CONTENTS

1. INTRODUCTION

Feynman's path integral plays a central rôle in present day quantum field theory. It has, in particular, shown its value in the development of non-abelian gauge theories, models that go back to the work of Yang and Mills in the fifties and that now dominate elementary particle theory.

Mathematically a quantum field can be understood as an infinite system of quantum oscillators. For a free field this means uncoupled harmonic oscillators. The physically more interesting case of a field with interaction is then a perturbation by non-linear anharmonic coupling terms. The effectiveness of the path integral method in field theory is strongly connected with this oscillator picture, which may therefore serve as an excellent vehicle for explaining its essential features.

It is with this in mind that we shall give an exposition of Feynman's path integral for non-relativistic, one dimensional quantum mechanics, more particularly for a one dimensional harmonic oscillator, free and with various perturbations. This will be done in such a way that it can indeed be used as an introduction to the subject of Feynman's path integral in quantum field theory.

Our discussion will stress aspects of mathematical structure, however, without much pretence of analytical rigour. This can hardly be otherwise as most of the path integral method even for non-relativistic quantum mechanics exists still only on a heuristic level.

The path integral concept in the traditional version that we shall present here applies only to boson theories. An analogue for the fermion case has been developed. Superficically it looks very similar but it is mathemathically very different. The fermion path integral is not an integral, not even in a heuristic sense. Although of great interest it will not be discussed here.

2. THE FORMALISM OF QUANTUM THEORY

The description of a physical system in quantum theory has the following main features:

1. There is a complex Hilbert space H. The unit vectors ψ in H represent possible *states* of the system (ψ and $e^{i\alpha}\psi$ represent the same state)

2. The self-adjoint or as the physicists say hermitian operators A in H represent *observable quantities* of the system. One has momentum operators, position operators, energy operators, etc. The theory gives predictions that can be verified by experiments and that are always of statistical nature: If the system is in a state ψ there is for every quantity A a probability distribution for the possible outcomes of measurement of A. The expectation or average value of the distribution is given by the inner product $(\psi,A\psi)$. The complete distribution can in principle be recovered from the pair ψ and A. It is roughly speaking such that its higher moments are equal to $(\psi,A^n\psi)$. (Note that we use the physicist's convention for the inner product $(\psi_1,\lambda\psi_2) = \lambda(\psi_1,\psi_2)$

3. The dynamics i.e. the *time evolution* of the system is described by a system of unitary evolution operators $U(t_1,t_0)$ in H. They have the characteristic properties:

$$
\begin{aligned}
&\text{a.} \quad U(t,t) = 1 \qquad\qquad \forall t \in \mathbb{R} \\
&\text{b.} \quad U(t_3,t_2)U(t_2,t_1) = U(t_3,t_1) \qquad \forall t_1, t_2, t_3 \in \mathbb{R} \\
&\text{c.} \quad \text{suitable continuity properties for the dependence} \\
&\qquad\quad \text{on } t_1,\ t_0
\end{aligned}
\tag{1}
$$

One may use the $U(t,t_0)$ to make the state vectors time dependent according to

$$\psi(t) = U(t,t_0)\psi(t_0) \tag{2}$$

The system of operators $U(t,t_0)$ then contains the general solution of the ordinary differential equation for the time evolution of state vectors that can be written as

$$\frac{d}{dt}\psi(t) = -\frac{i}{\hbar}H(t)\psi(t) \tag{3}$$

with $H(t)$ an in general time dependent operator which is self-adjoint, is called the Hamiltonian operator and is given ($\hbar$ is Planck's constant)
One may distinguish two cases:

a. The autonomous or "time independent" case. The operators $U(t,t_0)$ depend only on the difference $t - t_0$. This is equivalent with the fact that the Hamiltonian operator H does not depend on t. The evolution operators form a one-parameter group of unitary operators with as generator the constant operator H. They can be written as

$$U(t,t_0) = U(t-t_0) = e^{-\frac{i}{\hbar}(t-t_0)H} \tag{4}$$

b. The general, non-autonomous or "time dependent" case:
The Hamiltonian operator $H(t)$ depends on t. The operators $U(t,t_0)$ form what could be called a true two-parameter system of unitary operators. (Not much can be found on such systems in the standard Hilbert space literature. The properties, in particular the relation between $H(t)$ and the solution $U(t,t_0)$ are certainly much more complicated than those of one parameter groups. The complications that may occur are, however, of no concern to us here). For given $H(t)$ one may also in this case write down at least a formal expression for the $U(t,t_0)$. If all operators involved would commute one would obtain

$$U(t,t_0) = e^{-\frac{i}{\hbar}\int_{t_0}^{t} H(t')dt'} \tag{5}$$

In stead of this in general incorrect expression one has the following formula, very popular in theoretical physics

$$U(t,t_0) = T\left\{e^{-\frac{i}{\hbar}\int_{t_0}^{t} H(t')dt'}\right\} :=$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t}\dots\int_{t_0}^{t} dt_1\dots dt_n\, T\{H(t_1)\dots H(t_n)\} \quad (6)$$

with $T\{H(t_1)\dots H(t_n)\}$ the so-called time-ordered product of the operators $H(t_1), \dots, H(t_n)$ defined as

$$T\{H(t_1)\dots H(t_n)\} = H(t_{\sigma(1)})\dots H(t_{\sigma(n)}) \quad (7)$$

with for every $t_1, \dots, t_n$, σ the permutation of $1, \dots, n$ such that $t_{\sigma(1)} > t_{\sigma(2)} > \dots > t_{\sigma(n)}$. For given $H(t)$ with suitable properties this series converges indeed to $U(t,t_0)$.

There exist a number of different examples of quantum theories, i.e. models that each describe a certain area of physical phenomena and that possess the general features that we have just sketched. Of importance here are two such examples: Elementary (non-relativistic) quantum mechanics, suitable for the description of the phenomena of atomic physics, and the more advanced quantum field theory, necessary for understanding the deeper lying level of sub-nuclear physics reached by experiments at very high energies.

Quantum mechanics in the simplest case of a single particle in a given potential $V(x,y,z,t)$ has as Hilbert space H the function space $L_2(\mathbb{R}^3, dx\,dy\,dz)$, the vectors ψ are called wave functions. The most important operators are the momentum operators $P_x = \frac{\hbar}{i}\frac{\partial}{\partial x}$, $P_y = \frac{\hbar}{i}\frac{\partial}{\partial y}$, $P_z = \frac{\hbar}{i}\frac{\partial}{\partial z}$, the position operators Q_x, Q_y, Q_z which are operators of multiplication with x, y, z, and the Hamilton operator $H = \frac{\vec{p}^2}{2m} + V(Q_x,Q_y,Q_z,t)$.

The evolution equation (3) appears as a partial differential equation

$$\frac{\partial}{\partial t}\psi = -\frac{i}{\hbar}\{-\frac{\hbar^2}{2m}\Delta + V(x,y,z,t)\}\psi \tag{8}$$

with m the mass of the particle. This equation is called the *Schrödinger equation*.

Quantum field theory is a much more complicated theory. Its Hilbert space is not an elementary function space, but an infinite direct sum of tensor product spaces. Basic operators are so-called field operators, defined for each point in space. Quantum field theory is also less well established as a rigorous mathematical theory. As a physical theory it describes systems of a variable number of particles, with and without mutual interaction and with the possibility of creation and annihilation such as occurs in elementary particle physics. In these lectures we restrict ourselves to quantum mechanics, apart from a few occasional remarks. We discuss, however, in this relatively simple context the basic features of a method which at present has it main application in quantum field theory.

3. FEYNMAN'S PATH INTEGRAL FORMULA

The quantum mechanical model of a single particle can be simplified further by taking it one-dimensional. By this all formulas will become simpler and more transparant, while we loose nothing that is essential for our purpose. The Hilbert space H is $L_2(\mathbb{R},dx)$, with wave functions $\psi(x)$ or $\psi(x,t)$ with the time dependence included. There is one moment operator $P = \frac{\hbar}{i}\frac{\partial}{\partial x}$ and one position operator Q. The Schrödinger equation (8) becomes

$$\frac{\partial}{\partial t}\psi(x,t) = -\frac{i}{\hbar}\{-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x,t)\}\psi(x,t) \tag{9}$$

Solving this equation means in Hilbert space language finding the evolution operators $U(t,t_0)$. One may assume that the action of these operators can be written in terms of integration

$$(U(t,t_0)\psi)(x) = \psi(x,t) =$$

$$= \int_{-\infty}^{+\infty} K(x,t;\, x',t_0)\psi(x',t_0)dx' \tag{10}$$

Solving the Schrödinger equation then becomes finding the integral kernel $K(x,t;\, x',t_0)$, or as one says in the theory of linear partial differential equations, the *Green function* $K(x,t;\, x',t_0)$. Observe that one has relations for the Green functions corresponding to the first two operator relations of (1):

$$\lim_{t\downarrow 0} K(x,t;\, x',t_0) = \delta(x-x') \tag{11}$$

$$K(x_3,t_3;\, x_1,t_1) = \int_{-\infty}^{+\infty} K(x_3,t_3;\, x_2,t_2)K(x_2,t_2;\, x_1,t_1)dx_2 \tag{12}$$

The last relation is sometimes called the Chapman-Kolmogorov equation, a name which comes from a similar relation in the theory of stochastic processes.

In 1948 R.P. Feynman has given a general expression for the integral kernel of the evolution operator of a quantum mechanical system. This expression which is very heuristic gives the general solution of the Schrödinger equation in terms of the classical, i.e. non-quantum description of the system, as a kind of infinite dimensional integral over all possible classical "paths". For the one-dimensional quantum systems that we consider, this expression is, for $t_b > t_a$

$$K(x_b,t_b;\, x_a,t_a) = N \int_{\substack{q(t_a)=x_a \\ q(t_b)=x_b}} e^{\frac{i}{\hbar}\int_{t_a}^{t_b} L(q(t),\dot{q}(t),t)dt} \mathcal{D}[q(\cdot)] \qquad (13)$$

This is *Feynman's path integral*. Before we derive it, as far as this is possible for such a heuristic expression, a few remarks may be made.

a. The expression is an integral over a space of "paths", i.e. all possible ways in which a one-dimensional particle starting at time $t = t_a$ in the point x_a can reach a second point x_b at a later time $t = t_b$. A path is therefore a real function $g(\cdot)$ on the interval $[t_a, t_b]$ with fixed prescribed values x_a and x_b in t_a and t_b. One may think of such a function as specified in the limit $n \to \infty$ by values $x_j = q(t_j)$ in n discrete time points $t_1 < t_2 < \ldots < t_n$. The integration measure $\mathcal{D}[q(t)]$ is then a sort of limit for $n \to \infty$ of the ordinary Lebesgue measure $dx_1\, dx_2 \ldots dx_n$. However, as is well-known "$\lim_{n\to\infty} dx_1 \ldots dx_n$" does not exist in any rigorous mathematical sense. To compensate for the singular character of the "measure" $\mathcal{D}[q(\cdot)]$ an "infinite normalization factor" N is placed in front.

b. The integrand, a functional of the paths $q(\cdot)$ contains the classical action $\int_{t_a}^{t_b} L(q(t),\dot{q}(t),t)dt$. In classical mechanics one characterizes

the motion of a system of particles by means of the Lagrange function, which is a function of position coordinates, velocities, and possibly the time. The equations of motion of the system are then the Euler-Lagrange equations of a variational problem in which the action integral defined as $\int_{t_2}^{t_2} L(q_j(t),\dot{q}_j(t),t)dt$ must be minimized. For a one-dimensional system as considered here the Lagrange function is $L(q(t),\dot{q}(t),t) = \frac{1}{2}m\dot{q}(t)^2 - V(q(t),t)$. The Euler-Lagrange equation is $\frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) = 0$ or $m\frac{d^2}{dt^2}q = -\frac{\partial V}{\partial q}$, which is of course Newton's equation for the connection between force and acceleration.

c. The foregoing can be summarized in a very intuitive manner by saying that the Feynman path integral gives the function $K(x_b,t_b;\, x_a,t_a)$ as a "continuous superposition of phase factors. Every possible classical path contributes a phase determined by the classical action along the path.

d. An intuitively attractive aspect of the path integral which is, however, extremely difficult to make rigorous is the following: The path integral may be seen as an infinite dimensional generalization of a type of finite dimensional integral in which the integrand is a variable phase factor and which is therefore not absolutely convergent, but converges only as an improper integral, because of interference effects. The contribution to the integral in an area is smaller when the phase factor varies more rapidly. The action functional in the integrand of the Feynman path integral is stationary in the path $q_{c\ell}(\cdot)$, the classical path that represents the true classical motion, i.e. is the solution of the equation of motion with prescribed endpoints $q(t_a) = x_a$, $q(t_b) = x_b$. The main contribution to the integral comes from the neighbourhood of $q_{c\ell}(\cdot)$. One says that $K(x_b,t_b;\, x_a,t_a)$ consists of a "classical part" with added to that "higher order quantum corrections".

4. DERIVATION OF THE PATH INTEGRAL FORMULA

The possibility of writing $K(x_b,t_b;\, x_a,t_a)$ as a path integral, at least in a heuristic sense, is an immediate consequence of property b of (1) for the evolution operator $U(t_b,t_a)$, in the explicit form of the Chapman-Kolmogorov equation (12). Repeated application of this with intermediate times t_j, $t_a < t_1 < t_2 < \ldots < t_{n-1} < t_b$ gives

$$K(x_b,t_b;\, x_a,t_a) =$$
$$= \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} K(x_b,t_b;\, x_{n-1},t_{n-1}) K(x_{n-1},t_{n-1};\, x_{n-2},t_{n-2}) \ldots$$
$$\ldots K(x_1,t_1;\, x_a,t_a)\, dx_{n-1} \ldots dx_1 \qquad (14)$$

In the limit $n \to \infty$ the integration $dx_{n-1} \ldots dx_1$ over the coordinates $x_1, \ldots, x_{n-1}$ associated with the time points $t_1, \ldots, t_{n-1}$ goes over into an integration $\mathcal{D}[q(\cdot)]$ over paths $q(\cdot)$ with fixed endpoints $q(t_a) = x_a$, $q(t_b) = x_b$. The integrand, a function of $x_1 \ldots x_{n-1}$ becomes a functional of the paths $q(\cdot)$. The heuristic result of this way of thinking can then be written as

$$K(x_b,t_b;\, x_a,t_a) = \int_{\substack{q(t_a)=x_a \\ q(t_b)=x_b}} K[q(\cdot)]\, \mathcal{D}[q(\cdot)] \qquad (15)$$

This shows that $K(x_b,t_b;\, x_a,t_a)$ can be written as a path integral, but it does not tell us what the functional $K[q(\cdot)]$ is. That it indeed has the form

$$K[q(\cdot)] = N\, e^{\frac{i}{\hbar} \int_{t_a}^{t_b} L(q(t),\dot{q}(t),t)\, dt} \qquad (16)$$

will be the result of a more elaborate argument.

We consider first the "time-independent" case, i.e. with

$$U(t_b,t_a) = e^{-\frac{i}{\hbar}(t_b-t_a)(H_0+V)} \tag{17}$$

If the operators $H_0 = \frac{p^2}{2m}$ and $V = V(Q)$ would commute, which they definitely do not, $U(t_b,t_a)$ could be written as

$$U(t_b,t_a) = e^{-\frac{i}{\hbar}(t_b-t_a)H_0}\, e^{-\frac{i}{\hbar}(t_b-t_a)V} \tag{18}$$

The integral kernel of this product operator is not hard to find. One has

$$(e^{-\frac{i}{\hbar}(t_b-t_a)V}\psi)(x) = e^{-\frac{i}{\hbar}(t_b-t_a)V(x)}\psi(x) \tag{19}$$

and with the Fourier transform for the wave functions

$$\hat{\psi}(p) = \frac{1}{(2\pi\hbar)^{\frac{1}{2}}}\int_{-\infty}^{-\infty} e^{-\frac{i}{\hbar}px}\psi(x)dx \tag{20}$$

and its inverse one calculates

$$\begin{aligned}
&(e^{-\frac{i}{\hbar}(t_b-t_a)H_0}\psi)(x) = \\
&= \frac{1}{(2\pi\hbar)^{\frac{1}{2}}}\int_{-\infty}^{+\infty} e^{\frac{i}{\hbar}px}\, e^{-\frac{i}{\hbar}\frac{p^2}{2m}(t_b-t_a)}\hat{\psi}(p)dp = \\
&= \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty}\frac{1}{2\pi\hbar}\, e^{-\frac{i}{\hbar}p(x-x')-\frac{i}{\hbar}\frac{p^2}{2m}(t_b-t_a)}dp\right]\psi(x')dx' = \\
&= \int_{-\infty}^{+\infty}\left(\frac{m}{2\pi i\hbar(t_b-t_a)}\right)^{\frac{1}{2}} e^{\frac{im}{2\hbar}\frac{(x-x')^2}{t_b-t_a}}\psi(x')dx'
\end{aligned} \tag{21}$$

in which we have used the elementary improper integral

$$\int_{-\infty}^{+\infty} e^{-i(ay^2+by)}dy = \left(\frac{\pi}{a}\right)^{\frac{1}{2}} e^{i\frac{b^2}{2a}} e^{-i\frac{\pi}{4}} = \left(\frac{\pi}{ia}\right)^{\frac{1}{2}} e^{i\frac{b^2}{4a}} \tag{22}$$

for $a > 0$, $b \in \mathbb{R}$.

Combination of (19) and (21) gives

$$\left(e^{-\frac{i}{\hbar}(t_b-t_a)H_0} e^{-\frac{i}{\hbar}(t_b-t_a)V}\psi\right)(x) =$$

$$= \int_{-\infty}^{+\infty} \left(\frac{m}{2\pi i\hbar(t_b-t_a)}\right)^{\frac{1}{2}} e^{\frac{i}{\hbar}\{\frac{1}{2}m\left(\frac{x-x'}{t_b-t_a}\right)^2 - V(x')\}(t_b-t_a)} \psi(x')dx' \tag{23}$$

which gives us the integral kernel of the product operator $e^{-\frac{i}{\hbar}(t_b-t_a)H_0} e^{-\frac{i}{\hbar}(t_b-t_a)V}$ as

$$\left(\frac{m}{2\pi i\ (t_b-t_a)}\right)^{\frac{1}{2}} e^{\frac{i}{\hbar}\{\frac{1}{2}m\left(\frac{x-x'}{t_b-t_a}\right)^2 - V(x')\}(t_b-t_a)} \tag{24}$$

Because the product operator is *not* equal to $e^{-\frac{i}{\hbar}(t_b-t_a)(H_0+V)}$ this result is not the desired integral kernel of $U(t_b,t_a)$. We can use it nevertheless asymptotically, in the limit for short time intervals. The possibility of doing this is based on the so-called *Lie-Trotter formula* for operators

$$e^{i(A+B)} = \lim_{n\to\infty} \left(e^{\frac{iA}{n}} e^{\frac{iB}{n}}\right)^n \tag{25}$$

It holds as a strong operator limit, for arbitrary bounded operators A and B, or for self-adjoint possibly unbounded operators A and B if only certain conditions on the domain of definitions of A, B and $A+B$ are met. We shall sketch a proof of this formula in the simplest case at the end of this section and proceed now to its application. For this we divide the fixed time interval $[t_a, t_b]$ in n equal subintervals, by means of intermediate points t_j, $j = 1,2,\ldots n-1$,

$t_{j+1} - t_j = \frac{t_b - t_a}{n} = \Delta t$. (Call also $t_a = t_0$, $t_b = t_n$, and later $x_a = x_0$, $x_b = x_n$), Using (25) we have

$$U(t_b,t_a) = e^{-\frac{i}{\hbar}(t_b-t_a)(H_0+V)} =$$

$$= \lim_{n\to\infty}\left(e^{-\frac{i}{\hbar}(\frac{t_b-t_a}{n})H_0} e^{-\frac{i}{\hbar}(\frac{t_b-t_a}{n})V}\right)^n =$$

$$= \lim_{n\to\infty}\left(e^{-\frac{i}{\hbar}\Delta t H_0} e^{-\frac{i}{\hbar}\Delta t V}\right)^n =$$

$$= \lim_{n\to\infty}\left(e^{-\frac{i}{\hbar}\Delta t H_0} e^{-\frac{i}{\hbar}\Delta t V} \ldots e^{-\frac{i}{\hbar}\Delta t H_0} e^{-\frac{i}{\hbar}\Delta t V}\right) =$$

$$= \lim_{n\to\infty}\left(e^{-\frac{i}{\hbar}(t_n-t_{n-1})H_0} e^{-\frac{i}{\hbar}(t_n-t_{n-1})V} \ldots\right.$$

$$\left. \ldots e^{-\frac{i}{\hbar}(t_1-t_0)H_0} e^{-\frac{i}{\hbar}(t_1-t_0)V}\right) \qquad (26)$$

The integral kernel of the operator on the right-hand side before $\lim_{n\to\infty}$ is taken, can be obtained by using result (24) for every interval $[t_j, t_{j+1}]$. If one does not worry about the step from the L_2-convergence in the Lie-Trotter formula to pointwise convergence of integral kernels, one obtains $K(x_b,t_b;\, x_a,t_a)$ as limit of n-fold integration, in which terms of the integrand are combined in an obvious way

$$K(x_b,t_b;\, x_a,t_a) = \lim_{n\to\infty}\left(\frac{mn}{2\pi i\hbar(t_b-t_a)}\right)^{n/2} \qquad (27)$$

$$\int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty} e^{\frac{i}{\hbar}\sum_{j=0}^{n-1}\{\frac{1}{2}m\left(\frac{x_{j+1}-x_j}{t_{j+1}-t_j}\right)^2 - V(x_j)\}(t_{j+1}-t_j)} dx_{n-1}\ldots dx_1$$

The expression in the exponent of the integrand can be seen as a discrete

approximation of the classical action integral. This means that the sum

$$\sum_{j=0}^{n-1} \{\tfrac{1}{2}m\left(\frac{x_{j+1}-x_j}{t_{j+1}-t_j}\right)^2 - V(x_j)\}(t_{j+1}-t_j) \tag{28}$$

becomes in the limit $n \to \infty$ in which a discrete sequence $x_0, x_1, x_2, \ldots, x_n$ is supposed to go into a path $q(\cdot)$, and $dx_{n-1} \ldots dx_1$ into $\mathcal{D}[q(\cdot)]$ an integral

$$\int_{t_a}^{t_b} \{\tfrac{1}{2}m\left(\frac{dq}{dt}\right)^2 - V(q(t))\}dt = \int_{t_a}^{t_b} L(q(t),\dot{q}(t))dt \tag{29}$$

This brings us to the heuristic expression (13), the path integral of Feynman. Note that the derivation of (27) is rigorous or can be made rigorous for a large class of potentials $V(x)$. This means that the expression of the kernel $K(x_b,t_b;\, x_a,t_a)$ as a limit of finite dimensional integrals is a rigorous result. The interpretation of this by means of (29) as an infinite dimensional integral is heuristic. No such integral exists in any precise mathematical sense.

We make now a few remarks on the more general "time dependent" case. Starting from (1) b one obtains

$$U(t_b,t_a) = U(t_n,t_{n-1})U(t_{n-1},t_n) \ldots U(t_1,t_0) =$$

$$= \lim_{n\to\infty} U(t_n,t_{n-1}) \ldots U(t_1,t_0) \tag{30}$$

One then shows that for potentials $V(t)$ that depend in a smooth way on t the operator $U(t_{j+1},t_j)$ is for $t_{j+1}-t_j \to 0$ asymptotically equal to $e^{-\frac{i}{\hbar}(t_{j+1}-t_j)(H_0+V(t_j))}$. This gives

$$U(t_b,t_a) =$$

$$= \lim_{n\to\infty} e^{-\frac{i}{\hbar}(t_n-t_{n-1})(H_0+V(t_{n+1}))} \ldots e^{-\frac{i}{\hbar}(t_1-t_0)(H_0+V(t_0))} \tag{31}$$

The next step is to derive a generalization of the Lie-Trotter formula, which we shall not discuss here. When applied to (31) one finds

$$U(t_b,t_a) = \lim_{n\to\infty} e^{-\frac{i}{\hbar}(t_n-t_{n-1})H_0} e^{-\frac{i}{\hbar}(t_n-t_{n-1})V(t_{n-1})} \cdots \cdots e^{-\frac{i}{\hbar}(t_1-t_0)H_0} e^{-\frac{i}{\hbar}(t_1-t_0)V(t_0)} \tag{32}$$

From this one immediately obtains formula (27) with an extra t_j dependence in $V(x_j,t_j)$, which can be interpreted heuristically as the path integral formula (13). To fill in the details of this proof will probably be not very difficult if one starts from strong smoothness assumptions on the operator $V(t)$. Contrary to the easier time independent case not much can be found in the literature on this case.

REMARKS.

1. In the theory of stochastic processes there is a path integral concept, closely related to that from quantum mechanics, that has a well-defined meaning in the context of the mathematical theory of measure and integration. A specific example is Wiener measure and its associated integral, which occurs in the theory of Brownian motion. In the "time independent" case there is a clear mathematical connection. The one-parameter group of unitary operators $e^{-\frac{i}{\hbar}tH}$ can for $H \geqslant 0$ be seen as the (continuous) boundary value of an operator valued function $U(z) = e^{-\frac{i}{\hbar}zH}$, holomorphic in the lower half-plane $\mathrm{Im}\, z < 0$. For $z = -i\tau$, $\tau > 0$ this gives a so-called contractive semi-group $e^{-\frac{\tau}{\hbar}H}$. Instead of studying the integral kernel of $U(t)$ one may consider the kernel of the analytic continuation $U(-i\tau)$. For this one can derive a limit formula like (27), but without the imaginary i in the exponent. Lebesgue measure $dx_1 \ldots dx_{n-1}$ has no limit for $n \to \infty$, but the measure obtained by putting in front normalized probability densities

$\rho_j(x)$, such as for instance Gaussian densities $\rho(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, $\rho_1(x_1) \ldots \rho_{n-1}(x_{n-1})dx_1 \ldots dx_{n-1}$, does have a limit. Many authors look therefore at the quantum path integral exlusively as something obtained by analytic continuation from a much better defined "imaginary time" analogon. This point of view is of particular importance for Yang-Mills quantum fields in connection with the "instanton" concept.

2. The presentation of the path integral that we have given deviates considerably from the original one of Feynman himself. In Feynman's formulation of quantum mechanics there is no mentioning of Hilbert space, only of explicit wave functions and the Schrödinger equation. His aim is to use the path integral for a formulation of time evolution of wave functions directly based on the classical description, without using Schrödinger's equation, but nevertheless leading to the same results. For this he considers first ordinary probability theory, where formula (12) as Chapman-Kolmogorov equation holds for conditional probabilities or transition probabilities of stochastic processes such as Markov processes. He then observes that in quantum mechanics one also has transition probabilities, for instance the probability that a particle that was at time t_1 in x_1 will be found at time t_2 in x_2. These probabilities and the way they have to be combined differ from what one expects from ordinary probability theory. There is no additivity because of interference effects typical for quantum theory. The probabilities are squares of absolute values of complex "probability amplitudes". These are the fundamental objects and for these one should have relations such as (12). The function $K(x_b,t_b; x_a,t_a)$ is for Feynman not the integral kernel of an evolution operator but a probability amplitude in the above sense, with $|K(x_b,t_b; x_a,t_a)|^2$ the experimental transition probability. Relation (12) leads via (14) to a path integral (15). Using an earlier suggestion of Dirac he then postulates the form $N e^{i\int L dt}$ for the

functional integrand and finally he verifies in a very heuristic manner that this leads to solutions of the Schrödinger equation. The Lie-Trotter formula (25) which we have used for a partial mathematical justification of the path integral is a later contribution to the subject due to the mathematician E. Nelson.

3. To avoid conceptual confusion and paradoxical consequences one has to be careful when speaking of $|K(x_b,t_b;\, x_a,t_a)|^2$ as a transition probability. According to the general principles of quantum mechanics $|\psi(x)|^2$ is the probability density for the position x of the particle, i.e. $\int_a^b |\psi(x)|^2 dx$ is the probability for finding in a measurement the particle between a and b, if the state of the particle is given by the wave function $\psi(x)$. If the particle is at time t_a with probability one in a small interval $[x_0, x_0+\varepsilon]$, or equivalent has at t_a a wave function which vanishes outside this interval, then the wave function at a later time t_b will be

$$\psi(x_b,t_b) = \int_{-\infty}^{+\infty} K(x_b,t_b;\, x_a,t_a)\psi(x_a,t_a)dx_a \tag{33}$$

which is in an approximation for small ε

$$K(x_b,t_b;\, x_0,t_a)C_0 \tag{34}$$

with

$$C_0 = \int_{x_0}^{x_0+\varepsilon} \psi(x_a,t_a)dx_a \tag{35}$$

The probability density for the position at time t_b in this approximation is

$$|\psi(x_b,t_b)|^2 \approx |K(x_b,t_b;\, x_0,t_0)|^2 \;\; |C_0|^2 \tag{36}$$

If the initial distribution is sufficiently sharply peaked, then the final distribution is approximately proportional with $|K(x_b,t_b;\, x_0,t_0)|^2$. This justifies calling it a transition probability although it is strictly speaking not a transition probability density in the sense of ordinary stochastic theory.

We finally give a short derivation of the Lie-Trotter formula for the simplest case of bounded self-adjoint operators. This is not sufficient for the path integral because quantum mechanics involves unbounded operators, but it gives an idea why such a formula holds.

Let A and B be two bounded self-adjoint operators in the Hilbert space H. Define $C = e^{i\frac{A}{n}} e^{i\frac{B}{n}}$, $D = e^{i\frac{A+B}{n}}$. One has

$$\left(e^{i\frac{A}{n}} e^{i\frac{B}{n}}\right)^n - e^{i(A+B)} = C^n - D^n =$$

$$= C^n - C^{n-1}D + C^{n-1}D - C^{n-2}D^2 + C^{n-2}D^2 - \ldots + CD^{n-1} - D^n =$$

$$= \sum_{k=0}^{n-1} C^k (C-D) D^{n-1-k} \tag{37}$$

Using operator norms we have an estimate

$$\left\| \left(e^{i\frac{A}{n}} e^{i\frac{B}{n}}\right)^n - e^{i(A+B)} \right\| =$$

$$= \left\| \sum_{k=0}^{n-1} C^k (C-D) D^{n-1-k} \right\| \leq \sum_{k=0}^{n-1} \|C\|^k \|C-D\| \, \|D\|^{n-1-k} =$$

$$= n\|C-D\| = n \left\| e^{i\frac{A}{n}} e^{i\frac{B}{n}} - e^{i\frac{A+B}{n}} \right\| \tag{38}$$

(We have used also that C and D are unitary).

Because A and B are bounded the exponential expressions can be written as convergent series in powers of $\frac{1}{n}$

$$
\begin{aligned}
& n \left\| e^{i\frac{A}{n}} e^{i\frac{B}{n}} - e^{i\frac{A+B}{n}} \right\| = \\
& = n \left\| \left(1+i\frac{A}{n}+\frac{i^2}{2}\frac{A^2}{n^2}+\ldots\right)\left(1+i\frac{B}{n}+\frac{i^2}{2}\frac{B^2}{n^2}+\ldots\right) - \right. \\
& \left. - \left(1+i\frac{A+B}{n}+\frac{i^2}{2}\frac{(A+B)^2}{n^2}+\ldots\right) \right\| = \\
& = n \left\| -\frac{1}{2n^2}[A,B] + \text{higher order terms in } \frac{1}{n} \right\| \qquad (39)
\end{aligned}
$$

This goes obviously to zero for $n \to \infty$, which proves the Lie-Trotter formula (25) for this simple case.

5. PHYSICAL ASPECTS

The path integral formula as we have discussed it up to this point gives the integral kernel $K(x_b,t_b;\, x_a,t_a)$ of the time evolution operator of a one-dimensional quantum mechanical system in the "position representation", i.e. in the Hilbert space of wave functions $\psi(x)$, where $|\psi(x)|^2$ has the physical meaning of a probability density for measurement of the position variable x. From the kernel $K(x_b,t_b;x_a,t_a)$ one obtains $|K(x_b,t_b;\, x_a,t_a)|^2$ which is interpreted as the probability that the one-dimensional particle will be at time t_b at the point x_b, given that it was at x_a at the earlier time t_a. (For a precise formulation of this one should take into account the observation about the $|K(x_b,t_b;\, x_a,t_a)|^2$ made in section 4.)

In quantum mechanics the position x, or x, y, z in the three-dimensional case, is an important variable from a theoretical point of view, as a starting point in setting up the formalism. Considering the main area of application of quantum mechanics, atomic physics, the position is not a variable that is often measured. Of much greater experimental importance are such quantities as energy, momentum and angular momentum.

To understand what is of interest in actual experiments we make a few general remarks on the physical aspects of the sort of systems that can be described by elementary quantum mechanics, and for that matter by elementary classical mechanics, for the distinction is not very important for the points that we want to make.

The sort of physical systems that we have in mind can be divided into two classes. This is a rough division which neglects important special cases, but we nevertheless adopt it here for the sake of our argument. Each category contains a basic case which in itself is physically trivial, but which serves as reference point for the other cases in the same class.

These then appear as non-trivial deviations or perturbations of the basic situation.

The first class contains as basic case the situation of a particle on which no forces act, i.e. with potential $V(x)$ or $V(x,y,z) = 0$. The typical time evolution for this case is that such a free particle arrives in the area of observation, coming from infinity, i.e. a far away area in space, and then disappears again in the opposite direction in infinity without changing its momentum, velocity or energy. Nothing of physical interest has happened. The interesting cases which can be seen as perturbations of the basic situation are the motions in potentials the action of which in terms of forces drives the particle away from the origin of the coordinate system and becomes negligible at large distances of this origin. The typical time evolution for such a perturbed system is that a particle arrives at the origin, coming from special infinity, where it was at $t = -\infty$, a mathematical idealization, is influenced by the potential in the neighbourhood of the origin, is deflected in its motion, moves on, with its momentum changed and disappears at $t = +\infty$ to infinity as a free particle. What is physically interesting in this case is not the precise orbit of the particle, the positions at various times, but the total effect of the potential on the motion between $t = -\infty$ and $t = +\infty$. This total effect as deviation from the free motion is the aspect of the system that one measures, at least in atomic physics.

The second class of systems is described by potentials that drive the particle back to the origin, with forces that become stronger with increasing distance from the origin. The particle cannot escape to infinity and will be most of the time in the neighbourhood of the origin. The simplest example which for this class is the basic case is the harmonic oscillator, in one dimension given by a potential $V(x) = \frac{1}{2}Kx^2$. The harmonic oscillator

is both classically and in quantum theory a physical trivial system of which the equation of motion can be solved completely in an elementary manner. In the classical version the particle oscillates with fixed frequently $\nu = \frac{\omega}{2\pi}$, $\omega = \sqrt{\frac{K}{m}}$, m the mass. The only free parameter in the system is the total energy which is of course constant in time and which may have a arbitrary non-negative value. The perturbed cases for which the harmonic oscillator is the free, "reference" system will have similar potentials, for instance $V(x) = \frac{1}{2}Kx^2 - g(t)x$, a harmonic oscillator with a spatially homogeneous and possibly time dependent external force, or $V(x) = \frac{1}{2}Kx^2 + g(t)x^4$, an anharmonically perturbed oscillator. These are special examples. We should however realize that for a potential $V(x)$ with $V(0) = 0$, which is symmetric in x and which goes to ∞ for $|x| \to \infty$ an obvious approximation for small x, the area where the particle is most of the time, is given by the expansion $V(x) = ax^2 + bx^4 + \ldots$. Again one is not so much interested in the position x of the particle at different times t, but more in the total change over a large time interval. For an oscillator with time dependent perturbation the total change in energy will be of interest. In particular, if the time-dependence is such that there is only a perturbation present during a finite time interval, $g(t) = 0$, for $|t| \geqslant \tau_0$, or such that $g(t)$ goes to 0 fast when $|t| \to +\infty$, then one is in a similar situation as described in the first class of systems. For very early times $(t \to -\infty)$ the particle is a free oscillator with frequency ω and energy E_1, then the perturbation sets in, the energy changes until at a very late time $(t \to +\infty)$ the particle is again a free oscillator, with the same frequency ω but with a different energy E_2. It is of no interest to measure or calculate the properties of the oscillating particle at intermediate times. The important thing is the total effect of the perturbation. In the classical case this is the difference $E_2 - E_1$. In the quantum case it is

the probability of going from E_1 to E_2.

SUMMARIZING. In both type of systems one is interested in a change or transition between $t = -\infty$ and $t = +\infty$, with reference to a trivial or free system.

In quantum theory there is for this an elegant and quite general formalism. This will be discussed in the next section as an addition to our remarks on quantum theory and quantum mechanics in section 2.

6. THE S-OPERATOR

Consider a quantum system in the Hilbert space H with a system of unitary evolution operators $U(t_b,t_a)$. We are interested in the total effect of this time evolution from $t = -\infty$ to $t = +\infty$, relative to a free evolution described by a system $U_0(t_b,t_a)$, which has the form

$$U_0(t_b,t_a) = e^{-\frac{i}{\hbar}(t_b-t_a)H_0} \tag{40}$$

and which is chosen as a reference system. For this purpose one defines a sort of modified evolution operator system consisting of operators $U_I(t_b,t_a)$ defined as

$$U_I(t_b,t_a) = U_0(0,t_b)U(t_b,t_a)U_0(t_a,0) =$$

$$= e^{\frac{i}{\hbar}t_b H_0} H(t_b,t_a)\, e^{-\frac{i}{\hbar}t_a H_0} \tag{41}$$

This is usually called the *interaction picture*. On verifies immediately that the $U_I(t_b,t_a)$ satisfy the requirements (1) for a system of evolution operators. Note also that it is a true two-parameter system, even in the case where not only $U_0(t_b,t_a)$ but also $U(t_b,t_a)$ is a one-parameter group. The operators $U(t_b,t_a)$ will in general have no limit for $t_a \to -\infty$, $t_b \to +\infty$. In many cases such a limit does exist for the $U_I(t_b,t_a)$. It is also clear that $U_I(t_b,t_a)$ describe in a way the deviation of $U(t_b,t_a)$ from the free $U_0(t_b,t_a)$. If both are the same then $U_I(t_b,t_a) \equiv 1$. When the limit of the evolution operator in the interaction picture exists as a unitary operator we call it the *S-operator*.

$$S = \lim_{\substack{t_a \to -\infty \\ t_b \to +\infty}} U_I(t_b,t_a) \tag{42}$$

In the simplest case in which the evolution operators $U(t_b,t_a)$ are perturbed, i.e. different from $U_0(t_b,t_a)$ only in a finite time interval

$[-\tau_0, \tau_0]$ the existence of the S-operator is trivial: For $t_a < -\tau_0$, $t_b > \tau_0$ one has $U(t_b,t_a) = U(t_b,\tau_0)U(\tau_0,-\tau_0)U(-\tau_0,t_a) =$ $= U_0(t_b,\tau_0)U(\tau_0,-\tau_0)U_0(-\tau_0,t_a)$. Then $U_I(t_b,t_a) = U_0(0,t_b)U(t_b,t_a)U_0(t_a,0) =$ $= U_0(0,\tau_0)U(\tau_0,-\tau_0)U_0(-\tau_0,0) = U_I(\tau_0,-\tau_0)$, i.e. independent of t_b, t_a. One can prove that the S-operator also exists in a less trivial manner in many cases in which the perturbation is not totally absent outside a finite time interval but goes fast to zero for $t_a \to -\infty$, $t_b \to +\infty$. This applies in particular to the two main classes of physical systems describe in section 5. In the first class, scattering of a particle by a potential $V(x)$ or $V(x,y,z)$ the S-operator or scattering operator will exist even in the time independent case whenever $V(x)$ goes to zero fast enough for $|x| \to \infty$. This seems fairly reasonable from an intuitive physical point of view and is established rigorously in the abundant mathematical literature on this subject. For the second type of systems, perturbed oscillators, the S-operator exist only in time dependent cases. This is not the case in quantum field theory which in a way contains aspects of both types of physical systems.

In general the S-operator cannot be calculated in closed form, but only in approximation. This is based on the limit $t_0 \to -\infty$, $t \to +\infty$ of formula (6), applied on the modified or interaction picture time evolution $U_I(t,t_0)$. Because of the fundamental importance of this so-called *Dyson series*, in particular in quantum field theory we sketch a derivation of this formula for the S-operator.

Associated with the two-parameter system $U_I(t_b,t_a)$ there is on general grounds a "modified" time dependent Schrödinger equation

$$\frac{\partial}{\partial t}\psi_I(t) = -\frac{i}{\hbar} H_I(t)\psi_I(t) \tag{43}$$

or

$$\frac{\partial}{\partial t}U_I(t,t_0)\psi_I(t_0) = -\frac{i}{\hbar}H_I(t)U_I(t,t_0)\psi_I(t_0) \tag{44}$$

for every fixed t_0 and fixed Hilbert space vector $\psi_I(t_0)$. This is equivalent to a differential equation for the operators $U_I(t,t_0)$

$$\frac{\partial}{\partial t}U_I(t,t_0) = -\frac{i}{\hbar}H_I(t)U_I(t,t_0) \tag{45}$$

with initial condition

$$U_I(t_0,t_0) = 1 \tag{46}$$

These two together are then equivalent to the integral equation.

$$U_I(t,t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} H_I(t_1)U_I(t_1,t_0)dt_1 \tag{47}$$

(For sufficient smoothness in t and t_0 such operator-valued integrals are well-defined). This integral equation can be iterated to a formal series which in favourable cases converges to the desired solution $U_I(t,t_0)$

$$U_I(t,t_0) = \sum_{n=0}^{\infty}\left(\frac{-i}{k}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \ldots \int_{t_0}^{t_{n-1}} dt_n H_I(t_1)H_I(t_2)\ldots \ldots H_I(t_n) \tag{48}$$

In order to write the integrals as n-fold integrals over the simple integration area $t_0 \leqslant t_j \leqslant t$, $j = 1,2,\ldots,n$, one defines the time ordered product $T\{H_I(t_1) \ldots H_I(t_n)\}$ as was done in section 2, formula (7). By this definition one obtains a function that is equal to the product $H_I(t_1) \ldots H_I(t_n)$ on the area $t \geqslant t_1 \geqslant t_2 \geqslant \ldots \geqslant t_n \geqslant t_0$, but which as a symmetric function in $t_0, \ldots, t_n$ can be extended to $t \geqslant t_j \geqslant t_0$, $j = 1,2,\ldots,n$. The integral over the extended area is therefore the same if a compensating factor $\frac{1}{n!}$ is put in front

$$U_I(t,t_0) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{-i}{\hbar}\right)^n \int_{t_0}^{t} \cdots \int_{t_0}^{t} dt_1 \cdots dt_n T\{H_I(t_1) \cdots H_I(t_n)\} \tag{49}$$

This can be written as indicated in section 2 as a symbolic "time ordered" exponential

$$U_I(t,t_0) = T\left\{e^{-\frac{i}{\hbar}\int_{t_0}^{t} H_I(t_1)dt_1}\right\} \tag{50}$$

So far the derivation has been quite general, not restricted in fact to the interaction picture evolution but valid for an arbitrary two-parameter system of evolution operators satisfying (1) a,b,c. For the specific case of $U_I(t,t_0)$ which is our subject of interest here we have to express $H_I(t)$ in the given "true" Hamiltonian operator $H(t)$ and the free Hamiltonian H_0. We have the two equations

$$\begin{aligned} \frac{\partial}{\partial t}U(t,t_0) &= -\frac{i}{\hbar}H(t)U(t,t_0) \\ \frac{\partial}{\partial t}U_0(t,t_0) &= -\frac{i}{\hbar}H_0 U_0(t,t_0) \end{aligned} \tag{51}$$

If we call the perturbation part of the true Hamiltonian $H(t) - H_0 = V(t)$ we obtain

$$\begin{aligned} &\frac{\partial}{\partial t}U_I(t,t_0) = \frac{\partial}{\partial t}(U_0(0,t)U(t,t_0)U_0(t_0,0)) = \\ &= \frac{i}{\hbar}U_0(0,t)H_0 U(t,t_0)U_0(t_0,0) - \frac{i}{\hbar}U_0(0,t)H(t)U(t,t_0)U_0(t_0,0) = \\ &= -\frac{i}{\hbar}U_0(0,t)V(t)U(t,t_0)U_0(t_0,0) = \\ &= -\frac{i}{\hbar}U_0(0,t)V(t)U_0(t,0)U_0(0,t)U(t,t_0)U(t_0,0) = -\frac{i}{\hbar}H_I(t)U_I(t,t_0) \end{aligned} \tag{52}$$

with therefore

$$H_I(t) = U_0(0,t)V(t)U_0(t,0) = e^{\frac{i}{\hbar}t H_0} V(t)\, e^{-\frac{i}{\hbar}t H_0} \tag{53}$$

One should observe that the interaction picture Hamiltonian H_I depends always explicitly on t, even in the case $V = H - H_0$ does not depend on t.

Taking finally the limits $t_0 \to -\infty$, $y \to +\infty$ one arives at the standard expression for the S-operator on which perturbative calculations are based

$$S = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{-i}{\hbar}\right)^n \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} dt_1 \dots dt_n T\{H_I(t_1) \dots H_I(t_n)\} \qquad (54)$$

or in symbolic exponential form

$$S = T\left\{ e^{-\frac{i}{\hbar}\int_{-\infty}^{+\infty} H_I(t)dt} \right\} \qquad (55)$$

with

$$H_I(t) = e^{\frac{i}{\hbar} t H_0} V(t)\, e^{-\frac{i}{\hbar} t H_0} \qquad (56)$$

In quantum mechanics (54) is for many cases a mathematically rigorous formula, of which the convergence can be proved. (In potential scattering for time independent potentials one has to sum the series before taking the limits $t_0 \to -\infty$, $t \to +\infty$)

The S-operator is the basic theoretical concept in quantum field theory such as this is used in elementary particle physics; the perturbation series (54) is the point of departure for all calculations of processes. Nevertheless, the whole S-operator concept in quantum field theory abounds with fundamental mathematical problems that have not yet been solved. The convergence of the series is uncertain, moreover, it is not even clear whether the separate terms make sense as operators in Hilbert space.

The path integral gives a method of setting up the perturbation scheme for the S-operator in a heuristic, but very systematic and transparant manner.

Because of this many people prefer it over more traditional purely Hilbert space operator methods, especially in quantum field theory. It is the main purpose of these lectures to demonstrate this method in the context of elementary quantum mechanics of a single harmonic oscillator, as a prelude to a possible later study of the method in quantum field theory.

7. THE ONE-DIMENSIONAL HARMONIC OSCILLATOR

a. The classical oscillator

The classical harmonic oscillator is a particle bound to the origin by a force proportional to the distance from the origin. (K is the constant of proportionality). Such a particle oscillates around the origin with frequency $\nu = \frac{1}{2\pi}\omega$, $\omega = \sqrt{\frac{K}{m}}$, m the mass of the particle, and with constant energy $E \geqslant 0$. We shall put $m = 1$, and use $\omega = \sqrt{K}$. As a mechanical system it is characterized by the Lagrangian function

$$L_0(q,\dot{q}) = \tfrac{1}{2}\dot{q}^2 - \tfrac{1}{2}\omega^2 q^2 \tag{57}$$

which leads to the Euler-Lagrange equation

$$\ddot{q} + \omega^2 q = 0 \qquad (\ddot{q} = \frac{d^2}{dt^2}q) \tag{58}$$

This has the general solution

$$q(t) = q(0)\cos\omega t + \dot{q}(0)\omega^{-1}\sin\omega t \tag{59}$$

If there is a external force $j(t)$, constant in x, but time dependent, the Lagrangian function becomes

$$L(q,\dot{q},t) = \tfrac{1}{2}\dot{q}^2 - \tfrac{1}{2}\omega^2 q^2 + jq \tag{60}$$

with Euler-Lagrange equation

$$\ddot{q} + \omega^2 q = j \tag{61}$$

This equation can be solved by means of a Green function $G(t,t')$ according to

$$q(t) = \int_{t_a}^{t_b} G(t,t')j(t')dt' \qquad t_a < t < t_b \tag{62}$$

with $G(t,t')$ a solution of

$$\left(\frac{\partial^2}{\partial t^2} + \omega^2\right) G(t,t') = \delta(t-t') \tag{63}$$

An oscillator with a possibly time dependent anharmonic perturbation of quartic type has a Lagrangian

$$L(q,\dot{q},t) = \tfrac{1}{2}\dot{q}^2 - \tfrac{1}{2}\omega^2 q^2 + \tfrac{1}{4}gq^4 \tag{64}$$

with g is $g(t)$, a given function. The Euler-Lagrange equation is

$$\ddot{q} + \omega^2 q = -gq^3 \tag{65}$$

Explicit solutions can of course not be given for this non-linear equation.

b. The quantum oscillator.

In the standard quantum mechanical theory one usually starts from a classical description in Hamiltonian form, i.e. with canonical variables q and $p = \frac{\partial L}{\partial \dot{q}}$ and a Hamiltonian function $H = p\dot{q} - L$, considered as a function of q, p and possibly t.

The quantum mechanical description uses the Hilbert space $H = L_2(\mathbb{R},dx)$ with position operator $Q = x$, momentum operator $P = -i\frac{d}{dx}$ (we use from now on units such that $\hbar = 1$), and a Hamilton operator which for the three cases discussed above is

$$\begin{aligned} H_0 &= \tfrac{1}{2}P^2 + \tfrac{1}{2}\omega^2 Q^2 && \text{(free oscillator)} \\ H(t) &= H_0 - j(t)Q && \text{(oscillator in external force)} \\ H(t) &= H_0 + \tfrac{1}{4}g(t)Q^4 && \text{(anharmonically perturbed oscillator)} \end{aligned} \tag{66}$$

The problem of the quantum mechanical free oscillator has a simple and elementary solution. The self-adjoint operator H_0 has a pure discrete non-degenerate spectrum with an orthonormal base of eigenfunctions

$$
\begin{aligned}
&H_0 \phi_n = E_n \phi_n \\
&E_n = (n+\tfrac{1}{2})\omega \\
&\phi_n(x) = N_n H_n(\omega^{\frac{1}{2}} x) e^{-\frac{1}{2}\omega x^2} \qquad\qquad n = 0,1,2,\ldots
\end{aligned}
\tag{67}
$$

with $N_n = (\frac{\omega}{\pi})^{\frac{1}{4}} (2^n n!)^{-\frac{1}{2}}$ and $H_n(\cdot)$ the n^{th} Hermite polynomial.

The quantum mechanical oscillator with external homogeneous force can also be solved completely. i.e. $U(t_b,t_a)$ and for suitable $j(t)$ also S can be explicitly determined. This can be done either by operator methods or by path integral method; we shall give a mixed derivation. The anharmonic oscillator can be solved only approximatively. Doing this with the path integral method will be our main goal in the later sections.

We finally remark that the solution (67) can be easily obtained by the introduction of so-called creation and annihilation operators.

$$
\begin{aligned}
A &= 2^{-\frac{1}{2}}(\omega^{\frac{1}{2}} Q + i\,\omega^{-\frac{1}{2}} P) \\
A^* &= 2^{-\frac{1}{2}}(\omega^{\frac{1}{2}} Q - i\,\omega^{-\frac{1}{2}} P)
\end{aligned}
\tag{68}
$$

with inverse relations

$$
\begin{aligned}
Q &= 2^{-\frac{1}{2}}\,\omega^{-\frac{1}{2}}(A+A^*) \\
P &= -i\,2^{-\frac{1}{2}}\,\omega^{\frac{1}{2}}(A-A^*)
\end{aligned}
\tag{69}
$$

For the operators A, A^* one has the fundamental relation

$$
[A, A^*] = 1 \tag{70}
$$

which is an immediate consequence of $[Q, P] = i$. With these new operators the Hamiltonian H_0 takes a simple form

$$
H_0 = \omega A^* A + \tfrac{1}{2}\omega\,. \tag{71}
$$

With a single simple differential equation, viz. that for $\phi_0(x)$ and some

algebraic manipulations of an elementary sort one obtains the eigenvalues E_n and the eigenfunctions ϕ_n. These are in fact

$$\phi_n = \frac{1}{\sqrt{n!}}(A^*)^n \phi_0 \tag{72}$$

This form is for our purpose more important than that with the explicit Hermite polynomials. Creation and annihilation operators play also an important rôle in quantum field theory.

8. THE S-OPERATOR AND THE TRANSITION AMPLITUDES FOR PERTURBED OSCILLATORS. INTRODUCTORY REMARKS

The S-operator as it was introduced in section 6 describes the effect of a perturbation during a certain time interval of a free system. For oscillators such as discussed in section 7 one is interested in transition amplitudes $(\phi_m, S\phi_n)$, with ϕ_m and ϕ_n energy eigenstates of the free oscillator. According to general principles of quantum theory $|(\phi_m, S\phi_n)|^2$ is the probability that the system ends up in a final state with sharp energy $E_m = (m+\frac{1}{2})\omega$ if it was initially in a state with sharp energy $E_n = (n+\frac{1}{2})\omega$. For fixed n $P_m = |(\phi_m, S\phi_n)|^2$ is indeed a (discrete) probability distribution for the final energy E_m.

We have in our discussion the following systems:

a. The free oscillator with frequency ω. The classical Lagrange function is $L_0(q,\dot{q}) = \frac{1}{2}\dot{q}^2 + \frac{1}{2}\omega^2 q^2$, the quantum mechanical Hamiltonian operator $H_0 = \frac{1}{2}P^2 + \frac{1}{2}\omega^2 Q^2$. The properties of this system are completely known and were reviewed in section 7.

b. The harmonic oscillator in a given external spatially homogeneous force. The classical Lagrangian is $L(q,\dot{q},t) = L_0(q,\dot{q}) + j(t)q$, the quantum mechanical Hamiltonian operator is $H = H_0 - j(t)Q$. This system can also be solved exactly. In particular, the S-operator and therefore the transition amplitudes $(\phi_m, S\phi_n)$ can be obtained in closed form. This can be done either by operator methods or by using the path integral. We shall employ a mixture of both. We are not interested in the result for its own sake, but for its value in the approach to the next system, which is our main concern.

c. The anharmonically perturbed oscillator. The classical Lagrange function is $L(q,\dot{q},t) = L_0(q,\dot{q}) - g(t)q^4$ and the quantum mechanical Hamiltonian $H(t) = H_0 + g(t)Q^4$. The function $g(t)$ goes to zero $t \to \pm\infty$;

it determines the strength of the perturbation. In quantum field theory it is usually a constant and is then called the coupling constant. (Compared to (66) we have absorbed a factor $\frac{1}{4}$ in $g(t)$. The anharmonic oscillator is not exactly solvable because of non-linearity. The S-operator and the amplitudes $(\phi_m, S\,\phi_n)$ can be obtained only approximately, in a perturbation series, which in a certain sense is a power series in $g(t)$. There are two main methods, the standard operator method that starts from the Dyson series (54) and calculates the ℓ^{th} term by operator manipulations and secondly the path integral method, which also gives a series in orders of $g(t)$, in fact the same Dyson series, but differently obtained. Both methods lead to essential the same series for S. Calculation of the ℓ^{th} order in $S = \sum_{\ell=0}^{\infty} S^{(\ell)}$ is facilitated by a technique of diagrams, the so-called Feynman diagrams.

The oscillator systems that we discuss are meant as simple models for the much more complicated systems of quantum field theory. More specifically one may think of the following field theoretical situations:

a. The free scalar field. This is a field $\phi(x)$ satisfying the Klein-Gordon equation $(\partial_\mu \partial^\mu + m^2)\phi = 0$. In the simplest case it is on the classical level a real function on space-time and at the quantum level consists of self-adjoint operators defined at all points of space-time. Such a model describes a system of arbitrary numbers of non-interacting relativistic particles of mass m. It is completely solvable and may be considered as an infinite system of free harmonic oscillators with frequencies $\omega(\vec{k}) = (\vec{k}^2+m^2)^{\frac{1}{2}}$, $\vec{k}$ as wave vector running through all real vector values. The one-dimensional harmonic oscillator is mathematically the free scalar field in one-dimensional space-time, i.e. depending only on t.

b. The scalar field in an external source. Its field equation is

$(\partial_\mu \partial^\mu + m^2)\phi = j$, with $j(x)$ a given source function. This equation can again be solved by Green function methods.

c. The non-linear so-called ϕ^4 model with field equation $(\partial_\mu \partial^\mu + m^2)\phi = -g\,\phi^3$ (The Lagrangian function contains a ϕ^4 term). This is the simplest non-trivial model in quantum field theory which is fully relativistic and describes particles with interaction, i.e. with scallering, and creation or annihilation. Only approximate calculations are possible. For this the results from the external source models can be used as a tool.

9. THE S-OPERATOR AND ITS MATRIX ELEMENTS FOR AN OSCILLATOR WITH EXTERNAL FORCE

In this section we shall derive the S-operator and its matrix elements $(\phi_m, S\,\phi_n)$ for the external force model discussed in section 8, point b. The functional dependence of the result on the given force $j(t)$ is important for later applications, so we shall write explicitly $S[j]$ and $(\phi_m, S[j]\phi_n)$.

As a first step we determine the simplest transition amplitude, namely $(\phi_0, S[j]\phi_0)$, the amplitude for remaining in the ground state, in field theory parlance the vacuum-vacuum amplitude. For this we employ the path integral associated with

$$(\phi_0, S[j]\phi_0) = \lim_{\substack{t_b \to \infty \\ t_a \to -\infty}} (\phi_0, e^{i t_b H_0} U(t_b, t_a) e^{-i t_a H_0} \phi_0) =$$

$$= \lim_{\substack{t_b \to \infty \\ t_a \to -\infty}} e^{\frac{i}{2}\omega(t_b - t_a)} (\phi_0, U(t_b, t_a)\phi_0) \tag{73}$$

To avoid technicalities we assume that the function $j(t)$ is sufficiently smooth and has support inside a finite time interval $[-\tau_0, \tau_0]$. This means that the limit (73) exists trivially, as was argued in section 6. From this point on we assume $t_a < -\tau_0$, $\tau_0 < t_b$. Because of this (73) can be written as

$$(\phi_0, S[j]\phi_0) =$$

$$= e^{\frac{i}{2}\omega(t_b - t_a)} \int_{-\infty}^{+\infty} \overline{\phi_0(x_b)}\, K(x_b, t_b; x_a, t_a)\phi_0(x_a) dx_a\, dx_b \tag{74}$$

with for $K(x_b, t_b; x_a, t_a)$ the path integral

$$K(x_b,t_b;\, x_a,t_a) = N \int_{\substack{q(t_a)=x_a \\ q(t_b)=x_b}} e^{\,i\int_{t_a}^{t_b} (\frac{1}{2}\dot{q}^2-\frac{1}{2}\omega^2 q^2+jq)dt} \mathcal{D}[q(\cdot)] \qquad (75)$$

We calculate this path integral by a change of variables. For this we choose a fixed path $q_{c\ell}(\cdot)$, and write for an arbitrary path

$$q(t) = q_1(t) + q_{c\ell}(t) \qquad (76)$$

The path $q_1(\cdot)$ introduced in this manner becomes the new variable in the integral. For $q_{c\ell}(\cdot)$ we take the solution of the classical equation of motion $\ddot{q}_{c\ell} + \omega^2 q_{c\ell} = j$, with $q_{c\ell}(t) = 0$ for $t \leqslant -\tau_0$. This solution can be easily found and is equal to

$$q_{c\ell}(t) = \frac{1}{\omega}\int_{-\infty}^{t} \sin\omega(t-t')j(t')dt' \qquad (77)$$

We write the action functional $I[q] = \int_{t_a}^{t_b} L(q,\dot{q})dt$ in the new variable $q_1(\cdot)$. For this we substitute (76) in the Lagrangian function

$$\begin{aligned} &\tfrac{1}{2}\dot{q}^2 - \tfrac{1}{2}\omega^2 q^2 + jq = I + II + III \\ &I = \tfrac{1}{2}\dot{q}_1^2 - \tfrac{1}{2}\omega^2 q_1^2 \\ &II = \tfrac{1}{2}\dot{q}_{c\ell}^2 - \tfrac{1}{2}\omega^2 q_{c\ell}^2 - \tfrac{1}{2}jq_{c\ell} + \tfrac{1}{2}jq_{c\ell} \\ &III = \dot{q}_{c\ell}q_1 - \omega^2 q_1 q_{c\ell} + jq_1 \end{aligned} \qquad (78)$$

The term I gives the action of the free oscillator

$$I_0[q_1] = \int_{t_a}^{t_b} (\tfrac{1}{2}\dot{q}_1^2 - \tfrac{1}{2}\omega^2 q_1^2)dt \qquad (79)$$

Term II can be written as

$$\tfrac{1}{2}\frac{d}{dt}(q_{c\ell}\dot{q}_{c\ell}) - \tfrac{1}{2}q_{c\ell}(\ddot{q}_{c\ell}+\omega^2 q_{c\ell}-j) + \tfrac{1}{2}jq_{c\ell} \tag{80}$$

after integration over the interval and using the properties of $q_{c\ell}$ as a solution this gives a contribution

$$\tfrac{1}{2}q_{c\ell}(t_b)\dot{q}_{c\ell}(t_b) + \tfrac{1}{2}\int_{-\infty}^{+\infty} j(t)q_{c\ell}(t)dt \tag{81}$$

which can be calculated by means of (77). Term III can be written as

$$\frac{d}{dt}(q_1\dot{q}_{c\ell}) - q_1(\ddot{q}_{c\ell}+\omega^2 q_{c\ell}-j) \tag{82}$$

which contributes only a boundary term to the action, namely $q_1(t_b)\dot{q}_{c\ell}(t_b)$. We denote $q_{c\ell}(t_b) = \beta$, $\dot{q}_{c\ell}(t) = \beta'$, $\int_{-\infty}^{+\infty} j(t)q_{c\ell}(t)dt = \alpha$ and observe that the change of variables $q(\cdot)$ $q_1(\cdot)$ in the path integral is a translation. All this together gives us

$$K(x_b,t_b;\, x_a,t_a) =$$

$$= e^{\frac{i}{2}\alpha}\, e^{i x_b \beta' - \frac{i}{2}\beta\beta'} N \int_{\substack{q_1(t_a)=x_a \\ q_1(t_b)=x_b-\beta}} e^{i\int_{t_a}^{t_b} L_0(\dot{q}_1,q)dt}\, \mathcal{D}[q_1(\cdot)] =$$

$$= e^{\frac{i}{2}\alpha}\, e^{-\frac{i}{2}\beta\beta'}\, e^{i x_b \beta'}\, K_0(x_b-\beta,t_b;\, x_a,t_a) \tag{83}$$

There is no need to calculate K_0, the integral kernel of the free oscillator, because it is the kernel of the operator $e^{-i(t-t_0)H_0}$, of which ϕ_0 is an eigenfunction.

$$(\phi_0, S[j]\phi_0) = e^{\frac{i}{2}\omega(t_b-t_a)}\, e^{\frac{i}{2}\alpha - \frac{i}{2}\beta\beta'}$$

$$\int_{-\infty}^{+\infty}\int \overline{\phi_0(x_b)}\, e^{i x_\beta \beta'} K_0(x_b-\beta,t_b;\, x_a,t_a)\phi_0(x_a)dx_a\, dx_b =$$

$$= e^{\frac{i}{2}\alpha - \frac{i}{2}\beta\beta'} \int_{-\infty}^{+\infty} \overline{\phi_0(x_b)}\, e^{i x_b \beta'} \phi_0(x_b-\beta)\,dx \tag{84}$$

With $\phi_0(x) = (\frac{\omega}{\pi})^{\frac{1}{4}} e^{-\frac{1}{2}\omega x^2}$, see (67), this becomes

$$e^{\frac{i}{2}\alpha - \frac{i}{2}\beta\beta'} (\tfrac{\omega}{\pi})^{\frac{1}{2}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\omega x^2 + i x \beta' - \frac{1}{2}\omega(x-\beta)^2}\,dx \tag{85}$$

This is an elementary integral, absolutely convergent, the result of which is

$$e^{\frac{i}{2}\alpha - \frac{\omega}{2}\beta^2 - \frac{1}{4\omega}\beta'^2} \tag{86}$$

Finally, we calculate β, β' and α

$$\beta = \frac{1}{\omega} \int_{-\infty}^{+\infty} \sin\omega(t_b-t')j(t')dt' \tag{87}$$

Because the end result does not depend on t_b as long as $t_b > \tau_0$ we may choose t_b in a convenient manner, namely $t_b = n\frac{2\pi}{\omega}$, n sufficiently large. Then

$$\beta = -\frac{1}{\omega} \int_{-\infty}^{+\infty} \sin\omega t'dt' \tag{88}$$

Similarly

$$\beta' = \int_{-\infty}^{+\infty} \cos\omega t'j(t')dt' \tag{89}$$

Elementary recombination in (86) leads then to the result

$$(\phi_0, S[j]\phi_0) = e^{\frac{i}{2}G_F[j,\,j]} \tag{90}$$

with G_F the bilinear functional

$$G_F[j_1, j_2] = \iint_{-\infty}^{\infty} j_1(t')G_F(t'-t'')j_2(t'')dt'dt'' \tag{91}$$

in which

$$G_F(t) = \frac{1}{2\omega}(\varepsilon(t)\sin\omega t + i\cos\omega t) = -\frac{1}{2\omega i}(\theta(t)e^{-i\omega t}+\theta(-t)e^{i\omega t})$$
$$\varepsilon(t) = \theta(t) - \theta(-t); \quad \theta(t) = 1,\ t > 0; \quad \theta(t) = 0,\ t < 0. \tag{92}$$

REMARKS. 1. $G_F(t)$ is a complex Green function of the oscillator equation. 2. The derivation of (92) which is for suitable $j(t)$ an exact and rigorous result, is a heursitic application of the idea that the path integral is indeed an integral with the usual properties. It is, however, not clear from this argument why the final result is such a simple expression and why a Green function of $\frac{d^2}{dt^2} + \omega^2$ appears in it. The following rather vague reasoning may make this more plausible: Note first the one-dimensional integral

$$\int_{-\infty}^{+\infty} e^{i(\frac{a}{2}x^2+bx)}dx = e^{-\frac{ib^2}{2a}}\int_{-\infty}^{+\infty} e^{i\frac{a}{2}y^2}dy = e^{-\frac{ib^2}{2a}}e^{i\frac{a}{2}y^2} \tag{93}$$

For n variables

$$\int_{-\infty}^{+\infty}\!\!\ldots\!\int e^{i(\frac{1}{2}\sum_{j\ell}A_{j\ell}x^jx^\ell + \sum_j b_jx^j)}dx_1\ldots dx_n =$$
$$= e^{-\frac{i}{2}\sum_{j,\ell} b_j(A^{-1})_{j\ell}b_\ell}\int_{-\infty}^{+\infty}\!\!\ldots\!\int e^{\frac{i}{2}\sum_{j,\ell}A_{j\ell}y^jy^\ell}dy_1\ldots dy_n$$
$$= e^{-\frac{i}{2}\sum_{j,\ell} b_j(A^{-1})_{j\ell}b_\ell}\left(\frac{2\pi i}{\det\{A_{j\ell}\}}\right)^{n/2} \tag{94}$$

or in vector notation $x = x_1, \ldots, x_n$, $dx = dx_1 \ldots dx_n$

$$\int_{-\infty}^{+\infty} e^{i(\frac{1}{2}(x,Ax)+(b,x))}dx =$$

$$= e^{-\frac{i}{2}(b,A^{-1}b)} \int_{-\infty}^{+\infty} e^{\frac{i}{2}(y,Ay)}dy = e^{-\frac{i}{2}(b,A^{-1}b)}\left(\frac{2\pi i}{\det A}\right)^{n/2} \tag{95}$$

The path integral has as an integrand of a similar type; the action in the exponent of (75) can be written apart from boundary terms as

$$-\tfrac{1}{2}\int q(t)\left(\frac{d^2}{dt^2}+\omega^2\right)g(t)dt + \int j(t)q(t)dt \tag{96}$$

The differential operator $\frac{d^2}{dt^2} + \omega^2$ plays the rôle of A in (95); it has no inverse in a strict sense, but "one-sided inverses" or Green functions. The function $j(t)$ takes the place of $b = b_1, \ldots, b_n$. The $j(\cdot)$ dependence of the result has indeed a form that corresponds to $e^{-\frac{i}{2}(b,A^{-1}b)}$.

3. The result (90) is of fundamental importance and moreover very general. It will appear in the same form in quantum field theory: The function $G_F(t)$ becomes a distribution in space-time in field theory and is then called *Feynman propagator*.

The calculation of the general amplitude $(\phi_m, S[j]\phi_n)$ may be performed by means of the path integral method, however, we prefer here a more direct operator method. Instead of the $(\phi_m, S[j]\phi_n)$ it is more convenient to calculate a different set of matrix elements first. We introduce the so-called *coherent states*. For every $z \in \mathbb{C}$ we define the vector ψ_z

$$\psi_z = \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}}\phi_n \tag{97}$$

This is a well-defined vector in H because $\sum_{n=0}^{\infty} \frac{|z|^{2n}}{n!} = e^{|z|^2} < \infty$. In terms of the creation operator A^* one has

$$\psi_z = \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}}\frac{1}{\sqrt{n!}}(A^*)^n\phi_0 = e^{zA^*}\phi_0 \tag{98}$$

An important property of the coherent states is that they are eigenvectors of the annihilation operator A

$$A\psi_z = z\psi_z \tag{99}$$

Using the fundamental relation $[A, A^*] = 1$ one finds

$$(\psi_{z_1}, \psi_{z_2}) = e^{\bar{z}_1 z_2} \tag{100}$$

REMARKS. 1. The coherent state ψ_0 coincides with the ground state ϕ_0. 2. The uncountable family of coherent states can be used as a kind of "continuous" non-orthonormal base. It has very interesting properties.

The basic ingredient in the determination of $(\psi_{z_1}, S\psi_{z_2})$ is the commutation relation $[A, A^*] = 1$. Define the following operator valued function of t

$$F(t) = U_I(t_0,t) A\, U_I(t,t_0) \tag{101}$$

for fixed t_0. One has, in agreement with (44)

$$\frac{dF(t)}{dt} = i\, U_I(t_0,t)[H_I(t), A]\, U_I(t_0,t) \tag{102}$$

with

$$H_I(t) = e^{iH_0 t} V(t) e^{-iH_0 t} = -j(t)\, e^{itH_0} Q\, e^{-itH_0} =$$

$$= -j(t)(2\omega)^{-\frac{1}{2}} e^{itH_0}(A+A^*) e^{-itH_0} \tag{103}$$

Because

$$e^{itH_0} A\, e^{-itH_0} = A + it[H_0, A] + \frac{(it)^2}{2!}[H_0, [H_0, A]] + \ldots$$

$$= A + it\omega A + \frac{(-it)^2}{2!}\omega^2 A + \ldots = e^{-i\omega t} A \tag{104}$$

and similarly

$$e^{itH_0} A^* e^{-itH_0} = e^{i\omega t} A^* \tag{105}$$

formula (103) becomes

$$-j(t)(2\omega)^{-\frac{1}{2}}(e^{-i\omega t}A + e^{i\omega t}A^*) \tag{106}$$

and with this

$$[H_I(t), A] = (2\omega)^{-\frac{1}{2}} j(t) e^{i\omega t} \tag{107}$$

The differential equation (102) then simplifies to

$$\frac{dF(t)}{dt} = i(2\omega)^{-\frac{1}{2}} j(t) e^{i\omega t} \tag{108}$$

With the initial condition $F(t_0) = A$ this has the solution

$$F(t) = A + i(2\omega)^{-\frac{1}{2}} \int_{t_0}^{t} e^{i\omega t'} j(t')dt' \tag{109}$$

Going back to the definition (101) and writing $s_{tt_0}[j] = i(2\omega)^{-\frac{1}{2}} \int_{t_0}^{t} e^{i\omega t'} j(t')dt'$ one obtains

$$U_I(t_0,t)A\, U_I(t,t_0) = A + s_{tt_0}[j] \tag{110}$$

and from this the commutation relations

$$\begin{aligned} A\, U_I(t,t_0) &= U_I(t,t_0)(A + s_{tt_0}[j]) \\ U_I(t,t_0)A^* &= (A^* - \overline{s_{tt_0}[j]})\, U_I(t,t_0) \end{aligned} \tag{111}$$

With these commutation relations and the properties (98), (99) and (100) one finally calculates the matrix elements $(\psi_{z_1}, S[j]\psi_{z_2})$

$$(\psi_{z_1}, S[j]\psi_{z_2}) = \lim_{\substack{t\to\infty \\ t\to-\infty}} (\psi_{z_1}, U_I(t,t_0)\psi_{z_2}) \tag{112}$$

and

$$(\psi_{z_1}, U_I(t,t_0)\psi_{z_2}) = (e^{z_1 A^*}\phi_0, U_I(t,t_0)\psi_{z_2}) =$$

$$= (\phi_0, e^{\bar{z}A}U_I(t,t_0)\psi_{z_2}) = (\phi_0, U_I(t,t_0)e^{\bar{z}_1(A+s_{tt_0}[j])}\psi_{z_2}) =$$

$$= e^{\bar{z}_1 s_{tt_0}[j]}(\phi_0, U_I(t,t_0)e^{\bar{z}_1 A}\psi_{z_2}) =$$

$$= e^{\bar{z}_1 s_{tt_0}[j]}e^{\bar{z}_1 z_2}(\phi_0, U_I(t,t_0)\phi_0) =$$

$$= e^{\bar{z}_1 z_2}e^{s_{tt_0}[j]\bar{z}_1}(\phi_0, U_I(t,t_0)e^{z_2 A^*}\phi_0) =$$

$$= e^{\bar{z}_1 z_2}e^{s_{tt_0}[j]\bar{z}_1}(\phi_0, e^{z_2(A^*-s_{tt_0}[j])}U_I(t,t_0)\phi_0) =$$

$$= e^{\bar{z}_1 z_2}e^{s_{tt_0}[j]\bar{z}_1 - \overline{s_{tt_0}[j]}z_2}(\phi_0, U_I(t,t_0)\phi_0) \qquad (113)$$

The limit $t_0 \to -\infty$, $t \to \infty$ can be taken, we know already the vacuum-vacuum amplitude, so we find

$$(\psi_{z_1}, S[j]\psi_{z_2}) = e^{\bar{z}_1 z_2}e^{s[j]\bar{z}_1}e^{\overline{s[j]}z_2}e^{\frac{i}{2}G_F[j,j]} \qquad (114)$$

with $s[j]$ the linear functional

$$s[j] = i(2\omega)^{-\frac{1}{2}}\int_{-\infty}^{+\infty} e^{i\omega t}j(t)dt \qquad (115)$$

It is obvious that this result will also hold when $j(t)$ tends to zero sufficiently fast for $|t| \to \infty$, instead of support $(j(t))$ in a finite interval.

It is not hard to obtain from the result (114) S as an operator expression. From (99) one obtains immediately

$$(\psi_{z_1}, (A^*)^\ell A^s \psi_{z_2}) = \bar{z}_1^\ell z_2^s e^{\bar{z}_1 z_2} \tag{116}$$

For an operator of the form $p_1(A^*)p_2(A)$, with $p_1(\cdot)$, $p_2(\cdot)$ polynomials, or suitable limits of polynomials this gives

$$(\psi_{z_1}, p_1(A^*)p_2(A)\psi_{z_2}) = e^{\bar{z}_1 z_2} p_1(\bar{z}_1)p_2(z_2) \tag{117}$$

In general and apart from technical points that are irrelevant to the spirit of our discussion every operator B can be written as an expression in creation and annihilation operators, in a unique manner in *normal ordening*, i.e. with all creation operators in front of all annihilation operators.

$$B = \sum_{\ell,s=0}^{\infty} \beta_{\ell s} (A^*)^\ell A^s \tag{118}$$

one has

$$(\psi_{z_1}, B\psi_{z_2}) = e^{\bar{z}_1 z_2} \sum_{\ell,s=0}^{\infty} \beta_{\ell s} \bar{z}_1^\ell z_2^s \tag{119}$$

This means that an operator B is uniquely characterized by a function $\beta(z_1, z_2)$

$$\beta(z_1, z_2) := (\psi_{z_1}, B\psi_{z_2}) \tag{120}$$

This function is anti-holomorphic in z_1 (i.e. holomorphic in $\bar{z}_1$) and holomorphic in z_2. It is called the Bargmann kernel of the operator B. The expression (114) is obviously the Bargmann kernel for the operator S. It is, in normal form,

$$S[j] = e^{\frac{i}{2} G_F[j,j]} e^{s[j]A^*} e^{\overline{s[j]}A} \tag{121}$$

We derive finally the physically more interesting matrix elements $(\phi_m, S[j]\phi_n)$. For this we observe that from $A\phi_k = \sqrt{k}\,\phi_{k-1}$, $k = 1,2,\ldots$, and $A\phi_0 = 0$ follows that $A^r \phi_k = \sqrt{\frac{k!}{(k-r)!}}\,\phi_{k-r}$, for $k - r \geq 0$, and

from this

$$(\phi_m,\ (A^*)^{\ell} A^{s} \phi_n) \ = \ \left(\frac{m!n!}{(m-\ell)!(n-s)!}\right)^{\frac{1}{2}} \tag{122}$$

for $m-\ell = n-s \geqslant 0$, and otherwise = 0.

One verifies that the right-hand side, for all values of m, ℓ, n, s can be written as

$$\frac{1}{\sqrt{m!}} \frac{1}{\sqrt{n!}} \left\{ \left(\frac{d}{d\bar{z}_1}\right)^m \left(\frac{d}{dz_2}\right)^n \left(e^{\bar{z}_1 z_2} \bar{z}_1^{\ell} z_2^{s} \right) \right\}_{z_1=0,\ z_2=0} \tag{123}$$

This means that for an arbitrary operator B the matrix elements with respect to the ϕ_n can be obtained from the Bargmann kernel by differentiation

$$(\phi_m, B\,\phi_n) \ = \ \frac{1}{\sqrt{m!}} \frac{1}{\sqrt{n!}} \left\{ \left(\frac{d}{d\bar{z}_1}\right)^m \left(\frac{d}{d\bar{z}_2}\right)^n \beta(z_1, z_2) \right\}_{z_1=0,\ z_2=0} \tag{124}$$

For the S-operator one has in particular

$$(\phi_m, S[j]\phi_n) \ =$$

$$= \frac{1}{\sqrt{m!}} \frac{1}{\sqrt{n!}} \left\{ \left(\frac{d}{d\bar{z}_1}\right)^m \left(\frac{d}{d\bar{z}_2}\right)^n (\psi_{z_1}, S[j]\psi_{z_2}) \right\}_{z_1=0,\ z_2=0} \tag{125}$$

with $(\psi_{z_1}, S[j]\psi_{z_2})$ as given by (114). This completes the derivation of the S-operator and its matrix elements for the oscillator in a homogeneous force, a result that will be used in the next section.

10. THE ANHARMONIC OSCILLATOR

In this section we shall discuss the quartic anharmonic oscillator, as a simple analogon for a non-trivial quantum field theory. In particular, we shall set up a method for a successive calculation of the terms $S^{(\ell)}$ in the Dyson perturbation series for the S-operator

$$S = \sum_{\ell=0}^{\infty} S^{(\ell)}$$

$$S^{(\ell)} = \frac{1}{\ell!} (-i)^{\ell} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} dt_1 \dots dt_{\ell} \, T\{H_I(t_1) \dots H_I(t_{\ell})\} \qquad (126)$$

In this procedure we shall make full use of the path integral as a heuristic device.

The anharmonic oscillator was given by the classical Lagrange function $L(q,\dot{q},t) = L_0(q,\dot{q}) + L_1(q,t)$, with $L_0(q,\dot{q}) = \frac{1}{2}\dot{q}^2 - \frac{1}{2}\omega^2 q^2$ the Lagrange function of the free oscillator, and $L_1(q,t)$ an anharmonic perturbation of the form $-g(t)q^4$. (This fourth power is the most obvious non-trivial example; the method to be discussed is, however, applicable to general polynomial perturbations)

We suppose again for technical reasons that the function $g(t)$ has its support in a finite time interval. Initial and final times t_a and t_b are supposed to be taken outside this interval. The limit $t_a \to -\infty$, $t_b \to +\infty$ for the S-operator is then trivial

Let ψ and ψ' be arbitrary vectors in the Hilbert space H. We have

$$(\psi, S\psi') = \lim_{\substack{t_a \to -\infty \\ t_b \to +\infty}} (\psi, e^{i t_b H_0} U(t_b,t_a) e^{-i t_a H_0} \psi') = \qquad (127)$$

$$= (\psi, e^{i t_b H_0} U(t_b,t_a) e^{-i t_a H_0} \psi') = (\psi_{t_b}, U(t_b,t_a)\psi'_{t_a})$$

with

$$\psi_{t_b} = e^{-i\,t_b H_0}\psi$$
$$\psi'_{t_a} = e^{-i\,t_a H_0}\psi \tag{128}$$

By means of the path integral formula for $U(t_b,t_a)$ (127) becomes

$$\int_{-\infty}^{+\infty}\int \overline{\psi_{t_b}(x_b)}\left[N \int_{\substack{q(t_a)=x_a\\ q(t_b)=x_b}} e^{i\int_{t_a}^{t_b}(L_0+L_1)dt}\,\mathcal{D}[q(\cdot)]\right]\psi'_{t_a}(x_a)dx_b\,dx_a =$$

$$= N \int_{\substack{q(\cdot)\ \text{arbitrary}\\ \text{on } [t_a,t_b]}} \overline{\psi_{t_b}(q(t_b))}\psi'_{t_a}(q(t_a))e^{i\int_{t_a}^{t_b}(L_0+L_1)dt}\,\mathcal{D}[q(\cdot)] \tag{129}$$

Expanding $e^{i\int_{t_a}^{t_b} L_1\,dt}$ in this formula in a very naive way one obtains

$$\sum_{\ell=0}^{\infty}\frac{(i)^\ell}{\ell!}\int_{q(\cdot)\ \text{on}\ [t_a,t_b]} \overline{\psi_{t_b}(q(t_b))}\psi'_{t_a}(q(t_a))\left(\int_{t_a}^{t_b} L_1\,dt\right)^\ell e^{i\int_{t_a}^{t_b} L_0\,dt}\,\mathcal{D}[q(\cdot)] \tag{130}$$

We denote the functionals $\overline{\psi_{t_b}(q(t_b))}\psi'_{t_a}(q(t_a))$ as $R[q]$, $\int_{t_0}^{t_1} L_0(q(t),\dot{q}(t))dt$ as $I_0[q]$ and $\int_{t_a}^{t_1} L_1(q(t),t)dt$ as $I_1[q]$, and obtain then

$$(\psi, S\,\psi') = \sum_{\ell=0}^{\infty}\frac{i^\ell}{\ell!}\int_{q(\cdot)\ \text{on}\ [t_a,t_b]} R[q](I_1[q])^\ell e^{I_1[q]}\,\mathcal{D}[q(\cdot)] \tag{131}$$

This series can be identified with the Dyson perturbation series; we have for each term

$$(\psi, S^{(\ell)}\psi') = \frac{i^\ell}{\ell!} \int R[q](I_1[q])^\ell \, e^{I_1[q]} \, \mathcal{D}[q(\cdot)] \tag{132}$$

a path integral expression for the ℓ^{th} order term in the Dyson series for a general matrix element $(\psi, S\psi')$ of the S-operator. This expression has to be calculated.

To understand the essence of the method for this we consider for a moment an analogon in terms of one-dimensional integrals. Let $R(x)$, $I_0(x)$, and $I_1(x)$ be functions of a single real variable. Consider the integral

$$\int_{-\infty}^{+\infty} R(x) e^{i(I_0(x)+I_1(x))} dx =$$

$$= \sum_{\ell=0}^{\infty} \frac{i^\ell}{\ell!} \int_{-\infty}^{+\infty} R(x)(I_1(x))^\ell \, e^{i I_0(x)} dx \tag{133}$$

Define a generating function

$$Z(u) = \int_{-\infty}^{+\infty} R(x) e^{i(I_0(x)+ux)} dx \tag{134}$$

one has

$$\frac{d^k}{du^k} Z(u) = i^k \int_{-\infty}^{+\infty} R(x) x^u e^{i(I_0(x)+ux)} dx \tag{135}$$

and therefore

$$\int_{-\infty}^{+\infty} R(x) x^k e^{i I_0(x)} dx = (-i)^k \left(\frac{d^k}{du^k} Z(u)\right)_{u=0} \tag{136}$$

all this of course in a very heuristic sense, with formal power series.

This "moment" formula can be extended to polynomials in x

$$\int_{-\infty}^{+\infty} R(x)p(x)e^{i I_0(x)}dx = \left(p(-i\frac{d}{du})Z(u)\right)_{u=0} \tag{137}$$

The term $(I_1(x))^{\ell}$ in (132) is a polynomial, and this means that we have a method for calculating the terms in (133) by differentiation of the generating function in $u = 0$. It is important to observe that this makes sense as a practical method only when $Z(u)$ is explicitly known and is such that repeated differentiation is not to difficult.

We apply these ideas now to the path integral expression (132):

a. One defines a generating functional

$$Z[u] := N \int_{q(\cdot) \text{ on } [t_a,t_b]} R[q]e^{i(I_0[q] + \int_{t_a}^{t_b} u(t)q(t)dt)} \mathcal{D}[q(\cdot)] \tag{138}$$

which is a functional on functions $u(\cdot)$ on the interval $[t_a, t_b]$.

b. One observes that in the path integral (132) that has to be calculated the functional $(I_1[q])^{\ell}$ is a polynomial. For this to be clear we must briefly discuss what we mean by polynomials in the case of functionals.

As functional $F[q]$ defined on a linear space of functions V is a homogeneous polynomial of degree K if $F[q]$ can be written as $F[q] = M[q, \ldots, q]$, with M a symmetric K linear map from $V \times \ldots \times V$ to $\mathbb{R}$ or $\mathbb{C}$. In that case M is uniquely determined by F. A functional $F[q]$ is a polynomial if it is a finite sum of homogeneous polynomials.

Examples:

1. $$I_1[q] = -\int_{t_a}^{t_b} g(t)q^4(t)dt \tag{139}$$

This is a homogeneous polynomial of degree 4 with

$$M[q_1, \ldots, q_4] = -\int_{t_a}^{t_b} g(t)q_1(t) \ldots q_4(t)dt \tag{140}$$

2. The free action

$$I_0[q] = \int_{t_a}^{t_b} \tfrac{1}{2}(\dot{q}(t)^2 - \omega^2 q(t)^2)dt \tag{141}$$

is a homogeneous quadratic functional with

$$M[q_1, q_2] = \int_{t_a}^{t_b} \tfrac{1}{2}(\dot{q}_1(t)\dot{q}_2(t) - \omega^2 q_1(t)q_2(t))dt \tag{142}$$

With a and b we are in a position to develop a procedure for the calculation of the path integral (132). The generating functional $Z[u]$ is indeed known as a simple closed expression. If we write for a moment $j(\cdot)$ instead of $u(\cdot)$ in (137), we recognize easily the expression for $(\psi, S[j]\psi)$, the general matrix element for the oscillator in an external force that we have determined explicitly in the last section, and that in this manner becomes useful here. The generating functional depends on two arbitrarily chosen vectors ψ and ψ' which we now take as $\psi = \psi_{z_1}$, $\psi' = \psi_{z_2}$. We write then $Z_{z_1 z_2}[j]$ for the generating functional and have as explicit result

$$\begin{aligned} Z_{z_1 z_2}[j] &= (\psi_{z_1}, S[j]\psi_{z_2}) = \\ &= e^{\bar{z}_1 z_2} e^{s[j]\bar{z}_1} e^{\overline{-s[j]}z_2} e^{\frac{i}{2}G_F[j,j]} \end{aligned} \tag{143}$$

This is indeed as functional in $j(\cdot)$ a simple expression, exponential functions containing linear and quadratic functional.

To proceed further we need a suitable definition of *functional differentiation*. It is not difficult to give such a definition, partly heuristic, but sufficient for our purpose. We generalize the notion of directional derivative:

Let $F[q]$ be a functional defined on a linear space V of functions $q(\cdot)$. One defines $(D_h F)[q]$, the derivative of F in q in the direction

of h as

$$(D_h F)[q] = \lim_{\lambda \to 0} \frac{1}{\lambda}\{F[q + \lambda h] - F[q]\} \tag{144}$$

This definition makes sense certainly for polynomial functionals. For a homogeneous polynomial of degree n, $F[q] = M[q, \ldots, q]$, with M the associated symmetric n-linear functional, this derivative is clearly

$$(D_h F)[q] = n\, M[h, q, \ldots, q] \tag{145}$$

The derivative is again a functional in q *and* at the same time a *linear* functional in h. The usual properties, Leibniz rule, chain rule, etc. hold.

In the physics literature a heuristic "partial" derivative is used. It fits in the picture that a function $q(\cdot)$ stands for an infinite of variables $\{q(t)\}_{t\in\mathbb{R}}$. This heuristic functional derivative will also be employed in our discussion. It can be obtained from (144) by taking for $h(\cdot)$ the Dirac δ-function $\delta_{t_1}(\cdot) = \delta(t_1 - \cdot)$:

$$\frac{\delta F[q]}{\delta q(t_1)} := (D_{\delta_{t_1}} F)[q] \tag{146}$$

The right-hand side of this formula is of course very symbolic and non-rigorous although the result may be well-defined, as is clear from the following example

$$\begin{aligned} F[q] = I_1[q] &= -\int_{t_a}^{t_b} g(t) q^4(t)\, dt \\ (D_h F)[q] &= -4 \int_{t_a}^{t_b} g(t) h(t) q^3(t)\, dt \\ \frac{\delta F[q]}{\delta q(t_1)} &= -4 g(t_1) q^3(t_1) \end{aligned} \tag{147}$$

It is useful to express $(D_h F)[q]$ in the heuristic derivative

$$(D_h F)[q] = \int h(t_1) \frac{\delta F[q]}{\delta q(t_1)} dt_1 \tag{148}$$

In an analogous manner one defines higher functional derivatives. The second derivative is for instance

$$(D_{h_1h_2}F)[q] := \lim_{\lambda\to 0}\frac{1}{\lambda}\{(D_{h_2}F)[q+\lambda h_1] - (D_{h_2}F)[q]\} \tag{149}$$

Applied on polynomials $(D_{h_1h_2}F)[q]$ is again a functional in $q(\cdot)$ *and* a symmetric bilinear functional in $h_1(\cdot)$ and $h_2(\cdot)$. There is a heuristic second derivative

$$\frac{\delta^2 F[q]}{\delta q(t_1)\delta q(t_2)} := (D_{\delta_{t_1}\delta_{t_2}}F)[q] \tag{150}$$

with the relation

$$(D_{h_1h_2}F)[q] = \iint h_1(t_1)h_2(t_2)\,\frac{\delta^2 F[q]}{\delta q(t_1)\delta q(t_2)}dt_1dt_2 \tag{151}$$

EXAMPLE:

$$F[q] = I_1[q] = -\int_{t_a}^{t_b} g(t)q^4(t)dt$$

$$(D_{h_1h_2}F)[q] = -3\cdot 4\cdot\int_{t_a}^{t_b} g(t)h_1(t)h_2(t)q^2(t)dt$$

$$(D_{h_1h_2h_3}F)[q] = -2\cdot 3\cdot 4\cdot\int_{t_a}^{t_b} g(t)h_1(t)h_2(t)h_3(t)q(t)dt$$

$$\frac{\delta^2 F[q]}{\delta q(t_1)\delta q(t_2)} = -3\cdot 4g(t_1)\delta(t_1-t_2)q(t_1)^2$$

$$\frac{\delta^3 F[q]}{\delta q(t_1)\delta q(t_2)\delta q(t_3)} = -2\cdot 3\cdot 4\cdot\delta(t_1-t_2)\delta(t_1-t_3)q(t_1) \tag{152}$$

REMARK. It is of course possible to extend the idea of functional differentiation from polynomials to a wider class of functionals in a mathematically rigorous manner. The definition (143), essentially a Gâteaux derivative is too weak to have many useful properties. The Fréchèt derivative would be more appropriate, but would involve the introduction of topological proper-

ties, norms or seminorms in V, continuity of the functionals, etc. All this is unnecessary for our purpose here, a heuristic method of generating formal power series.

We apply this concept of functional differentiation on (138)

$$Z[j] = N \int_{q(\cdot) \text{ on } [t_a,t_b]} R[q] e^{i(I_0[q] + \int_{t_a}^{t_b} j(t)q(t)dt} \mathcal{D}[q(\cdot)] \tag{153}$$

One differentiates $Z[j]$ in $j(\cdot) = 0$

$$(D_h Z)(0) = N \int_{q(\cdot) \text{ on } [t_a,t_b]} R[q](i \int_{t_a}^{t_b} h(t)q(t)dt) e^{i I_0[q]} \mathcal{D}[q(\cdot)] \tag{154}$$

which gives

$$-i\left(\frac{\delta Z[j]}{\delta j(t_1)}\right)_{j=0} = N \int R[q]\, q(t_1) e^{i I_0[q]} \mathcal{D}[q(\cdot)] \tag{155}$$

More generally one has

$$(-i)^k \left(\frac{\delta^k Z[j]}{\delta j(t_1) \ldots \delta j(t_k)}\right)_{j=0} =$$

$$= N \int_{q(\cdot) \text{ on } [t_a,t_b]} R[q]\, q(t_1) \ldots q(t_k) e^{i I_0[q]} \mathcal{D}[q(\cdot)] \tag{156}$$

This gives us a formula for $(\psi_{z_1}, S^{(\ell)} \psi_{z_2})$

$$(\psi_{z_1}, S^{(\ell)} \psi_{z_2}) =$$

$$= \frac{i^\ell}{\ell!} N \int_{q(\cdot) \text{ on } [t_a,t_b]} R[q](I_1[q])^\ell e^{(I_0[q]} \mathcal{D}[q(\cdot)] =$$

$$= \frac{(-i)^{\ell}}{\ell!} N \int_{q(\cdot)} R[q]\{ \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} dt_1 \dots dt_{\ell}\, g(t_1) \dots g(t_{\ell})\, q(t_1)^4 \dots q(t_{\ell})^4 \}$$

$$e^{i I_0[q]} \mathcal{D}[q(\cdot)] = \frac{(-i)^{\ell}}{\ell!} N \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} dt_1 \dots dt_{\ell}\, g(t_1) \dots g(t_{\ell})$$

$$\{N \int_{q(\cdot)} R[q] q(t_1)^4 \dots q(t_{\ell})^4 e^{i I_0[q]} \mathcal{D}[q(\cdot)]\} \tag{157}$$

and finally with (156)

$$(\psi_{z_1}, S^{(\ell)} \psi_{z_2}) =$$

$$= \frac{(-i)^{\ell}}{\ell!} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} dt_1 \dots dt_{\ell}\, g(t_1) \dots g(t_{\ell}) \left((-i)^{4\ell} \frac{\delta^{4\ell} Z_{z_1 z_2}[j]}{\delta j(t_1)^4 \dots \delta j(t_{\ell})^4} \right)_{j=0} \tag{158}$$

with

$$Z_{z_1 z_2}[j] = e^{\bar{z}_1 z_2} e^{s[j]\bar{z}_1} e^{\overline{-s[j]} z_2} e^{\frac{i}{2} G_F[j,j]} \tag{159}$$

Our last task will be to carry out the functional differentiation in (158) in an explicit, but systematic manner. This will be the subject of the next sections.

REMARK. The result (158) can be written in a very symbolic way as

$$(\psi_{z_1}, S^{(\ell)} \psi_{z_2}) = \frac{i^{\ell}}{\ell} \left(\left(\int_{t_a}^{t_b} L_1(-i \frac{\delta}{\delta j(t)}, t) dt \right)^{\ell} Z_{z_1 z_2}[j] \right)_{j=0} \tag{160}$$

An even more symbolic, but also very compact formulation of the result for the total matrix element as

$$(\psi_{z_1}, S \psi_{z_2}) = \left(e^{i \int_{t_a}^{t_b} L_1(-i \frac{\delta}{\delta j(t)}, t) dt} Z_{z_1 z_2}[j] \right)_{j=0} \tag{161}$$

This formula suggests that the heuristic validity of the method is much more general than the particular $L_1(q,t) = -g(t)q^4$ that we have been discussing.

11. EXPLICIT CALCULATIONS OF THE MATRIX ELEMENTS

To summarize the results of the last section we may say that the ℓ^{th} order term in

$$(\psi_{z_1}, S\psi_{z_2}) = \sum_{\ell=0}^{\infty} (\psi_{z_1}, S^{(\ell)}\psi_{z_2}) \tag{162}$$

can be calculated as

$$(\psi_{z_1}, S^{(\ell)}\psi_{z_2}) =$$

$$= \frac{(-i)^{\ell}}{\ell!} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} g(t_1) \cdots g(t_\ell)(-i)^{4\ell} \left(\frac{\delta^{4\ell} Z_{z_1 z_2}[j]}{\delta j(t_1)^4 \cdots \delta j(t_\ell)^4} \right)_{j=0} dt_1 \cdots dt_\ell \tag{163}$$

with

$$Z_{z_1 z_2}[j] = e^{\bar{z}_1 z_2} e^{\frac{i}{2} G_F[j,j]} e^{s[j]\bar{z}_1} e^{\overline{-s[j]} z_2} \tag{164}$$

Note that (163) must be compared with the ℓ^{th} order term in the Dyson series

$$(\psi_{z_1}, S^{(\ell)}\psi_{z_2}) = \frac{(-i)^{\ell}}{\ell!} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} (\psi_{z_1}, T\{H_I(t_1) \cdots H_I(t_\ell)\}\psi_{z_2}) dt_1 \cdots dt_\ell \tag{165}$$

Differentiating the functional $Z_{z_1 z_2}[j]$ is not particularly difficult. One uses Leibniz's rule, the chain rule and the derivatives of the simple functionals $G_F[j,j]$ and $s[j]$

$$\begin{aligned} (D_h G)[j,j] &= 2G_F[h,j] \\ (D_{h_1 h_2} G)[j,j] &= 2G_F[h_1, h_2] \end{aligned} \tag{166}$$

(higher derivatives = 0)

$$(D_h s)[j] = s[h] \tag{167}$$

(higher derivatives = 0)

Using this one obtains

$$(D_h Z_{z_1 z_2})[j] =$$

$$(i G_F[h, j] + s[h]\bar{z}_1 - \overline{s[h]z_2}) Z_{z_1 z_2}[j] \tag{168}$$

In terms of the heuristic "partial" derivatives

$$\frac{\delta}{\delta j(t)} G_F[j, j] = 2 \int G_F(t-t')j(t')dt'$$

$$\frac{\delta^2}{\delta j(t_1)\delta j(t_2)} = G_F[j, j] = 2G_F(t_1-t_2) \tag{169}$$

$$\frac{\delta}{\delta j(t)} s[j] = i(2\omega)^{-\frac{1}{2}} e^{-i\omega t}$$

The separate steps are all very simple. Nevertheless for higher derivatives of $Z_{z_1 z_2}[j]$ things will become quite complicated, so what is needed is a systematic approach to the differentiation procedure.

For application of formula (163) we use higher "partial" derivatives of $Z_{z_1 z_2}[j]$ taken in $j = 0$

$$\left(\frac{\delta^n Z_{z_1 z_2}[j]}{\delta j(t_1) \ldots \delta j(t_n)} \right)_{j=0} \tag{170}$$

This is a function of $t_1, t_2, \ldots t_n$, with $n = 4\ell$. It is used for values $t_1 = t_2 = t_3 = t_4$, $t_5 = t_6 = t_7 = t_8$, etc. A way of finding such partial derivatives of a functional is to write the functional as a power series around $j = 0$.

In general a functional $F[j]$ can be developed as

$$F[j] = F[0] + (D_j F)[0] + \frac{1}{2!}(D_{jj} F)[0] + \ldots \tag{171}$$

If $F[j]$ can be written as

$$F[j] = \sum_{p=0}^{\infty} \frac{1}{p!} \int_{-\infty}^{-\infty} \dots \int_{-\infty}^{+\infty} M^{(p)}(t_1,\dots,t_p) j(t_1) \dots j(t_p) dt_1 \dots dt_p \tag{172}$$

with functions $M^{(p)}(t_1,\dots,t_p)$ that are symmetric in $t_1, \dots, t_p$, than these functions can be identified with the heuristic partial derivatives

$$M^{(p)}(t_1,\dots,t_p) = \left(\frac{\delta^p F[j]}{\delta j(t_1) \dots \delta j(t_p)} \right)_{j=0} \tag{173}$$

It is very easy to expand the generating functional $Z_{z_1 z_2}$ by expanding the three exponential factors, that depend on j.

$$Z_{z_1 z_2}[j] = \\ = e^{\bar{z}_1 z_2} \sum_{k_1,k_2,r=0}^{\infty} \frac{1}{k_1!\, k_2!\, r!} \left(\frac{i}{2} G_F[j,j] \right)^r \left(s[j]\bar{z}_1 \right)^{k_1} \left(\overline{-s[j]} z_2 \right)^{k_2} \tag{174}$$

We use temporarily

$$\begin{aligned} \alpha(t) &= i(2\omega)^{-\frac{1}{2}} e^{i\omega t} \\ \beta(t) &= i(2\omega)^{-\frac{1}{2}} e^{-i\omega t} \\ G'(t) &= \frac{i}{2} G_F(t) \end{aligned} \tag{175}$$

and write the integrals in (174) in an explicit way

$$Z_{z_1 z_2}[j] = \sum_{p=0}^{\infty} \frac{1}{p!} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{\bar{z}_1 z_2} \sum_{\substack{k_1,k_2,r \\ k_1+k_2+2r=p}} \frac{p!}{k_1!\, k_2!\, r!} (\bar{z}_1)^{k_1} z_2^{k_2}$$

$$\cdot \alpha(t_1) \dots \alpha(t_{k_1}) G'(t_{k_1+1} - t_{k_1+2}) \dots G'(t_{k_1+2r-1} - t_{k_1+2r})$$

$$\cdot \beta(t_{k_1+2r+1}) \dots \beta(t_p) j(t_1) \dots j(t_p) dt_1 \dots dt_p \tag{176}$$

After symmetrization in $t_1, \dots, t_p$ this gives the heuristic partial derivatives

$$\left(\frac{\delta^p Z_{z_1 z_2}[j]}{\delta j(t_1)\ldots\delta j(t_p)}\right)_{j=0} = e^{\bar{z}_1 z_2} \sum_{\substack{k_1,k_2,r \\ k_1+k_2+2r=p}} \frac{(\bar{z}_1)^{k_1} z_2^{k_2}}{k_1!\, k_2!\, r!}$$

$$\sum_{\sigma\in S_p} \alpha(t_{\sigma(1)}) \ldots \alpha(t_{\sigma(k_1)})\; G'(t_{\sigma(k_1+1)} - t_{\sigma(k_1+2)}) \ldots$$

$$\ldots\; G'(t_{\sigma(k_1+2r-1)} - t_{\sigma(k_1+2r)})\; \beta(t_{\sigma(t_1+2r+1)}) \ldots \beta(t_{\sigma(k_1+2r+k_2)}) \qquad (176)$$

Such a derivative consists of a finite sum of terms which can be divided into classes, characterized by 3 numbers k_1, k_2 and r, with $k_1 + k_2 + 2r = p$. The terms belonging to the same class are obtained from each other by permutation of the t-variables. (S_p is the group of permutations of the numbers 1, 2, ... p). To avoid terms that are indistinguishable one could restrict the permutations by some ordering principle and put an extra factor $k_1!\, k_2!\, r!\, 2^r$ in front for compensation.

In (163) one has derivatives of the special type

$$\left(\frac{\delta^{4\ell} Z_{z_1 z_2}[j]}{\delta j(t_1)^4 \ldots \delta j(t_\ell)^4}\right)_{j=0} \qquad (177)$$

and these are obtained from (176) by putting the t variables equal in groups of four. This gives again terms which are equal, and this can again be taken into account by a further restriction on the permutations together with a compensating combinatorical factor. We shall not be concerned with this and concentrate on the classification of the different terms and on a discussion of their properties. This will entail the introduction of *diagram* techniques.

12. FEYNMAN DIAGRAMS

The ℓ^{th} order term $(\psi_{z_1}, S^{(\ell)}\psi_{z_2})$ in the series for $(\psi_{z_1}, S\psi_{z_2})$ consists of finitely many terms that can, in the first instance, be classified according to numbers k_1, k_2, r with $k_1 + k_2 + 2r = 4\ell$.

Our final aim is not an approximative calculation of $(\psi_{z_1}, S\psi_{z_2})$, but of the transition amplitudes $(\phi_m, S\phi_n)$. These are the objects of physical interest, because $|(\phi_m, S\phi_n)|^2$ is the probability to find the system in a final state ϕ_m with energy E_m when it was initially in a state ϕ_n with energy E_n. The matrix elements $(\phi_m, S\phi_n)$ and $(\sigma_m, S^{(\ell)}\phi_n)$ are obtained from the coherent state expressions $(\psi_{z_1}, S\psi_{z_2})$ and $(\psi_{z_1}, S^{(\ell)}\psi_{z_2})$ by the general formula (123) which for the ℓ^{th} order term of S gives

$$(\phi_m, S^{(\ell)}\phi_n) =$$

$$= \frac{1}{\sqrt{m!}}\frac{1}{\sqrt{n!}}\left\{\left(\frac{d}{d\bar{z}_1}\right)^m\left(\frac{d}{dz_2}\right)^n(\psi_{z_1}, S^{(\ell)}\psi_{z_2})\right\}_{\bar{z}_1=z_2=0} \tag{178}$$

Combining (163) with (176) we see that not all terms in $(\psi_{z_1}, S^{(\ell)}\psi_{z_2})$ contribute to the physical matrix element $(\phi_m, S^{(\ell)}\phi_n)$. We obtain the explicit expression

$$(\phi_m, S^{(\ell)}\phi_n) = \frac{(-i)^\ell}{\ell!}(-i)^{4\ell}(m!n!)^{\frac{1}{2}}$$

$$\sum_{\substack{k_1,k_2,r \\ 0\leqslant k_1\leqslant m, 0\leqslant k_2\leqslant n \\ k_1-k_2=m-n \\ r\geqslant 0; k_1+k_2+2r=4\ell}} \frac{1}{((m-k_1)!(n-k_2)!)^{\frac{1}{2}}}\frac{1}{k_1!\,k_2!\,r!}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} dt_1\cdots dt_\ell\, g(t_1)\cdots g(t_\ell)$$

$$\Bigg(\sum_{\sigma\in S_p}\alpha(\tau_{\sigma(1)})\cdots\alpha(\tau_{\sigma(k_1)})\,G'(\tau_{\sigma(k_1+1)}-\tau_{\sigma(k_1+2)})\cdots$$

$$\cdots G'(\tau_{\sigma(k_1+2r-1)}-\tau_{\sigma(k_1+2r)})\,\beta(\tau_{\sigma(k_1+2r+1)})\cdots$$

$$\left. \cdots \beta(\tau_{\sigma(k_1+2r+k_2)}) \right)_{\substack{\tau_1=\tau_2=\tau_3=\tau_4=t_1 \\ \tau_5=\tau_6=\tau_7=\tau_8=t_2, \text{ etc.}}} \tag{179}$$

One way to understand this formula is to return to operator language. We have seen that matrix elements with respect to the coherent states correspond with normally ordered operators through the relation $(\bar{z}_1)^{k_1} z_2^{k_2} e^{\bar{z}_1 z_2} \leftrightarrow$ $\leftrightarrow (A^*)^{k_1} A^{k_2}$. The energy eigenstates have the form $\phi_n = \frac{1}{\sqrt{n!}} (A^*)^n \phi_0$. The k_1 factors $\alpha(\cdot)\bar{z}_1$ give $(A^*)^{k_1}$, the k_2 factors $\beta(\cdot)z_2$ lead to A^{k_2}. A state ϕ_n is a state with n "quanta" of energy ω (or properly speaking $\hbar\omega$). An operator $(A^*)^{k_1} A^{k_2}$ takes away from the initial state ϕ_n (if possible) k_2 quanta, and then adds again k_1 quanta. In order to end up in ϕ_m one must clearly have $k_1 - k_2 = m - n$, one of the restrictions in the summation in (179). The terms characterized by k_1, k_2, r which occur in the summation and are permitted by the restriction, correspond to those parts of the operator $S^{(\ell)}$ written in normally ordered form as terms proportional to $(A^*)^{k_1} A^{k_2}$ that contribute to the matrix element $(\phi_m, S^{(\ell)} \phi_n)$.

EXAMPLES. 1. $\ell = 1$. (first order perturbation term)

a. Transition from ground state ϕ_0 to state ϕ_1; $m = 1$, $n = 0$. Because of $k_1 - k_2 = m - n$, $0 \leqslant k_1 \leqslant m$, $0 \leqslant k_2 \leqslant n$, the only possible pair is $k_1 = 1$, $k_2 = 0$. The condition $k_1 + k_2 + 2r = 4\ell$ becomes $2r = 3$, so there is no r. This means that there are no non-zero terms, i.e. $(\phi_1, S^{(1)} \phi_0) = 0$. (For this q^4 theory it is obvious that $k_1 + k_2 = 4\ell - 2r$ is even, so $m - n = k_1 - k_2$ even: i.e. no transitions for $m - n$ odd)

b. ϕ_0 to ϕ_2; i.e. $m = 2$, $n = 0$.

One possibility: $k_1 = 2$, $k_2 = 0$, $r = 1$.

c. ϕ_0 to ϕ_4; i.e. $m = 4$, $n = 0$.

One possibility: $k_1 = 4$, $k_2 = 0$, $r = 0$.

d. ϕ_2 to ϕ_2; i.e. $m = n = 2$.

Three possibilities: $k_1 = 0$, $k_2 = 0$, $r = 2$

$k_1 = 1$, $k_2 = 1$, $r = 1$

$k_1 = 2$, $k_2 = 2$, $r = 0$.

2. $\ell = 2$ (second order perturbation term)

ϕ_2 to ϕ_2; i.e. $m = n = 2$.

Five possibilities, for example $k_1 = 0$, $k_2 = 0$, $r = 4$

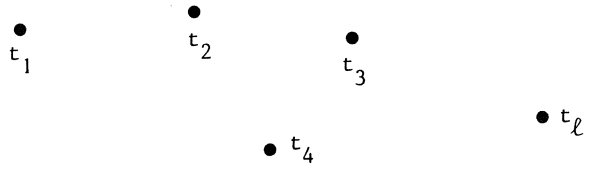
$k_1 = 2$, $k_2 = 2$, $r = 2$.

Formula (179) tells us which terms from $(\psi_{z_1}, S^{(\ell)}\psi_{z_2})$ contribute to the physical amplitude $(\phi_m, S^{(\ell)}\phi_n)$; the next thing to do is to discuss these terms themselves in more detail.

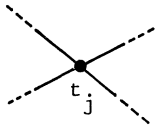
We choose a fixed set of integers k_1, k_2, r and ℓ such that $k_1 + k_2 + 2r = 4\ell$. If we do not bother about the combinatorial and other factors in front we have to discuss ℓ-fold integrals of the form

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} g(t_1) \dots g(t_\ell)\ \alpha(\cdot) \dots \alpha(\cdot)\ G'(\cdot,\cdot) \dots G'(\cdot,\cdot)\ \beta(\cdot) \dots \beta(\cdot)\ dt_1 \dots dt_\ell \qquad (180)$$

In this integral we have k_1 factors $\alpha(\cdot)$, k_2 factors $\beta(\cdot)$ and r factors $G'(\cdot,\cdot)$. On the $k_1 + k_2 + 2r$ dots there is a distribution of the ℓ variables $t_1, \dots, t_\ell$, in such a manner that every variable t_j is used 4 times. The possible different distributions, corresponding with the permutations σ in the second summation in (179) give different functions in the integral (180). These possibilities may be characterized by *diagrams* or graphs. Consider ℓ points, the *vertices*, denoted by the variables $t_1, \dots, t_\ell$. (Their relative positions are irrelevant)



Each vertex possesses 4 *lines*:



We construct a diagram by the following choices:

1. Choose k_1 lines that run to the left, and remain loose. (Several lines starting from the same point t_j may be selected).
These are the k_1 *outgoing lines*.

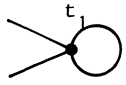
2. Choose k_2 lines that run in the same way to the right. These are the k_2 *incoming lines* (outgoing and incoming lines form together the *external* lines).

3. The remaining $4\ell - k_1 - k_2$ loose ends have to be connected. (It is permitted to connect some lines coming from the same point t_j). In this way we obtain $r = \frac{1}{2}(4\ell - k_1 - k_2)$ *internal lines*.

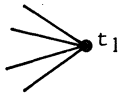
The choices 1, 2, 3 determine a diagram:

EXAMPLES

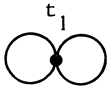
1. $\ell = 1$.

(a) 

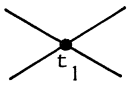
$k_1 = 2, \quad k_2 = 0, \quad r = 1.$

(b) 

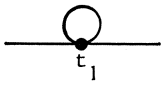
$k_1 = 4, \quad k_2 = 0, \quad r = 0.$

(c) 

$k_1 = 0, \quad k_2 = 0, \quad r = 2.$

(d) 

$k_1 = 2, \quad k_2 = 2, \quad r = 0.$

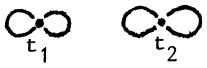
(e) 

$k_1 = 1, \quad k_2 = 1, \quad r = 1.$

2. $\ell = 2$.

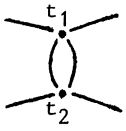
(a)  $k_1 = 0, \quad k_2 = 0, \quad r = 4.$

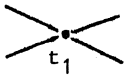
(b)  $k_1 = 0, \quad k_2 = 0, \quad r = 4.$

(c)  $k_1 = 0, \quad k_2 = 0, \quad r = 4.$

Note that 2a, b, c are different diagrams associated with the same set of numbers k_1, k_2, ℓ. Diagram 2c is to be considered as a single diagram. It is called disconnected, and is of course built from lower order diagrams. (The order of a diagram is the number of vertices ℓ.)

(d)  $k_1 = 2, \quad k_2 = 2, \quad r = 2.$

(e)  $k_1 = 2, \quad k_2 = 2, \quad r = 2.$

(f)  $k_1 = 2, \quad k_2 = 2, \quad r = 2.$

 (again a disconnected diagram)

Diagrams without external lines, such as 1c, 2a, 2b, 2c, are called vacuum diagrams.

Each diagram determines an integrand in (179) and (180) according to the following rules:

1. An outgoing line, starting at the vertex t_j gives a factor $\alpha(t_j) = i(2\omega)^{-\frac{1}{2}} e^{i\omega t_j}$.

2. An incoming line, ending at the vertex t gives a factor $\beta(t) = i(2\omega)^{-\frac{1}{2}} e^{-i\omega t}$.

3. An internal line from t_j to t_k gives a "propagator", i.e. a term $G'(t_k-t_j) = \frac{i}{2}G_F(t_k-t_j)$.

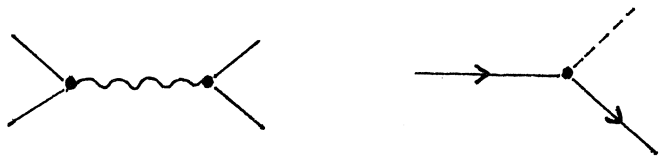
Finally there is a factor $g(t_j)$ for every vertex t_j. (In quantum field theory the "coupling function" becomes a coupling constant g; this results in an overall factor g^{ℓ} in front). To obtain the total contribution to the matrix element under discussion the appropriate combinatorial factor has to be supplied.

Although a *Feynman graph* or *Feynman diagram* as discussed above is really only a way of specifying an integral that contributes to the perturbation series of S-matrix elements, it is very often given a direct rather intuitive physical interpretation:

Consider as example diagram 2d. One reads it from right to left: In the initial state there are two quanta of energy. At time $t = t_2$ these disappear (are *annihilated*) because of the interaction the strength of which is given by $g(t_2)$. At the same moment a new pair of "virtual" quanta is created, these remain in existence until $y = t_1$, when they are annihilated but give at the same time rise to a new pair of non-virtual quanta which are observed in the final state. (We have supposed that $t_1 > t_2$, otherwise the interpretation is slightly different.)

In a system of N coupled oscillators, with different frequencies ω_k (and in the operator language described by N pairs of creation and annihilation operators A_k^*, A_k) we have a situation which can be described by a straightforward generalization of what we have discussed so far. We will find completely analogous results. There will be the same type of Feynman diagrams provided with extra indices. In the intuitive physical interpretation of the diagrams one thinks then of different types of quanta, with different energies $E_k = \hbar\omega_k$, that can be annihilated and created.

Quantum field theory can be seen as an infinite system of oscillators. The corresponding quanta not only have characteristic energies, but also momenta and various other possible properties such as for instance spin. These quanta are in fact the *particles* as we know them in elementary particle physics. The power of the formalism as we have explained it for a simple oscillator lies in the fact that the structure of the Feynman diagrams remains essentially the same in the much more complicated situation of quantum field theory. The same diagrams will carry more information. The vertices will signify space-time points (x,y,z,t) instead of time points t, there will be extra variables and indices, arrows, different types of lines, etc.



In quantum field theory it is also advantageous to employ instead of space-time variables a set of fourier-transformed variables with the same diagrams.

13. SELECTED REFERENCES

The path integral was introduced by Feynman in his Ph.D. thesis in 1942. His first publication on the subject was: R.P. Feynman: Space-time approach to non-relativistic quantum mechanics. Rev.Mod.Phys. 20, 367-387 (1948). His postulate on the form of the integrand of the path integral was inspired by among other things earlier work of Dirac: P.A.M. Dirac: The Lagrangian in quantum mechanics. Phys. Zeitschr.d. Sowjetunion 3, 64-72(1933). Feynman later wrote a book on path integrals, in collaboration with Hibbs: R.P. Feynman, A.R. Hibbs: Quantum mechanics and path integrals. New York 1965. This is now mainly of historical interest. The best modern book is: L.S. Schulman. Techniques and application of path integration. New York 1981. A book which gives the flavour of the use of path integral methods in advanced quantum field theory is: L.D. Faddeev, A.A. Slavnov: Gauge fields, introduction to quantum field theory.

Wiener's work on integration in function spaces started with: N. Wiener: The average of an analytical functional. Proc.Nat.Ac. of Sciences 7, 253-260(1921), a characteristic later publication is: N. Wiener: Differential space. J.Math. and Physics 2, 131-174(1923). The ideas of Wiener were developed further and brought into relation with the Feynman path integral by Kac, see e.g. M. Kac: On distributions of certain Wiener functionals. Trans.Amer.Math.Soc. 65, 1-13(1949). There are two books that give mathematically rigorous discussions of applications of path integrals on quantum physics through the method of analytical continuation to imaginary time, i.e. essentially the Wiener integral. These are: J. Glimm, A. Jaffe: Quantum physics, a functional integral point of view. New York 1981, and B. Simon: Functional integration and quantum physics. New York 1979.

INTRODUCTION TO SOLITONS IN (1+D) DIMENSIONS

by

E.A. de KERF
Institute for Theoretical Physics
University of Amsterdam

CONTENTS

1. INTRODUCTION

We discuss some examples of non-linear models in field theory in 1-time and D-spatial dimensions. These models have time-*in*dependent solutions which are non-singular, stable and localized in space and which have finite energy. These solutions are characterized by boundary conditions at infinity, they are called topological solitons.

The various models we are going to discuss share one feature, the ground-state is degenerate.

Examples in (1+1) dimensions are the ϕ^4-model and the sine-Gordon model for a scalar field $\phi(x,t)$. In (1+D) dimensions with $D > 1$ these models have no time-independent soliton solutions. (Derrick's theorem). A better insight in the relationship between the number of dimensions and the existence of soliton solutions is obtained by invoking arguments from homotopy theory. This gives firstly a general criterium for the existence of topological solitons, secondly one is lead in a natural way to the introduction of gauge fields. In section 2 we discuss the ϕ^4-model and the topological conservation law. In section 3 we consider briefly Derrick's theorem and a simple result from homotopy theory. In section 4 we consider and abelian gauge theory in (1+2) dimensions and the "topological quantization" of magnetic flux. Section 5 contains some general remarks.

The reader should be aware of the fact that this contribution to the seminar contains nothing new. It is meant to be a light-hearted interlude between the heavy mathematical contributions.

2. THE ϕ^4-MODEL (GOLDSTONE STRING) IN (1+1) DIMENSIONS

We consider first the Lagrangian density for the Klein-Gordon field $\phi(x,t)$.

$$L = \tfrac{1}{2}(\partial_\mu\phi\partial^\mu\phi - \frac{m^2c^2}{2}) = \tfrac{1}{2}\partial_\mu\phi\partial^\mu\phi - U(\phi)$$

$$(\partial_\mu = \frac{\partial}{\partial ct}, \frac{\partial}{\partial x}),\ (\partial^\mu = \frac{\partial}{\partial ct}, -\frac{\partial}{\partial x}),\ U(\phi) = \tfrac{1}{2}\frac{m^2c^2}{2}.$$

The equation of motion reads

$$\partial_\mu\partial^\mu\phi + \frac{m^2c^2}{2}\phi = 0. \tag{2.1}$$

This equation has solutions of the form

$$\phi(x,t) = A\sin(kx - \frac{\omega}{c}\cdot ct + \delta)$$

with $$(\omega)^2 = (kc)^2 + m^2c^4.$$

Comparing this with $E^2 = p^2c^2 + m^2c^4$, the relation between energy and momentum, one sees that m is a mass. The term $-\tfrac{1}{2}\frac{m^2c^2}{2}\phi^2$ in L is called the mass term. From now on we will use units for which $\hbar = 1$ and $c = 1$. Moreover we consider eq. (2.1) as a classical i.e. non-quantummechanical equation. Eq. (2.1) can be considered as the equation of motion of a string with potential energy density $\tfrac{1}{2}(\frac{\partial\phi}{\partial x})^2 + \tfrac{1}{2}m^2\phi^2$. The total energy, being the sum of kinetic energy and potential energy is given by

$$H = \int_{-\infty}^{\infty} \mathcal{H}(x,t)dx = \int_{-\infty}^{\infty} (\tfrac{1}{2}\dot{\phi}^2 + \tfrac{1}{2}\phi'^2 + U(\phi))dx \geqslant 0 \tag{2.2}$$

where $\phi(x,t)$ is a solution of (2.1), $\dot{\phi} = \frac{\partial\phi}{\partial t}$ and $\phi' = \frac{\partial\phi}{\partial x}$. From (2.2) it follows that $H = 0$ only for $\phi = 0$, there is no degeneracy of the ground state.

We now consider the so-called Goldstone string. The Lagrangian reads

$$L = \tfrac{1}{2}\partial_\mu\phi\partial^\mu\phi - U(\phi) \tag{2.3}$$

with $U(\phi)$ given by

$$(2.4) \qquad U(\phi) = \frac{\lambda^2}{4}\left(\phi^2 - \frac{m^2}{\lambda^2}\right)^2 = \frac{\lambda^2}{4}\phi^4 - \tfrac{1}{2}m^2\phi^2 + \frac{m^4}{4\lambda^2} \geqslant 0 \qquad (m,\lambda>0)$$

the equation of motion is

$$(2.5) \qquad \frac{\partial L}{\partial \phi} - \partial_\mu \frac{\partial L}{\partial \partial_\mu \phi} = 0 \Rightarrow \ddot{\phi} - \phi'' = -\lambda^2\phi\left(\phi^2 - \frac{m^2}{\lambda^2}\right).$$

Note that the term $\frac{1}{2}m^2\phi^2$ in (2.4) cannot be interpreted as a mass term, it has the wrong sign.

For solutions $\phi(x,t)$ of (2.5) the energy is given by

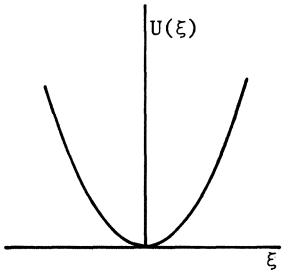
$$(2.6) \qquad H = \int_{-\infty}^{\infty} \mathcal{H}(x,t)dx = \int_{-\infty}^{\infty} (\tfrac{1}{2}\dot{\phi}^2 + \tfrac{1}{2}(\phi')^2 + U(\phi))dx \geqslant 0.$$

The ground state, i.e. the state with lowest energy, is that solution of the equation of motion for which H is minimal. From (2.5) and (2.6) it follows that $H = 0$ for solutions ϕ_0 with $\dot{\phi}_0 = \phi_0' = 0$ and $U(\phi_0) = 0$. This gives $\phi_{0,1} = m/\lambda$ and $\phi_{0,2} = -m/\lambda$. The Goldstone string has two ground states. Notice that $\phi = 0$ has infinite energy.

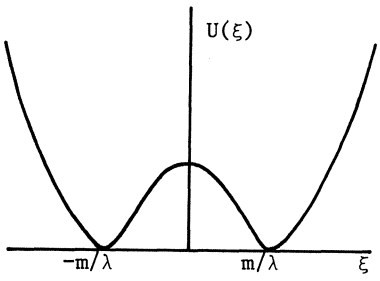
We denote the collection of ground states by

$$(2.7) \qquad V = \{v \in \mathbb{R} \mid U(v) = 0\} = \{\frac{m}{\lambda}, \frac{-m}{\lambda}\}.$$

In figures 2.1 and 2.2 we have drawn the term $U(\xi)$ for the Klein-Gordon string and the Goldstone string.



2.1. Klein-Gordon string



2.2. Goldstone string

For the Goldstone string one can study "small oscillations" around a ground state. Defining $\psi = \phi - m/\lambda$ we have $\psi = 0$ for $\phi = m/\lambda$. In terms of the shifted field ψ the Lagrangian (2.3) obtains the form

$$(2.8) \qquad L = \tfrac{1}{2}\partial_\mu\psi\partial^\mu\psi - \tfrac{1}{2}(2m^2)\psi^2 - m\lambda\psi^3 - \frac{\lambda^2}{4}\psi^4$$

$$= L_0 + L_{int.}$$

"Small oscillations" are given by $L_0 = \tfrac{1}{2}\partial_\mu\psi\partial^\mu\psi - \tfrac{1}{2}\mu^2\psi^2$, a Lagrangian with mass parameter $\mu = m\sqrt{2}$.

We now come to the real problem.

Does the system with $U(\phi)$ given by (2.4) possess non-trivial regular (i.e. non-singular) solutions with *finite* energy?

For solutions of the equation of motion the energy is given by expression (2.6):

$$H = \int_{-\infty}^{\infty} (\tfrac{1}{2}\dot{\phi}^2 + \tfrac{1}{2}\phi'^2 + U(\phi))dx$$

and H is *in*dependent of t.

Finiteness of the energy means that (2.6) must be finite for arbitrary but fixed t. As all terms in (2.6) are nonnegative we obtain as necessary conditions for convergence of the integral

$$\dot{\phi}(x,t) \xrightarrow[|x|\to\infty]{} 0 \qquad \forall t$$

$$(2.9) \qquad \phi'(x,t) \xrightarrow[|x|\to\infty]{} 0$$

$$U(\phi) \xrightarrow[|x|\to\infty]{} 0$$

These conditions imply

$$(2.10)\qquad \begin{cases} \lim\limits_{|x|\to\infty} \phi(x,t) := \phi(\pm\infty), \quad \text{independent of } t \\ U(\phi(\pm\infty)) = 0, \quad \text{so } \phi(\pm\infty) \in V \quad \text{(see 2.7)} \end{cases}$$

So $\phi(x,t)$ must tend to a zero of the potential U for $x \to \infty$ and $x \to -\infty$ for all t.

The possible solutions fall into four disjunct classes which are characterized by a "charge" Q.

$\phi(\infty)$	$\phi(-\infty)$	Q
m/λ	m/λ	0
$-m/\lambda$	$-m/\lambda$	0
$-m/\lambda$	$+m/\lambda$	$-2m/\lambda$
$+m/\lambda$	$-m/\lambda$	$+2m/\lambda$

A solution in one class cannot be deformed continuously into a solution in another class.

The charge Q is defined as follows. Consider

$$J_\mu(x,t) = \varepsilon_{\mu\nu}\partial^\nu\phi(x,t)$$

$$\varepsilon_{00} = \varepsilon_{11} = 0, \quad \varepsilon_{01} = -\varepsilon_{10} = -1$$

J_μ satisfies the local conservation law

$$(2.11)\qquad \partial^\mu J_\mu = 0.$$

The topological charge Q is given by

$$(2.12)\qquad Q = \int_{-\infty}^{\infty} J_0(x,t)dx = \int_{-\infty}^{\infty} \frac{\partial\phi}{\partial x}(x,t)dx = \phi(\infty) - \phi(-\infty).$$

Q is independent of t.

The local conservation law (2.11) has nothing to do with the dynamics of the system. This contrasts the conservation laws obtained by Noether's theorem.

We now look for time-*in*dependent solutions with *finite* energy. The equation of motion reduces to

$$\frac{d^2\phi}{dx^2} = \frac{dU}{d\phi}$$

and the expression for the energy is

$$H = \int_{-\infty}^{\infty} (\tfrac{1}{2}\phi'^2 + U(\phi))dx.$$

To obtain finite energy we must have $\phi(\pm\infty) \in V$. We use the Bogolmony decomposition to rewrite H.

$$H = \tfrac{1}{2}\int_{-\infty}^{\infty} \{(\phi' \pm \sqrt{2U})^2 \pm 2\phi'\sqrt{2U}\}dx \geq \left| \int_{-\infty}^{\infty} \sqrt{2U}\,\frac{d\phi}{dx}\,dx \right| = \int_{\phi(-\infty)}^{\phi(+\infty)} \sqrt{2U}\,d\phi$$

For $\phi(\infty) \neq \phi(-\infty)$ the energy is bounded from below by

$$\int_{\phi(-\infty)}^{\phi(\infty)} \sqrt{2U}\,d\phi = \frac{2}{3}m^3/\lambda^2.$$

This finite value is realized for solutions of the first order equations

$$\frac{d\phi}{dx} = \sqrt{2U} \quad \text{or} \quad \frac{d\phi}{dx} = -\sqrt{2U}\ .$$

The first equation has solutions given by

(2.13) $$\phi_{sol}(x) = m/\lambda \tanh \frac{m}{2}(x-a).$$

This is a non-singular soliton solution with finite energy. From "Lorentz" invariance it follows that

$$\phi_{sol}(x,t) = m/\lambda \tanh \gamma\frac{m}{2}(x-a-\beta t)$$

with $\beta = v/c$ and $\gamma = (1-\beta^2)^{-\frac{1}{2}}$

is a soliton with velocity v.

The energy density for (2.13) is strongly localized,

$$\underset{\text{sol}}{H(x)} = \frac{m^4}{4\lambda^2} \cdot \frac{1}{\cosh^4 \frac{m(x-a)}{2}} .$$

The soliton solution is "topological" stable. It is also a dynamically stable solution. An initially small perturbation $\rho(x,t)$ around $\phi_{sol}(x)$ remains small.

A second example in (1+1) dimensions is the sine-Gordon model with

$$U(\phi) = 1 - \cos b\phi .$$

The degenerate ground states are contained in

$$V = \{v \in \mathbb{R} \mid U(v) = 0\} = \{v = \frac{2\pi}{b} n \mid n \in \mathbb{Z}\} .$$

3. SCALAR FIELD THEORY IN (1+D) DIMENSIONS. DERRICK'S THEOREM

We start with the Lagrangian

$$L = \partial_\mu \phi^* \partial^\mu \phi - U(\phi)$$

with $$\partial_\mu = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2} \cdots \frac{\partial}{\partial x^D}\right) \quad \text{and} \quad U(\phi) \geqslant 0.$$

For later use we have taken a complex valued function ϕ. To obtain the energy for solutions of the equations of motion we must integrate the energy density over D-dimensional space.

$$H[\phi] = \int_{-\infty}^{\infty} \{|\dot{\phi}|^2 + \nabla\phi^* \cdot \nabla\phi + U(\phi)\} d^D x.$$

We now use Derrick's argument to show that there are no time-*in*dependent solutions with finite energy for $D \geqslant 2$.

For a time-*in*dependent solution we have

$$H[\phi] = \int_{-\infty}^{\infty} \nabla\phi^* \cdot \nabla\phi \, d^D x + \int_{-\infty}^{\infty} U(\phi) d^D x = H_1[\phi] + H_2[\phi].$$

Both terms are *non*negative and must be finite. For time-independent solutions $H(\vec{x}) = -L(\vec{x})$ and $L[\phi]$ must be an extremum. So $H[\phi]$ must be an extremum. Consider now the function ϕ_ρ defined by

$$\phi_\rho(\vec{x}) := \phi(\rho\vec{x}), \quad \rho \in \mathbb{R}^+$$

and calculate $H[\phi_\rho]$. One obtains

$$H[\phi_\rho] = \int_{-\infty}^{\infty} \nabla\phi^*(\rho\vec{x}) \cdot \nabla\phi(\rho\vec{x}) d^D x + \int_{-\infty}^{\infty} U(\phi(\rho\vec{x})) d^D x$$

$$= \rho^{2-D} H_1[\phi] + \rho^{-D} H_2[\phi] \, .$$

For $\rho = 1$. H must have an extremum, so

$$\frac{dH[\phi_\rho]}{d\rho} = 0 \quad \text{for} \quad \rho = 1.$$

This gives $(2-D)H_1[\phi] = DH_2[\phi]$ from which it follows that $H_1[\phi] = H_2[\phi] = 0$ for $D > 2$. Hence the time-independent solutions must be constants which are zeros of U. These are the trivial solutions. For $D = 2$ we must have $H_2[\phi] = 0$, i.e. $U(\phi) = 0$. The energy is then given by $H_1 = \int_{-\infty}^{\infty} \nabla\phi^* \cdot \nabla\phi \, d^2x$ where $\phi(\vec{x})$ is a zero of U for all $\vec{x}$. This leads again to ϕ = constant.

To obtain a better feeling for these results we consider the case $D = 2$ with U given by

$$U(\phi) = \frac{\lambda^2}{4}\left(|\phi|^2 - \frac{m^2}{\lambda^2}\right)^2, \quad m,\lambda > 0. \tag{3.1}$$

The expression for the energy is

$$H = \int_{-\infty}^{\infty} (\dot{\phi}^*\dot{\phi} + \nabla\phi^* \cdot \nabla\phi + U(\phi)) d^2x \geqslant 0 .$$

The ground states are given by $\phi(x,y,t)$ = constant, with $|\phi|^2 = m^2/\lambda^2$. We have now a continuum of ground states,

$$V = \{v \in \mathbb{C} \mid U(v) = 0\} = \{\frac{m}{\lambda} e^{i\alpha} \mid \alpha \in \mathbb{R}\} = S^1.$$

Introducing polar coordinates r, θ in the $x-y$ plane we have as necessary conditions for finite energy

$$\begin{aligned} \dot{\phi}(r,\theta,t) &\xrightarrow[r\to\infty]{} 0 \\ \nabla\phi &\xrightarrow[r\to\infty]{} 0 \\ U(\phi) &\xrightarrow[r\to\infty]{} 0 \end{aligned} \quad .$$

These conditions imply (see section 2)

$$\lim_{r\to\infty} \phi(r,\theta,t) = \phi(\infty,\theta) \quad \text{(independent of } t\text{)}$$

$$\phi(\infty,\theta) \in V .$$

Let us now consider a time-independent configuration $\phi(r,\theta)$. The energy is then given by

$$H = \int_{r=0}^{\infty} \int_{\theta=0}^{2\pi} \left\{ \left| \frac{\partial\phi}{\partial r} \right|^2 + \frac{1}{r^2} \left| \frac{\partial\phi}{\partial\theta} \right|^2 + U(\phi) \right\} r\,dr\,d\theta \,. \tag{3.2}$$

For $r \to \infty$ the function $\phi(r,\theta)$ must have a limit $\phi(\infty,\theta) \in V$. We must now consider the θ dependence of $\phi(\infty,\theta)$. For simplicity we assume $\phi(r,\theta)$ to have the property

$$\lim_{r\to\infty} \frac{\partial\phi(r,\theta)}{\partial\theta} = \frac{\partial}{\partial\theta} \lim_{r\to\infty} \phi(r,\theta) = \frac{d}{d\theta}\phi(\infty,\theta).$$

With this in mind we rewrite the middle term in (3.2)

$$\int_{r=0}^{\infty} \int_{\theta=0}^{2\pi} \frac{1}{r^2} \left| \frac{\partial\phi}{\partial\theta} \right|^2 r\,dr\,d\theta = \int_{r=0}^{R} \int_{0}^{2\pi} \left| \frac{\partial\phi}{\partial\theta} \right|^2 d\theta \frac{dr}{r} + \int_{r=R}^{\infty} \int_{0}^{\pi} \left| \frac{\partial\phi}{\partial\theta} \right|^2 d\theta \frac{dr}{r} .$$

Taking R sufficiently large we may, using our assumption, replace $\frac{\partial\phi(r,\theta)}{\partial\theta}$ in the second term on the right by $\frac{d}{d\theta}\phi(\infty,\theta)$ and we see that the integral diverges for $r \to \infty$ *unless* $\frac{d}{d\theta}\phi(\infty,\theta) = 0$. The function $\phi(\infty,\theta)$ can be considered as a mapping from the circle $S^1_\infty(\lim_{r\to\infty}\phi(r,\theta))$ into the set $V = S^1$ of degenerate ground states. (Recall that $\phi(\infty,\theta) \in V$).

These mappings can be divided in homotopy classes. All elements in one class can be transformed into each other by continuous deformations. The homotopy classes are labelled by an integer n, the winding number. The winding number tells us how many times we cover the image circle if θ runs from zero to 2π. A mapping with winding number n can be brought in the standard form

$$\phi_n(\infty,\theta) = \frac{m}{\lambda} e^{in\theta}.$$

From these considerations we learn that in order to obtain finite energy the solution $\phi(r,\theta)$ must for $r \to \infty$ take the limiting value $\phi(\infty,\theta)$ with $\frac{d}{d\theta}\phi(\infty,\theta) = 0$. This means that $\phi(r,\theta)$ has only one boundary value for all directions θ. Invoking the results of section 2 where we have seen that in order to have a soliton solution $\phi(x)$ must have different boundary value for $x \to \infty$ and $x \to -\infty$ we conclude that there are no time-indepenent finite energy solutions in this model. We have also discovered that it is the tangential derivative i.e. $\frac{1}{r}\frac{\partial\phi}{\partial\theta}$ which causes the trouble.

4. A GAUGE MODEL IN (1+2) DIMENSIONS. FLUX QUANTIZATION

In this section we show that the coupling of the complex ϕ field to the electromagnetic field modifies the tangential derivative $\frac{1}{r}\frac{\partial\phi}{\partial\theta}$ in such a way that solitonlike structures may appear. Before doing this we discuss some interesting properties of the model with $U(\phi)$ given by (3.1)
We start again with the Lagrangian

$$(4.1) \qquad L = \partial_\mu\phi^*\partial^\mu\phi - U(\phi); \quad \partial_\mu = \left(\frac{\partial}{\partial t},\frac{\partial}{\partial x},\frac{\partial}{\partial y}\right).$$

$$(4.2) \qquad U(\phi) = \frac{\lambda^2}{4}(|\phi|^2 - m^2/\lambda^2)^2 .$$

As we have seen in the previous section there is a continuum of ground states given by

$$(4.3) \qquad V = \{v\in\mathbb{C} \mid U(v) = 0\} = \{\frac{m}{\lambda}e^{i\alpha} \mid \alpha\in\mathbb{R}\} .$$

This model is an example of a field theory with *spontaneous symmetry breaking*. The Lagrangian and the equations of motion are *invariant* under the phase transformations

$$(4.4) \qquad \phi(x) \longmapsto \phi'(x) = e^{i\omega}\phi(x), \quad \omega\in\mathbb{R}$$

i.e. they are invariant under the unitary group $U(1)$.
Choosing, however, a ground state e.g. $\phi = m/\lambda$ we obtain a new ground state $\phi' = e^{i\omega}\cdot m/\lambda$. The ground states are not invariant under $U(1)$. This is called spontaneous symmetry breaking. Related to this is the appearance of massless bosons (Goldstone bosons). To show this we decompose the complex field in its real and imaginary parts

$$(4.5) \qquad \phi = \phi_1 + i\phi_2$$

Considering now small oscillations around a ground state, say $v = m/\lambda$, we have the shifted field

$$\psi = \phi - m/\lambda$$

with $\psi_1 = \phi_1 - m/\lambda$ and $\psi_2 = \phi_2$.

Substituting this in the Lagrangian (4.1) we obtain

(4.6) $$L = \partial_\mu\psi_1\partial^\mu\psi_1 - m^2\psi_1^2 + \partial_\mu\psi_2\partial^\mu\psi_2 + \text{higher powers of } \psi_1 \text{ and } \psi_2.$$

There is no mass term for the ψ_2 fields. Figure 4.1 shows why this is the case.

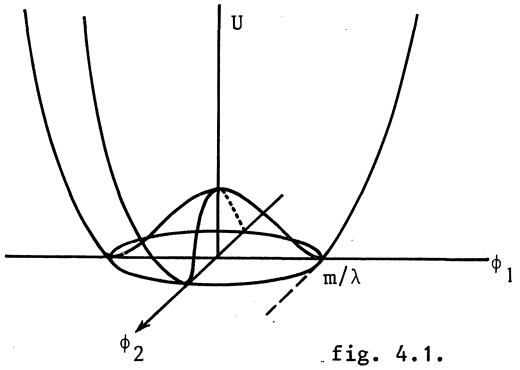


fig. 4.1.

The potential U has the shape of the bottom of a bottle. For small oscillations around m/λ in the ϕ_2 direction, in general tangent to the bottom circle, the force $m\phi^2$ is absent.

We now generalize the model to a field theory with gauge potentials $A_\mu(\vec{x},t)$. The gauge invariant Lagrangian is

(4.7) $$L(A,\phi) = [(\partial_\mu - ieA_\mu)\phi]^* (\partial^\mu - ieA^\mu)\phi - U(\phi) - \tfrac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

$$\partial_\mu = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}\right); \quad U(\phi) = \frac{\lambda^2}{4}(|\phi|^2 - m^2/\lambda^2)^2 .$$

The electromagnetic field tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

The Lagrangian (4.7) is invariant under the local U(1) group, with transformations

$$\phi(x) \longmapsto \phi'(x) = e^{ie\omega(x)}\phi(x) \qquad , \qquad (x = (t,x^1,x^2))$$

(4.8)

$$A_\mu(x) \longmapsto A'_\mu(x) = A_\mu(x) + \partial_\mu\omega(x).$$

The energy for solutions of the equations of motion is given by

$$H = \int\{(D_0\phi)^*(D_0\phi) + (\vec{D}\phi)^* \cdot (\vec{D}\phi) + U(\phi) + \tfrac{1}{2}(\vec{E}^2+B^2)\}dx^1dx^2$$

(4.9) $$\vec{D}\phi = (\nabla - ie\vec{A})\phi; \quad D_0\phi = (\frac{\partial}{\partial t} - ieA_0)\phi$$

$$E_1 = F_{01}, \quad E_2 = F_{02}, \quad B = F_{12}.$$

The ground states are given by

(4.10) $$V = \{v = \frac{m}{\lambda} e^{i\alpha} \mid \alpha \in \mathbb{R}\} \quad \text{and} \quad A_\mu = 0 .$$

We will show that there may be time-independent localized configurations with finite energy. Comparing (4.9) with (3.2) one sees that the gradient terms are modified. Recalling that the tangential derivative $\frac{1}{r}\frac{\partial\phi}{\partial\theta}$ was a source of infinite energy it might be that we obtain finite energy for (4.9). We simplify (4.9) by gauging away the component A_0. For time-*in*dependent configurations we have in that case $\dot{\phi} = 0$ and $\vec{E} = 0$ and the first term in H is no longer present. We get

(4.11) $$H = \int\{\tfrac{1}{2}B^2(\vec{x}) + (\vec{D}\phi)^* \cdot (\vec{D}\phi) + U(\phi)\}dx^1dx^2.$$

The condition of finite energy leads again to restrictions on the various terms in H. Using polar coordinates we have the conditions

$$\phi(r,\theta) \xrightarrow[r\to\infty]{} \phi(\infty,\theta) \in V$$

$$|\vec{D}\phi| = O\left(\frac{1}{r^{1+\varepsilon}}\right) \qquad (\varepsilon>0)$$

$$|B(r,\theta)| = O\left(\frac{1}{r^{1+\varepsilon}}\right)$$

We now consider the term which contains the derivative $\frac{1}{r}\frac{\partial}{\partial\theta}$

$$(4.12)\qquad D_\theta\phi = \frac{1}{r}\frac{\partial\phi(r,\theta)}{\partial\theta} - ieA_\theta\phi(r,\theta)$$

with $\qquad A_\theta = \vec{A}\cdot\vec{1}_\theta$.

Invoking again the assumption that we may interchange the limit $r \to \infty$ and the derivative with respect to θ we have

$$(4.13)\qquad \lim_{r\to\infty} D_\theta\phi = \lim_{r\to\infty}\left(\frac{1}{r}\frac{d\phi(\infty,\theta)}{d\theta} - ieA_\theta(r,\theta)\phi(\infty,\theta)\right).$$

In order that this expression is of order $\frac{1}{r^{1+\varepsilon}}$ for large r the $\frac{1}{r}$ behaviour of the first term must be compensated by the second term. Hence we must have

$$(4.14)\qquad \lim_{r\to\infty} ierA_\theta(r,\theta) = \frac{1}{\phi(\infty,\theta)}\frac{d\phi(\infty,\theta)}{d\theta} .$$

We now see that it is possible that this model has non-trivial configurations with finite energy. The mapping $\phi(\infty,\theta)\colon S^1_\infty \to S^1$ may have a winding number $n \neq 0$ provided the field equations have solutions $\vec{A}(r,\theta)$ and $\phi(r,\theta)$ with $A_\theta(r,\theta) \xrightarrow[r\to\infty]{} \frac{C}{r}$. We do not know explicit solutions. It is, however, possible, starting from an appropriate "Ansatz" to show that there are regular solutions.

Flux quantization.

We define the magnetic flux Φ as follows

$$\Phi = \iint B(\vec{x})dx^1dx^2 = \iint (\mathrm{rot}\,A)\,dx^1dx^2 = \oint \vec{A}\cdot d\vec{s}.$$

For a field satisfying (4.14) we have

$$\Phi = \lim_{r\to\infty} \oint_{\text{circle}} \vec{A}\cdot d\vec{s} = \lim_{r\to\infty} \int_0^{2\pi} A_\theta(r,\theta)rd\theta$$

$$= \frac{1}{ie}\int_0^{2\pi} \frac{1}{\phi(\infty,\theta)}\frac{d\phi(\infty,\theta)}{d\theta}d\theta .$$

Now $\phi(\infty,\theta) = \frac{m}{\lambda} e^{i\sigma(\theta)}$, hence

$$\Phi = \frac{1}{e} \int_0^{2\pi} \frac{d\sigma}{d\theta} d\theta = \frac{2\pi}{e} n$$

with n the winding number.

Taking into account the physical dimensions of the various quantities we obtain

$$(4.15) \qquad \Phi = n\frac{h}{q}, \qquad n \in \mathbb{Z}$$

(h is Planck's constant, q an electric charge).

We have here a remarkable result, the magnetic flux is quantized. As the winding number of a solution cannot change, the flux is a constant. Φ is just as the "charge" Q in section 2 a topological conserved quantity. The time-independency of Q can also be derived in the following way.

Starting with $F^{\mu\nu}$ we define k_μ by

$$k_\mu = \tfrac{1}{2}\varepsilon_{\mu\nu\lambda}F^{\nu\lambda} \qquad (\mu,\nu,\lambda = 0,1,2).$$

We then have $\partial^\mu k_\mu = 0$, independent of the equations of motion. This local conservation law yields the conserved flux

$$\Phi = \iint k_0 d^2x = \iint B d^2x.$$

We conclude this section by some remarks on gauge transformations. According to (4.10) $\phi_0(r,\theta) = \frac{m}{\lambda}$, $\vec{A} = 0$ is a ground state. All ground states have winding number zero. Performing a gauge transformation we obtain a different ground state

$$\begin{cases} \phi' = e^{ie\omega(\vec{x})}\phi_0 \\ \vec{A}' = \nabla\omega(\vec{x}) \end{cases} .$$

Taking now $\omega(\vec{x}) = \frac{1}{e}\theta$ with $tg\theta = \frac{x^2}{x^1}$ we have

$$(4.16) \qquad \begin{cases} \phi'(r,\theta) = \frac{m}{\lambda} e^{i\theta(\vec{x})} \\ \vec{A}'(r,\theta) = \frac{1}{e}\nabla\theta \end{cases}$$

and it looks as if we have now a ground state with winding number $n = 1$. This is, however, not true. The transformation (4.16) is not allowed because $\theta(\vec{x})$ has to make a jump of 2π e.g. along the negative x^1-axis. See figure 4.2

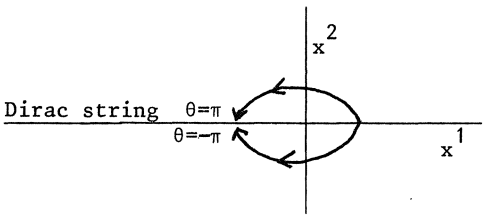


fig. 4.2 Discontinuity of θ.

Taking (4.16) and θ as in figure 4.2 we obtain

$$A'_\theta = \frac{1}{er}\frac{d}{d\theta}(\theta - 2\pi H(\theta-\pi)) = \frac{1}{er}(1-2\pi\delta(\theta-\pi))$$

and $$\oint A'_\theta r d\theta = 0.$$

(H is the Heavyside step function).

5. SOME GENERAL REMARKS

In the preceding sections we have shown that topological solitons occur in field theory for $D = 1$. For $D = 2$ we need gauge fields. In both cases we need a degeneracy of the ground state to obtain different boundary values at spatial infinity.

For $D = 3$ and gauge group $SO(3,\mathbb{R})$ one has the well-known 't Hooft-Polyakov magnetic monopole solution.

In four-dimensional Euclidean space-time and gauge group $SU(N)$ one has the instanton solutions for the pure Yang-Mills field. A lot more can be said about the classification of monopole solutions with special symmetry and about the classification of instantons. This goes far beyond the scope of this introductory talk.

REFERENCES

ATIYAH, M.F., *Geometry of Yang-Mills Fields*, Lezioni Fermiane, Pisa 1979.

COLEMAN, S.R., Erice Lectures (1975).

DERRICK, G.H., J.Math.Phys. 5, 1252 (1964)

't HOOFT, G., Nucl.Phys. B79, 276, (1974).

't HOOFT, G., Banff Lectures (1977).

't HOOFT, G., Scientific American 242, nr.6 (1980).

HUANG, K., *Quarks, Leptons and Gauge fields*. World Scientific, Singapore 1982, ISBN 9971-950-03-0.

NIELSEN, H.B. - OLESON, P., Nucl.Phys. B61, 45 (1973)

GEOMETRICAL METHODS IN QUANTUM FIELD THEORY

by

J.N. BUUR* and E.A. de KERF
Institute for Theoretical Physics
University of Amsterdam

CONTENTS

* present address: KSEPL
Volmerlaan 6,
2288 GD Rijswijk(Z.H.),
The Netherlands.

1. INTRODUCTION

The standard models for electro-weak and colour interactions are all based on relativistic quantum field theory. These models provide accurate descriptions of many physical processes that first appear at high energies. The central quantity in relativistic quantum field theory is the generating functional for Green's functions. These Green's functions are related to transition amplitudes for scattering processes. The generating functional for Green's functions can be expressed as a functional integration over classical fields only. The purpose of this paper is to introduce in a geometrical way the notions necessary to write down such a functional integral. Therefore we discuss the notions of field and Lagrangian. In particular, the Dirac field and the Yang-Mills field are put in their proper settings.

2. JET BUNDLES AND FIELDS

One of the most important concepts in relativistic field theory is the Lagrangian. We first point out why one needs the jet bundle formalism to give a geometrical definition of a Lagrangian. Consider as an example the Lagrangian for a real scalar field $\phi(x)$:

$$L = \partial_\mu \phi \partial^\mu \phi - m^2 \phi^2 \tag{1}$$

This Lagrangian L is a function of the variables ϕ and $\partial_\mu \phi$. We would like to interpret these variables as coordinates on a manifold. This manifold turns out to be a jet bundle.

Let P and F be differentiable manifolds. Define $C(P,F)$ to be the set of maps

$$f: U \to F \tag{2}$$

where U is an arbitrary open subset of P. Two maps $f, g \in C(P,F)$ are said to have contact up to order k in $x \in P$ if: 1) $f(x) = g(x)$; 2) There exist charts (U,ϕ) on P and (V,ψ) on F, with $x \in U$ and $f(x) \in V$, such that $\psi \circ f \circ \phi^{-1}$ and $\psi \circ g \circ \phi^{-1}$ have the same k-th order Taylor expansion. The notion contact up to order k is chart independent and thus introduces an equivalence relation on $C(P,F)$. The k-jet of f in x is defined to be the equivalence class $j^k_x(f)$ that belongs to $f \in C(P,F)$.

If $J^k_x(P,F)$ is the set of k-jets in x, then

$$J^k(P,F) := \bigcup_{x \in P} J^k_x(P,F) \tag{3}$$

is called the k-jet bundle. Using coordinates on P and F, the local coordinates of $j^k_x(f)$ are given by

(4) $$(x^j, y^i(x), \partial_j y^i(x), \ldots, \partial_{j_1 \ldots j_k} y^i(x))$$

with $\psi \circ f \circ \phi^{-1} = y(x)$. Hence it is possible to provide the k-jet bundle with a differentiable structure, [1].

Let $\pi: P \to M$ be a principle fibre bundle with structure group G. Let F be a differentiable manifold.

DEFINITION. A field on P is a map

(5) $$f: P \to F.$$

This rather abstract definition of fields will be clarified in the sequal when the Dirac field and the Yang-Mills field will be discussed.

REMARK. Fields representing leptons and quarks are called matter field. Fields mediating the interactions between leptons and quarks are called gauge fields.

3. THE LAGRANGIAN

Consider a jet bundle $J^k(P,F)$, where P is a principle fibre bundle. Define Q to be a set of fields $f: P \to F$. Sometimes Q is called a configuration space.

DEFINITION. A Lagrangian on Q is a map

$$L: J^k(P,F) \to \mathbb{R} \tag{6}$$

such that for fields $f \in Q$

$$L(j^k_u(f)) = L(j^k_{ug}(f)) \tag{7}$$

where $u \in P$, $g \in G$.

This definition implies that $L(f)$ is constant along fibres, hence we can consider $L(f)$ to be a real valued function on the base manifold M.

This property of the Lagrangian is used in the following definition. Let μ be a volume form on M. Then the action S belonging to the physical system described by the Lagrangian L is defined to be

$$S(f) := \int L(f)\mu \tag{8}$$

where the integration is over an appropriate subset of M.

A simple example may help to clarify things. Let $M = \mathbb{R}^4$ and $G = \{e\}$. Let the configuration space Q be the set of maps $f: M \to \mathbb{R}$. Then the Lagrangian for a real scalar field can be written as

$$L(f) = df \cdot df - m^2 f \cdot f, \tag{9}$$

where the dot stands for the metric on forms induced by the Minkowski metric η on M. For the 1-jet of f in $x \in M$ we have

$$j^1_x(f) = (x, f(x), df(x)). \tag{10}$$

Hence it is clear that this Lagrangian is nothing but a function on the 1-jet bundle $J^1(M,\mathbb{R})$.

REMARK. In (9) the Lagrangian is defined in a coordinate free way. Only for the 1-jet bundle it is possible to give a coordinate free definition of the 1-jet as done in (10).

4. GAUGE TRANSFORMATIONS. GAUGE INVARIANT LAGRANGIANS

Let $\pi: P \to M$ be a principle fibre bundle with structure group G.

DEFINITION. A gauge transformation is a diffeomorphism $y: P \to P$ such that

1. $y(ug) = y(u)g$;
2. $\pi \circ y = \pi$

for all $u \in P$, $g \in G$.

REMARK. Note that gauge transformations leave fibres invariant.

The set of gauge transformations provided with the composition law of mappings forms a group: the gauge group Y.

Given a field $f \in Q$ and a gauge transformation $y \in Y$ the composition $f \circ y$ is clearly also a field. Hence the gauge transformations introduce an equivalence relation on the configuration space Q. In physics the assumption is made that the gauge equivalent fields f and $f \circ y$ represent the same physical configuration. Thus the space of physical configurations is defined to be the quotient space

$$\Phi := Q/Y. \tag{11}$$

Moreover in physics only Lagrangians are considered that are constant on equivalence classes. Hence in the sequel it is assumed that Lagrangians are gauge invariant:

$$L(f) = L(f \circ y). \tag{12}$$

In sections 2,3,4 we have discussed all the basic notions necessary to introduce the generating functional for Green's functions, which will be the subject of the next section.

5. THE GENERATING FUNCTIONAL: path integral approach

Let L be a Lagrangian on a configuration space Q, and Φ the space of physical configurations. Assume that, up to technicalities, $\Pi: Q \to \Phi$ is a principle fibre bundle with structure group Y, where Q, Φ and Y are Banach manifolds, [2]. Let Σ be a section of $\Pi: Q \to \Phi$:

$$\Sigma: \Phi \to Q , \quad \Pi \circ \Sigma = id_{\Phi}. \tag{13}$$

In physics such a section is called a gauge fixing condition, or a choice of gauge. Usually it is formulated in a somewhat different way. To a section Σ belongs a functional $H: Q \to \mathbb{C}$ defined by

$$H(f) = 0 \leftrightarrow f \in \Sigma \tag{14}$$

and H is used in stead of Σ.

REMARK. Consider the Maxwell theory of the electromagnetic field. Then the Lorentz gauge for the 4-potential A^{μ}:

$$\partial_{\mu} A^{\mu} = 0 \tag{15}$$

is an example of a gauge fixing condition given by a functional.

Assume there exists a metric on F, denoted with a dot, and a configuration space Q, such that $f_1 \cdot f_2$ is constant along fibres of P for all fields $f_1, f_2 \in Q$. Let μ be a volume form on M. Then there exists a metric on Q defined by

$$(f_1, f_2) := \int f_1 \cdot f_2 \mu \tag{16}$$

where the integration is over an appropriate subset of M. Note that this metric on Q is gauge invariant:

$$(f_1, f_2) = (f_1 \circ y, f_2 \circ y) \tag{17}$$

where $y \in Y$.

Consider the following functional integration over configurations $f \in Q$:

$$Z(j) := N \int \delta(H(f)) \det(\Sigma_*) \exp\{i S(f) + i(f,j)\} Df \tag{18}$$

with the "source" $j \in Q$ arbitrary but fixed.

The δ-distribution in the integrand is meant to take from each equivalence class of Q only one member. Without the δ-distribution the functional integral would be meaningless from the very beginning, because the contribution from the subset of fields $\{f \circ Y \mid y \in Y\}$ would give a factor proportional to the "volume" of the gauge group Y and this is in general infinite. Then the so-called Faddeev-Popov jacobian determinant $\det(\Sigma_*)$ ensures that under a change of section $\Sigma \to \Sigma'$ the integral is invariant. The exponent consists of the action S and a source term. The integration measure Df is assumed to be gauge invariant: $Df = D(f \circ y)$.

Finally, the normalisation constant N is such that $Z(0) = 1$.

The map $Z: j \in Q \longmapsto Z(j) \in \mathbb{C}$ is called the generating functional for Green's functions.

REMARK. The n-point Green's function is defined as

$$\left. \frac{\delta^n Z(j)}{\delta j^n} \right|_{j=0}.$$

This cursory discussion of the generating functional circumvents at least two difficulties.

1) A rigorous definition of functional integration is not given. We refer to the literature where functional integrals are also known as path integrals, [3]. For an interesting discussion of the path integral applied in quantum mechanics, see Dr. P.M.J. Bongaarts in this volume.

2) In general a choice of gauge, or section Σ of Q, does not exist globally. This problem is known as the Gribov ambiguity, [4].

In spite of these difficulties it is possible to compute the generating functional in a perturbative way. Key words in such a procedure are Feynman diagrams and renormalisation theory, [5].

The generating functional is important in quantum field theory, because all quantities of physical interest can be derived from it. The point is that once a Lagrangian has been specified the computation of the generating functional goes along more or less known paths. The physicist is thus faced with the problem to suggest the Lagrangian that describes best his experiments. To achieve this two important building blocks are used: the Dirac field and the Yang-Mills field.

In what follows these fields and their corresponding Lagrangians will be discussed. The construction of models for electro-weak and strong interactions based on these Lagrangians can be found in the literature, e.g. [6].

6. THE DIRAC FIELD: an intermezzo

We briefly review the Dirac equation:

$$(19) \qquad (i\gamma^{\mu}\partial_{\mu}-m)\psi(x) = 0,$$

with $\psi(x) \in \mathbb{C}^4$. The Dirac matrices $\gamma^{\mu} \in GL(4,\mathbb{C})$ satisfy the anti-commutation relations

$$(20) \qquad \gamma^{\mu}\gamma^{\nu} + \gamma^{\nu}\gamma^{\mu} = 2\eta^{\mu\nu},$$

where $\eta^{\mu\nu} := \mathrm{diag}(1,-1,-1,-1)$ is the Minkowski metric. A useful realization of these Dirac matrices is given by the "chiral" representation:

$$\gamma^{\mu} = \begin{pmatrix} 0 & -\sigma_{\mu} \\ -\sigma^{\mu} & 0 \end{pmatrix}$$

with $\sigma^{\mu} = (\mathbb{1},\sigma^{i})$, where $\{\sigma^{i}\}$ are the Pauli matrices

$$\sigma^{1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^{2} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^{3} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The raising and lowering of indices is done with respect to $\eta^{\mu\nu}$.

Consider the proper orthochronous Lorentz group $L_{+}^{\uparrow}$. The covering map $\Lambda: SL(2,\mathbb{C}) \to L_{+}^{\uparrow}$ is defined as follows:

$$(22) \qquad \Lambda: A \longmapsto \Lambda(A)^{\mu}{}_{\nu} := \tfrac{1}{2}\mathrm{tr}(\sigma^{\mu}A\sigma_{\nu}A^{\dagger}).$$

Note that this homomorphism is such that $\Lambda(A) = \Lambda(-A)$. We now explain the meaning of relativistic invariance. Under a Lorentz transformation $\Lambda(A): x \longmapsto \hat{x}$ the function $\psi(x)$ transforms as follows:

$$(23) \qquad \psi(x) \longmapsto \hat{\psi}(\hat{x}) = \rho(A)\psi(x).$$

If $\psi(x)$ is a solution of the Dirac equation then $\hat{\psi}(x) = \rho(\Lambda(A))\psi(\Lambda^{-1}(x))$ is also a solution. The matrix $\rho(A) \in GL(4,\mathbb{C})$ is a reducible representa-

tion of $SL(2,\mathbb{C})$, which (in the "chiral" representation) reads

$$\rho(A) = \begin{pmatrix} A & 0 \\ 0 & \bar{A} \end{pmatrix} \tag{24}$$

with $\bar{A} := \sigma^2 A^* \sigma^2$, where the asterix denotes complex conjugation.

The irreducible representations of $SL(2,\mathbb{C})$ are labelled by pairs (s,s') with $s, s' = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$. Representations $(s,0)$ and $(0,s)$ with s half integer are called spinor representations. The representation ρ is the direct sum $(\frac{1}{2},0) + (0,\frac{1}{2})$. Note that the representations $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$ are inequivalent.

The covering map Λ restricted to the subgroup $SU(2)$ is a covering map of $SO(3)$ (as a subgroup of $L_+^\uparrow$). A representation (s,s') of $SL(2,\mathbb{C})$ restricted to $SU(2)$ decomposes into a direct sum of irreducible representations of $SU(2)$, such that $2(s+s') + 1$ is the maximal dimension of the representation spaces that occur in this sum. The number $s + s'$ determines the spin of the representation (s,s'). For an interpretation of spin as intrinsic angular momentum in quantum mechanics, see: [7]. Our purpose is to give the geometrical setting for the Dirac field. To achieve this we need the notions of frame bundle and spinor structure.

7. THE FRAME BUNDLE

Let M be a manifold: $\dim M = n$. A linear frame v in $x \in M$ is defined to be an ordered basis for the tangent space $T_x(M)$, with $X_i \in T_x(M)$ for $i = 1...n$:

$$v = (X_1,...,X_n). \tag{25}$$

If $L_x(M)$ is the set of linear frames in x, then

$$L(M) := \bigcup_{x\in M} L_x(M) \tag{26}$$

is called the frame bundle. Using any chart $(U,x^1...x^n)$ on M, a linear frame v is characterized by the local coordinates:

$$(x^j, A_i^k) \tag{27}$$

with $X_i = \sum_k A_i^k \partial_k$ and $(A_i^k) \in GL(n,\mathbb{R})$.

Recall that $GL(n,\mathbb{R})$ has a natural differentiable structure as an open subset of $\mathbb{R}^{n^2}$. Hence it is possible to provide the frame bundle with a differentiable structure. With the projection $\pi: L_x(M) \to x$ and the right action of $GL(n,\mathbb{R})$ on $L(M)$, defined by:

$$(v,g) \longmapsto vg := (\sum_i X_i g_k^i) \tag{28}$$

with $g = (g_k^i) \in GL(n,\mathbb{R})$, the frame bundle is a principle fibre bundle $\pi: L(M) \to M$ with structure group $GL(n,\mathbb{R})$, [10].

REMARK. Note that a frame bundle is uniquely determined by its base manifold alone.

The tangent bundle $T(M)$ can be seen as a vector bundle associated to $L(M)$ with typical fibre $\mathbb{R}^n$, because to each linear frame $v \in L(M)$ there corresponds a linear isomorphism $v: \mathbb{R}^n \to T_x(M)$ defined through

(29) $\quad v\colon \xi \longmapsto \sum_i X_i \xi^i$

where $\quad \xi = (\xi^i) \in \mathbb{R}^n$.

A vector field X on M, i.e. a section of $T(M)$, can be identified with a function $f\colon L(M) \to \mathbb{R}^n$ of type id: $f(vg) = g^{-1}f(v)$ for all $g \in GL(n,\mathbb{R})$, as follows:

(30) $\quad X_x = vf(v)$.

Note that such a function f is a field (in the sense of section 2) on $L(M)$ with values in $\mathbb{R}^n$.

The orthonormal frame bundle.

Consider a manifold M provided with a metric η of signature (a,b). Denote the natural metric on $\mathbb{R}^n$ with signature (a,b) by η_{ij}. Define $F(M)$ to be the set of linear frames $u \in L(M)$ such that

(31) $\quad \eta(X_i,X_j) = \eta_{ij}$

where $v = (X_i)$. Then $F(M)$ is a principle fibre bundle $\pi\colon F(M) \to M$ with structure group $O(a,b)$. $F(M)$ is called the orthonormal frame bundle. The orthonormal frame bundle $F(M)$ is a subbundle of the frame bundle $L(M)$.

REMARK. In physics $\dim M = 4$ and the metric η has signature $(1,3)$. Then the structure group of the orthonormal frame bundle $F(M)$ has structure group $O(1,3)$: the full Lorentz group. Moreover, the orthonormal frame bundle $F(M)$ can be considered in more physical terms as the set of all local Lorentz frames.

8. SPINOR STRUCTURES

Consider the frame bundle $L(M)$ over a manifold M. Let $\dim M = 4$, and η be a metric on M of signature $(1,3)$. A fibre of the orthonormal frame bundle $F(M)$ is isomorphic to the full Lorentz group $O(1,3)$ and therefore consists of four disjunct components. If $F(M)$ also consists of four components, a choice of one of these four, denoted by $F_0(M)$, is called a time and space orientation. Clearly $\pi_F: F_0(M) \to M$ is a principle fibre bundle with structure group $L_+^\uparrow$: the proper orthochronous Lorentz group.

Let M be time and space orientable. Let $\pi_S: S(M) \to M$ be a principle fibre bundle with structure group $SL(2,\mathbb{C})$.

DEFINITION. A spinor structure is a map $\lambda: S(M) \to F_0(M)$ such that

1. $\pi_F \circ \lambda = \pi_S$
2. $\lambda \circ R_A = R_{\Lambda(A)} \circ \lambda$

for all $A \in SL(2,\mathbb{C})$.

The covering map $\Lambda: SL(2,\mathbb{C}) \to L_+^\uparrow$ is defined in section 6, formula (22).

REMARK. Let M be non-compact. Then a spinor structure on M exists if and only if $F_0(M)$ is trivializable. The proof of this assertion is not elementary: [8]. If a spinor structure exists it may not be unique: the number of non-equivalent spinor structures equals $\dim H^1(M,Z_2)$: [9].

9. DIRAC FIELD

Let $\lambda: S(M) \to F_0(M)$ be a spinor structure, with respect to a metric η. Consider the $(\frac{1}{2},0) + (0,\frac{1}{2})$ representation ρ of $SL(2,\mathbb{C})$ on $GL(4,\mathbb{C})$: $\rho(A) = \mathrm{diag}(A,\bar{A})$, which we encountered in section 6.

DEFINITION. A Dirac field is a function $f: S(M) \to \mathbb{C}^4$ such that

$$f(uA) = \rho(A^{-1})f(u) \tag{32}$$

for all $u \in S(M)$, $A \in SL(2,\mathbb{C})$.

Vierbein fields.

Let $\sigma: U \to S(M)$ be a section. Then the spinor structure induces a section $\lambda\circ\sigma: U \to F_0(M)$. The vector fields $\{\varepsilon_\mu\}$, $\mu = 0 \ldots 3$ defined through

$$(\varepsilon_\mu) := \lambda\circ\sigma(x), \quad x \in U \tag{33}$$

span an orthonormal ordered basis in $T_x(M)$ with respect to η, and are called vierbein fields.

Define $\psi := \sigma^* f: U \to \mathbb{C}^4$, then under a change of section $\sigma \to R_A\sigma$ the function ψ transforms as $\psi \to \rho(A^{-1})\psi$. The vector fields transform as

$$\varepsilon_\mu \to \Lambda(A)_\mu{}^\nu \varepsilon_\nu \tag{34}$$

which is nothing but a local Lorentz transformation. Clearly ψ has the transformation character of a Dirac spinor.

To define a Lagrangian for the Dirac field a connection on $S(M)$ is required. Thereto the spinor structure is used and a connection ω on $F_0(M)$. It is not difficult to see from the properties of a spinor structure that when ω is a connection on $F_0(M)$

$$\tilde{\omega} := \Lambda_*^{-1} \lambda^* \omega \tag{35}$$

is a connection on $S(M)$

Note that the representation ρ is orthogonal with respect to the metric on $\mathbb{C}^4$ defined through:

$$\xi \cdot \zeta := \tfrac{1}{2}\mathrm{Re}(\xi^{\dagger}\gamma^0\zeta), \tag{36}$$

where $\xi, \zeta \in \mathbb{C}^4$. Finally, let $\{e_\mu\}$, $\mu = 0 \ldots 3$ denote the natural basis of $\mathbb{R}^4$, and let D be the exterior covariant derivative on $S(M)$ with respect to $\tilde{\omega}$. On functions the exterior covariant reads $Df = df + \rho_*\tilde{\omega}f$. Then the Dirac Lagrangian can be defined.

THEOREM. *The map* $L_D: J^1(S(M), \mathbb{C}^4) \to \mathbb{R}$ *defined via*

$$L_D(f) := (i\gamma^\mu Df(E_\mu) - mf) \cdot f \tag{37}$$

is a Lagrangian on the set of Dirac fields Q, *where* E_μ *is an arbitrary vector from the linear space*

$$\pi_{S_*}^{-1}(\lambda(u)e_\mu) \subset T_u(S(M)) \tag{38}$$

with $x \in M$, $\pi_S(u) = x$ *and* $m \in \mathbb{R}$.

PROOF. 1) For L_D to be constant along fibres it is sufficient to check that

$$\gamma^\mu Df(E'_\mu) = \rho(A^{-1})\gamma^\mu Df(E_\mu) \tag{39}$$

where

$$E'_\mu \in \pi_S^{-1}(\lambda(uA)e_\mu) \subset T_{uA}(SM)) \tag{40}$$

and $A \in SL(2,\mathbb{C})$, because the representation ρ is orthogonal. Note that Df is of type ρ. Use one of the defining properties of the spinor structure λ to establish:

(41) $\lambda(uA)e_\mu = \Lambda(A)_\mu{}^\nu \lambda(u)e_\nu$.

Hence it follows that

(42) $\gamma^\mu Df(E'_\mu) = \gamma^\mu \rho(A^{-1})\Lambda(A)_\mu{}^\nu Df(E_\nu)$,

because

(43) $R_{A^{-1}}{}^*(E'_\mu) = \Lambda(A)_\mu{}^\nu E_\nu$ + vertical vector.

Finally, use one of the properties of the Dirac matrices:

(44) $\gamma^\mu \rho(A^{-1})\Lambda(A)_\mu{}^\nu = \rho(A^{-1})\gamma^\nu$.

2) Invariance under gauge transformations is ensured because Df is of type ρ.

REMARK. In high energy physics space-time is taken to be such that: $M = \mathbb{R}^4$, η the Minkowski metric and ω the canonical flat connection on $L_0(M) = M \times L_+^\uparrow$. Then there is a natural choice for the spinor structure: $S(M) = M \times SL(2,\mathbb{C})$ and λ is defined as $\lambda(x,A) = (x,\Lambda(A))$. Let (U,ϕ) be a chart on M such that $\eta = \eta^{\mu\nu}\partial_\mu\partial_\nu$. Then there exists a section (unique up to a sign) $\sigma: U \to S(M)$ such that $\lambda\circ\sigma = (\partial_\mu)$. With $\psi := f\circ\sigma$ the coordinate expression of the Dirac Lagrangian reads

(45) $L_D = (i\gamma^\mu\partial_\mu\psi - m\psi)\cdot\psi$.

The Dirac Lagrangian is used to describe so-called spin-$\frac{1}{2}$ particles. All elementary constituents of matter, leptons and quarks, are spin-$\frac{1}{2}$ particles.

10. YANG-MILLS FIELD

Let $P(M,G)$ be a principle fibre bundle. To each element A of the Lie algebra G' corresponds a fundamental vector field A^* on P. Define F_u to be the set of linear maps $T_u(P) \to G'$ such that $A^*_u \longmapsto A$ for all $A \in G'$. Then the set

$$F := \bigcup_{u \in P} F_u \tag{46}$$

can be given a differentiable structure that makes F into a vector bundle over P with projection $\pi_F: F \to P$. The representation ρ of G on F is defined through

$$\rho(g)\xi = \mathrm{Ad}(g^{-1})\xi \circ R_{g^{-1}*} \tag{47}$$

where $g \in G$, $\xi \in F$ and Ad the adjoint representation.

Note that if $\xi \in F_u$ then $\rho(g)\xi \in F_{ug}$.

DEFINITION. A Yang-Mills field is a function $f: P \to F$ such that:

1. $\pi_F \circ f = \mathrm{id}_P$
2. $f(ug) = \rho(g^{-1})f(u)$

for all $u \in P$, $g \in G$.

REMARK. A Yang-Mills field can be regarded as a connection form on P.

Let η be a metric on M and denote the trace on the Lie algebra G' with tr. Then η and tr induce a metric on the space of horizontal forms of type Ad: $[10]$. Denote this metric with a dot. Then the Yang-Mills Lagrangian can be defined.

THEOREM. *The map* $L_{YM}: J^1(P,F) \to \mathbb{R}$ *defined through*

$$L_{YM}(f) := \Omega \cdot \Omega \tag{48}$$

is a Lagrangian on the set of Yang-Mills fields Q, *where* Ω *is the curvature form:*

$$\Omega := df + [f \wedge f]. \tag{49}$$

PROOF. The map L_{YM} is constant along fibres as well as invariant under gauge transformations, because Ω is a horizontal form of type Ad.

REMARK. If μ is the volume form associated with the metric η then the Yang-Mills action can be written as

$$S = \int \Omega \cdot \Omega \mu = \int \mathrm{tr}(*\Omega \wedge \Omega) \,. \tag{50}$$

REMARK. Let η be the Minkowski metric, (U,ϕ) a chart on M such that $\eta = \eta^{\mu\nu}\partial_\mu \partial_\nu$, and $\sigma: U \to P$ a section. With the notation $\sigma_* \Omega \equiv \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu$ the coordinate expression of the Yang-Mills Lagrangian reads

$$L_{YM} = \tfrac{1}{4}\mathrm{tr}(F_{\mu\nu}F^{\mu\nu}). \tag{51}$$

The Yang-Mills Lagrangian is used to describe the spin-1 particles that mediate gauge interactions. All interactions encountered in high energy physics are gauge interactions. In the electro-weak interaction the structure group G is $SU(2) \times U(1)$. The colour interaction has structure group $SU(3)$.

11. COMMENTS

With the geometrical description of fields and Lagrangians the use of differential geometry in relativistic quantum field theory has not come to an end.

Non-trivial field configurations that are important in the generating functional can be best understood in geometrical terms: e.g. think of instanton solutions. In addition the geometrical description gives the physicist a clue how to define the various fields and Lagrangians when gravitation is present.

LITERATURE

[1] PIRANI, F.E.A., e.a.: *Local jet bundle formulation of Bäcklund transformations*. Reidel, 1974.

[2] DANIEL, M. - VIALLET, C.M.: Rev. of Modern Physics 52(1980) 175 "The geometrical setting of gauge theories of the Yang-Mills type".

[3] FEYNMAN, R.P. - HIBBES, A.R.: *Quantum mechanics and path integrals*. McGraw-Hill, New York, 1965.

MARINOV, M.S.: Physics Reports 60(1980) 1 "Path integrals in quantum theory: an outlook of basic concepts".

GLIMM, J. - JAFFE, A.: *Quantum physics: a functional point of view*. Springer, New York, 1981.

[4] GRIBOV, V.N.: Nuclear Physics B139(1978) 1 "Quantization of non-Abelian gauge theories".

SINGER, I.M.: Comm.Math.Phys. 60(1978) 7 "Some remarks on the Gribov ambiguity".

[5] FADDEEV, L.D. - SLAVNOV, A.A.: *Gauge fields: introduction to quantum theory*. Benjamin, London, 1980.

ITZYKSON, C. - ZUBER, J.: *Quantum field theory*. McGraw-Hill, New York, 1980.

RAMOND, P.: *Field theory: a modern primer*. Benjamin, London, 1981.

[6] ABERS, E.S. - LEE, B.J.: Physics Reports 9C(1973) 1 "Gauge theories".

MARCIANO, W. - PAGELS, H.: Physics Reports 36C(1978) 138 "Quantum Chromodynamics".

[7] KERF, E.A. de: Representations of locally compact groups with applications. Part 1. M.C. Syllabus 38.1. T.H. Koornwinder(ed.), Mathematical Centrum, Amsterdam, 1979.

[8] GEROCH, R.: Jour.Math.Phys. 9(1968) 739 "Spinor structure of space-time in General Relativity I".

[9] ISHAM, C.: Proc.Roy.Soc. A364(1978) 591 "Spinor fields in four dimensional space time".

GENERAL REFERENCES

[10] KOBAYASHI, S. - NOMIZU, K., *Foundations of differential geometry I.* Interscience, New York, 1963.

BLEEKER, D., *Gauge theory and variational principle.* Addison-Wesley, London, 1981.

CHOQUET-BRUHAT, Y. e.a., *Analysis, manifolds and physics.* North-Holland, Amsterdam, 1982: revised edition.

BUUR, J.N., *Differential geometry and classical field theory.* Doctoraalscriptie, Amsterdam, December 1982.

SELF-DUAL YANG-MILLS EQUATIONS

by

H.G.J. PIJLS
Mathematical Institute
University of Amsterdam

1. THE YANG-MILLS EQUATIONS

(1.1) We consider a 4-dimensional oriented (pseudo)-Riemannian manifold M. Let $\pi: P \to M$ be a principal G-bundle.

Let $\mathcal{C}(P)$ be the space of the connection forms ω for which the Yang-Mills functional (the action)

$$L(\omega) := -\int \mathrm{tr}(\Omega \wedge *\Omega)$$

is finite (Ω is the curvature of ω).

The stationary points ω of L satisfy the Yang-Mills equation

$$D(*\Omega) = 0$$

where D is the covariant derivative associated with ω. Since $D\Omega = 0$ (identity of Bianchi) we see that ω satisfies the Yang-Mills equations if $\Omega = \pm *\Omega$. The equation

$$\Omega = *\Omega$$

is called the *self-dual Yang-Mills equation*.

(1.2) In physics one considers the case that $M = \mathbb{R}^4$ with Euclidean metric and P is the trivial $SU(2)$-bundle. If $s: \mathbb{R}^4 \to P$ is a global section, then we can write

$$s^*\omega = \sum_{\mu=0}^{3} A_\mu \, dx^\mu,$$

$$s^*\Omega = \sum_{\mu<\nu} F_{\mu\nu} \, dx^\mu \wedge dx^\nu .$$

The A_μ are called gauge potentials and the $F_{\mu\nu}$ gauge fields. We have

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] .$$

The Euclidean Yang-Mills equations can be written as

$$\sum_\mu \partial_\mu F_{\mu\nu} + [A_\mu, F_{\mu\nu}] = 0,$$

where $F_{\mu\nu} = -F_{\nu\mu}$.

The Euclidean self-dual Yang-Mills equations are

$$F_{01} = F_{23}, \quad F_{02} = -F_{13}, \quad F_{03} = F_{12}.$$

REMARK. $A_\mu(x)$ and $F_{\mu\nu}(x)$ belong to the Lie algebra of $SU(2)$. If we write $A_\mu = i\tilde{A}_\mu$ and $F_{\mu\nu} = i\tilde{F}_{\mu\nu}$, then $\tilde{A}_\mu$ and $\tilde{F}_{\mu\nu}$ are traceless hermitian 2×2-matrices and

$$\tilde{F}_{\mu\nu} = \partial_\mu \tilde{A}_\nu - \partial_\nu A_\mu + i[\tilde{A}_\mu, \tilde{A}_\nu] .$$

If we apply a gauge transformation g, then the gauge potentials A_μ are transformed into gauge potentials $B_\mu = i\tilde{B}_\mu$ with

$$B_\mu = g^{-1} A_\mu g + g^{-1} \partial_\mu g$$

or

$$\tilde{B}_\mu = g^{-1} \tilde{A}_\mu g - i \, \partial g^{-1} \partial_\mu g .$$

(1.3) Given a solution of the Euclidean Yang-Mills equations on $\mathbb{R}^4$ which has finite action, then it can be extended to a solution on some principal $SU(2)$-bundle on S^4. This fact is a consequence of the following theorem.

THEOREM (K. Uhlenbeck). *Let* (A_μ) *be a solution of the Euclidean Yang-Mills equations on* $\mathbb{R}^4$ *with finite action. If* (A_μ) *has an isolated singularity, then this singularity can be removed by applying a gauge transformation.*

Precisely: if (A_μ) *has an isolated singularity in* p, *then there is a ball* B *with centre* p *and a gauge transformation defined on* $B \setminus \{p\}$ *so that the transformed potentials* $(\hat{A}_\mu)$ *can be extended to* B.

Hence a solution of the Euclidean Yang-Mills equations on $\mathbb{R}^4$ can be considered as a connection on some principal SU(2)-bundle P on S^4.

THEOREM. *The self-dual solutions minimise the Yang-Mills functional on* C(P).

Let k be the Chern number of the bundle P. Then self-dual solutions with finite action exist for all $k \geqslant 0$; they are called k-*instantons*.

(1.4) Also of interest are the so-called monopole solutions.

DEFINITION. (A_μ) is called a monopole if

(i) $\partial_0 A_\mu = 0$,

(ii) (A_μ) satisfies the self-dual Euclidean Yang-Mills equations,

(iii) (A_μ) has finite energy.

2. THE PENROSE TRANSFORM

(2.1) If we complexify S^4 we obtain Q_4 (Klein's quadric). The equation of Q_4 in homogeneous coordinates on $P_5(\mathbb{C})$ is

$$(1) \qquad \sum_{i=1}^{5} z_i^2 = z_6^2 .$$

The quadric Q_4 can be considered as the Grassmann variety of all lines in $P_3(\mathbb{C})$: the line through x and y in $P_3(\mathbb{C})$ has Plücker coordinates

$$(2) \qquad p_{\alpha\beta} := x_\alpha y_\beta - x_\beta y_\alpha$$

and these satisfy

$$p_{01}p_{23} - p_{02}p_{13} + p_{03}p_{12} = 0.$$

Using a suitable coordinate transformation one can transform (2) into (1).

(2.2) Consider the complexification $\mathbb{C}^4$ of $\mathbb{R}^4$ with standard coordinates (x^μ) and define new coordinates x^{PQ} $(P,Q = 1,2)$ by

$$\begin{pmatrix} x^{11} & x^{12} \\ x^{21} & x^{22} \end{pmatrix} = \begin{pmatrix} x^0 - ix^3 & -ix^1 - x^2 \\ -ix^1 + x^2 & x^0 + ix^3 \end{pmatrix}$$

Let $Z^\alpha = (\omega^P, \pi_Q)$ be complex coordinates on "twistor space" $\mathbb{C}^4$, so $Z^0 = \omega^1$, $Z^1 = \omega^2$, $Z^2 = \pi_1$, $Z^3 = \pi_2$. A point (x^{PQ}) determines a line in projective twistor space $P_3(\mathbb{C})$ by

$$\omega^P = x^{PQ}\pi_Q .$$

This equation expresses the relationship between Z^α and (x^μ).

The "reality" structure on $P_3(\mathbb{C})$ is given by

$$\sigma: Z^\alpha \longmapsto (Z^*)^\alpha := (\overline{Z^1}, -\overline{Z^0}, \overline{Z^3}, -\overline{Z^2}).$$

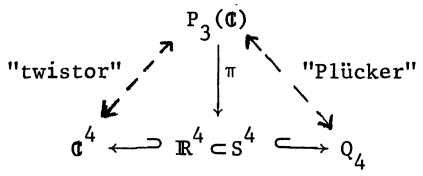
So

$$\omega^P \longmapsto (\omega^*)^P = (\overline{\omega^2}, -\overline{\omega^1})$$

$$\pi^P \longmapsto (\pi^*)^P = (\overline{\pi^2}, -\overline{\pi^1}).$$

A line in $P_3(\mathbb{C})$ is called real if $\sigma(\ell) = \ell$. The real lines correspond exactly with the points in $\mathbb{R}^4$.

(2.3) We have now the following situation:



The map $\pi: P_3(\mathbb{C}) \to S^4$ is a fibre bundle.

If one consider some problem on $\mathbb{R}^4$ or S^4, then the procedure is as follows:

- first complexify
- go to $P_3(\mathbb{C})$.

So a problem on $\mathbb{R}^4$ is translated into another problem on $P_3(\mathbb{C})$. This process is called the *Penrose transformation*.

3. THE CONSTRUCTION OF WARD

(3.1) In order to motivate the construction we make some introductory remarks (cf. [1]).

Consider gauge potentials (A_μ) (associated with a connection ω) on $\mathbb{C}^4$ with self-dual field $(F_{\mu\nu})$ (corresponding to the curvature Ω of ω)

With a point $\theta = (\omega,\pi) \in P_3(\mathbb{C})$ we can associate an α-plane α_θ (anti-self-dual plane) in $\mathbb{C}^4$. If v,w are tangent to α_θ, then $\Omega(v,w) = 0$ because $v \wedge w = -*(v \wedge w)$ and so

$$\begin{aligned}\Omega(v,w) &= \langle \Omega, v \wedge w \rangle = -\langle \Omega, *(v \wedge w) \rangle \\ &= -\langle *\Omega, v \wedge w \rangle = -\langle \Omega, v \wedge w \rangle \\ &= -\Omega(v,w).\end{aligned}$$

So α_θ is an integral manifold of $\Omega = 0$. This implies that for a given α-plane α_θ the following equations are integrable:

$$(3) \qquad \begin{cases} A_\mu \, dx^\mu = -i \, g^{-1} \partial_\mu g \, dx^\mu \\ (dx)\pi = 0. \end{cases}$$

The parallel transport of a 2-spinor from x to y $(x,y \in \alpha_\theta)$ is independent of the chosen path in α_θ. If the parallel transport is denoted by $g_\theta(x,y)$, then g_θ satisfies equation (3). We introduce the vector space V_θ of 2-spinor fields ψ_θ defined over α_θ which satisfy

$$\psi_\theta(x) = g_\theta(x,y)\psi_\theta(y) \qquad (x,y \in \alpha_\theta).$$

V_θ is a two-dimensional vector space. The family of spaces V_θ forms a two-dimensional analytic vector bundle over $P_3(\mathbb{C})$. If $\theta = (\omega^1,\omega^2,\pi^1,\pi^2)$, then we define

$$x_\theta^1 = \begin{pmatrix} \omega_1/\pi_1 & 0 \\ \omega_2/\pi_2 & 0 \end{pmatrix} \in \alpha_\theta \qquad \text{if} \qquad \pi_1 \neq 0,$$

$$x_\theta^2 = \begin{pmatrix} 0 & \omega_1/\pi_2 \\ 0 & \omega_1/\pi_2 \end{pmatrix} \in \alpha_\theta \qquad \text{if} \qquad \pi_2 \neq 0.$$

If $\pi_1 \neq 0$ and $\pi_2 \neq 0$, then we have two charts for V_θ, namely

$$\psi_\theta \longmapsto \psi_\theta(x_\theta^1) \quad \text{and} \quad \psi_\theta \longmapsto \psi_\theta(x_\theta^2).$$

They are related by

$$\psi_\theta(x_\theta^1) = g_\theta(x_\theta^1, x_\theta^2)\psi_\theta(x_\theta^2).$$

Hence on $P_3(\mathbb{C}) \setminus \{\pi_1 = \pi_2 = 0\}$ the vector bundle is completely determined by $g(\theta) := g_\theta(x_\theta^1, x_\theta^2)$.

Ward pointed out that the potentials (A_μ) may be regained from $g(\theta)$ by exploiting the fact that

$$g(\theta) = g_\theta(x_\theta^1, x)g\,(x, x_\theta^2) \quad \text{for each} \quad x \in \alpha_\theta.$$

Thus we write

$$h(x,\zeta) = g_\theta(x_\theta^1, x)$$

$$k(x,\zeta) = g_\theta(x_\theta^2, x),$$

where $\zeta = \pi_1/\pi_2$. Note that $h(x,\zeta)$ is defined only for $\zeta \neq 0$ $(\pi_1 \neq 0)$ and that $k(x,\zeta)$ is defined only for $\zeta \neq \infty$ $(\pi_2 \neq 0)$. Furthermore

$$g(x\pi, \pi) = h(x,\zeta)k(x,\zeta)^{-1}. \tag{4}$$

If we assume that $h(x,\zeta)$ is analytic away from $\zeta = 0$ and $k(x,\zeta)$ is analytic away from $\zeta = \infty$, it follows from Liouville's theorem that the factorization (4) is unique up to a gauge transformation.

We know that h and k also satisfy equation (3). Since $(dx)\pi = 0$ we have $dx_{P_2} = -\zeta\, dx_{P_1}$. If we write $A_\mu\, dx^\mu = A_{PQ}\, dx^{PQ}$, then it follows that

$$A_{P_1} - \zeta A_{P_2} = i\, h(x,\zeta)^{-1}\left(\frac{\partial}{\partial x_{P_1}} - \zeta \frac{\partial}{\partial x_{P_2}}\right) h(x,\zeta)$$

$$= -i\, k(x,\zeta)^{-1}\left(\frac{\partial}{\partial x_{P_1}} - \zeta \frac{\partial}{\partial x_{P_2}}\right) k(x,\zeta).$$

Thus the isomorphism class of the bundle determines the gauge potentials (A_μ) up to gauge equivalence.

(3.2) Now we will describe Wards's construction (cf. [2]). Consider again twistor space $\mathbb{C}^4$ with coordinates $Z^\alpha = (\omega^P, \pi_Q)$. Remove the subspace $\pi_1 = \pi_2 = 0$ and introduce in the π-space the equivalence relation $\pi_P \sim \lambda\pi_P$ $(0 \neq \lambda \in \mathbb{C})$. Then we obtain the space $\mathbb{C}^2 \times P_1(\mathbb{C})$ ($\mathbb{C}^2$ with coordinates ω^P and $P_1(\mathbb{C})$ with coordinate $\zeta = \pi_1/\pi_2$). Cover $P_1(\mathbb{C})$ with two patches U and $\hat{U}$ such that U contains the point $\zeta = 0$, $\hat{U}$ contains the point $\zeta = \infty$, $U \cap \hat{U}$ is an annular region containing the circle $|\zeta| = 1$, and U and $\hat{U}$ are conjugate to each other (i.e. the antipodal map $\zeta \longmapsto \bar{\zeta}^{-1}$ maps U onto $\hat{U}$).

Let $g(Z^\alpha)$ now be a (2×2)-matrix function of Z^α such that

(i) g is homogeneous of degree zero in Z^α, i.e. $g(\lambda Z^\alpha) = g(Z^\alpha)$ for all $\lambda \neq 0$;

(ii) $\det g = 1$;

(iii) $g(Z)^* = g(Z^*)$, where $*$ on the left-hand side denotes complex conjugate transpose and $Z^{*\alpha} = (\overline{Z^1}, -\overline{Z^0}, \overline{Z^3}, -\overline{Z^2})$;

(iv) $g(x^{PQ}\pi_Q, \pi_R)$ is analytic for $(x,\zeta) \in \mathcal{R} \times (U \cap \hat{U})$, where $\mathcal{R}$ is some region in $\mathbb{R}^4$ (this makes sense because g, for a fixed value of x, is homogeneous of degree zero in π_P and so is a function of $\zeta = \pi_1/\pi_2$);

(v) g can be factorized

$$g(x\pi,\pi) = h(x,\zeta)k(x,\zeta)^{-1}$$

where h and k are (2×2)-matrices which are analytic on $\mathcal{R} \times \hat{U}$ and $\mathcal{R} \times U$ respectively.

Let D_P denote the differential operator $D_P = \partial_{P_1} - \zeta\partial_{P_2}$, where $\partial_{PQ} = \dfrac{\partial}{\partial x^{PQ}}$.

THEOREM (Ward).

(i) *There exist real-analytic functions* $A_{PQ}(x)$ $(x\in\mathcal{R})$ *such that*

$$A_{P_1} - \zeta A_{P_2} = -i\, h^{-1} D_P h$$
$$= -i\, k^{-1} D_P k.$$

(ii) *The field* $A_\mu(x)$, *defined by* $A_\mu\, dx = A_{PQ}\, dx^{PQ}$, *satisfies the self-duality equations on* $\mathcal{R}$.

REMARK 1. By replacing h and k by $h\Lambda$ and $k\Lambda$ respectively where $\Lambda(x) \in G\ell(2,\mathbb{C})$, one can be sure that A_μ takes values in the Lie algebra of $SU(2)$, i.e. that $\operatorname{trace}(A_\mu) = 0$ and $A_\mu^* = A_\mu$.

REMARK 2. This procedure produces all solutions of the $SU(2)$ self-duality equations on $\mathcal{R} \subset \mathbb{R}^4$.

PROOF. (i) Since $D_P(x^R\pi_R) = 0$, it follows from (v) that

$$h^{-1} D_P h = k^{-1} D_P k.$$

Now the left-hand side of this equation is analytic on $\mathcal{R} \times U$ and the right-hand side is analytic on $\mathcal{R} \times (\hat{U}\backslash\{\infty\})$, so both sides must be analytic on $\mathcal{R} \times (P_1(\mathbb{C})\backslash\{\infty\} = \mathcal{R} \times \mathbb{C}$. Since both sides are $O(|\zeta|)$ as $|\zeta| \to \infty$, it follows from Liouville's theorem that there exist functions A_{PQ} such that

$$A_{P_1} - \zeta A_{P_2} = -i\, h^{-1} D_P h.$$

(ii) We compute $D_2(A_{11}-\zeta A_{12}) - D_1(A_{21}-\zeta A_{22})$. This is equal to $(\partial_{21}-\zeta\partial_{22})(A_{11}-\zeta A_{12}) - (\partial_{11}-\zeta\partial_{12})(A_{21}-\zeta A_{22})$. On the other hand it is also equal to $-i\, D_2(h^{-1}D_1 h) + i\, D_1(h^{-1}D_2 h) =$

$-i\, h^{-1}(D_2 h)h^{-1}D_1 h - i\, h^{-1}D_2D_1 h - i\, h^{-1}(D_1 h)h^{-1}D_2 h + i\, h^{-1}D_1D_2 h =$

$-i(A_{21}-\zeta A_{22})(A_{11}-\zeta A_{12}) + i(A_{11}-\zeta A_{21})(A_{21}-\zeta A_{22})$.

It follows that $\partial_{21}A_{11} - \partial_{11}A_{21} = i[A_{11},A_{21}]$

$$-\partial_{22}A_{11} - \partial_{21}A_{12} + \partial_{11}A_{22} + \partial_{12}A_{21} = -i[A_{11},A_{22}] - i[A_{12},A_{21}]$$

$$\partial_{22}A_{12} - \partial_{12}A_{22} = i[A_{12},A_{22}].$$

We have to rewrite these equations using

$$A_{11} = \tfrac{1}{2}(A_0+i\,A_3) \qquad \partial_{11} = \tfrac{1}{2}(\partial_0+i\,\partial_3)$$

$$A_{22} = \tfrac{1}{2}(A_0-i\,A_3) \qquad \partial_{22} = \tfrac{1}{2}(\partial_0-i\,\partial_3)$$

$$A_{12} = \tfrac{1}{2}(i\,A_1-A_2) \qquad \partial_{12} = \tfrac{1}{2}(-\partial_2+i\,\partial_1)$$

$$A_{21} = \tfrac{1}{2}(i\,A_1+A_2) \qquad \partial_{21} = \tfrac{1}{2}(\partial_2+i\,\partial_1)$$

Then one obtains the self-dual Yang-Mills equations

$$F_{01} = F_{23}, \quad F_{02} = -F_{13}, \quad F_{03} = F_{12}$$

where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu].$$

REFERENCES

[1] CORRIGAN, E.F. - FAIRLIE, D.B. - YATES, R.G. and GODDARD, P, *The construction of self-dual solutions to* SU(2) *gauge theory*, Comm. Math.Phys. 58, 223-240 (1970).

[2] WARD, R.S., *Ansätze for self-dual Yang-Mills fields*, Comm.Math.Phys. 80, 563-574 (1981).

COMPUTATION OF THE YANG-MILLS POTENTIALS

by

J.W. de ROEVER
Department of Mathematics
Twente University of Technology

1. THE CONSTRUCTION OF WARD (continued)

We continue the construction of the potentials A_{PQ} mentioned in Ward's theorem, p.123 of the foregoing section.

So we start with a function g defined in (or possibly in a subset W of) $\mathbb{C}^4 \setminus \{Z_3 = Z_4 = 0\}$

$$g: W \subset \mathbb{C}^4 \to SL(2,\mathbb{C})$$

with the following properties (p.122 and 123 of the foregoing section.)

(i) g is homogeneous of degree zero

(iv) g is analytic for $(x,\zeta) \in R \times (U \cap \hat{U})$, i.e. for fixed $\pi = (\pi_1,\pi_2) \in \mathbb{C}^2$, $g(\tilde{x}\pi,\pi)$ is an analytic function of $x = (x^0,x^1,x^2,x^3) \in R \subset \mathbb{R}^4$, where $\tilde{x}$ denotes the matrix

$$(1) \qquad \tilde{x} = x^{PQ} = \begin{pmatrix} x^{11} & x^{12} \\ x^{21} & x^{22} \end{pmatrix} = \begin{pmatrix} x^0 - ix^3 & -ix^1 - x^2 \\ -ix^1 + x^2 & x^0 + ix^3 \end{pmatrix},$$

and for fixed $x \in R$ $g(\tilde{x}\pi,\pi)$ is a holomorphic function of $\zeta = \pi_1/\pi_2 \in U \cap \hat{U}$, where $U \cap \hat{U}$ is an annular domain in $\mathbb{C}$ containing the circle $|\zeta| = 1$, U a domain in $\mathbb{C}$ containing this circle and the point $\zeta = 0$, and where $\hat{U}$ is a domain in the Riemann sphere containing this circle and the point $\zeta = \infty$, such that $U \cup \hat{U}$ covers the complete Riemann sphere $P_1(\mathbb{C})$. (Thus the domain $W \subset \mathbb{C}^4 \setminus \{Z_3 = Z_4 = 0\}$ is determined by R and $U \cap \hat{U}$).

(v) g can be factorized in the following way

(2) $$g(\tilde{x}\pi,\pi) = \hat{h}(x,\zeta)h(x,\zeta)^{-1}, \qquad (x,\zeta) \in \mathcal{R} \times (U \cap \hat{U})$$

with analytic functions

$$h\colon \mathcal{R} \times U \to G\ell(2,\mathbb{C})$$

$$h\colon \mathcal{R} \times \hat{U} \to G\ell(2,\mathbb{C})$$

By Ward's theorem analytic Yang-Mills potentials A_μ in $\mathcal{R}$ (this means that the associated fields $F_{\mu\nu}$ are self-dual) are obtained from

(3) $$-ih(x,\zeta)^{-1}D_P(\zeta)h(x,\zeta) = A_{P1}(x) - \zeta A_{P2}(x),$$

where $D_P = D_P(\zeta) = \frac{\partial}{\partial x^{P1}} - \zeta\frac{\partial}{\partial x^{P2}}$, $P = 1,2$ and where $A_\mu dx^\mu = A_{PQ}dx^{PQ}$.

The factorization (2) is not unique. For let $\Lambda\colon \mathcal{R} \to G\ell(2,\mathbb{C})$ be analytic, then also the functions $h\Lambda$ and $\hat{h}\Lambda$ satisfy (2). The associated potentials $\tilde{A}_{PQ}$ follow from (3):

$$\tilde{A}_{P1} - \zeta\tilde{A}_{P2} = -i\Lambda^{-1}h^{-1}D_P(h\Lambda) = -i\Lambda^{-1}h^{-1}(D_Ph)\Lambda - i\Lambda^{-1}D_P\Lambda =$$

$$= \Lambda^{-1}(A_{P1}-\zeta A_{P2})\Lambda - i\Lambda^{-1}(\partial_{P1}-\zeta\partial_{P2})\Lambda,$$

hence

$$\tilde{A}_{PQ} = \Lambda^{-1}A_{PQ}\Lambda - \Lambda^{-1}\partial_{PQ}\Lambda.$$

Furthermore, A_μ is the same linear combination of the functions A_{PQ} as ∂_μ is of ∂_{PQ} (compare the formulae on p.124 of the foregoing section), namely

(4)
$$A_0 = A_{11} + A_{22}$$

$$A_1 = -i(A_{12}+A_{21})$$

$$A_2 = A_{21} - A_{12}$$

$$A_3 = i(A_{22}-A_{11}).$$

So $\tilde{A}_\mu$ follows from A_μ by the gauge transformation Λ.

If g satisfies moreover condition (iii), then by the following theorem the field $F_{\mu\nu}$ satisfies $iF_{\mu\nu} \in su(2)$ for $x \in R$, which is required by physics.

THEOREM. *If* g *satisfies moreover*

(iii) $g(Z)^* = g(Z^*)$

then $\Lambda: R \to SL(2,\mathbb{C})$ *can be chosen such that* $iA_\mu(x) \in su(2)$ *for* $x \in R$ *where* A_μ *are the associated potentials.*

PROOF. 1) Since $\det g = 1$ arbitrary h and $\hat{h}$ satisfying (2) have the same determinant, which as a consequence is holomorphic in $U \cup \hat{U}$ for $x \in R$. Hence by Liouville $\det h$ is a function of x alone which can then be included in the function Λ. Therefore it is possible to consider functions

$$h: R \times U \to SL(2,\mathbb{C})$$

$$\hat{h}: R \times \hat{U} \to SL(2.\mathbb{C})$$

$$\Lambda: R \longrightarrow SL(2,\mathbb{C})$$

only. Writing $h = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $h^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ (for $\det h = ad - bc = 1$) we get

$$\operatorname{tr}(h^{-1}D_P h) = dD_P a - cD_P b - bD_P c + aD_P d = D_P(ad-bc) = 0.$$

Hence $\operatorname{tr}(A_{P1} - \zeta A_{P2}) = 0$ for all ζ, thus $\operatorname{tr} A_{PQ} = 0$ and from (4) $\operatorname{tr} A_\mu = 0$.

2) Properties (iii) and (v) of g imply

$$h(x,\zeta)^{*-1}\hat{h}(x,\zeta)^* = \hat{h}(x,w)h(x,w)^{-1}$$

with $w = -1/\bar{\zeta}$. This yields the first equality in the following formula

$$\hat{h}(x,\zeta)^{*}h(x,w) = h(x,\zeta)^{*}\hat{h}(x,w) = \lambda(x)$$

where the left-hand side is holomorphic for w away from ∞ and the middle term for w away from 0. Then the last equality follows from Liouville. Altogether this implies that $\lambda(x)$ is hermitian

$$\lambda^{*}(x) = \hat{h}(x,w)^{*}h(x,\zeta) = \lambda(x).$$

Hence $\lambda(x)$ has real eigenvalues. The replacement of h and $\hat{h}$ by $h\Lambda$ and $\hat{h}\Lambda$ with $\Lambda(x) \in SL(2,\mathbb{C})$ would have resulted in the replacement of λ by $\Lambda^{*}\lambda\Lambda$. Therefore we can choose $\Lambda(x) \in SU(2)$ such that

$$\Lambda^{*}\lambda\Lambda = \Lambda^{-1}\lambda\Lambda = \begin{pmatrix} \mu & 0 \\ 0 & 1/\mu \end{pmatrix}$$

with $\mu = \mu(x) \neq 0$ real, because $\det\lambda(x) = 1$. Finally, we choose

$$\Lambda' = \begin{pmatrix} \dfrac{\operatorname{sign}(\mu)}{\sqrt{|\mu|}} & 0 \\ 0 & \operatorname{sign}(\mu)\sqrt{|\mu|} \end{pmatrix} : R \longrightarrow SL(2,\mathbb{C}) .$$

Then we have

$$\Lambda'^{*}\begin{pmatrix} \mu & 0 \\ 0 & 1/\mu \end{pmatrix}\Lambda' = \begin{pmatrix} \operatorname{sign}(\mu) & 0 \\ 0 & \operatorname{sign}(\mu) \end{pmatrix} = \pm I$$

Thus by choosing a suitable gauge transformation $\Lambda: R \to SL(2,\mathbb{C})$ it can be achieved that we deal with h and $\hat{h}$ with the property

$$h(x,\zeta)^{*}\hat{h}(x,w) = \pm I$$

or

$$\hat{h}(x,w) = \pm h(x,\zeta)^{*-1}.$$

It follows from (3) that

$$A^*_{11} - \bar{\zeta}A^*_{12} = i(\bar{D}_1 h^*)h^{*-1} = i(\bar{D}_1 \hat{h}(x,w)^{-1})\hat{h}(x,w).$$

Since

$$\overline{D_1(\zeta)} = \overline{\partial_{11} - \zeta\partial_{12}} = \tfrac{1}{2}(\partial_0 - i\partial_3) - \bar{\zeta}\,\tfrac{1}{2}(-\partial_2 - i\partial_1) = \partial_{22} + \bar{\zeta}\partial_{21} =$$

$$= \bar{\zeta}D_2(w),$$

we get

$$-A^*_{12} - wA^*_{11} = i(D_2(w)\hat{h}(x,w)^{-1})\hat{h}(x,w) = -i\hat{h}(x,w)^{-1}D_2(w)\hat{h}(x,w) =$$

$$= A_{21} - wA_{22}.$$

This implies

$$A_{21} = -A^*_{12} \quad \text{and} \quad A_{22} = A^*_{11},$$

or

$$A_2 + iA_1 = A^*_2 + iA^*_1 \quad \text{and} \quad A_0 - iA_3 = A^*_0 - iA^*_3,$$

so that $A^*_\mu = A_\mu$. ∎

Finally, it should be remarked that it is possible that different functions g yield the same potentials A_μ. For the matrix g is a transition matrix of a vector bundle over the union of two domains in the twistor space $\mathbb{C}^4$ (cf. p.121 of the foregoing section). An equivalent vector bundle (that is, so to say, one with the same "kind of non-triviality") is determined by the transition matrix

$$\tilde{g} = \hat{m}gm \quad \text{(valid in the intersection of the two domains)}$$

where the matrices $\hat{m}$ and m can each be extended to the whole of one of the two domains in $\mathbb{C}^4$, i.e. more precisely $\hat{m}$: (possibly subset of) $\mathbb{C}^4 \setminus \{Z_3 = 0\} \to SL(2,\mathbb{C})$ and m: (possibly subset of) $\mathbb{C}^4 \setminus \{Z_4 = 0\} \to SL(2,\mathbb{C})$

are homogeneous of degree zero such that $\hat{m}(\tilde{x}\pi,\pi)$ is analytic for $(x,\zeta) \in R \times \hat{U}$ and $m(\tilde{x}\pi,\pi)$ for $(x,\zeta) \in R \times U$. The function $\tilde{g}$ yields the same potentials A_μ, for

$$\hat{\tilde{h}}(x,\zeta) = \hat{m}(\tilde{x}\pi,\pi)\hat{h}(x,\zeta) \quad \text{and} \quad \hat{h}(x,\zeta)\, m(\tilde{x}\pi,\pi)^{-1}h(x,\zeta),$$

hence (since $D_p(\tilde{x}\pi) = 0$)

$$\tilde{h}^{-1}D_p h = \tilde{h}^{-1}mD_p(m(\tilde{x}\pi,\pi)^{-1}h) = h^{-1}mm^{-1}D_p h = h^{-1}D_p h.$$

Thus for the computation of the A_μ one may start with a function g not satisfying (iii) provided that there exist functions $\hat{m}$ and m such that $\tilde{g}$ satisfies (iii).

2. COMPUTATION OF THE POTENTIALS FROM THE "ANSÄTZE" A_n

The greatest difficulty of the construction of section 1 is to find the functions h and $\hat{h}$ in (2). In order to get explicite formulae for h and $\hat{h}$ (and hence also for A_μ) we make a (possibly restricting) assumption on the form of the matrix $g(Z)$, namely, we assume that it is a triangular matrix.

Let f and Γ be two functions of $Z \in W$, homogeneous of degree zero and analytic for $(x,\zeta) \in R \times (U \cap \hat{U})$. We will start our construction with a matrix g of the form

$$(5) \qquad g(Z) = \begin{pmatrix} \zeta^{n_e} f(Z) & \Gamma(Z) \\ 0 & \zeta^{-n_e - f(Z)} \end{pmatrix} , \quad \text{here} \quad \zeta = Z_3/Z_4 \neq 0.$$

The function $g: W \subset \mathbb{C}^4 \to SL(2,\mathbb{C})$ satisfies the properties (i), (ii) and (iv) on p.122. If we moreover (later on) impose conditions on the functions f and Γ such that there are functions $\hat{m}$ and m with the property that $\tilde{g} = \hat{m}gm$ satisfies (iii), then it is for sure that $iA_\mu(x) \in su(2)$ can be satisfied. We will now express the potentials A_μ in terms of f and Γ by the construction of section 1. As a special case (namely, $n = 1$ and $f = 0$) we will get the 't Hooft solutions with arbitrary instanton number k.

In order to factorize g as in (2) we first split f:

$$(6) \qquad f(\tilde{x}\pi,\pi) = \hat{\mu}(x,\zeta) - \mu(x,\zeta) , \qquad (x,\zeta) \in R \times (U \cap \hat{U})$$

where the functions μ and $\hat{\mu}$ are analytic in $R \times U$ and in $R \times \hat{U}$, respectively. Such functions can be obtained in the following way

$$\mu(x,\zeta') = \frac{-1}{2\pi i} \oint \frac{f(\tilde{x}\pi,\pi)}{\zeta - \zeta'} d\zeta \quad \text{with contour} \quad \text{around } 0 \text{ and } \zeta',$$

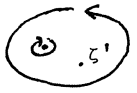
and

$$\hat{\mu}(x,\zeta') = \frac{-1}{2\pi i} \oint \frac{f(\tilde{x}\pi,\pi)}{\zeta-\zeta'} d\zeta \quad \text{with contour} \quad$$

around 0 but not around ζ'.

By Cauchy's formula we have

$$\hat{\mu}(x,\zeta') - \mu(x,\zeta') = \frac{1}{2\pi i} \int \frac{f(\tilde{x}\pi,\pi)}{\zeta-\zeta'} d\zeta = f(\tilde{x}\pi',\pi'), \quad \zeta' = \pi_1'/\pi_2'$$

with contour 

.

Furthermore, the function $\rho = \Gamma e^{-\mu-\hat{\mu}}$ is written in a Laurent series

$$\rho(x,\zeta) = \Gamma(\tilde{x}\pi,\pi) e^{-\mu(x,\zeta)-\hat{\mu}(x,\zeta)} = \sum_{k=-\infty}^{\infty} \Delta_{-k}(x)\zeta^k, \tag{7}$$

where

$$\Delta_r(x) = \frac{1}{2\pi i} \oint \zeta^{r-1}\rho(x,\zeta)d\zeta, \qquad x \in R. \tag{8}$$

We look for functions $\hat{a}$, $\hat{b}$, $\hat{c}$ and $\hat{d}$, analytic in $R\times\hat{U}$, and for functions a, b, c and d, analytic in $R\times U$, such that

$$g = \begin{pmatrix} \zeta^n e^{\hat{\mu}-\mu} & \Gamma \\ 0 & \zeta^{-n} e^{-\hat{\mu}+\mu} \end{pmatrix} = \begin{pmatrix} \hat{a} & \hat{b} \\ \hat{c} & \hat{d} \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

and such that $\det \hat{h} = \hat{a}\hat{d} - \hat{b}\hat{c} = \det h = ad - bc = 1$.

So we begin with solving the following system

$$\begin{pmatrix} \zeta^n e^{\hat{\mu}-\mu} & \Gamma \\ 0 & \zeta^{-n} e^{-\hat{\mu}+\mu} \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \hat{a} & \hat{b} \\ \hat{c} & \hat{d} \end{pmatrix} .$$

First this yields in $R\times(U\cap\hat{U})$

$$\zeta^{-n} e^{-\hat{\mu}} e^{\mu} c = \hat{c}, \quad \text{or} \quad e^{\mu} c = \zeta^n e^{\hat{\mu}} \hat{c}$$

and

$$\zeta^{-n} e^{-\hat{\mu}} e^{\mu} d = \hat{d}, \quad \text{or} \quad e^{\mu} d = \zeta^{n} e^{\hat{\mu}} \hat{d}.$$

Since $c(x,\zeta) \exp \mu(x,\zeta)$ is analytic in $R \times U$ and $\hat{c}(x,\zeta) \exp \hat{\mu}(x,\zeta)$ in $R \times \hat{U}$, c and $\hat{c}$ have the following form

$$e^{\mu(x,\zeta)} c(x,\zeta) = \zeta^{n} e^{\hat{\mu}(x,\zeta)} \hat{c}(x,\zeta) = \sum_{k=0}^{n} c_k(x) \zeta^{k} \tag{9}$$

where $c_0, \ldots, c_n$ are analytic functions in R.
Similarly,

$$e^{\mu(x,\zeta)} d(x,\zeta) = \zeta^{n} e^{\hat{\mu}(x,\zeta)} \hat{d}(x,\zeta) = \sum_{k=0}^{n} d_k(x) \zeta^{k}. \tag{10}$$

We assume $n \geqslant 1$ and we set

$$e^{-\hat{\mu}(x,\zeta)} \hat{a}(x,\zeta) = \sum_{k \leqslant 0} \theta_k(x) \zeta^{k}$$

$$e^{-\hat{\mu}(x,\zeta)} \hat{b}(x,\zeta) = \sum_{k \leqslant 0} \phi_k(x) \zeta^{k}$$

and

$$-\zeta^{n} e^{-\mu(x,\zeta)} a(x,\zeta) = \sum_{k \geqslant 1} \theta_k(x) \zeta^{k}$$

$$-\zeta^{n} e^{-\mu(x,\zeta)} b(x,\zeta) = \sum_{k \geqslant 1} \phi_k(x) \zeta^{k}$$

with $\theta_1 = \ldots = \theta_{n-1} = 0$ (or no condition if $n = 1$) and $\phi_1 = \ldots = \phi_{n-1} = 0$ (or nothing if $n = 1$). The other two equations yield

$$e^{\hat{\mu}} \zeta^{n} e^{-\mu} a + \Gamma = \hat{a} \quad \text{or} \quad \zeta^{n} e^{-\mu} a + \rho e^{\mu} c = e^{-\hat{\mu}} \hat{a}$$

and similarly

$$\zeta^n e^{-\mu} b + \rho e^{\mu} d = e^{-\hat{\mu}} \hat{b}.$$

Together with (7) and (9) this implies

$$\sum_{k=-\infty}^{\infty} \Delta_{-k} \zeta^k \sum_{\ell=0}^{n} c_\ell \zeta^\ell = \sum_{k=-\infty}^{\infty} \theta_k \zeta^k,$$

hence

$$\theta_k(x) = \sum_{\ell=0}^{n} \Delta_{\ell-k}(x)\, c_\ell(x), \qquad k \in \mathbb{Z}. \tag{11}$$

Similarly,

$$\phi_k(x) = \sum_{\ell=0}^{n} \Delta_{\ell-k}(x)\, d_\ell(x), \qquad k \in \mathbb{Z}. \tag{12}$$

In the solution (9), (10), (11) and (12) $2n+2$ analytic functions $c_0, \ldots, c_n, d_0, \ldots, d_n$ occur which are not arbitrary but satisfy the $2n-2$ conditions

$$\theta_k = \phi_k = 0\,, \qquad 1 \leqslant k \leqslant n-1 \tag{13}$$

and the condition $\det h = 1$

$$-\phi_0 c_n + \theta_0 d_n = c_0 \phi_n - d_0 \theta_n = 1, \tag{14}$$

cf. formula (15) below, where it will turn out that the first equality automatically holds, so that the last equality appears as a condition. The freedom which remains is three arbitrary analytic functions which is exactly the choice of a function $\Lambda\colon \mathbb{R} \to SL(2,\mathbb{C})$, the freedom we had in determining the matrix $h = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Furthermore, it is clear that the functions Δ_r with $|r| \leqslant n$ have to satisfy certain conditions, thus conditions for the functions f and Γ, in order that the conditions (13) and (14) can be satisfied. This will be discussed later.

It has already been shown that $\det h(x,\zeta)$ is independent of ζ; directly, this can be seen as follows

$$-bc = \sum_{k=n}^{\infty} \phi_k \zeta^{k-n} \sum_{\ell=0}^{n} c_\ell \zeta^\ell = \sum_{m=1}^{\infty} \left(\sum_{\ell=0}^{n} \phi_{m+n-\ell} c_\ell \right) \zeta^m + \phi_n c_0$$

$$\text{(because } \phi_1 = \ldots = \phi_{n-1} = 0\text{)}$$

(15)

$$= \sum_{m=1}^{\infty} \left(\sum_{\ell=0}^{n} \sum_{k=0}^{n} \Delta_{\ell+k-m-n} d_k c_\ell \right) \zeta^m + \phi_n c_0 = \text{(similarly)}$$

$$= \sum_{k=n}^{\infty} \theta_k \zeta^{k-n} \sum_{\ell=0}^{n} d_\ell \zeta^k + \phi_n c_0 - \theta_n d_0 = -ad + \phi_n c_0 - \theta_n d_0 .$$

If in this calculation the sum $\sum_m$ had ranged from $m = 0$ until $m = \infty$ then the correction term $\phi_n c_0$ should have been replaced by $-\phi_0 c_n$. In this way one finds

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \phi_n c_0 - \theta_n d_0 = -\phi_0 c_n + \theta_0 d_n .$$

Similar reasonings will be met frequently in the computation of the potentials, thus of $h^{-1} D_P h$.

The functions $\Delta_k(x)$ given by (8) occurring in the above found expression for h are not arbitrary but have to satisfy certain relations. Since $D_P(\tilde{x}\pi) = 0$ it follows from (6) that $D_P \mu = D_P \hat{\mu}$ and (as in the proof of Ward's theorem on p.123 as a consequence there must be analytic functions B_{PQ} with

$$B_{P1}(x) - \zeta B_{P2}(x) = D_P(\zeta)\mu(x,\zeta), \qquad P = 1,2.$$

(In the next section we will see that B_μ is the potential of a self-dual Maxwell field). Furthermore, on the one hand (because $D_P \Gamma = 0$)

$$D_P \rho = D_P \Gamma e^{-\mu-\hat{\mu}} = -2(D_P \mu)\rho = -2B_{P1}\rho + \zeta\, 2B_{P2}\rho$$

holds, and on the other hand

$$D_P\rho = \sum_k \left(\partial_{P1}\Delta_{-k}\right)\zeta^k - \sum_k \left(\partial_{P2}\Delta_{-k}\right)\zeta^{k+1}$$

holds. Hence the following relation between the Laurent coefficients Δ_r is obtained

$$(\partial_{P1}+2B_{P1})\Delta_r = (\partial_{P2}+2B_{P2})\Delta_{r+1}\ , \qquad P = 1,2. \tag{16}$$

The computation of $h^{-1}D_P h$ is now straightforward.

$$h^{-1}D_P h = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} D_P \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} dD_P a - bD_P c & dD_P b - bD_P d \\ -cD_P a + aD_P c & -cD_P b + aD_P d = \\ & -(dD_P a - bD_P c) \end{pmatrix}$$

for $D_P(ad-bc) = D_P 1 = 0$. First, we consider $dD_P a - bD_P c$ (we have shown already that this function depends linearly on ζ):

$$dD_P a - bD_P c = -\sum_{\ell=0}^{n} d_\ell \zeta^\ell \Big((D_P\mu) + D_P\Big) \sum_{k=n}^{\infty} \theta_k \zeta^{k-n} + \sum_{k=n}^{\infty} \phi_k \zeta^{k-n} \Big((-D_P\mu) + D_P\Big) \sum_{\ell=0}^{n} c_\ell \zeta^\ell =$$

$$\text{(according to (15))} = -D_P\mu - \sum_{\ell=0}^{n} d_\ell \zeta^\ell \Big(2(D_P\mu) + D_P\Big) \sum_{k=n}^{\infty} \theta_k \zeta^{k-n} +$$

$$+ \sum_{k=n}^{\infty} \phi_k \zeta^{k-n} \sum_{\ell=0}^{n} (D_P c_\ell)\zeta^\ell =$$

$$= -D_P\mu - \sum_{m=1}^{\infty} \left\{ \sum_{k=0}^{n} \sum_{\ell=0}^{n} d_k (2(D_P\mu) + D_P)(\Delta_{\ell+k-m-n} c_\ell) \right\} \zeta^m - d_0 (2(D_P\mu) + D_P)\theta_n +$$

$$+ \sum_{m=1}^{\infty} \left\{ \sum_{k=0}^{n} \sum_{\ell=0}^{n} d_k \Delta_{\ell+k-m-n} (D_P c_\ell) \right\} \zeta^m + \phi_n D_P c_0 =$$

$$= -D_P\mu + \phi_n D_P c_0 - d_0 (2(D_P\mu) + D_P)\theta_n - \sum_{m=1}^{\infty} \left\{ \sum_{k=0}^{n} \sum_{\ell=0}^{n} d_k c_\ell (2(D_P\mu) + D_P)\Delta_{\ell+k-m-n} \right\} \zeta^m =$$

$$= -D_P\mu + \phi_n D_P c_0 - d_0 (2(D_P\mu) + D_P)\theta_n - \zeta \sum_{k=0}^{n} \sum_{=0}^{n} d_k c_\ell (2B_{P1} + \partial_{P1})\Delta_{\ell+k-m-n} -$$

$$- \sum_{m=2}^{\infty} \left\{ \sum_{k=0}^{n} \sum_{\ell=0}^{n} d_k c_\ell (2B_{P1} + \partial_{P1})\Delta_{\ell+k-m-n} \right\} \zeta^m + \sum_{m=1}^{\infty} \left\{ \sum_{k=0}^{n} \sum_{\ell=0}^{n} d_k c_\ell (2B_{P2} + \partial_{P2})\Delta_{\ell+k-m-n} \right\} \zeta^{m+1}.$$

According to (16) in the last line the two terms cancel so that indeed the linear term in ζ in the line before remains. We write the term with ζ in a different way after first applying (16) another time:
since $\phi_1 = \ldots \phi_{n-1} = 0$ it follows from (12) (and similarly from (11)) that

$$\sum_{k=0}^{n} \sum_{\ell=0}^{n} d_k c_\ell \Delta_{\ell+k-n} = \phi_n c_0 + \phi_0 c_n = \theta_n d_0 + \theta_0 d_n .$$

By differentiation this formula implies

$$\sum_{\ell,k=0}^{n} d_k c_\ell \partial_{P2} \Delta_{\ell+k-n} = d_0 \partial_{P2} \theta_n + \theta_n \partial_{P2} d_0 + \theta_0 \partial_{P2} d_n + d_n \partial_{P2} \theta_0 -$$

$$- \sum_{\ell,k=0}^{n} d_k \Delta_{\ell+k-n} \partial_{P2} c_\ell - \sum_{\ell,k=0}^{n} c_\ell \Delta_{\ell+k-n} \partial_{P2} d_k . \tag{17}$$

Again using (11), (12) and (13) we can write the last two terms as

$$-\left\{ \sum_{\ell=0}^{n} \phi_{n-\ell} \partial_{P2} c_\ell + \sum_{k=0}^{n} \theta_{n-k} \partial_{P2} d_k \right\} = -\left\{ \phi_n \partial_{P2} c_0 + \phi_0 \partial_{P2} c_n + \theta_n \partial_{P2} d_0 + \theta_0 \partial_{P2} d_n \right\} .$$

Altogether we have obtained

$$dD_P a - bD_P c = -D_P \mu + \phi_n D_P c_0 - d_0 (2(D_P \mu) + D_P) \theta_n -$$

$$-\zeta (2B_{P2} (\phi_n c_0 + \phi_0 c_n) + d_0 \partial_{P2} \theta_n + d_n \partial_{P2} \theta_0 - \phi_0 \partial_{P2} c_n - \phi_n \partial_{P2} c_0) =$$

$$= -B_{P1} + \phi_n \partial_{P1} c_0 - 2B_{P1} d_0 \theta_n - d_0 \partial_{P1} \theta_n - \zeta (-B_{P2} + \phi_n \partial_{P2} c_0 - 2B_{P2} d_0 \theta_n -$$

$$-d_0 \partial_{P2} \theta_n + 2B_{P2} c_0 \phi_n + 2B_{P2} c_n \theta_0 + d_0 \partial_{P2} \theta_n + d_n \partial_{P2} \theta_0 - \phi_0 \partial_{P2} c_n - \phi_n \partial_{P2} c_0)$$

which in virtue of (14) is finally expressed as

$$= -B_{P1} + \phi_n \partial_{P1} c_0 - 2B_{P1} d_0 \theta_n - d_0 \partial_{P1} \theta_n - \zeta (B_{P2} + 2B_{P2} c_n \phi_0 + d_n \partial_{P2} \theta_0 - \phi_0 \partial_{P2} c_n) . \tag{18}$$

Here we have got the upper left corner (and the lower right corner) of the matrices A_{PQ}.

The right upper corner is obtained by calculating in a similar way

$$dD_P b - bD_P d = -\sum_{\ell=0}^{n} d_\ell \zeta^\ell ((D_P\mu) + D_P) \sum_{k=n}^{\infty} \phi_k \zeta^{k-n} + \sum_{k=n}^{\infty} \phi_k \zeta^{k-n} (-(D_P\mu) + D_P) \sum_{\ell=0}^{n} d_\ell \zeta^\ell =$$

$$= -d_0(2(D_P\mu) + D_P)\phi_n - \sum_{m=1}^{\infty} \left\{ \sum_{k=0}^{n} \sum_{\ell=0}^{n} d_\ell (2(D_P\mu) + D_P) \Delta_{\ell+k-m-n} d_k \right\} \zeta^m +$$

$$+ \phi_n D_P d_0 + \sum_{m=1}^{\infty} \left\{ \sum_{k=0}^{n} \sum_{\ell=0}^{n} d_\ell \Delta_{\ell+k-m-n} D_P d_k \right\} \zeta^m =$$

just as before with the aid of (16)

$$= \phi_n D_P d_0 - d_0 D_P \phi_n - 2(D_P\mu) d_0 \phi_n - \zeta \sum_{k=0}^{n} \sum_{\ell=0}^{n} d_k d_\ell (2B_{P2} + \partial_{P2}) \Delta_{\ell+k-n} =$$

just as in (17)

$$= \phi_n \partial_{P1} d_0 - d_0 \partial_{P1} \phi_n - 2B_{P1} d_0 \phi_n - \zeta(\phi_n \partial_{P2} d_0 - d_0 \partial_{P2} \phi_n - 2B_{P2} d_0 \phi_n + 2B_{P2} \phi_n d_0 + 2B_{P2} \phi_0 d_n +$$

$$+ d_0 \partial_{P2} \phi_n + d_n \partial_{P2} \phi_0 - \phi_0 \partial_{P2} d_n - \phi_n \partial_{P2} d_0) =$$

$$= \phi_n \partial_{P1} d_0 - d_0 \partial_{P1} \phi_n - 2B_{P1} d_0 \phi_n - \zeta(2B_{P2} \phi_0 d_n + d_n \partial_{P2} \phi_0 - \phi_0 \partial_{P2} d_n). \tag{19}$$

The left lower corner is obtained from (19) by a minus sign and by interchanging ϕ and θ and also d and c

$$aD_P c - cD_P a = -\theta_n \partial_{P1} c_0 + c_0 \partial_{P1} \theta_n + 2B_{P1} c_0 \theta_n -$$

$$- \zeta(-2B_{P2} \theta_0 c_n - c_n \partial_{P2} \theta_0 + \theta_0 \partial_{P2} c_n). \tag{20}$$

Formulae (18), (19) and (20) determine the potentials A_{PQ}. They appear to depend only on θ_0, θ_n, ϕ_0 and ϕ_n, hence on the coefficients Δ_r with $|r| \leqslant n$. These coefficients have to satisfy certain relations, cf. (16). We will now express A_{PQ} in terms of the coefficients Δ_r in a special gauge. The gauge is determined by choosing three relations between the functions $c_0, \ldots, c_n, d_0, \ldots, d_n$. We take $c_n = d_0 = 0$ and $c_0 = d_n$ (which corresponds to Yang's R-gauge). In (18) only the functions c_0, ϕ_n,

θ_0 and d_n occur then. They will be expressed in the coefficients Δ_r which are determined by the functions f and Γ.

Let M denote the $n\times n$-matrix

$$M = \begin{pmatrix} \Delta_{-n+1} & \cdots & \Delta_0 \\ \vdots & & \vdots \\ \Delta_0 & \cdots & \Delta_{n-1} \end{pmatrix}, \quad \text{so} \quad M_{rs} = \Delta_{r+s-n-1}$$

We assume that $\det M \neq 0$, which is a condition on the functions f and Γ (for example $\Gamma \neq 0$). In the next section we will show that this condition enables us to solve (13) and (14), hence to guarantee the factorization (2) of the matrix (5).

Represent the inverse of M as

$$M^{-1} = \begin{pmatrix} E & \cdots & F \\ \vdots & & \vdots \\ F & \cdots & G \end{pmatrix}.$$

Here E, F and G are certain analytic functions of $x \in \mathcal{R} \setminus \{x \in \mathbb{R}^4 \mid \det M(x) = 0\}$. If $c_n = 0$ then (11) becomes

$$M \begin{pmatrix} c_0 \\ \vdots \\ c_{n-1} \end{pmatrix} = \begin{pmatrix} \theta_{n-1} \\ \vdots \\ \theta_0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \theta_0 \end{pmatrix}$$

because of (13). Hence

$$M^{-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \theta_0 \end{pmatrix} = \begin{pmatrix} c_0 \\ \vdots \\ c_{n-1} \end{pmatrix},$$

so that

$$(21) \qquad c_0 = F\theta_0, \quad c_{n-1} = G\theta_0.$$

If $d_0 = 0$, by (12) and (13) we have

$$M\begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \begin{pmatrix} \phi_n \\ \vdots \\ \phi_1 \end{pmatrix} = \begin{pmatrix} \phi_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Hence

$$M^{-1}\begin{pmatrix} \phi_n \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix},$$

so that

$$(22) \qquad d_1 = E\phi_n\,, \qquad d_n = F\phi_n.$$

It is possible that $F(x) = 0$ while $\det M(x) \neq 0$. From (21) it follows that for such an x $c_0(x) = 0$ and hence with our choice $d_0 = 0$ condition (14) cannot hold. For this x the choice $c_n = d_0 = 0$ cannot be made and the potentials A_{PQ} have to be calculated in another gauge!

Our third choice $c_0 = d_n$ implies with (14) that $\theta_0 = \phi_n$ and with (21) that $F\phi_n^2 = 1$. Thus

$$(23) \qquad \theta_0 = \phi_n, \quad \phi_n^2 = \frac{1}{F}\,.$$

In (18) the terms $\phi_n \partial_{P1} c_0$ and $d_n \partial_{P2} \theta_0$ can be expressed in F now

$$(24) \qquad \phi_n \partial_{PQ} c_0 = -d_n \partial_{PQ} \theta_0 = \frac{1}{2F}\,\partial_{PQ} F,$$

(remember that $F(x) \neq 0$).

To use (22) in (19) we first have to rewrite the term with ζ (the other term vanishes because $d_0 = 0$): as in (17) with $d_0 = 0$ we get

$$2B_{P2}\phi_0 d_n + \partial_{P2}(d_n\phi_0) - 2\phi_0\partial_{P2}d_n = (2B_{P2}+\partial_{P2}) \sum_{k,\ell=0}^{n} d_\ell d_k \Delta_{\ell+k-n} - 2\phi_0\partial_{P2}d_n =$$

$$= \sum_{\ell,k=0}^{n} d_\ell d_k (2B_{P1}+\partial_{P1})\Delta_{\ell+k-n-1} + 2\sum_{\ell,k=0}^{n} d_\ell \Delta_{\ell+k-n}\partial_{P2}d_k - 2\phi_0\partial_{P2}d_n =$$

$$= 2B_{P1}(\phi_{n+1}d_0+\phi_n d_1) + \sum_{\ell,k=0}^{n} d_k\partial_{P1}d_\ell\Delta_{\ell+k-n-1} - \sum_{\ell,k=0}^{n} d_k\Delta_{\ell+k-n-1}\partial_{P1}d_\ell =$$

$$= 2B_{P1}\phi_n d_1 + d_1\partial_{P1}\phi_n - \phi_n\partial_{P1}d_1$$

because $d_0 = 0$ and with (22) and (23) this yields

$$= 2B_{P1}\frac{E}{F} - \frac{1}{F}\partial_{P1}E. \tag{25}$$

Similarly, now with $c_n = 0$ the term in (20) independent of ζ becomes

$$2B_{P1}c_0\theta_n - \theta_n\partial_{P1}c_0 + c_0\partial_{P1}\theta_n = 2B_{P2}c_{n-1}\theta_0 + c_{n-1}\partial_{P2}\theta_0 - \theta_0\partial_{P2}c_{n-1} =$$

(with (21) and (23))

$$= 2B_{P2}\frac{G}{F} - \frac{1}{F}\partial_{P2}G. \tag{26}$$

If we finally substitute (24), (25) and (26) into (18), (19) and (20), by (3) we find the formulae for the potentials

$$A_{P1} = \frac{-i}{2F}\begin{pmatrix} \nabla_{P1}F & 0 \\ -2\nabla_{P2}G & -\nabla_{P1}F \end{pmatrix}$$

$$A_{P2} = \frac{-i}{2F}\begin{pmatrix} -\nabla_{P2}F & -2\nabla_{P1}E \\ 0 & \nabla_{P2}F \end{pmatrix} \qquad P = 1,2 \tag{27}$$

with $\nabla_{PQ} = \partial_{PQ} - 2B_{PQ}$.

Formulae (27) are Ward's "Ansätze" A_n: for given functions f and Γ and an integer $n \geqslant 1$ formulae (27) yield the potentials of a self-dual Yang-Mills field. Since the functions f and Γ are homogeneous, for each

n the class of gauge equivalent solutions depends on two freely eligible functions of three variables (provided that the condition $\det M(x) \neq 0$ holds).

In the next section we will consider the potentials A_{PQ} as depending on the functions B_{PQ}, E, F and G according to (27) and we will discuss the conditions these functions have to satisfy. Thus in the region $R' \subset \mathbb{R}^4$ where these conditions hold there exist (maybe only locally) analytic potentials. In the points $x \in R'$ where $F(x) \neq 0$ formulae (27) are an example of this. The singularities in (27) in the points $x \in R'$ where $F(x) = 0$ can be removed by choosing another gauge!

3. VALIDITY OF THE SOLUTION (27)

In this section we will investigate more directly the conditions for the functions E, F, G and B_{PQ} in order that the potentials A_{PQ} (27) yield a self-dual Yang-Mills field.

A. *The condition* $\det M \neq 0$.

Since E, F and G are the corner elements of the inverse of the $n\times n$-matrix M the condition $\det M \neq 0$ has been used already in the determination of the functions E, F and G. Furthermore, it has been needed in the derivation of (27) that conditions (13) and (14) could be satisfied. We will show now that the condition $\det M \neq 0$ is sufficient and necessary for (13) and (14) to hold so that we have the following theorem.

THEOREM. *In order that the matrix* (5) *can be factorized as in* (2) *it is necessary and sufficient that the* $n\times n$*-matrix*

$$M = \begin{pmatrix} \Delta_{-n+1} & \cdots & \Delta_0 \\ \vdots & & \vdots \\ \Delta_0 & \cdots & \Delta_{n-1} \end{pmatrix},$$

with coefficients Δ_r *determined by* (7) *and* (8), *is non-singular.*

PROOF. If we set $\vec{c} = (c_0, c_1, \ldots, c_n)\colon \mathbb{R} \to \mathbb{C}^{n+1}$ and $\vec{d} = (d_0, \ldots, d_n)\colon \mathbb{R} \to \mathbb{C}^{n+1}$ then with

$$N = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \Delta_{-n} & \Delta_{-n+1} & \cdots & \Delta_0 \\ \vdots & \vdots & & \vdots \\ \Delta_{-1} & \Delta_0 & \cdots & \Delta_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \cdot & & & \\ \cdot & & M & \\ \cdot & & & \end{pmatrix}$$

condition (13) can be written as

$$N\cdot\vec{c} = \begin{pmatrix} c_0 \\ \theta_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{and} \quad N\cdot\vec{d} = \begin{pmatrix} d_0 \\ \phi_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

So we have to look for two vectors $\vec{c}$ and $\vec{d}$ in $\mathbb{C}^{n+1}$ whose images under the linear map induced by N lie in a given complex 2-dimensional subspace of $\mathbb{C}^{n+1}$. Moreover, condition (14) means that the (complex) volume V between the image-vectors $V = \det \begin{pmatrix} c_0 & d_0 \\ \theta_n & \phi_n \end{pmatrix} = c_0\phi_n - d_0\theta_n = 1$.
Hence if $\det M \neq 0$ (thus also $\det N \neq 0$) then the vectors $\vec{c}$ and $\vec{d}$ follow from the choice of two 2-dimensional functions $p: \mathbb{R} \to \mathbb{C}^2$ and $q: \mathbb{R} \to \mathbb{C}^2$ with $\det \begin{pmatrix} p_1 & q_1 \\ p_2 & q_2 \end{pmatrix} = 1$, namely

$$(28) \qquad \vec{c} = N^{-1}\begin{pmatrix} p_1 \\ p_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{and} \quad \vec{d} = N^{-1}\begin{pmatrix} q_1 \\ q_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Now we assume that the complex 2-dimensional subspace of $\mathbb{C}^{n+1}$, spanned by the vectors $(01,0,\ldots,0)$ and $(1,0,\ldots,0)$, belongs to the range of the linear map induced by N (which is necessary for (2) to hold) and that $\det M = 0$. So there exists a vector $u \in \mathbb{C}^{n+1}$ with

$$N\cdot u = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

hence $u_0 = 1$ and

$$M\cdot\tilde{u} = \begin{pmatrix} \Delta_{-n} \\ \vdots \\ \Delta_{-1} \end{pmatrix}$$

for some vector $\tilde{u} \in \mathbb{C}^n$ and there exists a vector $v \in \mathbb{C}^{n+1}$ with

$$N \cdot v = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

hence $v_0 = 0$ and

$$M \cdot \tilde{v} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

for some vector $\tilde{v} \in \mathbb{C}^n$. Thus the n-dimensional vectors

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ \Delta_{-n+1} \\ \vdots \\ \Delta_{-1} \end{pmatrix}$$

depend linearly on the columns of the matrix M. Then the $n \times (n+2)$-matrix

$$\begin{pmatrix} 1 & 0 & \Delta_{-n+1} & \cdots & \Delta_0 \\ 0 & \Delta_{-n+1} & \vdots & & \vdots \\ \vdots & \vdots & \Delta_{-1} & & \vdots \\ 0 & \Delta_{-1} & \Delta_0 & \cdots & \Delta_{n-1} \end{pmatrix}$$

has rank $\leqslant n-1$, because $\operatorname{rank}(M) \leqslant n-1$ according to the assumption $\det M = 0$, and moreover in this matrix the first two columns depend linearly on the other n. Now the first row is independent of the others, so the remaining (n-1) rows have rank $\leqslant n-2$, thus

$$\operatorname{rank} \begin{pmatrix} \Delta_{-n+1} & \Delta_{-n+2} & \cdots & \Delta_0 & \Delta_1 \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ \Delta_{-1} & \Delta_0 & \cdots & \Delta_{n-2} & \Delta_{n-1} \end{pmatrix} \leqslant n-2.$$

Hence also the rank of the first n columns is not larger than $n-2$. By adding two columns which depend linearly on these n columns the rank will not increase, in particular

$$\operatorname{rank}\begin{pmatrix} 1 & 0 & \Delta_{-n+1} & \Delta_{-n+2} & \cdots & \Delta_0 \\ 0 & \Delta_{-n+1} & \vdots & \vdots & \iddots & \vdots \\ \vdots & \vdots & \vdots & \iddots & & \\ 0 & \Delta_{-2} & \Delta_{-1} & \Delta_0 & \cdots & \Delta_{n-2} \end{pmatrix} \leqslant n-2.$$

Repeating this process we finally get

$$\operatorname{rank}\begin{pmatrix} 1 & 0 & \Delta_{-n+1} & \cdots & \Delta_{-1} & \Delta_0 \\ 0 & \Delta_{-n+1} & \Delta_{-n+2} & & \Delta_0 & \Delta_1 \end{pmatrix} \leqslant 1.$$

Hence $\Delta_{-n+1} = 0 \Rightarrow \Delta_{-n+2} = 0 \Rightarrow \ldots \Rightarrow \Delta_{-1} = 0 \Rightarrow \Delta_0 = 0$ and the first row becomes $(1\ 0\ \ldots\ 0)$. However, in that case the first column cannot depend linearly on the last n. So the assumption $\det M = 0$ cannot be valid. ∎

REMARK. In part C we will show that the coefficients in the matrix N are analytic functions. Thus if the gauge $(p,q)\colon R \to SL(2,\mathbb{C})$ is analytic, then also the functions $\vec{c}$ and $\vec{d}$ determined by (28) are analytic.

B. *Conditions on the functions* B_{PQ} *- Introduction to the Penrose transformation.*

The field associated to the potentials B_{PQ} is a self-dual Maxwell field. Indeed, from $D_1D_2\mu = D_2D_1\mu$ it follows that

$$\partial_{21}B_{11} - \zeta(\partial_{22}B_{11}+\partial_{21}B_{12}) + \zeta^2\partial_{22}B_{12} =$$
$$= \partial_{11}B_{21} - \zeta(\partial_{12}B_{21}+\partial_{11}B_{22}) + \zeta^2\partial_{12}B_{22},$$

hence

$$\partial_{21}B_{11} = \partial_{11}B_{21}, \quad \partial_{12}B_{22} = \partial_{22}B_{12} \tag{29}$$

and

$$(30) \qquad \partial_{22}B_{11} - \partial_{12}B_{21} = \partial_{11}B_{22} - \partial_{21}B_{12}.$$

These equations for the potentials are exactly those implying the self-duality of the associated field,(with values in a commutative Lie algebra) cf. the formulae on p.124 with vanishing commutators.

REMARK. To get the Maxwell field $F_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ one should change to the Minkowski metric by defining the potentials B_μ by

$$B_{PQ}(x^o,-ix^a)dx^{PQ} = B_o(x)dx^o + iB_a(x)dx^a.$$

Then (29) and (30) imply the self-duality equations

$$F_{01} = iF_{23}, \quad F_{02} = -iF_{13}, \quad F_{03} = iF_{12}$$

or $\vec{E} = -i\vec{H}$, with $\vec{E} = (F_{01},F_{02},F_{03})$ the electric fieldstrength and with $\vec{H} = (-F_{23},F_{13},-F_{12})$ the magnetic fieldstrength. Such a field satisfies automatically the Maxwell equations $dF = d*F = 0$.

This method for obtaining the potentials of a self-dual Maxwell field (the Penrose transform) is very similar to the method of Ward for finding the potentials of a self-dual Yang-Mills field. In order to illustrate this similarity we repeat the construction for getting the potentials B_{PQ}: starting with a function

$$g: W \subset P^3(\mathbb{C}) \to G\ell(1,\mathbb{C}) \cong \mathbb{C} \setminus \{0\}$$

we apply the following transformations

$$g = e^{if} = e^{i\hat{\mu}} \cdot (e^{i\mu})^{-1} \to -i(e^{-i\mu})D_P(e^{i\mu}) = D_P\mu = B_{P1} - \zeta B_{P2} \to F_{\mu\nu}.$$

We will show that (at least locally) the correspondence $g \to F_{\mu\nu}$ is surjective, i.e. for a given field $F_{\mu\nu}(x)$ and a given point $\bar{x} \in \mathbb{R}^4$ there is a

neighbourhood R of $\bar{x}$ in $\mathbb{R}^4$, a domain $W \subset P_3(\mathbb{C})$ and a function g in W such that the above given transformation applied on g yields $F_{\mu\nu}(x)$ for $x \in R$. Thus to a globally defined field $F_{\mu\nu}$ there corresponds a collection $\{g_j\}$ of functions g_j defined in domains W_j and if this collection is interpreted in the right manner, the correspondence $\{g_j\} \to F_{\mu\nu}$ will be bijective. This correspondence is called the *Penrose transformation* and will be studied in a later chapter.

Firstly, we check that the above given map $g \to F_{\mu\nu}$ is well-defined. The ambiguity in writing $f = \hat{\mu} - \mu$ is the addition of a function α_0 of x only: $f(\tilde{x}\pi,\pi) = (\hat{\mu}(x,\zeta) + \alpha_0(x)) - (\mu(x,\zeta) + \alpha_0(x))$. Hence the result of adding $D_P\alpha_0$ to $D_P\mu$ is the addition of a gradient to the potentials B_μ, which does not contribute to the field $F_{\mu\nu}$. Thus the choice of the splitting of f determines the gauge.

Secondly, just as in the case of the potentials A_{PQ} we will express the potentials B_{PQ} in the Laurent coefficients of $f(\tilde{x}\pi,\pi)$. If we write

$$f(\tilde{x}\pi,\pi) = \sum_{k=-\infty}^{\infty} \alpha_k(x)\zeta^k$$

we have immediately a splitting of f:

$$\mu(x,\zeta) = \sum_{k=0}^{\infty} \alpha_k(x)\zeta^k, \qquad \hat{\mu}(x,\zeta) = \sum_{k=-1}^{-\infty} \alpha_k(x)\zeta^k,$$

where the term α_0 could have been included in $\hat{\mu}$, too. The function μ (and also $\hat{\mu}$) cannot be arbitrary but must be such that $D_P\mu$ depends only linearly on ζ for $P = 1,2$. This is the only property that matters and we do not even have to bother about the convergence of the series μ. So we may consider a formal power series $\mu = \sum_{k\geq 0} \alpha_k\zeta^k$ as well such that in the power series $D_P\mu$ all terms except the first two vanish for $P = 1,2$. This means that the following conditions for the coefficients α_k have to hold

(31) $\partial_{P1}\alpha_{k+1} = \partial_{P2}\alpha_k$, $k = 1,2,\ldots$, $P = 1,2.$

This is an overdetermined system which can have a solution only if α_k satisfies

$$\partial_{21}\partial_{12}\alpha_k = \partial_{21}\partial_{11}\alpha_{k+1} = \partial_{11}\partial_{21}\alpha_{k+1} = \partial_{11}\partial_{22}\alpha_k,$$

thus

(32) $(\partial_{11}\partial_{22}-\partial_{21}\partial_{12})\alpha_k = 0$ $k = 1,2,\ldots$.

For real x this equation is

$$\Box\alpha_k = 0, \quad \Box = \partial_\mu\partial^\mu \text{ the Laplace operator.}$$

So all coefficients α_k must be harmonic, thus certainly analytic, in R for $k = 1,2, \ldots$. Summarizing, we have found that a formal power series whose coefficients α_k satisfy (31) is such that

$$D_P \sum_{k\geq 0} \alpha_k\zeta^k = \partial_{P1}\alpha_0 - \zeta(\partial_{P2}\alpha_0-\partial_{P1}\alpha_1), \qquad P = 1,2.$$

We have seen that then the potentials

(33) $B_{P1} = \partial_{P1}\alpha_0$, $B_{P2} = \partial_{P2}\alpha_0 - \partial_{P1}\alpha_1$, $P = 1,2$

satisfy the self-duality equations (29) and (30).

Conversely, can all solutions of (29) and (30) be obtained in this way? In other words, the question is whether, for given B_{PQ} satisfying (29) and (30), the overdetermined systems (33) and (31) can be solved. In order that the first system in (33) has a solution α_0 the equation

$$\partial_{21}B_{11} = \partial_{21}\partial_{11}\alpha_0 = \partial_{11}\partial_{21}\alpha_0 = \partial_{11}B_{21}$$

should hold and this is just the first equation in (29) for B_{PQ}. If α_0 is a solution then for the second system in (33) to have a solution α_1 we

must have

$$\partial_{21}\partial_{11}\alpha_1 = \partial_{21}\partial_{12}\alpha_0 - \partial_{21}B_{12} = \partial_{12}B_{21} - \partial_{21}B_{12} =$$

$$= \partial_{11}\partial_{21}\alpha_1 = \partial_{11}\partial_{22}\alpha_0 - \partial_{11}B_{22} = \partial_{22}B_{11} - \partial_{11}B_{22}$$

and this is just equation (30). Since B_{PQ} also satisfies the second equation in (29) a solution α_1 will be harmonic:

$$\partial_{22}\partial_{11}\alpha_1 = \partial_{22}\partial_{12}\alpha_0 - \partial_{22}B_{12} = \partial_{12}(\partial_{22}\alpha_0 - B_{22}) = \partial_{12}\partial_{21}\alpha_1.$$

This is the compatibility condition for the system (31). This system can be solved for successive k, for if α_k satisfies the compatibility condition (32) then also a solution α_{k+1} of (31):

$$\partial_{22}\partial_{11}\alpha_{k+1} = \partial_{22}\partial_{12}\alpha_k = \partial_{12}\partial_{22}\alpha_k = \partial_{12}\partial_{21}\alpha_{k+1}.$$

Summarizing, we can say that (29) and (30) are just the compatibility conditions for the overdetermined systems (33) and (31). For a given B_{PQ} satisfying (29) and (30) the systems (33) and (31) can be solved at least locally (namely in simply connected domains). Hence all soluti ns of (29) and (30) are given, at least locally, by (33) for an arbitrary function α_0 and a harmonic function α_1.

REMARK. Of course, one can also check immediately that functions B_{PQ} given by (33) with harmonic α_1 satisfy the equations (29) and (30). Then one does not need to consider the system (31) in order to show that (33) can be obtained from a formal power series μ, which guarantees that (29) and (30) is satisfied. However, it is just this method that will be applied to the potentials A_{PQ} in part C of this section.

We have not yet shown the surjectivity of the (local) Penrose transformation $g \to F_{\mu\nu}$. For that purpose one needs a solution $\mu = \Sigma \alpha_k(x)\zeta^k$ of the systems (33) and (31) which is a convergent power series for $|\zeta| < r_1$ in a

neighbourhood of a point $\bar{x} \in \mathbb{R}^4$, and also one needs a solution

$$\hat{\mu} = \sum_{k=0}^{-\infty} \hat{\mu}_k(x)\zeta^k$$

of the system

$$\begin{aligned} \partial_{P2}\hat{\alpha}_0 &= B_{P2} \\ \partial_{P2}\hat{\alpha}_{-1} &= \partial_{P1}\hat{\alpha}_0 - B_{P1} \qquad\qquad P = 1,2 \\ \partial_{P2}\hat{\alpha}_{k-1} &= \partial_{P1}\hat{\alpha}_k, \quad k = -1, -2, \ldots \end{aligned}$$

which is a convergent power series for $|\zeta| > r_2$ in a neighbourhood of the point $\bar{x}$ with $r_2 < r_1$. We will now show that this indeed is possible.

A solution α_1 of (33) is harmonic so certainly analytic. Assume that $\alpha_1(z)$ is holomorphic in an open neighbourhood of the closed neighbourhood $\{z \mid |z^{PQ} - \tilde{z}^{PQ}| \leqslant \delta_{PQ}\} \subset \mathbb{C}^4$ of the point $\tilde{z}$. Then in this neighbourhood

$$\alpha_{k+1}(z) = \int_{\tilde{z}^{11}}^{z^{11}} \partial_{12}\alpha_k(z_1^{11}, z^{12}, z^{21}, z^{22})dz_1^{11} + \int_{\tilde{z}^{21}}^{z^{21}} \partial_{22}\alpha_k(\tilde{z}^{11}, z^{12}, z_1^{21}, z^{22})dz_1^{21}$$

is a solution of (31) if α_k satisfies (32) there. So starting with α_1 we find

$$\alpha_{k+1}(z) = \int_{\tilde{z}^{11}}^{z^{11}} \cdots \int_{\tilde{z}^{11}}^{z_k^{11}} \partial_{12}^k\alpha_1(z_1^{11}, z^{12}, z^{21}, z^{22})dz_1^{11}dz_2^{11}\cdots +$$

$$+ \sum_{j=0}^{k-1} \frac{(z^{11} - \tilde{z}^{11})^j}{j!} \int_{\tilde{z}^{21}}^{z^{21}} \cdots \int_{\tilde{z}^{21}}^{z_{k-j}^{21}} \partial_{12}^j\partial_{22}^{k-j}\alpha_1(\tilde{z}^{11}, z^{12}, z_1^{21}, z^{22})dz_1^{21}dz_2^{21}\cdots .$$

Let

$$\gamma = \min\left\{\frac{\delta_{12}}{\delta_{11}}, \frac{\delta_{22}}{\delta_{21}}\right\}, \qquad \gamma' = \max\left\{\frac{\delta_{12}}{\delta_{11}}, \frac{\delta_{22}}{\delta_{21}}\right\}$$

and let $m > 1$ be such that $\frac{1}{m}\gamma' < m\gamma$. For

$$z \in \omega = \{z \mid |z^{PQ} - \tilde{z}^{PQ}| \leqslant \frac{\delta_{PQ}}{m+1}\} \subset \mathbb{C}^4$$

we will estimate the holomorphic function $\alpha_{k+1}(z)$ by using Cauchy's formula

$$|\partial^n_{PQ}\alpha_1(z)| \leqslant \frac{n!}{((1-\frac{1}{m+1})\delta_{PQ})^n} \max_{|z^{PQ}-\tilde{z}^{PQ}|\leqslant\delta_{PQ}} |\alpha_1(z)|.$$

We get

$$|\alpha_{k+1}(z)| \leqslant K\left(\frac{1}{m}\frac{\delta_{11}}{\delta_{12}}\right)^k + K\sum_{j=0}^{k-1} \frac{\left(\frac{\delta_{11}}{m+1}\right)^j}{j!} \frac{\left(\frac{\delta_{21}}{m+1}\right)^{k-j}}{(k-j)!} \frac{j!}{\left(\frac{m}{m+1}\delta_{12}\right)^j} \frac{(k-j)!}{\left(\frac{m}{m+1}\delta_{22}\right)^{k-j}} =$$

$$= K\left(\frac{1}{m}\right)^k \sum_{j=0}^{k} \left(\frac{\delta_{11}}{\delta_{12}}\right)^j \left(\frac{\delta_{21}}{\delta_{22}}\right)^{k-j} \leqslant K(k+1)\left(\frac{1}{m\gamma}\right)^k$$

with $K = \max\{|\alpha_1(z)| \mid |z^{PQ} - \tilde{z}^{PQ}| \leqslant \delta_{PQ},\ P = 1,2,\ Q = 1,2\}$.

Hence $\mu = \sum_{k\geqslant 0} \alpha_k(z)\zeta^k$ is convergent for $|\zeta| < r_1 = m\gamma$ for z in the neighbourhood ω of $\tilde{z}$ in $\mathbb{C}^4$. Similarly, we can construct a solution $\hat{\mu} = \sum_{k\leqslant 0} \hat{\alpha}_k(z)\zeta^k$ which converges for $|\zeta| > r_2 = \frac{1}{m}\gamma'$ for $z \in \omega$. Finally, we define a holomorphic function f, homogeneous of degree zero, by

$$f(\tilde{x}\pi,\pi) = \hat{\mu}(x,\zeta) - \mu(x,\zeta)$$

and we set $g = \exp i f$; (for more details cf. the chapter on the Penrose transformation).

REMARK. To get the self-dual (Minkowski) Maxwell field from the potentials B_{PQ} given by (33) the function α_1 has to satisfy the wave equation. For $(x^o, -ix^a)$ is no longer real and (32) transforms to $\Box\alpha_k = 0$ now with $\Box = \partial_\mu\partial^\mu$ with respect to the Minkowski metric, which is the wave operator.

We have seen already that in the solution (33) for the potentials the function α_0 determines the gauge. So essentially the solution depends on one harmonic function α_1. If we choose the gauge α_0 analytic then it follows that B_{PQ} is analytic (if x is real). Hence any solution is es-

sentially real-analytic, i.e. if B_{PQ} is not analytic then one can choose a gauge $\Lambda = \exp(i\alpha_0)$ such that the transformed $B_{PQ} + \partial_{PQ}\alpha_0$ is analytic.

C. *Conditions for the functions* E, F *and* G.

Let $h(x,\zeta) \in S\ell(2,\mathbb{C})$ be an analytic function of $x \in \mathbb{R}$ which is also holomorphic for ζ in a neighbourhood of zero in $\mathbb{C}$, thus h is a (convergent) power series in ζ with values in $S\ell(2,\mathbb{C})$ whose coefficients are analytic functions in $S\ell(2,\mathbb{C})$ of x. We may even consider a formal power series h in ζ with coefficients analytic in x. In both cases the power series h must be such that for $P = 1,2$ the series $h^{-1}D_P h$ contains the zero'th and first order term in ζ only. Then in proof (ii) of Ward's theorem in section 3.2 of the preceeding chapter it has been pointed out that the potentials A_{PQ} given by (3) yield a self-dual field. Conversely, it has been shown in section 3.1 of the preceeding chapter that any solution A_{PQ} of the self-duality equations can be obtained in this way (even with a converging power series h).

If μ denotes a (formal) power series in ζ with analytic coefficients in x such that $D_P\mu = B_{P1} - \zeta B_{P2}$, $P = 1,2$, cf. part B of this section, and if in $h = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ we substitute for a, b, c and d the (formal) power series

$$a = -e^{\mu} \sum_{k=n}^{\infty} \theta_k \zeta^{k-n}$$

$$b = -e^{\mu} \sum_{k=n}^{\infty} \phi_k \zeta^{k-n}$$

$$c = e^{-\mu} \sum_{k=0}^{n} c_k \zeta^{k}$$

$$d = e^{-\mu} \sum_{k=0}^{n} d_k \zeta^{k}$$

(cf. formulae (9), (10) and further), then we have seen in section 2 that $h^{-1}D_P h$ is linear in ζ provided that:

1°. conditions (11) and (12) are satisfied, i.e. the functions θ_k and ϕ_k are linear combinations of the functions c_ℓ and d_ℓ with coefficients Δ_j,

2°. condition (13) is satisfied,

3°. the functions Δ_{n-1}, Δ_{n-2}, ... arising in 1° and 2° satisfy condition (16) for $r = n-2, n-3, \ldots$.

In part A of this section we have seen that c and d can be chosen such that 2° is satisfied if $\det M \neq 0$, cf. (28). So it remains to pay attention to condition (16).

Note that the procedure described here, is not the most general one in order that $h^{-1}D_P h$ is linear in ζ. Hence we do not get all solutions A_{PQ} as in part B of this section.

With the notation $d_{PQ} = \partial_{PQ} + 2B_{PQ}$ condition (16) is

$$d_{P1}\Delta_{k-1} = d_{P2}\Delta_k, \quad k = n-1, n-2, \ldots, \quad P = 1,2. \tag{34}$$

Here the functions B_{PQ} satisfy (29) and (30). It follows from (29) that the operators d_{11} and d_{21} commute and also the operators d_{12} and d_{22}. Just as (31) the overdetermined system (34) can have a solution Δ_{k-1} for given Δ_k only if Δ_k satisfies

$$d_{11}d_{22}\Delta_k = d_{11}d_{21}\Delta_{k-1} = d_{21}d_{11}\Delta_{k-1} = d_{21}d_{12}\Delta_k$$

thus the compatibility condition

$$(d_{11}d_{22} - d_{21}d_{12})\Delta_k = 0, \tag{35}$$

(compare (32)). Condition (30) implies the equality of the operators

$$d_{11}d_{22} - d_{21}d_{12} = d_{22}d_{11} - d_{12}d_{21}.$$

Hence a solution Δ_{k-1} of (34) satisfies (35), too:

$$(d_{11}d_{22}-d_{21}d_{12})\Delta_{k-1} = (d_{22}d_{11}-d_{12}d_{21})\Delta_{k-1} = d_{22}d_{12}\Delta_k - d_{12}d_{22}\Delta_k = 0.$$

So (29), (30) and (35) are the compatibility conditions for the overdetermined system (34).

For a given Δ_{n-1} satisfying (35) the system (34) can be solved: let B_{PQ} be given by (33) then a local solution (in a neighbourhood of the point $\bar{x}$) is

$$\Delta_{k-1} = e^{-2\alpha_0}\int_{\bar{x}^{11}}^{x^{11}} e^{2\alpha_0} d_{12}\Delta_k + e^{-2\alpha_0}\int_{\bar{x}^{21}}^{x^{21}} (e^{2\alpha_0} d_{22}\Delta_k)\Big|_{x^{11}=\bar{x}^{11}}.$$

Hence, besides $\det M \neq 0$, the only condition for the functions E, F and G in (27) is condition (34) (and therefore automatically (35)) for the functions Δ_k from which E, F and G are derived. In the solution (18), (19) and (20) another function Δ_n has been introduced by means of formulae (11) and (12) with $k = 0$, which also satisfies (34) for $k = n$. Thus this function Δ_n has to satisfy (35), too.

For real x equation (35) is elliptic (the principal term is again the Laplace operator) with analytic coefficients, if B_{PQ}, or in fact α_0, has been taken analytic. Hence the functions Δ_k are always analytic. If we choose an analytic gauge $\begin{pmatrix} p_1 & q_1 \\ p_2 & q_2 \end{pmatrix}: \mathcal{R} \to S\ell(2,\mathbb{C})$ in (28), the functions $\vec{c}$ and $\vec{d}$ and hence θ_k and ϕ_k will be analytic. Thus the Yang-Mills potentials A_{PQ} are essentially analytic, i.e. if there is a non-analytic solution then there is a gauge-equivalent solution which is analytic.

Actually, the choice of α_0 is not that important, because the dependence on α_0 can be removed by the gauge. To see this we will trace the effect of writing $f = (\hat{\mu}+\alpha_0) - (\mu+\alpha_0)$ instead of $f = \hat{\mu} - \mu$. From (7)

and (8) it follows that the functions Δ_k are multiplied by $\exp(-2\alpha_0)$. (This follows also directly: if in the operator d_{PQ} the term $\partial_{PQ}\alpha_0$ is added to B_{PQ}, then in (34) this is compensated if Δ_k is multiplied there by $\exp(-2\alpha_0)$). The matrix M is multiplied by $\exp(-2\alpha_0)$; (thus the functions E, F and G by $\exp(2\alpha_0)$ so that α_0 cancels in (27) if there B_{PQ} is replaced by $B_{PQ} + \partial_{PQ}\alpha_0$). If we choose another gauge in (28), namely

$$\begin{pmatrix} e^{\alpha_0}p_1 & e^{\alpha_0}q_1 \\ e^{-\alpha_0}p_2 & e^{-\alpha_0}q_2 \end{pmatrix},$$

the vectors $\vec{c}$ and $\vec{d}$ are multiplied by $\exp(\alpha_0)$. So according to (11) and (12) θ_k and ϕ_k are multiplied by $\exp(-\alpha_0)$. Then in the power series of a, b, c and d the function α_0 does no longer occur. Hence also the solution A_{PQ} given by (18), (19) and (20) is independent of α_0.

Summarizing, we have got the following conditions for the validity of the solution (18), (19) and (20) (hence also (27)):

1^o. the functions B_{PQ} should satisfy (29) and (30)

2^o. the functions $\Delta_{-n}, \Delta_{-n+1}, \ldots, \Delta_n$ should satisfy the overdetermined system (34) for $k = -n+1, \ldots, n$ (whose compatibility condition is (35)).

3^o. $\Delta_{-n+1}, \ldots, \Delta_{n-1}$ should be such that $\det M \neq 0$.

4^o. the vectors c and d should be determined by (28) with $\begin{pmatrix} p_1 & q_1 \\ p_2 & q_2 \end{pmatrix} \in Sl(2,\mathbb{C})$.

REMARK. Contrarily to the case of the self-dual Maxwell fiel in part B of this section the direct verification of the self-duality equations of the solution (18), (19) and (20) under conditions 1^o, 2^o, 3^o and 4^o is much more difficult.

4. EXAMPLES AND SU(2)-FIELDS

A. *Examples.*

To get a simple example of (27) we investigate whether the functions Δ_k can be taken constant so that they do not occur in the potentials A_{PQ}. From (34) it follows that the ratio $\Delta_k/\Delta_{k-1} = \gamma$ must be constant. Then $\Delta_k = \Delta_0\gamma^k$, hence the coefficients Δ_k cannot be those of a convergent Laurent series ρ as in (7) which is holomorphic in an annular domain (for in that case on the one hand one would have $|\zeta| < \gamma$ and on the other hand $|\zeta| > \gamma$). Furthermore the condition $\det M \neq 0$ can hold only if $n = 1$ (and $\Delta_0 \neq 0$) and the functions B_{PQ} have to satisfy (34):

$$B_{P1} = \gamma B_{P2}\,, \qquad P = 1,2.$$

With (29) and (30), hence (33), we get

$$B_{P1} = \partial_{P1}\alpha_0 = \gamma B_{P2} = \gamma\partial_{P2}\alpha_0 - \gamma\partial_{P1}\alpha_1.$$

So we must have $\partial_{P1}(\frac{1}{\gamma}\alpha_0+\alpha_1) = \partial_{P2}\alpha_0$, which is possible for a harmonic α_1 if α_0 is harmonic, too. Then we find

(36) $$B_{P1} = \gamma B_{P2} = \partial_{P1}\alpha_0, \quad \alpha_0 \text{ harmonic.}$$

Hence there exists a solution!

Let us now be a little less restrictive by assuming, again for $n = 1$, only $\Delta_0 = 1$. In order that Δ_{-1} and Δ_1 can be determined from (34) the function Δ_0 has to satisfy (35) which in this case is

(37) $$\partial_{11}B_{22} - \partial_{21}B_{12} + 2B_{11}B_{22} - 2B_{21}B_{12} = 0.$$

Of course, the functions (36) are a solution of (37) but there are more general ones. Note that besides (37) the functions B_{PQ} have to satisfy (29) and (30), too. The solution (27) (with $n = 1$, $F = G = E = 1/\Delta_0$) becomes

$$A_{P1} = i\begin{pmatrix} B_{P1} & 0 \\ -2B_{P2} & -B_{P1} \end{pmatrix}, \quad A_{P2} = i\begin{pmatrix} -B_{P2} & -2B_{P1} \\ 0 & B_{P2} \end{pmatrix}, \quad P = 1,2.$$

As a special case we take $\alpha_1 = 0$ in (33), thus $B_{PQ} = \partial_{PQ}\alpha_0$ for a function α_0 which satisfies (37) or

$$(38) \qquad (\partial_{11}\partial_{22} - \partial_{21}\partial_{12})\alpha_0 + 2\partial_{11}\alpha_0 \cdot \partial_{22}\alpha_0 - 2\partial_{21}\alpha_0 \cdot \partial_{12}\alpha_0 = 0.$$

Then we find the potentials

$$A_{P1} = i\begin{pmatrix} \partial_{P1} & 0 \\ -2\partial_{P2} & -\partial_{P1} \end{pmatrix}\alpha_0, \quad A_{P2} = i\begin{pmatrix} -\partial_{P2} & -2\partial_{P1} \\ 0 & \partial_{P2} \end{pmatrix}\alpha_0, \quad P = 1,2.$$

We have already pointed out that the solution A_{PQ} can also be given in a form independent of the gauge function α_0 of the potentials B_{PQ}. This form of the above given solution is obtained by taking $B_{PQ} = 0$ (hence $f = 0$), $n = 1$ and $E = F = G = 1/\Delta_0 = 1/\phi$. Then the function ϕ has to satisfy (35) which in this case means that ϕ must be harmonic. The potentials are

$$(39) \qquad A_{P1} = \frac{i}{2\phi}\begin{pmatrix} \partial_{P1} & 0 \\ -2\partial_{P2} & -\partial_{P1} \end{pmatrix}\phi, \quad A_{P2} = \frac{i}{2\phi}\begin{pmatrix} -\partial_{P2} & -2\partial_{P1} \\ 0 & \partial_{P2} \end{pmatrix}\phi, \quad P = 1,2$$

ϕ harmonic.

This is the same formula as above, for equation (38) can be written as

$$\tfrac{1}{2}e^{-2\alpha_0}(\partial_{11}\partial_{22} - \partial_{21}\partial_{12})e^{2\alpha_0} = 0,$$

so $\phi = \pm \exp(2\alpha_0)$ must be harmonic. Note that the singularities in the points where ϕ vanishes are due to the gauge.

B. SU(2)-*fields*.

Condition (14) assures us that the trace of the matrices A_{PQ} vanishes.

We moreover want that A_μ is hermitian so that $iA_\mu \in su(2)$. If the matrix g is equivalent with a matrix $\tilde{g}$ (i.e. $\tilde{g} = \hat{m}gm$, cf. the end of section 1) which satisfies condition (iii) of the theorem in section 1, then the matrices A_{PQ} in (27) satisfy the required properties (perhaps after performing a gauge transformation $\Lambda: R \to SL(2,\mathbb{C})$). We will give two sets of conditions on the functions f and Γ in order that such $\hat{m}$ and m can be found. Each set of conditions is sufficient for this but it is certainly not necessary that one of these sets of conditions has to hold.

I. The conditions are:

1^o. f imaginary, i.e. $\overline{f(Z)} = -f(Z^*)$.

2^o. Γ real, i.e. $\overline{\Gamma(Z)} = \Gamma(Z^*)$

3^o. n odd.

Then we set $\hat{m} = I$ and $m = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ so that

$$\begin{pmatrix} \zeta^n e^f & \Gamma \\ 0 & \zeta^{-n} e^{-f} \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \Gamma & -\zeta^n e^f \\ \zeta^{-n} e^{-f} & 0 \end{pmatrix} = \tilde{g}(Z)$$

which satisfies (iii) $\tilde{g}(Z)^* = \tilde{g}(Z^*)$, for $-\overline{\zeta^n} = (\zeta^{-n})^*$ if n is odd.

Condition 1^o means that the Maxwell potentials B_μ must be real. To see this one reasons similarly to part 2) of the proof of the theorem in section 1:

$$\overline{f(\tilde{x}\pi,\pi)} = \overline{\hat{\mu}(x,\zeta)} - \overline{\mu(x,\zeta)} = -f(\tilde{x}\pi^*,\pi^*) = -\hat{\mu}(x,w) + \mu(x,w)$$

with $w = \zeta^* = -1/\bar{\zeta}$, hence

$$\overline{\hat{\mu}(x,\zeta)} - \mu(x,w) = \overline{\mu(x,\zeta)} - \hat{\mu}(x,w) = i\lambda(x)$$

where λ is a real function of x alone. Choosing new $\hat{\mu}$ and μ by adding $\frac{1}{2}i\lambda$ we get

$$\overline{\mu(x,\zeta) + \tfrac{1}{2}i\lambda} - (\hat{\mu}(x,w) + \tfrac{1}{2}i\lambda) = i\lambda - \tfrac{1}{2}i\lambda - \tfrac{1}{2}i\lambda = 0;$$

thus we may assume that

$$\overline{\mu(x,\zeta)} = \hat{\mu}(x,w).$$

Then for the potentials B_{PQ} we have

$$\overline{B_{11}} - \bar{\zeta}\overline{B_{12}} = \overline{D_1(\zeta)\mu(x,\zeta)} = \bar{\zeta}D_2(w)\hat{\mu}(x,w) = \bar{\zeta}B_{12} + B_{22}.$$

Hence $\overline{B_{11}} = B_{22}$ and $B_{21} = -\overline{B_{12}}$ and from (4) it follows that $B_\mu = \bar{B}_\mu$ is real. If in (33) we take $\alpha_1 = 0$ the foregoing says that the function α_0 is real except for an additive imaginary constant.

Condition 2^o implies that also Δ_0 is a real function:

$$\Delta_0(x) = \frac{-1}{2\pi i}\oint \frac{e^{-\overline{\hat{\mu}(x,\zeta)}-\overline{\mu(x,\zeta)}}\overline{\Gamma(\tilde{x}\pi,\pi)}}{\bar{\zeta}}d\bar{\zeta} =$$

$$= \frac{1}{2\pi i}\oint e^{-\mu(x,w)-\hat{\mu}(x,w)}\Gamma(\tilde{x}\pi^*,\pi^*)w\frac{dw}{w^2} = \Delta_0(x).$$

Thus in (39) ϕ must be real. Then this yields the Corrigan-Fairlie-'t Hooft-Wilczek "Ansatz":

$$iA_{P1} = -\tfrac{1}{2}\begin{pmatrix} \partial_{P1} & 0 \\ -2\partial_{P2} & -\partial_{P1}\end{pmatrix}\log|\phi|, \quad iA_{P2} = -\tfrac{1}{2}\begin{pmatrix} -\partial_{P2} & -2\partial_{P1} \\ 0 & \partial_{P2}\end{pmatrix}\log|\phi|,$$

$$P = 1,2$$

for a real harmonic function ϕ. With (4) the potentials iA_μ are obtained.

Since the action integral must be finite we let ϕ vanish at infinity, but then R' (where ϕ is defined) cannot be the whole $\mathbb{R}^4$. For example, choose

$$\phi(x) = \frac{1}{x_\mu x^\mu} = \frac{1}{(x^0)^2+(x^1)^2+(x^2)^2+(x^3)^2}, \quad R' = \mathbb{R}^4\setminus\{0\}.$$

The potentials A_{PQ} then have a singularity in 0, but this is not found in the associated field $F_{\mu\nu}$. The field vanishes and the potentials are

pure gauge:

$$A_{PQ} = -i\Lambda^{-1}\partial_{PQ}\Lambda$$

with

$$(40) \qquad \Lambda = \frac{i}{r}\begin{pmatrix} -x^{21} & x^{22} \\ x^{11} & -x^{12} \end{pmatrix} \in SU(2), \quad r = \sqrt{x_\mu x^\mu} = \sqrt{x^{11}x^{22} - x^{12}x^{21}}\,.$$

We have to find such $su(2)$-fields $iF_{\mu\nu}$ that the action

$$S = -\int_{S^4} \operatorname{trace}(\Omega \wedge *\Omega) = \sum_{\mu,\nu} \int_{\mathbb{R}^4} \det iF_{\mu\nu}\, dx^0 dx^1 dx^2 dx^3$$

is finite, where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu]$.
With the above given potentials A one finds

$$S = \int -\tfrac{1}{2}\Box\Box \log|\phi|\,.$$

The 't Hooft-solution is obtained by taking

$$\phi(x) = 1 + \sum_{j=1}^{k} \frac{\lambda_j^2}{(x-x_j)^2}$$

which is singular in the points $x_j = (x_j^0, x_j^1, x_j^2, x_j^3) \in \mathbb{R}^4$, $j = 1,\ldots,k$.
Since for $r \to \infty$ $\Box\Box \log r \sim \Box\, 1/r^2 \sim 1/r^4$ this ϕ will give a finite action S. According to Uhlenbeck's theorem in the preceding chapter the singularities must be pure gauge; thus the associated field $F_{\mu\nu}$ is smooth in $\mathbb{R}^4$. The gauge Λ_j that removes the singularity at the point x_j is given by (40) for $x_j = 0$.

The action S can be calculated by integrating over $\mathbb{R}^4$ minus small balls around the points x_j and by letting the radii of the balls tend to zero. In virtue of Gauss' formula it remains to integrate the normal derivatives of $-\tfrac{1}{2}\Box \log|\phi|$ over the surfaces of the balls (pointed inwards) and only the contributions with $1/r_j^3$ in the j^{th} ball yield a non-van-

ishing term. We get

$$S = \sum_{j=1}^{k} 2\pi^2 r_j^3 \frac{-\partial}{\partial r_j}(-\tfrac{1}{2})\left(\frac{\partial^2}{\partial r^2} - \frac{3}{r_j}\frac{\partial}{\partial r_j}\right) \log\left(1 + \frac{\lambda_1^2}{r_1^2} + \ldots + \frac{\lambda_k}{r_k^2}\right)\Big|_{r_j=0} = 8\pi^2 k.$$

Hence for any instanton number k we have an explicit solution.

II. 1^o. f real.

2^o. $\Gamma = H^{-1}(e^f + (-1)^n e^{-f})$ with $H = (\pi_1\pi_2)^{-n}P(Z)$ where P is a homogeneous polynomial of degree $2n$ with $\overline{P(Z)} = (-1)^n P(Z^*)$.

Then we take $m = \begin{pmatrix} 0 & -1 \\ 1 & \zeta^n H \end{pmatrix}$ and

$$\tilde{g} = \begin{pmatrix} \zeta^n e^f & \Gamma \\ 0 & \zeta^{-n} e^{-f} \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & \zeta^n H \end{pmatrix} = \begin{pmatrix} \Gamma & (-1)^n \zeta^n e^{-f} \\ \zeta^{-n} e^{-f} & H e^{-f} \end{pmatrix}$$

satisfies condition (iii) of the theorem in section 1. Here $m(\tilde{x}\pi,\pi)$ must be holomorphic for $\zeta \in U$ and indeed the function $\zeta^n H(\tilde{x}\pi,\pi) = (\pi_2)^{-2n} P(\tilde{x}\pi,\pi)$ is holomorphic for $\zeta \in U$ (because $\pi_2 = 0$ in $\hat{U}$).

With these conditions multi-monopole solutions can be obtained. For monopole solutions the fields are required to be independe t of x_0. If $g(\omega_1,\omega_2,\pi_1,\pi_2)$ is a function of $\zeta = \pi_1/\pi_2$ and of $\gamma = -i\omega_1/\pi_1 + i\omega_2/\pi_2$, then $g(\tilde{x}\pi,\pi)$ does not depend on x_0.

Substitution of $\omega = \tilde{x}\pi$ in γ yields

$$\gamma(x,y,z,\zeta) = \xi\zeta - 2z - \bar{\xi}\zeta^{-1}$$

with $\xi = x + iy$, $x = x^1$, $y = x^2$, $z = x^3$. The function γ is then real in the sense that

$$\overline{\gamma(x,y,z,\zeta)} = \gamma(x,y,z,w) \quad \text{with} \quad w = -\frac{1}{\bar{\zeta}}.$$

Hence for getting monopole solutions we only need to consider functions g of γ and ζ. It turns out that the "Ansatz" A_n yield monopole solutions

with charge n (Ward) and even every n-monopole solution can be obtained from A_n (Hitchin).

REFERENCES

Besides the two references of the preceding chapter we mention

PRASAD, M.K., *Instantons and monopoles in Yang-Mills gauge field theories*. Physica 1D(1980), p.167-191.

LINEAR FIELDS AND YANG-MILLS THEORY

by

R. MARTINI
Department of Mathematics
Twente University of Technology

1. BACKGROUND IN TWISTOR THEORY

M is the Minkowski space: $\|x\|^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$.

$M_{\mathbb{C}}$ is the complexified Minkowski space with holomorphic metric:
$\|z\|^2 = (z^0)^2 - (z^1)^2 - (z^2)^2 - (z^3)^2$.

We have introduced 2-dimensional complex spaces $\mathbb{C}^A$, $\mathbb{C}^{A'}$ with duals $\mathbb{C}_A$ and $\mathbb{C}_{A'}$ respectively. In these spaces we have complex conjugation

$$\begin{array}{ll} \mathbb{C}^A \longleftrightarrow \mathbb{C}^{A'} & \mathbb{C}_A \longleftrightarrow \mathbb{C}_{A'} \\ \xi^A \longmapsto \bar{\xi}^{A'} & \xi_A \longmapsto \bar{\xi}_{A'} \\ \bar{\xi}^A \longmapsto \xi^{A'} & \bar{\xi}_A \longmapsto \xi_A \end{array} .$$

The complexified Minkowski space may be identified with the space $\mathbb{C}^{AA'} = \mathbb{C}^A \otimes \mathbb{C}^{A'}$ by the isomorphism

$$(z^\mu) \longmapsto (z^{AA'}) = \frac{1}{\sqrt{2}} \begin{pmatrix} z^0+z^1 & z^2+iz^3 \\ z^2-iz^3 & z^0-z^1 \end{pmatrix} .$$

We introduce skew-symmetric forms ε_{AB} and $\varepsilon_{A'B'}$ on $\mathbb{C}^A$ and $\mathbb{C}^{A'}$ by defining

$$\varepsilon_{AB} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \varepsilon_{A'B'} .$$

We can use these forms to raise and lower indices, such as

$$\xi_A = \xi^B \varepsilon_{BA}, \quad \xi^A = \varepsilon^{AB} \xi_B, \quad \xi_{A'} = \xi^{B'} \varepsilon_{B'A'}, \quad \xi^{A'} = \varepsilon^{A'B'} \xi_{B'} ,$$

with $\varepsilon^{AB} = \varepsilon^{A'B'} = \varepsilon_{AB}$.

The twistor space $\mathbb{T} = \mathbb{C}_{A'} \oplus \mathbb{C}^{A'}$ and the projective twistor space is $\mathbb{P} = P(\mathbb{T})$. On $\mathbb{T}$ there is defined the Hermitian form Φ with $\Phi(Z) = \omega^{A}\bar{\pi}_{A} + \bar{\omega}^{A'}\pi_{A'}$ if $Z = (\omega^{A}, \pi_{A'})$.

$\mathbb{T}$ is divided by Φ into three parts $\mathbb{T}^{+}$, $\mathbb{T}^{0}$, $\mathbb{T}^{-}$ and $\mathbb{P}$ in corresponding parts $\mathbb{P}^{+}$, $\mathbb{P}^{0}$, $\mathbb{P}^{-}$.

The matrix $(z^{AA'})$ may be identified with a linear transformation $\mathbb{C}_{A'} \to \mathbb{C}^{A}$ given by $\pi_{A'} \longmapsto iz^{AA'}\pi_{A'}$. A point $(z^{AA'})$ of the complexified Minkowski space give rise to a 2-plane in $\mathbb{T}$, given by

$$\{(\omega^{A}, \pi_{A'}) \in \mathbb{T} \mid \omega^{A} = iz^{AA'}\pi_{A'}\} .$$

2. MASSLESS FIELDS

Prototype of the massless field equations are Maxwell's equations

$$(2.1)\qquad \begin{cases} \operatorname{rot} E + \dfrac{\partial B}{\partial t} = 0 \quad , & E = \text{electric field} \\ \operatorname{rot} B - \dfrac{\partial E}{\partial t} = 0 \quad , & B = \text{magnetic field} \\ \operatorname{div} E = 0 \\ \operatorname{div} B = 0 \end{cases}$$

We rewrite these equations as follows $(t=x^0)$. Define the skew-symmetric matrix

$$F_{ab} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix}$$

and set

$$F = F_{ab}\, dx^a \wedge dx^b .$$

Then it is not difficult to see that

i) $dF = 0 \iff \operatorname{rot} E + \dfrac{\partial B}{\partial t} = 0$ and $\operatorname{div} B = 0$

ii) $d{*}F = 0 \iff \operatorname{rot} B - \dfrac{\partial E}{\partial t} = 0$ and $\operatorname{div} E = 0$

Here $*$ is the Hodge $*$-operator: $*: \Lambda^2 M \to \Lambda^2 M$. Note that $*^2 = -1$ in Minkowski space M in contrast with $*^2 = 1$ in the Euclidean case. $\Lambda^2 M$ denotes the space of complex-valued 2-forms on M. So Maxwell's equations can be written in the form

$$(2.3)\qquad \begin{cases} dF = 0 \\ d{*}F = 0 \end{cases}$$

Splitting up: $\Lambda^2 M = \Lambda^2_+ M \oplus \Lambda^2_- M$

with

$$\Lambda^2_\pm M = \{F \in \Lambda^2 M \mid *F = \pm iF\}$$

and writing correspondingly:

$$F = \tfrac{1}{2}(F - i*F) + \tfrac{1}{2}(F + i*F) = F_+ + F_- .$$

We can write Maxwell's equations in the form

$$(2.4) \qquad \begin{cases} dF_+ = 0 \\ \\ dF_- = 0 \end{cases}$$

We express the last equations (2.4) into spinor coordinates $(x^{AA'})$. In terms of spinors F takes the form

$$(2.5) \qquad F = F_{AA'BB'}\, dx^{AA'} \wedge dx^{BB'}$$

Some elementary linear algebraic calculations yield the following relation between $[F_{ab}]$ and $[F_{AA'BB'}]$

$$(2.6) \qquad \begin{cases} F_{00'11'} = -F_{01} \\ F_{00'01'} = \tfrac{1}{2}\left[F_{02} + F_{12} - i(F_{03}+F_{13})\right] \\ F_{01'11'} = \tfrac{1}{2}\left[F_{12} - F_{02} + i(F_{03}-F_{13})\right] \end{cases}$$

$$(2.6a) \qquad \begin{cases} F_{01'10'} = iF_{23} \\ F_{00'10'} = \tfrac{1}{2}\left[F_{02} + F_{12} + i(F_{03}+F_{13})\right] \\ F_{10'11'} = \tfrac{1}{2}\left[F_{12} - F_{02} - i(F_{03}-F_{13})\right] \end{cases}$$

and of course

$$F_{AA'BB'} = -F_{BB'AA'} .$$

So $$F_{AA'BB'} = \tfrac{1}{2}(F_{AA'BB'} - F_{BB'AA'}) = \tfrac{1}{2}(F_{AA'BB'} - F_{BA'AB'} + F_{BA'AB'} - F_{BB'AA'}).$$

It is easy to check that:

$$F_{AA'BB'} - F_{BA'AB'} = \varepsilon_{AB} F_{CA'}{}^{C}{}_{B'}$$

$$F_{BA'AB'} - F_{BB'AA'} = \varepsilon_{A'B'} F_{BC'A}{}^{C'}$$

and setting

$$\phi_{AB} = \tfrac{1}{2} F_{BC'A}{}^{C'}, \quad \psi_{A'B'} = \tfrac{1}{2} F_{CA'}{}^{C}{}_{B'}$$

we have

$$\phi_{AB} = \phi_{BA}, \quad \psi_{A'B'} = \psi_{B'A'}$$

and moreover

(2.7) $$F_{AA'BB'} = \varepsilon_{AB}\psi_{A'B'} + \varepsilon_{A'B'}\phi_{AB}$$

In this setting we have the following result.

THEOREM 1. i) F *is real if and only if* $\psi_{A'B'} = \bar{\phi}_{AB}$

ii) F *is selfdual if and only if* $\phi_{AB} = 0$

iii) F *is anti-selfdual if and only if* $\psi_{A'B'} = 0$.

The proof is computational and depends on the following formula for the *-operator in terms of spinors

(2.8) $$(*F)_{AA'BB'} = i(\varepsilon_{AB}\psi_{A'B'} - \varepsilon_{A'B'}\phi_{AB})$$

So in the splitting $F = F_+ + F_-$ of F is selfdual and anti-selfdual part we have

(2.9) $$F_+ = \varepsilon_{AB}\psi_{A'B'}, \quad F_- = \varepsilon_{A'B'}\phi_{AB}$$

The six independent components of (F_{ab}) are replaced by the six indepen-

dent spinor quantities ϕ_{00}, ϕ_{01}, ϕ_{11}, $\psi_{0'0'}$, $\psi_{0'1'}$, $\psi_{1'0'}$, $\psi_{1'1'}$ and the relations between the (F_{ab}) and the $(\phi_{AB},\psi_{A'B'})$ is of linear algebraic nature.

Now we can rewrite Maxwell's equations in terms of spinors. Let $(x^{AA'})$ be spinor coordinates on M. We set

$$(2.10) \qquad \nabla_{AA'} = \frac{\partial}{\partial x^{AA'}} \quad , \quad \nabla^{AA'} = \frac{\partial}{\partial x_{AA'}} \, ,$$

where $x_{AA'} = x^{BB'}\varepsilon_{BA}\varepsilon_{B'B'}$

so that $\nabla^{AA'} = \varepsilon^{BA}\varepsilon^{B'A'}\nabla_{BB'}$.

THEOREM 2. *In terms of spinors Maxwell's equations take the form*

$$(2.11.a) \qquad \nabla^{AA'}\phi_{AB} = 0 \quad , \quad \phi_{AB} \text{ symmetric}$$

$$(2.11.b) \qquad \nabla^{AA'}\psi_{A'B'} = 0, \quad \psi_{A'B'} \text{ symmetric}$$

This theorem is proved by direct computations. As observed before there are six independent quantities $(\phi_{AB},\psi_{A'B'})$ and there are eight independent equations corresponding to the original form (2.1). Note that we are considering complex solutions. Written out (2.11.a) becomes

$$\begin{cases} \nabla^{00'}\phi_{00} + \nabla^{10'}\phi_{10} = 0 \\ \nabla^{00'}\phi_{01} + \nabla^{10'}\phi_{11} = 0 \\ \nabla^{01'}\phi_{00} + \nabla^{11'}\phi_{10} = 0 \\ \nabla^{01'}\phi_{01} + \nabla^{11'}\phi_{11} = 0 \end{cases}$$

and similar form for (2.11.b). In the form (2.11) Maxwell's equations are easy to generalize to

$$(2.12.a) \qquad \nabla^{AA'}\phi_{AB\ldots D} = 0$$

(2.12.b) $\nabla^{AA'}\psi_{A'B'\ldots D} = 0$

where $\phi_{AB\ldots D}$ and $\psi_{A'B'\ldots D'}$ are symmetric spinor quantities with $2s$ indices, $s = \frac{1}{2}, 1, \frac{3}{2}, \ldots$.

(2.12) are the massless field equations for positive helicities.

For $s = \frac{1}{2}$ this equations describe a neutrino and are known as the Dirac-Weyl equations of an electron with mass zero.

For $s = 1$ the equations describe a photon and as we have already seen these are Maxwell's equations above.

For $s = 2$ the equations describe a graviton and these are the linearized Einstein equations.

3. SOME HOLOMORPHIC LINE BUNDLES OVER $P_n(\mathbb{C})$. THE DOLBEAULT COMPLEX

Let X be a complex manifold and $V \xrightarrow{\pi} X$ a holomorphic vector bundle of rank r. Then there exists an open covering $\{U_\alpha\}$ of X and local trivializations (gauges) $\phi_\alpha: \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{C}^r$ and corresponding holomorphic transition functions (gauge transformations) $g_{\alpha\beta} = \phi_\alpha \circ \phi_\beta^{-1}$ with $g_{\alpha\beta}: U_\alpha \cap U_\beta \to GL(r,\mathbb{C})$ satisfying the cocycle condition

$$g_{\alpha\beta}(x)g_{\beta\gamma}(x) = g_{\alpha\gamma}(x), \quad \forall x \in U_\alpha \cap U_\beta \cap U_\gamma \tag{3.1}$$

Conversely, a covering $\{U_\alpha\}$ and a system $\{g_{\alpha\beta}\}$ consisting of holomorphic mappings $g_{\alpha\beta}: U_\alpha \cap U_\beta \to GL(r,\mathbb{C})$ satisfying the cocycle condition determine a (up to equivalence) unique holomorphic vector bundle.

Now consider the complex projective space $P_n(\mathbb{C})$ and let $[z] = [z^0, z^1, \ldots, z^n]$ be homogeneous coordinates of a point in $P_n(\mathbb{C})$. Set $U_\alpha = \{[z] \mid z_\alpha \neq 0\}$.

Over $P_n = P_n(\mathbb{C})$ we construct the canonical line bundle L as follows: the fibre $L_{[z]}$ over the point $[z] \in P_n$ consists of all points on the line through zero and the point $z = (z^0, z^1, \ldots, z^n)$ in $\mathbb{C}^n$

$$\pi(z^0, z^1, \ldots, z^n) = [z^0, z^1, \ldots, z^n], \quad \phi_\alpha: \pi^{-1}(U_\alpha) \to \mathbb{C} .$$

is defined by

$$\phi_\alpha(\lambda(z^0, z^1, \ldots, z^n)) = \phi_\alpha\left(\lambda z^\alpha\left(\frac{z^0}{z^\alpha}, \frac{z^1}{z^\alpha}, \ldots, 1, \ldots, \frac{z^n}{z^\alpha}\right)\right) = \lambda z^\alpha .$$

For the transition function we have: $g_{\alpha\beta}([z]) = \dfrac{z^\alpha}{z^\beta}$.

Evidently these $\{g_{\alpha\beta}\}$ satisfy the cocycle condition.

More generally, the transition functions $\{g_{\alpha\beta}^m\}$ given by

$$g_{\alpha\beta}^m([z]) = \left(\frac{z^\alpha}{z^\beta}\right)^m \tag{3.2}$$

define a holomorphic line bundle over $P_n(\mathbb{C})$, denoted by L^m. L^0 is the trivial line bundle $\cong P_n \times \mathbb{C}$. L^{-1} is called the *hyperplane bundle* and is denoted by H.

On a complex manifold X there are C^∞ differential forms of type (p,q). In local holomorphic coordinates they are described by

$$(3.3) \qquad \phi = \sum_{\substack{i_1<\ldots<i_p \\ j_1<\ldots<j_q}} \phi_{i_1\ldots i_p j_1\ldots j_q}\, dz^{i_1} \wedge \ldots \wedge dz^{i_p} \wedge d\bar{z}^{j_1} \wedge \ldots \wedge d\bar{z}^{j_q},$$

where $\phi_{i_1\ldots i_p j_1\ldots j_q}$ are C^∞ functions and $dz^j = dx^i + idy^j$, $d\bar{z}^j = dx^j - idy^j$.

$E^{p,q}(X)$ denotes the differential forms of type (p,q) on X and $E^r(X)$ the complex-valued differential forms of degree r on X. Then there is a natural projection $\pi_{p,q}: E^{p+q}(X) \to E^{p,q}(X)$. This induces the following differential operator complex

$$(3.4) \qquad \ldots \to E^{p,q-1}(X) \xrightarrow{\bar{\partial}} E^{p,q}(X) \xrightarrow{\bar{\partial}} E^{p,q+1}(X) \to \ldots$$

with $\bar{\partial}\phi = (\pi_{p,q+1} \circ d)\phi$, where d is the usual exterior derivative of the differential form ϕ of degree p+q. There holds $\bar{\partial}^2 = 0$. The complex (3.4) is called the *Dolbeault complex*.

The holomorphic functions on X are denoted by $\mathcal{O}(X)$ and are the kernel of the mapping $\bar{\partial}$, acting on functions (differential forms of type (0,0)), i.e. the sequence

$$(3.5) \qquad 0 \to \mathcal{O}(X) \to E^{0,0}(X) \xrightarrow{\bar{\partial}} E^{0,1}(X)$$

is exact. The obstruction to (3.5) being exact for $(p,q) \neq (0,0)$ is part of the cohomology which we will meet below.

More generally we may consider differential forms of type (p,q) with coefficients in a holomorphic vector bundle V over X, denoted by $E^{p,q}(X)$

and the $\bar{\partial}$-operator can be extended in a natural way to act on this more general forms. We obtain the more general Dolbeault-complex

$$(3.6) \qquad \ldots \to E^{p,q-1}(X,V) \xrightarrow{\bar{\partial}} E^{p,q}(X,V) \xrightarrow{\bar{\partial}} E^{p,q+1}(X,V) \to \ldots$$

We have $\bar{\partial}^2 = 0$. So we can define

$$(3.7) \qquad H^{p,q}(X,V) = \frac{\operatorname{Ker} \bar{\partial} : E^{p,q}(X,V) \to E^{p,q+1}(X,V)}{\operatorname{Im} \bar{\partial} : E^{p,q-1} \to E^{p,q}(X,V)}$$

and we call $H^{p,q}(X,V)$ the (Dolbeault-)cohomology group of type (p,q) with coefficients in the holomorphic vector bundle V. If $V = X \times \mathbb{C}$, i.e. we consider the usual Dolbeault complex we write $H^{p,q}(X)$ instead of $H^{p,q}(X,V)$. By definition $H^{0,0}(X,V)$ consists of globally defined holomorphic sections of V over X.

The cohomology groups $H^{p,q}(X)$ are the complex analogue of the *De Rham cohomology groups* $H^r(X)$, constructed with the so-called *De Rham complex*

$$(3.8) \qquad E^0(X) \xrightarrow{d} E^1(X) \xrightarrow{d} E^2(X) \xrightarrow{d} \ldots \xrightarrow{d} E^r(X),$$

defined on a C^∞ manifold X. In this case $H^r(X)$ is given by

$$H^r(X) = \frac{\operatorname{Ker} d : E^r(X) \to E^{r+1}(X)}{\operatorname{Im} d : E^{r-1}(X) \to E^r(X)} .$$

The dimension of $H^r(X)$ is a topological invariant of X and is called the r-th *Betti number*.

To a certain extent the cohomology groups $H^{p,q}(X,V)$ describe the global complex-analytic behaviour of the complex manifold X, equipped with a vector bundle. Triviality of $H^{p,q}(X,V)$ for some p,q and V means that certain obstructions to solve certain global-analytic problems are vanished.

4. REPRESENTATIONS OF HOLOMORPHIC SOLUTIONS OF THE MASSLESS FIELD EQUATIONS

THEOREM 3. *Let* $\pi: \mathbb{C}^{n+1}\backslash\{0\} \to P_n$ *be the projection with* $\pi z = [z]$ and U *an open set in* P_n. *Then sections over* U *of the bundle* $L^m \to P_n$ *correspond with holomorphic functions on* $\hat{U} = \pi^{-1}(U)$ *being homogeneous of degree* $-m$ *in the coordinates* $(z^0, z^1, \ldots, z^n)$.

PROOF. Let $s_\alpha = \phi_\alpha \cdot s : U \cap U_\alpha \to \mathbb{C}$. Then on $U_\alpha \cap U_\beta \cap U_\gamma$ we have

$$s_\alpha[z] = g_{\alpha\beta}([z]) s_\beta[z] = \left(\frac{z^\alpha}{z^\beta}\right)^m s_\beta[z]$$

and thus

$$\frac{s_\alpha[z]}{(z^\alpha)^m} = \frac{s_\beta[z]}{(z^\beta)^m} \ .$$

Therefore the function $\hat{s}: \hat{U} \to \mathbb{C}$ defined on $U_\alpha \cap U$ by

$$\hat{s}(z) = \frac{s_\alpha[z]}{(z^\alpha)^m}$$ is independent of α and homogeneous of degree $-m$ in z.

Conversely, such a function $\hat{s}: \hat{U} \to \mathbb{C}$ defines a section s of the bundle $L^m \to P_n$ over U.

THEOREM 4. $\pi^*: E^p(U) \to \{f \in E^p(\hat{U}) \mid i_Y f = i_{\bar{Y}} f = 0,\ i_Y df = i_{\bar{Y}} df = 0\}$ *is an isomorphism. Here* $Y = \Sigma z^\alpha \frac{\partial}{\partial z^\alpha}$ *is the Euler vector field and* $\bar{Y} = \Sigma \bar{z}^\alpha \frac{\partial}{\partial \bar{z}^\alpha}$.

PROOF. This is a consequence of the fact that Y and $\bar{Y}$ are vector fields tangent to the fibres of $\pi: \mathbb{C}^{n+1}\backslash\{0\} \to P_n$ and linearly independent at each point $z \neq 0$.

From Theorem 4 it follows immediately that for holomorphic p-forms on U we have

THEOREM 5. $\Omega^p(U) \cong \{f \in \Omega^p(\hat{U}) \mid i_Y f = 0,\ i_Y \partial f = 0\}$.

Considering forms with coefficients in L^m yields

THEOREM 6. $\Omega^0(U,L^m) = \mathcal{O}(U,L^m) \cong \{f \in \mathcal{O}(\hat{U}) \mid i_Y \partial f = -mf\}$

$\Omega^p(U,L^m) \cong \{f \in \Omega^p(\hat{U}) \mid i_Y f = 0,\ i_Y \partial f = -mf\}$

EXAMPLE. $\theta = I_{\alpha\beta} z^\alpha dz^\beta$ with I skew-symmetric, represents an element of $\Omega^1(P_n, L^{-2})$.

PROOF. $i_Y\theta = I_{\alpha\beta} z^\alpha z^\gamma i_{\frac{\partial}{\partial z^\gamma}} dz^\beta = I_{\alpha\beta} z^\alpha z^\gamma \delta^\beta_\gamma = I_{\alpha\beta} z^\alpha z^\beta = 0$ for I is skew-symmetric. $\partial\theta = I_{\alpha\beta} dz^\alpha \wedge dz^\beta$ and so

$$i_Y \partial\theta = I_{\alpha\beta} z^\gamma i_{\frac{\partial}{\partial z^\gamma}} dz^\alpha \wedge dz^\beta = I_{\alpha\beta} z^\gamma (\delta^\alpha_\gamma dz^\beta - \delta^\beta_\gamma dz^\alpha) =$$

$$= I_{\alpha\beta} z^\alpha dz^\beta - I_{\alpha\beta} z^\beta dz^\alpha = \theta + \theta = 2\theta.$$

Now we consider the twistor space $\mathbb{T} \cong \mathbb{C}^4$ and we use coordinates $(Z^0,Z^1,Z^2,Z^3) = (\omega^A, \pi_{A'})$ and we consider the differential form

(4.1) $\qquad \psi = h \cdot f \wedge \theta,$

here each of the quantities h, f, θ is defined on $\hat{U} = \pi^{-1}(U)$, U an open set in $\mathbb{P} \cong P_3$ and where

i) h is a homogeneous holomorphic polynomial of degree n

ii) f is a C^∞ form of type $(0,1)$ with $\bar{\partial} f = 0$ and $i_{\bar{Y}} f = 0$, $i_Y \partial f = (-n-2)f$

iii) θ is a holomorphic form of type $(1,0)$ with $i_Y\theta = 0$, $i_Y \partial\theta = 2\theta$.

Then ψ is a well-defined form of type $(1,1)$ on $\hat{U}$, which satisfies the conditions of Theorem 4 and represents a form of type $(1,1)$ on U. Indeed,

$$i_Y\psi = hi_Y(f\wedge\theta) = h(i_Yf)\wedge\theta - hf\wedge i_Y\theta = 0-0 = 0$$

$$i_{\bar{Y}}\psi = hi_{\bar{Y}}(f\wedge\theta) = h(i_{\bar{Y}}f)\wedge\theta - hf\wedge(i_{\bar{Y}}\theta) = 0-0 = 0$$

$$\begin{aligned} i_Yd\psi &= i_Y(dh\wedge f\wedge\theta + hdf\wedge\theta - hf\wedge d\theta) = \\ &= Y(h)f\wedge\theta - dh\wedge(i_Yf)\wedge\theta + dh\wedge f\wedge i_Y\theta + \\ &+ h(i_Ydf)\wedge\theta + hdf\wedge i_Y\theta - h(i_Yf)\wedge d\theta + hf\wedge i_Yd\theta = \\ &= nhf\wedge\theta - 0 + 0 + h(i_Y\partial f)\wedge\theta + 0 - 0 + hf\wedge i_Y\partial\theta = \\ &= nhf\wedge\theta + h(-n-2)f\wedge\theta + hf\wedge 2\theta = 0. \end{aligned}$$

Similarly:

$$\begin{aligned} i_{\bar{Y}}d\psi &= i_{\bar{Y}}(dh\wedge f\wedge\theta + hdf\wedge\theta - hf\wedge d\theta) = \\ &= -dh\wedge(i_{\bar{Y}}f)\wedge\theta + h(i_{\bar{Y}}df)\wedge\theta + hf\wedge i_{\bar{Y}}d\theta = \\ &= 0 + h(i_{\bar{Y}}\partial f)\wedge\theta + hf\wedge i_{\bar{Y}}\partial\theta = 0 + h(\partial i_{\bar{Y}}f)\wedge\theta + 0 = 0. \end{aligned}$$

Now suppose we have taken an element of $H^{0,1}(U,L^{n+2})$ represented by $f \in E^{0,1}(\hat{U})$ with

$$\bar{\partial}f = 0, \quad i_{\bar{Y}}f = 0, \quad i_Y\partial f = (-n-2)f$$

We take h of the form $h(Z) = \pi_{A'}\dots\pi_{D'}$ (n factors) and $\theta = I_{\alpha\beta}z^{\alpha}dz^{\beta}$ with

$$(I_{\alpha\beta}) = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & \varepsilon^{A'B'} \end{array}\right) \quad \text{and} \quad U = P_3^+ .$$

So $\theta = -\pi^{E'}d\pi_{E'} = \pi_1d\pi_2 - \pi_2d\pi_1 = \pi_{E'}d\pi^{E'}$ and $\hat{U} = \mathbb{T}^+$.
Now let $z^{BB'}$ be a point in

$$M_{\mathbb{C}}^+ = \{z^{BB'} \mid [iz^{BB'}\pi_{B'},\pi_{B'}] \in \mathbb{P}^+, \ \forall\pi_{B'}\}$$

and define

$$\psi_{A'\dots D'}(z^{BB'}) = \int_{P_z} \pi_{A'},\dots\pi_{D'}f\wedge\theta, \tag{4.2}$$

where $P_z = \{[\omega^A, \pi_{A'}] \in \mathbb{P}_3 \mid \omega^A = iz^{AA'}\pi_{A'}\}$

The integral (4.2) is well-defined and moreover is symmetric in the n spinor indices $A' \ldots D'$.

Now we have the following result

THEOREM 7. *The function* $\psi_{A'\ldots D'}$, *defined in* (4.2) *is holomorphic in* $M^+_{\mathbb{C}}$, *is symmetric in the indices* $A', \ldots, D'$ *and it satisfies the massless field equations of helicity* $s = \frac{n}{2}$

$$\nabla^{AA'}\psi_{A'\ldots D'} = 0$$

on $M^+_{\mathbb{C}}$. *Furthermore,* $\psi_{A'\ldots D'}$ *depends only on the Dolbeault-cohomology class represented by* f *in* $H^{0,1}(P_3^+, L^{n+2})$.

PROOF. Let

$$I_{A'\ldots D'}(Z) = \pi_{A'}\ldots\pi_{D'} f(Z) \wedge \theta(Z)$$

be the integrand of (4.2), where as before $Z = (\omega^A, \pi_{A'})$. Let $z^{BB'} \in M^+_{\mathbb{C}}$, then the 2-plane in $\mathbb{PT}$ parametrized by $z^{BB'}$ with respect to the Penrose transformation consists of all points $(\omega^B, \pi_{B'})$ with $\omega^B = z^{BB'}\pi_{B'}$ i.e. it is the graph of the linear mapping $(z^{BB'}) : \mathbb{C}_{B'} \to \mathbb{C}^B$.

First we will check that $\nabla^{AA'}\psi_{A'\ldots D'} = 0$, i.e. $\varepsilon^{BA}\varepsilon^{B'A'}\nabla_{BB'}\psi_{A'\ldots D'} = 0$.

Let $\alpha_0 : \mathbb{C}_{A'} \times M \to \mathbb{PT}$ be defined by

$$\alpha_0(\pi_{A'}, z^{AA'}) = (iz^{AA'}\pi_{A'}, \pi_{A'})$$

and $\alpha : P_1 \times M \to \mathbb{P}_3$ the mapping induced by α_0. So

$$\alpha([\pi_{A'}], z^{AA'}) = [iz^{AA'}, \pi_{A'}]$$

Then we can write (4.2) in the form

$$\psi_{A'\ldots D'}(z^{BB'}) = \int_{P_1\times\{z^{BB'}\}} \alpha^* I_{A'\ldots D'}([\pi_{B'}, z^{BB'}) .$$

The last integral represents an integral over $\mathbb{P}_1$ parametrized by $z^{BB'}$.

Now we investigate the expression $\nabla_{BB'}\psi_{A'\ldots D'}$.

If we differentiatie under the integral sign we obtain

$$\nabla_{BB'}\psi_{A'\ldots D'}(z^{BB'}) = \int_{P_1} \frac{\partial}{\partial z^{BB'}} \alpha^* I_{A'\ldots D'}([\pi_{B'}, z^{BB'}).$$

We write

$$f = f_\alpha d\bar{z}^\alpha = f_{M'} d\bar{\omega}^{M'} + f^M d\bar{\pi}_M .$$

Then

$$\alpha_0^* I_{A'\ldots D'}(\pi_{B'}, z^{BB'}) = -\pi_{A'}\ldots\pi_{D'}(f^\alpha d\bar{z}^\alpha) \wedge \pi^{E'} d\pi_{E'} =$$

$$= -\pi_{A'}\ldots\pi_{D'}\{f_{M'} d(-i\bar{z}^{M'M}\bar{\pi}_M) + f^M d\bar{\pi}_M\} \wedge \pi^{E'} d\pi_E .$$

Differentiating with respect to $z^{BB'}$ we find

$$\frac{\partial}{\partial z^{BB'}}(\alpha_0^* I_{A'\ldots D'})(\pi_B, z^{BB'}) =$$

$$= \pi_{A'}\ldots\pi_{D'}\pi^{E'} d\pi_{E'} \wedge \frac{\partial}{\partial z^{B'B'}}\{ f_{M'} d(-i\bar{z}^{M'M}\bar{\pi}_M) + f^M d\pi_M\} =$$

$$= (i\pi_{B'})\pi_{A'}\ldots\pi_{D'}\pi^{E'} d\pi_{E'} \wedge \{\frac{\partial f_{M'}}{\partial \omega^B} d\bar{\omega}^{M'} + \frac{\partial f^M}{\partial \omega^B} d\bar{\pi}_M\} .$$

So we see that the last expression is symmetric in the indices $A', B', \ldots, D'$. Therefore $\frac{\partial}{\partial z^{BB'}}\psi_{A'\ldots D'}$ is symmetric in the indices $A', \ldots, D'$.

This implies that

$$\varepsilon^{B'A'} \frac{\partial}{\partial z^{BB'}} \psi_{A'\ldots D'} = 0$$

by skew-symmetry of $\varepsilon^{B'A'}$ and so

$$\nabla^{AA'}\psi_{A'\ldots D'} = \varepsilon^{BA}\varepsilon^{B'A'}\frac{\partial}{\partial z^{BB'}} = 0 .$$

Next we show that $\psi_{A'\ldots D'}$ depends only on the cohomology class represented by f in $H^{0,1}(P_3^+,L^{n+2})$. Suppose that f represents $F \in E^{0,1}(P_3^+,L^{n+2})$ with $F = \bar\partial H$ and $H \in E(P_3^+,L^{n+2})$. Then H is represented by an element $h \in E^0(\mathbf{T}^+)$ with $i_Y dh = (-n-2)h$ and $f = \bar\partial h$. Then we have

$$\int_{P_z} \pi_{A'}\ldots\pi_{D'} f\wedge\theta = \int_{P_z} \pi_{A'}\ldots\pi_{D'}(\bar\partial h)\wedge\theta =$$

$$= \int_{P_z} \bar\partial(\pi_{A'}\ldots\pi_{D'} h\wedge\theta) = \int_{P_z} d(\pi_{A'}\ldots\pi_{D'} h\wedge\theta) = 0$$

by Stokes' theorem and because the integral of a (2,0)-form vanishes on the 1-dimensional complex manifold P_z.

Finally we show that $\psi_{A'\ldots D'}$ is holomorphic on $M_{\mathbb{C}}^+$. We have

$$\frac{\partial}{\partial\bar z^{B'B}} f = \frac{\partial}{\partial\bar z^{B'B}}(f_{M'}d\bar\omega^{M'} + f^M d\bar\pi_M) =$$

$$= -i\,\bar\pi_B\left(\frac{\partial f_{M'}}{\partial\bar\omega^{B'}}d\bar\omega^{M'} + \frac{\partial f}{\partial\bar\omega^{B'}}d\bar\pi_M\right) - i\,f_{B'}d\bar\pi_B$$

Since $\bar\partial f = 0$ we know that $\frac{\partial f_{M'}}{\partial\bar\omega^{B'}} = \frac{\partial f_{B'}}{\partial\bar\omega^{M'}}$ and $\frac{\partial f^M}{\partial\bar\omega^{B'}} = \frac{\partial f_{B'}}{\partial\bar\pi_M}$ and so we have

$$\frac{\partial}{\partial\bar z^{BB'}} f = -i\,\bar\pi_B\left(\frac{\partial f_{B'}}{\partial\bar\omega^{M'}}d\bar\omega^{M'} + \frac{\partial f_{B'}}{\partial\bar\pi_M}\right) - i\,f_{B'}d\bar\pi_B = -i\,\bar\partial(\bar\pi_B f_{B'})$$

Hence

$$\frac{\partial}{\partial\bar z^{B'B}}\psi_{A'\ldots D'}(z^{BB'}) = \int_{P_z}\frac{\partial}{\partial\bar z^{B'B}}\pi_{A'}\ldots\pi_{D'} f\wedge\theta =$$

$$= \int_{P_z}\pi_{A'}\ldots\pi_{D'}\left(\frac{\partial}{\partial\bar z^{B'B}}f\right)\wedge\theta = -i\int_{P_z}\pi_{A'}\ldots\pi_{D'}\bar\partial(\bar\pi_B f_{B'})\wedge\theta =$$

$$= \int_{P_z}\bar\partial(\pi_{A'}\ldots\pi_{D'}\bar\pi_B f_B\wedge\theta = 0$$

by the same arguments as in the proof that $\psi_{A'\ldots D'}$ depends only on the cohomology class.

Now we define the quantity with spinor indices $A, \ldots, D$

$$(4.3) \qquad \phi_{A\ldots D}(z^{BB'}) = \int_{P_z} \frac{\partial}{\partial\omega^A} \cdots \frac{\partial}{\partial\omega^D} f \wedge \theta ,$$

where f is a C^∞ form of type $(0,1)$ defined on $\mathbb{PT}^+$ with $\bar{\partial} f = 0$ and $i_Y f = 0$, $i_Y \partial f = (n-2)f$ and θ and P_z are the same as in 4.2. Then we have the following result, which is proved along the same lines as the proof of Theorem 7.

THEOREM 8. *The function* $\phi_{A\ldots D}$ *defined in* (4.3) *is holomorphic in* $M^+_{\mathbb{C}}$, *is symmetric in the indices* $A, \ldots, D$ *and it satisfies the massless field equations of helicity* $s = -\frac{n}{2}$

$$\nabla^{AA'}\phi_{A\ldots D} = 0$$

on $M^+_{\mathbb{C}}$. $\phi_{A\ldots D}$ *depends only on the Dolbeault cohomology class represented by* f *in* $H^{0,1}(P_3^+, L^{-n+2})$.

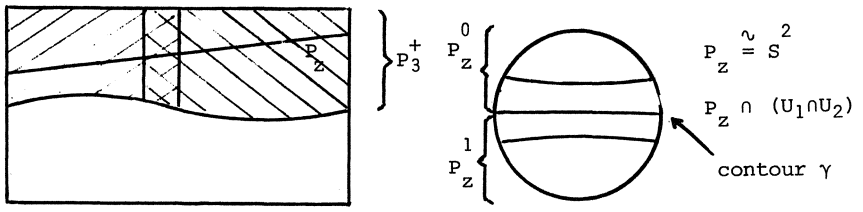
The mapping $P : H^{0,1}(P_3^+, L^{n+2}) \to \left\{\begin{matrix}\text{holomorphic massless fields} \\ \text{on } M^+_{\mathbb{C}} \text{ of helicity } \frac{n}{2}\end{matrix}\right\}$

is called the *Penrose transformation* in the literature. Here P is given by (4.2) for $n > 0$ and by 4.3 for $n < 0$.

One can prove that the transformation P is an isomorphism. Usually this is done using Čech cohomology.

To show how Čech cohomology comes into the game we give the following heuristic account.

Suppose that P_3^+ is covered by two open sets U_0 and U_1 such that $U_1 \cap U_2$ intersects any projective line $P_2, z \in M^+_{\mathbb{C}}$ according to an annular domain.



and that we may write for the restrictions $f\,|_{U_0}$ and $f\,|_{U_1}$ of the (0,1)-form f from (4.2):

$$f\,|_{U_0} = \bar{\partial} h_0 \quad , \quad f\,|_{U_1} = \bar{\partial} h_1$$

with $h_0 \in E^0(\hat{U})$, $h_1 \in E^1(\hat{U}_1)$ and both homogeneous and of the same degree as f. Then by Stokes' theorem we can write for 4.2

$$\psi_{A'\ldots D'}(z^{BB'}) = \int_{P_z} \pi_{A'}\ldots\pi_{D'} f\wedge\theta = \int_{P_z^0} \pi_{A'}\ldots\pi_{D'} f\wedge\theta + \int_{P_z^1} \pi_{A'}\ldots\pi_{D'} f\wedge\theta =$$

$$= \int_{P_z^0} \pi_{A'}\ldots\pi_{D'} (\bar{\partial} h_0)\wedge\theta + \int_{P_z^1} \pi_{A'}\ldots\pi_{D'} (\bar{\partial} h_1)\wedge\theta =$$

$$= \int_{P_z^0} \bar{\partial}(\pi_{A'}\ldots\pi_{D'} h_0\wedge\theta) + \int_{P_z^1} \bar{\partial}(\pi_{A'}\ldots\pi_{D'} h_1\wedge\theta) =$$

$$= \int_{P_z^0} d(\pi_{A'}\ldots\pi_{D'} h_0\wedge\theta) + \int_{P_z^1} d(\pi_{A'}\ldots\pi_{D'} h_1\wedge\theta) =$$

$$= \int_{\gamma} \pi_{A'}\ldots\pi_{D'} h_0\wedge\theta - \int_{\gamma} \pi_{A'}\ldots\pi_{D'} h_1\wedge\theta = \int_{\gamma} \pi_{A'}\ldots\pi_{D'} (h_0 - h_1)\theta\,.$$

Setting $h = h_0|_{U_0\cap U_1} - h_1|_{U_0\cap U_1}$ we have the following contour integral representation for $\psi_{A'\ldots D'}$

$$(4.4) \qquad \psi_{A'\ldots D'}(z^{BB'}) = \int_{\gamma} \pi_{A'}\ldots\pi_{D'} h\theta$$

In the same way we find,

$$(4.5) \qquad \psi_{A\ldots D}(z^{BB'}) = \int_{\gamma} \frac{\partial}{\partial\omega^A}\ldots\frac{\partial}{\partial\omega^D} h\theta$$

Contour integrals of type (4.4) and (4.5) were introduced by Penrose at the outset of twistor theory to solve the massless field equations. Later one realized that the functions appearing in the integrands may be interpreted in terms of Čech cohomology.

Formulas of type (4.4) and (4.5) were already known in the literature (Bateman, 1904).

REFERENCES

[1] WELLS, R.O., Jr. *Complex manifolds and mathematical physics*, Bull. Amer.Math.Soc. 1(1979), p. 296-336.

[2] WELLS, R.O., Jr. *Complex geometry in mathematical physics*, Les presses de l'université de Montréal, 1982.

C O H O M O L O G Y T H E O R Y

by

J.W. de ROEVER
Department of Mathematics
Twente University of Technology

Several times we have met already cohomology classes, cf. for example, the definition of Chern-class (4^{th} chapter in the proceedings of this seminar held in 1981-82) and Dolbeault-cohomology especially in the Penrose transformation (theorems 7 and 8 of the preceding chapter). In the next chapter we will study the Penrose transformation more closely and we will point out the relation with Ward's construction of solutions of the self-dual Yang-Mills equations. In order to do that a more profound and systematic discussion of cohomology theory is needed.

Problems which can be locally solved (cf. for example the local Penrose transformation discussed in section 3.B of the chapter "Computation of the Yang-Mills potentials") often cause a new problem when solved globally. In such situations cohomology groups appear in a natural way. Every problem would have a cohomology group of its own if not these groups, defined in completely different ways, were often the same thing. Cohomology theory tries to order this chaos. Before discussing it more rigorously we will give some examples.

1. EXAMPLES

A. Let $\Omega \subset \mathbb{C}^n$ be an open set, $\mathcal{O}(\Omega)$ the (Abelian) additive group of holomorphic functions in Ω and $\mathcal{O}^*(\Omega)$ the (Abelian) multiplicative group of non-vanishing holomorphic functions in Ω. The problem is when the map

$$e: \mathcal{O}(\Omega) \to \mathcal{O}^*(\Omega),$$

defined by

$$e(g)(z) = e^{2\pi i\, g(z)}, \qquad z \in \Omega,$$

is surjective. This will not always be true: if $\Omega = \mathbb{C} \setminus \{0\}$ and $h \in \mathcal{O}^*(\Omega)$ is defined by $h(z) = z$, then a function g with $e(g) = h$ should satisfy $g(z) = (\log z)/2\pi i$, but this cannot be a holomorphic function in Ω. However, the problem does have a solution locally: every $z_j \in \Omega$ has a (simply connected) neighbourhood $\omega_j \subset \Omega$ such that the map $e: \mathcal{O}(\omega_j) \to \mathcal{O}^*(\omega_j)$ is surjective (if $h_j \in \mathcal{O}^*(\omega_j)$, take $g_j = (\log h_j)/2\pi i \in \mathcal{O}^*(\omega_j)$).

Now let $h \in \mathcal{O}^*(\Omega)$ be a globally defined function and let $g_j \in \mathcal{O}(\omega_j)$ and $g_k \in \mathcal{O}(\omega_k)$ be such that $e(g_j) = h$ in ω_j and $e(g_k) = h$ in ω_k. Then in $\omega_j \cap \omega_k$ we have

$$g_j - g_k \overset{d}{=} f_{jk} \in \mathbb{Z}.$$

So actually we are dealing with the following object: a collection $f_{jk} \in \mathbb{Z}(\omega_j \cap \omega_k)$ with $f_{jk} = -f_{kj}$ and with

$$f_{k\ell} - f_{j\ell} + f_{jk} = 0 \quad \text{in} \quad \omega_j \cap \omega_k \cap \omega_\ell.$$

Such a collection $\{f_{jk}\}$ of integers is called a 1-*cocycle* of integers. Some of these cocycles could have the following form: in every $\omega_j \cap \omega_k$ $f_{jk} = f_k - f_j$ for certain $f_j \in \mathbb{Z}(\omega_j)$. Such 1-cocycles are called a 1-*coboundary*. The cohomology group associated to this problem is

$$H^1(\mathcal{U},\Omega,\mathbb{Z}) = \frac{\text{1-cocycles}}{\text{1-coboundaries}},$$

where $\mathcal{U} = \bigcup_j \omega_j$ is the covering of Ω. We will see that this is an example of a Čech-cohomology group.

If this group vanishes, i.e. if every 1-cocycle is a 1-coboundary, the map e is surjective: take the new function $\tilde{g}_j = g_j + f_j$ in ω_j; then we still have

$$e(\tilde{g}_j) = e^{2\pi i g_j} e^{2\pi i f_j} = e^{2\pi i g_j} = h_j \quad \text{in} \quad \omega_j$$

and moreover

$$\tilde{g}_j - \tilde{g}_k = g_j + f_j - g_k - f_k = f_{jk} + f_{kj} = 0 \quad \text{in} \quad \omega_j \cap \omega_k.$$

Hence the collection $\tilde{g}_j \in \mathcal{O}(\omega_j)$ determines a globally defined function $g \in \mathcal{O}(\Omega)$ with $e(g) = h$.

B. When does a globally defined field (of 2-forms) in X originate from globally defined potentials?

Let us represent the field by $F = F_{\mu\nu} dx^\mu \wedge dx^\nu$ with $F_{\mu\nu} \in \mathcal{F}(X)$, where $\mathcal{F}(X)$ denotes the additive group of C^∞-functions, locally L^2-functions or distributions in X. As part of the field equations $dF = 0$ holds. Locally, i.e. in small neighbourhoods ω_i, there are potentials $A_i = A^i_\mu dx^\mu$ with $A^i_\mu \in \mathcal{F}(\omega_i)$ and with $dA_i = F$ in ω_i.

Hence in $\omega_i \cap \omega_j$ $d(A_i - A_j) = 0$. We assume that the neighbourhoods are such (for example convex) that there are gauges $\alpha_{ij} \in \mathcal{F}(\omega_i \cap \omega_j)$ with

$$d\alpha_{ij} = A_i - A_j \quad \text{in} \quad \omega_i \cap \omega_j$$

(otherwise we take a refinement of the covering, cf. [Hörmander, lemma 7.3.6]). Now $\alpha_{ij} = -\alpha_{ji}$ and $d(\alpha_{jk} - \alpha_{ik} + \alpha_{ij}) = 0$ holds in $\omega_i \cap \omega_j \cap \omega_k$. It follows that there are constants $r_{ijk} \in R(\omega_i \cap \omega_j \cap \omega_k)$ with

$$\alpha_{jk} - \alpha_{ik} + \alpha_{ij} = r_{ijk}.$$

Thus the collection $\{\alpha_{ij}\}$ is not a 1-cocycle. The collection $\{r_{ijk}\}$ of constants is antisymmetric in the indices i, j and k, and satisfies

$$r_{jk\ell} - r_{ik\ell} + r_{ij\ell} - r_{ijk} = 0 \quad \text{in} \quad \omega_i \cap \omega_j \cap \omega_k \cap \omega_\ell.$$

Such a collection is called a 2-*cocycle* of constants. It might be that any 2-cocycle $\{r_{ijk}\}$ is a 2-*coboundary*: i.e. in $\omega_i \cap \omega_j \cap \omega_k$ $r_{ijk} = r_{jk} - r_{ik} + r_{ij}$ for certain $r_{ij} \in R(\omega_i \cap \omega_j)$. The associated cohomology group

$$H^2(U,X,R) = \frac{\text{2-cocycles}}{\text{2-coboundaries}}$$

vanishes in that case.

Then new gauges $\tilde{\alpha}_{ij} \in F(\omega_i \cap \omega_j)$ can be defined by

$$\tilde{\alpha}_{ij} = \alpha_{ij} - r_{ij}.$$

The collection $\{\tilde{\alpha}_{ij}\}$ now is a 1-cocylce of C^∞-functions, L^2_{loc}-functions or distributions, for

$$\text{in} \quad \omega_i \cap \omega_j \cap \omega_k: \quad \tilde{\alpha}_{jk} - \tilde{\alpha}_{ik} + \tilde{\alpha}_{ij} = r_{ijk} - r_{jk} + r_{ik} - r_{ij} = 0.$$

We will see that a 1-cocycle of C^∞-functions, L^2_{loc}-functions or distributions is always a 1-coboundary:

$$\text{in} \quad \omega_i \cap \omega_j: \quad \tilde{\alpha}_{ij} = \alpha_j - \alpha_i \quad \text{for certain} \quad \alpha_i \in F(\omega_i).$$

Finally, we define the new potentials $\tilde{A}_i = A_i + d\alpha_i$ in ω_i. Then still $d\tilde{A}_i = dA_i = F$ in ω_i, but also

$$\tilde{A}_i - \tilde{A}_j = A_i - A_j + d\alpha_i - d\alpha_j = d\alpha_{ij} - d\tilde{\alpha}_{ij} = dr_{ij} = 0 \quad \text{in} \quad \omega_i \cap \omega_j.$$

Hence the collection $\{\tilde{A}_i\}$ defines a globally defined potential in X

(provided that $H^2(\mathcal{U},X,R) = 0$).

The problem can also be formulated in the following way. Let

$$F(X) = F^0(X) \xrightarrow{d_0} F^1(X) \xrightarrow{d_1} F^2(X) \xrightarrow{d_2} \ldots$$

be the complex of mappings between p-forms $\in F^p(X)$ and (p+1)-forms determined by the d-operator (for $F = E$, the C^∞-functions, this is the de Rham-complex). When is the kernel of d_2 equal to the image of d_1? Apparently, instead of requiring that the Čech-cohomology group $H^2(\mathcal{U},X,R)$ vanishes (which is independent of the case with C^∞- functions, L^2_{loc}-functions or distributions), one could have required that the de Rham-cohomology group $H^2(X) = 0$ (in case of C^∞-functions) or (in the other cases) that the 2^{nd} cohomology group = $\operatorname{Ker} d_2/\operatorname{Im} d_1$ of the above given complex vanishes.

As an example, let $X = S_2$ be the earth surface. We divide the earth surface in three domains: E_1 = {all points between 0^o and 120^oE}, E_2 = {all points between 120^o and 180^oE and between 120^o and 180^oW}, E_3 = {all points between 0^o and 120^oW}. By taking narrow strips around these sets we obtain the covering $\mathcal{U}$ of the earth surface with three open neighbourhoods ω_i of the sets E_i. The set $\omega_1 \cap \omega_2 \cap \omega_3$ consists of two small domains, one near the north pole, the other near the south pole. A constant 2-cocycle r_{123} can attain a (constant) value in the north pole different from the one in the south pole. Such a cocycle can never be written as a coboundary: $r_{123} \neq r_{23} - r_{13} + r_{12}$, because r_{ij} is a constant defined in a neighbourhood of a meridian and hence has the same value in the north pole as in the south pole. Thus $H^2(\mathcal{U},S_2,R) \neq 0$. So in general, on the earth surface there are no global potentials to describe a field.

2. SHEAFS

In the foregoing examples we have used several properties of the space $F(X)$ of certain functions or distributions on the topological space X. These properties turn F into what is called a sheaf.

DEFINITION 1 *(Presheaf)*.

Let for every open set Ω in a topological space X an Abelian group (linear space, ring, module, ...) $F(\Omega)$ be given and let for every inclusion $\Omega_1 \subset \Omega_2$ a homomorphism (linear map, ...): $F(\Omega_2) \to F(\Omega_1)$, called restriction and denoted by $s \in F(\Omega_2) \longmapsto s\big|_{\Omega_1}$, be given with the property

$$\forall \Omega_1 \subset \Omega_2 \subset \Omega_3, \quad s \in F(\Omega_3) \Rightarrow s\big|_{\Omega_1} = (s\big|_{\Omega_2})\big|_{\Omega_1}.$$

Then F is called a presheaf on X.

EXAMPLES. $F(\Omega)$ =

1^o. = {the additive group of all functions: $\Omega \to \mathbb{R}, \mathbb{C}$ or $\mathbb{Z}$}

2^o. = {all continuous functions: $\Omega \to \mathbb{R}$ or $\mathbb{C}$}

3^o. = {all constant functions: $\Omega \to \mathbb{R}$, $\mathbb{C}$ or $\mathbb{Z}$}, denoted by $R(\Omega)$, $C(\Omega)$ and $\mathbb{Z}(\Omega)$ respectively.

4^o. (if X is a C^∞-manifold) = {all C^∞-functions in Ω}, denoted by $E(\Omega)$

5^o. (if X is a complex manifold) = {all holomorphic functions in Ω} denoted by $\mathcal{O}(\Omega)$

6^o. (if X is a complex manifold) = {the multiplicative group of all non-vanishing holomorphic functions}, denoted by $\mathcal{O}^*(\Omega)$.

7^o. (if $X = \mathbb{R}^n$) = {all bounded C^∞-functions with bounded derivatives in Ω which decrease rapidly at infinity if Ω is unbounded}, denoted by $S(\Omega)$

8^o. (if $X = \mathbb{R}^n$) = {all L^2-functions in Ω}, denoted by $L^2(\Omega)$

9^{o}. = {all bounded real functions}, denoted by $B(\Omega)$.

10^{o}. (if X is a manifold) = {continuous sections over Ω of a line bundle over X}.

In all these examples the restriction mappings are the usual ones.

The spaces $\mathcal{D}(\Omega)$ = {all C^{∞}-functions with compact support in Ω} do not form a presheaf.

A sheaf is a presheaf where local equality everywhere means global equality and where "small pieces can be glued together":

DEFINITION 2 *(Sheaf)*.

A presheaf F is called a sheaf if for every open $\Omega \subset X$ and for every open covering $\Omega = \bigcup_i \Omega_i$, $\Omega_i \subset X$ open, the following two properties hold:

i) if $s,s' \in F(\Omega)$ are such that $s\big|_{\Omega_i} = s'\big|_{\Omega_i}$ for any i, then $s = s'$.

ii) if $s_i \in F(\Omega_i)$ are such that $s_i\big|_{\Omega_i \cap \Omega_j} = s_j\big|_{\Omega_i \cap \Omega_j}$ for any i and j, then there is an $s \in F(\Omega)$ with $s\big|_{\Omega_i} = s_i$.

EXAMPLES. The examples 1^{o}, 2^{o}, 4^{o}, 5^{o}, 6^{o} and 10^{o} of presheafs.

If F is a presheaf on X the elements of $F(\Omega)$ are called sections over Ω and the set of all sections is also denoted by $\Gamma(\Omega,F) = F(\Omega)$. Sections over closed sets $A \subset X$ are defined by the direct limit:

$$\Gamma(A,F) = \varinjlim_{\Omega \supset A} \Gamma(\Omega,F),$$

i.e. $f \in \Gamma(A,F)$ is an equivalence class determined by $f_1 \in \Gamma(\Omega_1,F) \sim f_2 \in \Gamma(\Omega_2,F)$ with $A \subset \Omega_1$ and $A \subset \Omega_2 \iff$ there is an open neighbourhood Ω of A with $A \subset \Omega \subset \Omega_1 \cap \Omega_2$ and with $f_1\big|_{\Omega} = f_2\big|_{\Omega}$.

If in particular $A = \{x\}$ for $x \in X$ the set $\Gamma(\{x\},F)$ is called the *stalk* of the (pre)sheaf F at the point x; notation $\Gamma(\{x\},F) = F_x$. If $x \in \omega \subset X$ and $f \in F(\omega)$ then f determines an element $f_x \in F_x$ which is called the *germ* of f at x.

A presheaf F can be turned into a sheaf $\underline{F}$ as follows: Let an open covering $\Omega = \bigcup_\alpha \omega_\alpha$ of Ω and sections $f_\alpha \in F(\omega_\alpha)$ be such that $f_\alpha\big|_{\omega_\alpha \cap \omega_\beta} = f_\beta\big|_{\omega_\alpha \cap \omega_\beta}$ (hence $(f_\alpha)_x = (f_\beta)_x$ for $x \in \omega_\alpha \cap \omega_\beta$) then define the germ $f_x = (f_\alpha)_x$ if $x \in \omega_\alpha \subset \Omega$ and the section $f \in \underline{F}(\Omega)$ by $\bigcup_{x\in\Omega} f_x$ with the obvious restriction map:

$$f\big|_{\Omega'} = \bigcup_{x\in\Omega'} f_x \qquad (\Omega' \subset \Omega).$$

The presheaf obtained in this way is a sheaf $\underline{F}$ that equals F if the presheaf F is itself a sheaf.

EXAMPLES.

1°. From the presheafs R, C and $\mathbb{Z}$ we get the sheafs (also denoted by R, C and $\mathbb{Z}$, respectively) of functions which are constant on the components of Ω.

2°. The presheaf S yields the sheaf E.

3°. The presheaf L^2 yields the sheaf L^2_{loc} of locally square integrable functions.

4°. The presheaf B yields the sheaf of all real functions.

5°. From two presheafs F and G one can make the quotient presheaf H by $H(\Omega) = G(\Omega)/F(\Omega)$; even if F and G are sheafs H need not be a sheaf, cf. the example of section 1.A with $F = \mathbb{Z}$ and $G = e(\mathcal{O})$, where not always $\mathcal{O}^*(\Omega) = e(\mathcal{O}(\Omega))/\mathbb{Z}(\Omega)$. The sheaf $\underline{H}$ associated to H is called the *quotient sheaf*. The sheaf $\underline{H}$ is characterized by

$$\underline{H}_x = G_x/F_x, \qquad x \in X.$$

DEFINITION 3. A sheaf F is called *soft* if the induced restriction map:

$$F(\Omega) \to \Gamma(A,F), \qquad A \subset \Omega$$

is surjective for every closed subset A of any open set $\Omega \subset X$.

EXAMPLES.

1°. The sheafs $\mathcal{D}'$ of distributions, L^2_{loc} and $\mathcal{E}$ are soft. Every section f over a closed set A can be extended to a section over a larger set: let f be determined by $f \in \Gamma(\Omega', \mathcal{F})$, $A \subset \Omega'$. Take a C^∞-function ϕ equal to one in a neighbourhood of A and vanishing outside a neighbourhood Ω'' with $A \subset \Omega'' \subset \overline{\Omega''} \subset \Omega'$. Then ϕf (extended by zero outside Ω') is a section in $\mathcal{F}(\Omega)$ with $(\phi f)_x = f_x$ for $x \in A$.

2°. The sheafs $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{Z}$ are not soft: if Ω' is not connected while Ω is, then a function having different constant values on the various components of Ω' cannot be the restriction of one constant in Ω.

3°. The sheafs $\mathcal{O}$ and $\mathcal{O}^*$ are not soft: a holomorphic function in Ω' need not be extendable as a holomorphic function to a larger set Ω.

A *sheafmorphism* between two sheafs $\mathcal{F}$ and $\mathcal{G}$ is a group-homomorphism between the presheafs which commutes with the restriction mappings. For every $x \in X$ we then have a homomorphism: $\mathcal{F}_x \to \mathcal{G}_x$.

In the example in section 1.A the complex of mappings

$$0 \to \mathbb{Z}(\Omega) \xrightarrow{\text{injection}} \mathcal{O}(\Omega) \xrightarrow{e} \mathcal{O}^*(\Omega) \to 0\,,$$

which is exact at the first two places (i.e. the injection map is injective and its image is the kernel of the map e), is not always exact at the third place (i.e. e need not be surjective). As a complex of sheafmorphisms the sequence

$$0 \to \mathbb{Z} \to \mathcal{O} \xrightarrow{e} \mathcal{O}^* \to 0$$

is exact, i.e. for every $x \in X$ the sequence

$$0_x \to \mathbb{Z}_x \to \mathcal{O}_x \xrightarrow{e} \mathcal{O}^*_x \to 0_x$$

of maps between stalks is exact. It should be remarked that the sheaf $\mathbb{Z}$ is not soft and exactly this causes the non-exactness of the first sequence, as follows from the following theorem.

In the sequel we assume that the space X is paracompact (i.e. X is Hausdorff and every open covering has a locally finite open refinement), although some theorems remain true for non-paracompact X.

THEOREM 1. *Let*

$$0 \to F \to G \to H \to 0$$

be an exact sequence of sheafmorphisms and let the sheaf F *be soft. Then for every open* $\Omega \subset X$ *the complex of corresponding maps between sections*

$$0 \to F(\Omega) \to G(\Omega) \to H(\Omega) \to 0$$

is exact.

PROOF.

Exactness at the first two places follows immediately also if F is not soft. So we only have to show that the map: $G(\Omega) \to H(\Omega)$ is surjective.

Since the sequence of maps between the sheafs is exact, a section in $H(\Omega)$ comes locally from sections in G over small open sets covering Ω. We take a shrinking and a locally finite refinement so that we can work with a covering consisting of closed sets $\overline{\omega_i}$ (the union of arbitrary many of such sets $\overline{\omega_i}$ is closed then). Just as in the example of section 1.A we have to show that a 1-cocycle $\{f_{ij} \in \Gamma(\bar{\omega}_i \cap \bar{\omega}_j, F)\}$ is a 1-coboundary. We will do this by induction (assuming that I is well-ordered where I is the index set of the covering $\bigcup_{i \in I} \bar{\omega}_i = \Omega$ of Ω).

Define $f_0 = 0 \in \Gamma(\bar{\omega}_0, F)$. Assume that we have found $f_i \in \Gamma(\bar{\omega}_i, F)$ for $i < k$ with

$$f_j\big|_{\bar\omega_i\cap\bar\omega_j} - f_i\big|_{\bar\omega_i\cap\bar\omega_j} = f_{ij} \quad \text{if} \quad i,j < k$$

(this means $(f_j)_x - (f_i)_x = (f_{ij})_x$ for $x \in \bar\omega_i \cap \bar\omega_j$). We will construct a $f_k \in \Gamma(\bar\omega_k,F)$ with $f_k - f_i = f_{ik}$ in $\bar\omega_i \cap \bar\omega_k$ for $i < k$.

Since $\{f_{ij}\}$ is a cocycle:

$$(f_{jk}-f_{ik}+f_{ij})\big|_{\bar\omega_i\cap\bar\omega_j\cap\bar\omega_k} = 0,$$

we have in $\bar\omega_i \cap \bar\omega_j \cap \bar\omega_k$

$$f_{ik} + f_i = f_{jk} + f_j, \quad \text{for} \quad i,j < k\,.$$

Hence

$$\bigcup_{i<k} (f_{ik}+f_i)\big|_{\bar\omega_i\cap\bar\omega_k}$$

determines a section over the closed set $\bar\omega_k \cap \bigcup_{i<k} \bar\omega_i$. The sheaf F is soft, so we can extend this section to a section over an open neighbourhood of $\bar\omega_k$ determining a section $f_k \in \Gamma(\bar\omega_k,F)$ with

$$f_k\big|_{\bar\omega_i\cap\bar\omega_k} - f_i\big|_{\bar\omega_i\cap\bar\omega_k} = f_{ik}.$$

Hence $\{f_{ij}\}$ is a 1-coboundary. ∎

THEOREM 2. *If* $0 \to F \to G \to H \to 0$ *is exact and the sheafs* F *and* G *are soft then* H *is soft.*

PROOF.

Let $A \subset X$ be a closed set and let $h \in \Gamma(A,H)$ be determined by a section (also denoted by h) $h \in H(\Omega)$ where Ω is an open neighbourhood of A. According to theorem 1 the sequence $0 \to F(\Omega) \to G(\Omega) \to H(\Omega) \to 0$ is exact. Hence there is a $g \in G(\Omega)$ which is mapped onto h. This g determines a section $\in \Gamma(A,G)$ over A. Since G is soft there is an extension $g \in \Gamma(X,G)$ whose image $\bar h$ is an element of $H(X)$ coinciding with h in a neighbourhood of A. ∎

THEOREM 3. *Let*

$$0 \to F \to F_0 \to F_1 \to F_2 \to \dots$$

be an exact sequence of sheafmorphisms between soft sheafs. Then for any open $\Omega \subset X$ *the corresponding sequence*

$$0 \to \Gamma(\Omega,F) \to \Gamma(\Omega,F_0) \to \Gamma(\Omega,F_1) \to \dots$$

is exact.

PROOF.

Let Z_i, $i \geqslant 0$, be the sheaf

$$Z_i = \mathrm{Ker}(F_i \to F_{i+1}).$$

Then $Z_0 = F$ is soft and $Z_i = \mathrm{Im}(F_{i-1} \to F_i)$, $i > 0$.
The long exact sequence can be divided into short exact sequences

$$0 \to Z_i \to F_i \to Z_{i+1} \to 0, \qquad i \geqslant 0.$$

It follows from theorem 2 that the sheafs Z_i are soft. Theorem 1 implies that the short sequences

$$0 \to \Gamma(\Omega,Z_i) \to \Gamma(\Omega,F_i) \to \Gamma(\Omega,Z_{i+1}) \to 0$$

are exact. Hence for $i > 0$

$$\mathrm{Im}[\Gamma(\Omega,F_{i-1}) \to \Gamma(\Omega,F_i)] = \Gamma(\Omega,Z_i) =$$
$$= \mathrm{Ker}[\Gamma(\Omega,F_i) \to \Gamma(\Omega,F_{i+1})]. \quad \blacksquare$$

3. SHEAF COHOMOLOGY

DEFINITION 4. A *resolution* of the sheaf $\mathcal{F}$ on X is an exact sequence of sheaf morphisms

$$0 \to \mathcal{F} \to \mathcal{F}_0 \to \mathcal{F}_1 \to \mathcal{F}_2 \to \dots .$$

The resolution is called soft, if the sheafs $\mathcal{F}_0, \mathcal{F}_1, \dots$ are all soft.

EXAMPLES.

1°. $\mathcal{F} = \mathcal{R}$, $X = \mathbb{R}^n$.

$$0 \to \mathcal{R} \to \mathcal{E} \xrightarrow{d} \mathcal{E}^1 \xrightarrow{d} \mathcal{E}^2 \to \dots \mathcal{E}^n \to 0 \to 0 \dots \quad \text{(de Rham)}$$

$$0 \to \mathcal{R} \to L^2_{loc} \xrightarrow{d} \left(L^2_{loc}\right)^1 \xrightarrow{d} \left(L^2_{loc}\right)^2 \to \dots \to \left(L^2_{loc}\right)^n \to 0$$

$$0 \to \mathcal{R} \to \mathcal{D}' \xrightarrow{d} (\mathcal{D}')^1 \xrightarrow{d} (\mathcal{D}')^2 \to \dots \to (\mathcal{D}')^n \to 0$$

Here d is the d-operator from p-forms $\in \mathcal{G}^p$ to (p+1)-forms $\in \mathcal{G}^{p+1}$ with coefficients in $\mathcal{G} = \mathcal{E}$, L^2_{loc} or $\mathcal{D}'$ (in the last two cases differentiation is considered in the distributional sense). So it is possible that a sheaf has several resolutions. All three resolutions of $\mathcal{R}$ given above are soft. If $\mathcal{F} = \mathcal{C}$ then $\mathcal{E}$, L^2_{loc} or $\mathcal{D}'$ must be complex valued.

2°. $\mathcal{F} = \mathcal{O}$, $X = \mathbb{C}^n$.

$$0 \to \mathcal{O} \to \mathcal{E} \xrightarrow{\bar\partial} \mathcal{E}^{0,1} \xrightarrow{\bar\partial} \mathcal{E}^{0,2} \to \dots \mathcal{E}^{0,n} \to 0 \quad \text{(Dolbeault)}$$

$$0 \to \mathcal{O} \to L^2_{loc} \xrightarrow{\bar\partial} \left(L^2_{loc}\right)^{0,1} \xrightarrow{\bar\partial} \dots \to \left(L^2_{loc}\right)^{0,n} \to 0$$

$$0 \to \mathcal{O} \to \mathcal{D}' \xrightarrow{\bar\partial} (\mathcal{D}')^{0,1} \xrightarrow{\bar\partial} \dots \to (\mathcal{D}')^{0,n} \to 0.$$

Also here we have three soft resolutions of the sheaf $\mathcal{O}$.

It is not always easy to find a useful (soft) resolution of an arbitrary sheaf $\mathcal{F}$. We will show that at least one soft resolution exists for every sheaf. However, this one is of theoretical interest only and not very useful in practical problems. Then we will define the cohomology groups associated to the corresponding complex of sections and we will show that they

are independent of the resolution provided that this is soft.

DEFINITION 5. (*cohomology groups of a sheaf over an open set*). Let $0 \to \mathcal{F} \to \mathcal{F}_0 \to \mathcal{F}_1 \to \dots$ be a soft resolution of the sheaf $\mathcal{F}$ and let $0 \to \Gamma(\Omega,\mathcal{F}) \to \Gamma(\Omega,\mathcal{F}_0) \to \Gamma(\Omega,\mathcal{F}_1) \to \dots$ be the induced complex of mappings between sections over an open set $\Omega \subset X$. Then the p^{th} cohomology group of $\mathcal{F}$ over Ω is defined by

$$H^0(\Omega,\mathcal{F}) = \operatorname{Ker}[\Gamma(\Omega,\mathcal{F}_0) \to \Gamma(\Omega,\mathcal{F}_1)] = \Gamma(\Omega,\mathcal{F}), \qquad p = 0$$

$$H^p(\Omega,\mathcal{F}) = \frac{\operatorname{Ker}[\Gamma(\Omega,\mathcal{F}_p)\to\Gamma(\Omega,\mathcal{F}_{p+1})]}{\operatorname{Im}[\Gamma(\Omega,\mathcal{F}_{p-1})\to\Gamma(\Omega,\mathcal{F}_p)]}, \qquad p \geqslant 1.$$

EXAMPLES.

1°. $H^p(\Omega,R)$ are the de Rham cohomology groups $H^p(\Omega)$ of Ω.

2°. It follows from theorem 3 that $H^p(\Omega,\mathcal{F}) = 0$, $p \geqslant 1$, if $\mathcal{F}$ is soft.

Now we construct a soft resolution of an arbitrary sheaf $\mathcal{F}$. For $\Omega \subset X$ define

$$\hat{\mathcal{F}}_0(\Omega) = \prod_{x\in\Omega} \mathcal{F}_x.$$

With the restriction mapping $\hat{\mathcal{F}}_0(\Omega) \to \hat{\mathcal{F}}_0(\Omega')$, $\Omega' \subset \Omega$, defined by $\prod_{x\in\Omega} f_x \longmapsto \prod_{x\in\Omega'} f_x$, $f_x \in \mathcal{F}_x$, $\hat{\mathcal{F}}_0$ becomes a sheaf on X (here the subscript x in f_x labels the germ f chosen in $\mathcal{F}_x$ for all $x \in \Omega$). For $\mathcal{F} = R$ or $\mathbb{C}$ the sheaf $\hat{\mathcal{F}}_0$ can be identified with the sheaf of all real or complex functions on X and for $\mathcal{F} = E$, $\hat{\mathcal{F}}_0$ is even larger than this. Therefore, $\hat{\mathcal{F}}_0$ is sometimes called the *sheaf of discontinuous sections*. It is clear that $\hat{\mathcal{F}}_0$ is soft.

Let $\mathcal{Z}_1$ be the quotient sheaf $\hat{\mathcal{F}}_0/\mathcal{F}$ determined by the short exact sequence

$$0 \to \mathcal{F} \to \hat{\mathcal{F}}_0 \to \mathcal{Z}_1 \to 0.$$

We repeat the procedure with the sheaf $Z_1 : \hat{F}_1$ is the sheaf of discontinuous sections of Z_1, $\hat{F}_1(\Omega) = \prod_{x\in\Omega} (Z_1)_x$ and define Z_2 as the quotient sheaf determined by

$$0 \to Z_1 \to \hat{F}_1 \to Z_2 \to 0 \quad \text{(exact)}$$
$$\vdots \quad \text{etc.}$$
$$0 \to Z_p \to \hat{F}_p \to Z_{p+1} \to 0.$$

This yields a soft resolution of F:

$$0 \to F \to \hat{F}_0 \to \hat{F}_1 \to \hat{F}_2 \to \dots \to \hat{F}_p \to \dots ,$$

called the *canonical resolution* of F. Hence definition 5 can be applied to any sheaf F.

Next we will show that it does not matter which soft resolution of F in definition 5 is used. We will even show something stronger.

DEFINITION 6. A resolution $0 \to F \to F_0 \to F_1 \to \dots$ of F is called *acyclic on* Ω if $H^p(\Omega,F_i) = 0$, $p \geqslant 1$, $i \geqslant 0$.

EXAMPLE. A soft resolution is acyclic on every open $\Omega \subset X$.

THEOREM 4. *Let* $0 \to F \to F_0 \to F_1 \to \dots$ *be an acyclic resolution* F_* *on* Ω *of the sheaf* F *and let* $H^p(\Omega,F_*)$ *be the* p^{th} *cohomology group of the induced complex of mappings between sections over* Ω. *Furthermore, let* $H^p(\Omega,\hat{F})$ *be the* p^{th} *cohomology group of the complex of mappings between sections over* Ω *induced by the canonical resolution* $\hat{F}$. *Then*

$$H^p(\Omega,\hat{F}) \cong H^p(\Omega,F_*).$$

PROOF

A complex $F \to G \to H$ of sheaf morphisms induces the complex $\hat{F}_0 \to \hat{G}_0 \to \hat{H}_0$ which is exact if the first sequence is exact. Then also the sequence $\hat{F}_0/F \to \hat{G}_0/G \to \hat{H}_0/H$ is exact. It follows by recurrence that for

every p there is an exact sequence of mappings between the sheafs of the canonical resolutions: $\hat{F}_p \to \hat{G}_p \to \hat{H}_p$. In particular, the resolution F_* induces long exact sequences of sheafmorphisms between soft sheafs:

$$0 \to \hat{F}_p \to (F_0)\hat{}_p \to (F_1)\hat{}_p \to \dots \qquad p \geqslant 0,$$

where $(F_i)\hat{}_p$ is the p^{th} term in the canonical resolution of F_i.

Denote

$$A = \Gamma(\Omega,F)$$
$$A_i = \Gamma(\Omega,F_i)$$
$$A^p = \Gamma(\Omega,\hat{F}_p)$$
$$A_i^p = \Gamma(\Omega,F_i)\hat{}_p),$$

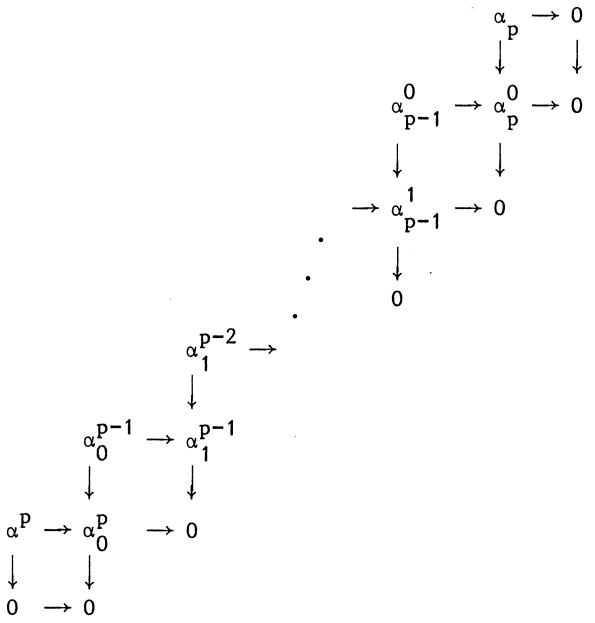
then we get the double commuting complex

$$\begin{array}{ccccccc}
 & & 0 & 0 & 0 & & \\
 & & \downarrow & \downarrow & \downarrow & & \\
0 & \to & A & \to A_0 & \to A_1 & \to & \dots \\
 & & \downarrow & \downarrow & \downarrow & & \\
0 & \to & A^0 & \to A_0^0 & \to A_1^0 & \to & \dots \\
 & & \downarrow & \downarrow & \downarrow & & \\
0 & \to & A^1 & \to A_0^1 & \to A_1^1 & \to & \dots \\
 & & \downarrow & \downarrow & \downarrow & & \\
 & & \vdots & \vdots & \vdots & &
\end{array}$$

Here the first row is the sequence of mappings between sections corresponding to the resolution F_* and according to theorem 3 the other sequences are exact. By passing to sections the first column is induced by the canonical resolution $\hat{F}$ of F. The other columns are induced by the canonical resolutions $\hat{F}_i$ of the sheafs F_i in the resolution F_* and they are exact because $H^p(\Omega,\hat{F}_i) = 0$, $p \geqslant 1$, $i \geqslant 0$, for the resolution F_* is acyclic (as long as theorem 4 is not proved we take this as definition of an acyclic

resolution; then the formulation of definition 6 follows as soon as theorem 4 has been proved).

Now we only have to chase diagrams. Let $\alpha^p \in \mathrm{Ker}(A^p \to A^{p+1})$ determine a class in the p^{th} cohomology group $H^p(\Omega,\hat{F})$ of the first column. The image of this class in the p^{th} cohomology group $H^p(\Omega,F_*)$ of the first sequence will be the class there determined by α_p, where α_p is obtained from α^p by chasing diagrams schematically indicated as follows:



with $\alpha_i^p \in A_i^p$.

Another choice in the same class as α^p or other choices of $\alpha_0^{p-1}, \alpha_1^{p-2}, \ldots, \alpha_{p-1}^0$ with the same properties yield another α_p, but in the same class as the original. Moreover, the way back yields the inverse: $H^p(\Omega,F_*) \to H^p(\Omega,\hat{F})$. ∎

So we can take the canonical resolution in determining cohomology groups and then, just as in the first part of the above proof, it follows that a sheafmorphism $F \to G$ induces in a natural way a map: $H^p(\Omega,F) \to H^p(\Omega,G)$

for $\Omega \subset X$ and $p \geqslant 0$.

Working with an acyclic resolution on an open set of a sheaf F is in fact an ideal situation. In practice useful resolutions are not always acyclic, cf. for example the following chapter on the Penrose transformation where resolutions are constructed using field equations. Then theorem 4 can not directly be applied. However, actually the requirements in the theorem are somewhat too strong: for the conclusion $H^1(\Omega,\hat{F}) \cong H^1(\Omega,F_*)$ only $H^1(\Omega,F_0) = 0$ is needed and for $H^2(\Omega,\hat{F}) \cong H^2(\Omega,F_*)$ only $H^1(\Omega,F_0) = H^2(\Omega,F_0) = H^1(\Omega,F_1) = 0$, etc. This also follows from the theorem in the following section, which gives more information than theorem 4.

REMARK.

The sheafs $\hat{F}_i$ in a canonical resolution satisfy a stronger property than being soft, namely for every open $\Omega' \subset \Omega$ the restriction map $\hat{F}_i(\Omega) \to \hat{F}_i(\Omega')$ is surjective. Such a sheaf $\hat{F}_i$ is called flabby. This enables us to use flabby instead of soft resolutions and especially theorem 1 is true for a flabby sheaf F and a space X which no longer need to be paracompact. Hence the theory of this section is also valid for spaces X which are not paracompact. However, except sheafs of discontinuous sections we will not meet other flabby sheafs.

4. THE LONG EXACT SEQUENCE OF SECTIONS

If we have a short exact sequence of sheafmorphisms we have seen (for example in section 1.A) that the corresponding sequence of mappings between sections over an open set might not be exact at the third place. It is possible to continue the sequence not with zero but with maps to other spaces so that an (in principle infinite long) exact sequence is obtained. This is done in the following theorem.

THEOREM 5. *Let* $0 \to F \to G \to H \to 0$ *be a short exact sequence of sheaf morphisms on an open set* Ω. *Then there are canonical (i.e. defined in a natural way) maps* δ *such that the following sequence of mappings between sections over* Ω *is exact*

$$0 \to F(\Omega) = H^0(\Omega,F) \to H^0(\Omega,G) \to H^0(\Omega,H) \xrightarrow{\delta} H^1(\Omega,F) \to$$
$$\to H^1(\Omega,F) \to H^1(\Omega,H) \xrightarrow{\delta} H^2(\Omega,F) \to H^2(\Omega,G) \to \ldots$$
$$\ldots \to H^p(\Omega,H) \xrightarrow{\delta} H^{p+1}(\Omega,F) \to \ldots .$$

The theorem can be proved with diagram chasing as in the proof of theorem 4. The maps δ follow in a natural way and we will discuss it in more detail after the following section where yet another way of computing cohomology groups is given.

What is the connection between theorems 4 and 5? Let

$$0 \to F \xrightarrow{i} F_0 \xrightarrow{d_0} F_1 \xrightarrow{d_1} F_2 \xrightarrow{d_2} F_3 \xrightarrow{d_3} \ldots$$

be a resolution of F. It follows from theorem 5 that for an open set Ω

$$0 \to F(\Omega) \xrightarrow{i} F_0(\Omega) \xrightarrow{d_0} \Gamma(\Omega,\operatorname{Ker} d_1) \xrightarrow{\delta} H^1(\Omega,F) \to H^1(\Omega,F_0)$$

is an exact sequence. Hence, if $H^1(\Omega,F_0) = 0$ then

$$H^1(\Omega,F) \cong \Gamma(\Omega,\operatorname{Ker} d_1)/\operatorname{Im} d_0(\Omega) ,$$

which is exactly the conclusion of theorem 4 (cf. the discussion following the proof of theorem 4 in section 3). For determining $H^2(\Omega,\mathcal{F})$ we go somewhat further in the sequence

$$\ldots \to H^1(\Omega,\mathcal{F}_0) \xrightarrow{d_0} H^1(\Omega,\operatorname{Ker} d_1) \xrightarrow{\delta} H^2(\Omega,\mathcal{F}) \to H^2(\Omega,\mathcal{F}_0).$$

Hence, if $H^1(\Omega,\mathcal{F}_0) = H^2(\Omega,\mathcal{F}_0) = 0$, then $H^2(\Omega,\mathcal{F}) \cong H^1(\Omega,\operatorname{Ker} d_1)$.

Now again according to theorem 5

$$0 \to \Gamma(\Omega,\operatorname{Ker} d_1) \to \Gamma(\Omega,\mathcal{F}_1) \xrightarrow{d_1} \Gamma(\Omega,\operatorname{Ker} d_2) \xrightarrow{\delta} H^1(\Omega,\operatorname{Ker} d_1) \to H^1(\Omega,\mathcal{F}_1)$$

is exact. Thus if moreover $H^1(\Omega,\mathcal{F}_1) = 0$ then

$$H^2(\Omega,\mathcal{F}) \cong \Gamma(\Omega,\operatorname{Ker} d_2)/\operatorname{Im} d_1(\Omega),$$

which is again the conclusion of theorem 4.

Therefore, in the calculation of a cohomolohy group $H^p(\Omega,\mathcal{F})$ for a fixed $p \geqslant 1$ with the aid of a resolution of $\mathcal{F}$ the requirement in theorem 4 that the resolution must be acyclic on Ω is a little too strong. In such cases theorem 5 is often more useful than theorem 4. In the following chapter we will deal with something like this. There in the calculation of left-handed fields we have a choice of using theorem 4 or theorem 5, but for right-handed fields we depend on theorem 5 only.

EXAMPLE. As an application of theorem 5 we will discuss $H^p(\Omega,\mathcal{O}^*)$ for $\Omega \subset \mathbb{C}^n$, cf. the example in section 1.A. From $0 \to \mathbb{Z} \to \mathcal{O} \to \mathcal{O}^* \to 0$ we obtain the exact sequence

$$\ldots \to H^p(\Omega,\mathcal{O}) \to H^p(\Omega,\mathcal{O}^*) \xrightarrow{\delta} H^{p+1}(\Omega,\mathbb{Z}) \to H^{p+1}(\Omega,\mathcal{O}) \to \ldots .$$

Assume that $H^p(\Omega,\mathcal{O}) = 0$ for $p \geqslant 1$. Then

$$H^p(\Omega,\mathcal{O}^*) \cong H^{p+1}(\Omega,\mathbb{Z}), \qquad p \geqslant 1.$$

So if $H^{p+1}(\Omega,\mathbb{Z})$ has been calculated, one has $H^p(\Omega,\mathcal{O}^*)$, too.

Let us consider the condition $H^p(\Omega,\mathcal{O}) = 0$, $p \geqslant 1$. From the Dolbeault resolution of $\mathcal{O}$ it follows that this is true for $p \geqslant n+1$ and any open $\Omega \subset \mathbb{C}^n$. Moreover, for $n = 1$ also $H^1(\Omega,\mathcal{O}) = 0$ for any open $\Omega \subset \mathbb{C}$; this follows from solving the differential equation induced by the Dolbeault resolution, cf. [Hörmander, theorem 1.4.4]. Hörmander's method is not valid if Ω = Riemann sphere $\cong P_1(\mathbb{C})$, but in the following section we will find a method of showing that $H^1(P_1(\mathbb{C}),\mathcal{O}) = 0$. (Again, that $H^p(P_1(\mathbb{C}),\mathcal{O}) = 0$ for $p \geqslant 2$ follows from the Dolbeault resolution that is valid on the Riemann sphere, too). The method performed at the end of section 1.B shows that $H^2(P_1(\mathbb{C}),\mathbb{Z}) \cong \mathbb{Z}$ (for the justification of this method cf. the next section). Hence $H^1(P_1(\mathbb{C}),\mathcal{O}^*) \cong \mathbb{Z}$.

Let us now assume $n > 1$. Then it is no longer true that $H^p(\Omega,\mathcal{O}) = 0$, $p \geqslant 1$, for every open set $\Omega \subset \mathbb{C}^n$. The open sets Ω for which $H^p(\Omega,\mathcal{O}) = 0$ for all $p \geqslant 1$ are the so-called *pseudoconvex* sets. We will not give the usual definition of pseudoconvexity (that this is equivalent with the vanishing of the cohomology groups $H^p(\Omega,\mathcal{O})$, $p \geqslant 1$, is far from trivial; it follows from [Hörmander, theorems 4.2.8 and 4.2.9, corollary 4.2.6 and theorem 7.4.1]), but we will merely give some properties of pseudoconvex open sets $\Omega \subset \mathbb{C}^n$:

1^o. convex sets are pseudoconvex

2^o. all open sets in $\mathbb{C}^1$ are pseudoconvex

3^o. all products $\Omega_1 \times \ldots \times \Omega_k \subset \mathbb{C}^{n_1+\ldots+n_k}$ of pseudoconvex open sets $\Omega_j \in \mathbb{C}^{n_j}$ are pseudoconvex.

4^o. intersections of pseudoconvex sets are pseudoconvex.

So we have

$$H^p(\Omega,\mathcal{O}^*) \cong H^{p+1}(\Omega,\mathbb{Z}), \quad p \geqslant 1, \quad \Omega \subset \mathbb{C}^n \text{ pseudoconvex.}$$

5. ČECH-COHOMOLOGY

We will discuss another useful representation of cohomology groups as has been performed in section 1.

Let $U = (\Omega_i)_{i\in I}$ be an open covering of a topological space X and denote

$$\Omega_{i_0\dots i_p} = \Omega_{i_0} \cap \dots \cap \Omega_{i_p}, \qquad p \geqslant 0.$$

If F is a sheaf on X, by $C^p(U,X,F)$ we mean the set (additive group) of alternating (i.e. antisymmetric in their indices) elements of

$$\prod_{(i_0,\dots,i_p)} F(\Omega_{i_0\dots i_p})$$

where $F(\emptyset) = \{0\}$. If I is totally ordered $C^p(U,X,F)$ is isomorphic with

$$\prod_{i_0<i_1<\dots<i_p} F(\Omega_{i_0\dots i_p}).$$

The elements are called *alternating* p-*cochains* of F on X with respect to U.

We define a map δ, the coboundary operator, as follows

$$\delta : C^p(U,X,F) \longrightarrow C^{p+1}(U,X,F);$$

let $f \in C^p(U,X,F)$ be determined by the alternating family $f_{i_0\dots i_p} \in F(\Omega_{i_0\dots i_p})$, then $\delta f \in C^{p+1}(U,X,F)$ is defined by

$$(\delta f)_{i_0\dots i_{p+1}} = \sum_{j=0}^{p+1} (-1)^j f_{i_0\dots \overset{\vee}{i_j}\dots i_{p+1}},$$

where

$$f_{i_0\dots \overset{\vee}{i_j}\dots i_{p+1}} = f_{i_0\dots i_{j-1}\, i_{j+1}\dots i_{p+1}} \Big|_{\Omega_{i_0\dots i_{p+1}}}.$$

By calculation it follows that $\delta \circ \delta = 0$.

If $\Omega \subset X$ is an open set, then U also yields a covering $(\Omega_i \cap \Omega)_{i \in I}$ of Ω, which we denote by U, too. In this way $C^p(U,\Omega,F)$ is determined and with the usual restriction mappings

$$\Omega \to C^p(U,\Omega,F)$$

defines a sheaf $C^p(U,F)$ on X.

THEOREM 6. *The complex*

$$0 \to F \xrightarrow{j} C^0(U,F) \xrightarrow{\delta} C^1(U,F) \xrightarrow{\delta} C^2(U,F) \xrightarrow{\delta} \dots ,$$

where j *is the restriction map, is a resolution of* F *(this means that the sequence is exact).*

PROOF.

Definition 2.i) (sheaf definition) implies that j is injective and definition 2.ii) that $\operatorname{Im} j = \operatorname{Ker}[C^0(U,F) \xrightarrow{\delta} C^1(U,F)]$. So let $p \geqslant 1$ and let ω be a neighbourhood of $x \in X$ so small that $\omega \subset \Omega_i$ for some $i \in I$. Hence $\omega \cap \Omega_i \cap \Omega_{i_0 \dots i_{p-1}} = \omega \cap \Omega_{i_0 \dots i_{p-1}}$. For $\alpha \in C^p(U,\omega,F)$ we define $\beta \in C^{p-1}(U,\omega,F)$ by

$$\beta_{i_0 \dots i_{p-1}} = \alpha_{i i_0 \dots i_{p-1}}.$$

Then

$$(\delta\beta)_{i_0 \dots i_p} = \sum_{j=0}^{p} (-1)^j \alpha_{i i_0 \dots \overset{\vee}{i_j} \dots i_p}$$

and if $\delta\alpha = 0$, moreover

$$\alpha_{i_0 \dots i_p} = \sum_{j=0}^{p} (-1)^j \alpha_{i i_0 \dots \overset{\vee}{i_j} \dots i_p} \quad \text{in} \quad \omega \cap \Omega_i \cap \Omega_{i_0 \dots i_p} = \omega \cap \Omega_{i_0 \dots i_p}.$$

Thus $\delta\alpha = 0$ implies $\delta\beta = \alpha$. ∎

We denote the cohomology groups (Čech-cohomology groups) of the derived complex of mappings between sections over Ω by

$$H^p(\mathcal{U},\Omega,\mathcal{F}) \overset{\text{def}}{=} \frac{\operatorname{Ker}[C^p(\mathcal{U},\Omega,\mathcal{F}) \xrightarrow{\delta} C^{p+1}(\mathcal{U},\Omega,\mathcal{F})]}{\operatorname{Im}[C^{p-1}(\mathcal{U},\Omega,\mathcal{F}) \xrightarrow{\delta} C^p(\mathcal{U},\Omega,\mathcal{F})]} \overset{\text{def}}{=}$$

$$= \frac{\{p\text{-cocycles with values in } \mathcal{F} \text{ with respect to } \mathcal{U}\}}{\{p\text{-coboundaries with values in } \mathcal{F} \text{ with respect to } \mathcal{U}\}}, \quad p \geqslant 1.$$

DEFINITION 7. The open covering $\mathcal{U}$ of Ω is called *acyclic* for $\mathcal{F}$ if

$$H^q(\Omega_{i_0 \dots i_p} \cap \Omega, \mathcal{F}) = 0, \quad \forall q \geqslant 1, \quad \forall p \geqslant 0, \quad \forall \Omega_i \in \mathcal{U}.$$

EXAMPLES.

1°. If $\mathcal{F}$ is soft (for example $\mathcal{E}$) every covering is acyclic.

2°. If $\mathcal{F} = \mathcal{O}$ a covering $\mathcal{U}$ of Ω is acyclic if every Ω_i is pseudo-convex and $\Omega_i \subset \Omega \subset \mathbb{C}^n$.

3°. The covering of the sphere S_2 used at the end of section 1.B is acyclic for $\mathbb{Z}$ and $\mathbb{R}$; for if $\Omega \subset \mathbb{R}^2$ is simply connected then $H^p(\Omega,\mathbb{Z}) = H^p(\Omega,\mathbb{R}) = 0$, $p \geqslant 1$. This can be shown by other means of calculating cohomology groups not discussed in this chapter.

THEOREM 7. *If the open covering $\mathcal{U}$ of Ω is acyclic for $\mathcal{F}$ then*

$$H^p(\Omega,\mathcal{F}) \cong H^p(\mathcal{U},\Omega,\mathcal{F}).$$

PROOF. The canonical resolution of $\mathcal{F}$ induces in a natural way a soft resolution of $C^p(\mathcal{U},\mathcal{F})$: (if $\mathcal{U}$ is not locally finite, use that the canonical resolution is flabby)

$$0 \to C^p(\mathcal{U},\mathcal{F}) \to C^p(\mathcal{U},\hat{\mathcal{F}}_0) \to C^p(\mathcal{U},\hat{\mathcal{F}}_1) \to \dots .$$

From definition 5 (definition of cohomology groups) and from $H^q(\Omega_{i_0 \dots i_p} \cap \Omega, \mathcal{F}) = 0$ if follows that the corresponding complex of mappings

between sections over Ω is exact. Then it follows from theorem 4 (or again from definition 5) that $H^q(\Omega, C^p(\mathcal{U},\mathcal{F})) = 0$, $q \geqslant 1$, $p \geqslant 0$. Hence the resolution of theorem 6 is acyclic on Ω and the theorem follows again from theorem 4. ∎

EXAMPLES.

1°. $X = P_1(\mathbb{C})$, $\mathcal{F} = \mathcal{O}$. We cover $P_1(\mathbb{C})$ with $U = \{z \in \mathbb{C} \mid |z| < 2\}$ and $\hat{U} = \{z \in \mathbb{C} \cup \{\infty\} \mid |z| > 1\}$. Since any open set in $\mathbb{C}$ is pseudoconvex this covering is acyclic. A function f holomorphic in $U \cap \hat{U}$ can be written as a Laurent series and by splitting this sum into a term with positive powers of z and a term with negative powers of z we have written f as a 1-coboundary. Hence $H^p(P_1(\mathbb{C}),\mathcal{O}) = 0$, $p = 1$ and for $p > 1$ this follows from the fact that all p-cocycles with respect to the covering $U \cup \hat{U}$ vanish.

2°. Example 1.B dealt explicitly with $H^2(X,R) \cong H^2(\mathcal{U},X,R)$ (note that without mentioning we used the acyclicity of the covering $\mathcal{U}$ in defining the constants r_{ijk}).

3°. As can be shown similarly to the end of section 1.B $H^2(S_2,\mathbb{Z}) = \mathbb{Z}$.

We will now compare the cohomology groups belonging to coverings which become finer and finer. A refinement $\mathcal{W}$ of an open covering $\mathcal{U}$ is an open covering such that $W_j \in \mathcal{W} = (W_j)_{j\in J}$ is contained in some $U_{\rho(j)} \in \mathcal{U} = (U_i)_{i\in I}$ for every $j \in J$. The map $\rho: J \to I$ induces a map (also called ρ): $C^p(\mathcal{U},X,\mathcal{F}) \to C^p(\mathcal{W},X,\mathcal{F})$ by

$$(\rho c)_{j_0\ldots j_p} = c_{\rho(j_0)\ldots\rho(j_p)} \Big|_{W_{j_0}\cap\ldots\cap W_{j_p}}.$$

This map commutes with the coboundary operators and hence induces a map: $H^p(\mathcal{U},X,\mathcal{F}) \to H^p(\mathcal{W},X,\mathcal{F})$ that appears to be independent of the choice of ρ

which describes the refinement $\mathcal{W}$ of $\mathcal{U}$. For, let $\rho' : J \to I$ be another map with $W_j \subset U_{\rho'(j)}$, $j \in J$. We calculate the difference $\rho'c - \rho c$ for $c \in C^p(\mathcal{U},X,F)$: let the map $k : C^{p+1}(\mathcal{U},X,F) \to C^p(\mathcal{W},X,F)$ be defined by

$$(kc)_{j_0 \dots j_p} = \sum_{\ell=0}^{p} (-1)^{\ell} c_{\rho(j_0)\dots\rho(j_\ell)\rho'(j_\ell)\dots\rho'(j_p)}$$

if $j_0 < j_1 < \dots < j_p$ and J is totally ordered. By computing it one gets

$$k\delta c + \delta k c = \rho' c - \rho c \in C^p(\mathcal{W},X,F)$$

Hence if $\delta c = 0$ $\rho' c - \rho c$ is a p-coboundary in $C^p(\mathcal{W},X,F)$.

Every two open coverings $\mathcal{U}$ and $\mathcal{V}$ of X have a common refinement $\mathcal{W}$. In this way we obtain maps

$$\begin{array}{ccc} H^p(\mathcal{U},X,F) & & \\ & \searrow & \\ & & H^p(\mathcal{W},X,F). \\ & \nearrow & \\ H^p(\mathcal{V},X,F) & & \end{array}$$

If the images of this map coincide for a refinement $\mathcal{W}$ of $\mathcal{U}$ and $\mathcal{V}$, we identify $H^p(\mathcal{U},X,F)$ with $H^p(\mathcal{V},X,F)$. This process defines the direct limit

$$\varinjlim_{\mathcal{U}} H^p(\mathcal{U},X,F).$$

THEOREM 8. *For paracompact* X

$$\varinjlim_{\mathcal{U}} H^p(\mathcal{U},X,F) \cong H^p(X,F), \qquad p \geqslant 0.$$

PROOF.

For any two open coverings $\mathcal{U}$ and $\mathcal{W}$ such that $\mathcal{W}$ is a refinement of $\mathcal{U}$ we choose a particular map ρ describing this refinement. Depending on these choices a direct limit of cochains can be defined. So we can consider the complex

$$0 \to \Gamma(X,F) \to \varinjlim_{U} C^0(U,X,F) \xrightarrow{\delta} \varinjlim_{U} C^1(U,X,F) \xrightarrow{\delta} \dots$$

where the maps δ are induced by the coboundary operators δ between cochains. The cohomology groups of this complex are independent of the maps ρ and they are equal to $\varinjlim_{U} H^p(U,X,F)$, which follows from the definition of direct limit.

Just as in theorem 6 it follows that the complex is derived from a resolution of F with the sheafs $\Omega \to \varinjlim_{U} C^p(U,\Omega,F)$. These sheafs are soft, because any covering U has a refinement $W = (W_j)_{j\in J}$ such that either $W_j \subset \Omega'$ or $W_j \cap A = \emptyset$ for a given closed set A in an open set Ω' and such that W is locally finite (which can be assumed by the paracompactness of X). Then according to definition 5 the cohomology groups of the complex equal $H^p(X,F)$, too. ∎

Hence the condition that a covering is acyclic can be replaced by the condition that it is arbitrarily fine. For example, if there are arbitrarily fine coverings U with $H^p(U,X,F) = 0$ then $H^p(X,F) = 0$, $p \geqslant 1$ and $H^p(X,F) = 0$ implies that there is at least one covering U with $H^p(U,X,F) = 0$ for $p \geqslant 1$. If $p = 1$ we have more:

$$H^1(X,F) = 0 \Rightarrow H^1(U,X,F) = 0 \text{ for every covering } U \text{ of } X.$$

This follows from:

THEOREM 9. *For an arbitrary open covering U the map:*

$$H^1(U,X,F) \to H^1(X,F)$$

is injective.

PROOF. Let $W = (W_s)_{s\in J}$ be a refinement of $U = (U_i)_{i\in I}$. We will show that every 1-cocycle $c \in C^1(U,X,F)$ (thus $\delta c = 0$) is a coboundary if $\rho c \in C^1(W,X,F)$ is a coboundary. The theorem then follows from theorem 8.

Let $\rho c = \delta\gamma$ with $\gamma \in C^0(W,X,\mathcal{F})$. Hence $c_{\rho(s)\rho(t)} = \gamma_t - \gamma_s$ in $W_s \cap W_t$ for all $s, t \in J$. Since $\delta c = 0$ it follows that

$$\gamma_t + c_{\rho(t)i} = \gamma_s + c_{\rho(s)i} \quad \text{in} \quad U_i \cap W_s \cap W_t.$$

Hence there is a section $c_i \in \Gamma(U_i,\mathcal{F})$ with $c_i = \gamma_s + c_{\rho(s)i}$ in $U_i \cap W_s$ for every s. For each s in $U_i \cap U_j \cap W_s$ we obtain

$$c_j - c_i = \gamma_s + c_{\rho(s)j} - \gamma_s - c_{\rho(s)i} = c_{\rho(s)j} + c_{i\rho(s)} = c_{ij}.$$

Thus c is a 1-coboundary. ∎

EXAMPLE *(Line bundles).*

If X is a complex manifold $H^1(X,\mathcal{O}^*)$ classifies the (equivalent) holomorphic line bundles over X. This follows from Čech cohomology: any holomorphic line bundle can be represented by an open covering $\mathcal{U} = (U_\alpha)$ and by transition functions $g_{\alpha\beta} \in \mathcal{O}^*(U_\alpha \cap U_\beta)$ satisfying the cocycle condition, thus by a 1-cocycle $\in C^1(\mathcal{U},X,\mathcal{O}^*)$ (note that the abelian group $\mathcal{O}^*$ is written multiplicative). The bundle is trivial if this 1-cocycle is a 1-coboundary. Hence any holomorphic line bundle determines (injectively) an element of $H^1(X,\mathcal{O}^*)$ by theorem 9. Conversely, a class in $H^1(X,\mathcal{O}^*)$ can be represented by a 1-cocycle with respect to some covering of X (theorem 8) and interpreted as transition functions this determines a holomorphic line bundle over X.

According to theorem 5 we have the exact sequence

$$\begin{array}{ccccccc} H^1(X,\mathcal{O}) & \to & H^1(X,\mathcal{O}^*) & \xrightarrow{\delta} & H^2(X,\mathbb{Z}) & \to & H^2(X,\mathcal{O}) \\ & & & & \downarrow j & & \\ & & & & H^2(X,R) & & \end{array}$$

where the vertical map j is induced by the sheafmorphism $\mathbb{Z} \to R$. Hence there is a map $j\delta : H^1(X,\mathcal{O}^*) \to H^2(X,R) \subset H^2(X,C)$, where the last inclusion follows from regarding real differential forms as complex differential forms. One can show that the first Chern class $c_1(E)$ of a line bundle E over X (as defined on p.177 in the proceedings 1981-82 of this seminar) is exactly the image of $j\delta$ if E is represented by its cohomology class in $H^1(X,\mathcal{O}^*)$ (cf. [Wells, prop.3.5 and theorem 4.5]).

If X is such that $H^1(X,\mathcal{O}) = H^2(X,\mathcal{O}) = 0$ then $H^1(X,\mathcal{O}^*) \cong H^2(X,\mathbb{Z})$, for example if $X = P_1(\mathbb{C})$ then $H^1(P_1(\mathbb{C}),\mathcal{O}^*) \cong H^2(P_1(\mathbb{C}),\mathbb{Z}) \cong \mathbb{Z}$, and the holomorphic line bundles are characterized by $H^2(X,\mathbb{Z})$. In general, this gives more information than its first Chern class, since the map j need not be injective.

Thus over such complex manifolds X where j is injective, the first Chern class $c_1(E)$ of a holomorphic line bundle can be represented by the map $\delta: H^1(X,\mathcal{O}^*) \to H^2(X,\mathbb{Z})$. The question whether j is injective depends purely on topological properties of the manifold X.

Differentiable complex line bundles over a differentiable manifold X are always characterized by their Chern class in $H^2(X,\mathbb{Z}) \cong H^1(X,\mathcal{E}^*)$ (here $\mathcal{E}$ (or $\mathcal{E}^*$) denotes the sheaf of complex valued (non-vanishing) C^∞-functions), because $H^p(X,\mathcal{E}) = 0$, $p \geq 1$, for any X. Similarly, by considering real C^∞-functions one finds that $H^2(X,\mathbb{Z})$ characterizes also the circle bundles over X.

Again, some information might get lost by looking at the real Chern class in $H^2(X,R)$; for example $H^2(P_n(\mathbb{R}),\mathbb{Z}) = \mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$ for $n \geq 2$, but $H^2(P_n(\mathbb{R}),R) = 0$.

4. (*continued*) The operators δ.

We will construct the operator $\delta = \tilde{\delta}$ in theorem 5 explicitly:

$$0 \to H^0(X,F) \to H^0(X,G) \to H^0(X,H) \xrightarrow{\tilde{\delta}} H^1(X,F) \xrightarrow{\tilde{\phi}} H^1(X,G).$$

Although we do not give the diagram, the construction actually follows from diagram chasing. In section 1.A we have already seen that $H^1(X,F)$ is the obstruction for obtaining global sections in H from those in G. The construction follows the pattern performed in section 1.

Let $h \in H^0(X,H) = H(X)$. Since the sequence

$$0 \to F \xrightarrow{\phi} G \xrightarrow{\psi} H \to 0$$

is exact every point $x_i \in X$ has a neighbourhood ω_i such that there is a section $g_i \in G(\omega_i)$ with $\psi(g_i) = h|_{\omega_i} = h_i$. Let $\mathcal{U} = (\omega_i)_i$ be the covering determined by this process and let g and h be the 0-cochains in $C^0(\mathcal{U},X,G)$ determined by g_i and h_i (so we identify a section $h \in H(X)$ with the 0-cocycle derived from it). We have $\psi(\delta g) = \delta\psi(g) = \delta h = 0$, where δ is the coboundary operator. Since ϕ is injective there exists an $f \in C^1(\mathcal{U},X,F)$ with $\phi(f) = \delta g$ and with $\delta f = 0$. Hence the section h determines a 1-cocycle $f \in C^1(\mathcal{U},X,F)$. If h comes from a section in $G(X)$ then we can choose the 0-cochain g such that $\delta g = 0$. Hence $f = 0$ then.

Let g_i' be another section in $G(\omega_i)$ with $\psi(g_i') = h_i$. Then there is another $f' \in C^1(\mathcal{U},X,F)$ with $\phi(f') = \delta g'$, or $\phi(f_{ij}') = g_j' - g_i'$ in $\omega_i \cap \omega_j$ for $f_{ij}' \in F(\omega_i \cap \omega_j)$. Thus the construction is not unique; the difference $f' - f$ is a 1-coboundary: for $\psi(g_i' - g_i) = 0$ so that there exists $\tilde{f}_i \in F(\omega_i)$ with $\phi(\tilde{f}_i) = g_i' - g_i$ and we have $\phi(\tilde{f}_j - \tilde{f}_i) = g_j' - g_j - g_i' + g_i = \phi(f_{ij}' - f_{ij})$ in $\omega_i \cap \omega_j$. Since ϕ is injective it follows that $f' - f$ is 1-coboundary. Hence the section h determines an element in $H^1(\mathcal{U},X,F)$, thus in $H^1(X,F)$.

In a similar way the other maps δ in theorem 5 can be constructed. In fact, the important thing is to find an appropriate covering. In the above construction this depends on the particular section h. If there is

a covering $U = (\omega_i)_i$ of X which is acyclic for F then for each section h we can take this covering: for

$$0 \to F(\omega_i) \xrightarrow{\phi} G(\omega_i) \xrightarrow{\psi} H(\omega_i) \to H^1(\omega_i,F) = 0$$

is exact (cf. the proof of theorem 1 and theorem 9) so that every $h \in H(\omega_i)$ comes from a $g \in G(\omega_i)$

We will now discuss the exactness of the sequence of theorem 5. We have already seen that $\psi(G(X)) \subset \operatorname{Ker} \tilde{\delta}$. Conversely, let now $f = \tilde{\delta}h \in C^1(U,X,F)$ be a 1-coboundary with respect to some refinement W of U. Then $\rho f = \delta \tilde{f}$ for some $\tilde{f} \in C^0(W,X,F)$ and from $\phi(f) = \delta g$ it follows that $\delta \rho g = \phi(\rho f) = \phi(\delta \tilde{f}) = \delta \phi(\tilde{f})$. Thus $\delta(\rho g - \phi \tilde{f}) = 0$ so that this determines a section $\tilde{g} \in G(X)$ with $\psi(\tilde{g}) = \psi(\rho g) - \psi(\phi(\tilde{f})) = \psi(\rho g) = \rho \psi(g) = h$ in the sets of the covering W. Hence $\psi(G(X)) = \operatorname{Ker} \tilde{\delta}$ and the sequence is exact at the third place.

For exactness at the next place we argue as follows. Since for the 1-cocycle f constructed in the definition of the map $\tilde{\delta}$ the formula $\phi(f) = \delta g$ holds $\phi(f)$ is a 1-coboundary with respect to the covering U with values in G. Thus $\phi(f)$ determines the zero class in $H^1(U,X,G)$ hence in $H^1(X,G)$ so that $\tilde{\phi}(f) = 0$. Conversely, let now $f \in \operatorname{Ker} \tilde{\phi} \subset H^1(X,F)$. This means that there is a covering U of X with $f \in H^1(U,X,F)$ (theorem 8) and that there is a refinement $W = (\omega_i)_i$ such that $\phi(\rho f)$ is a 1-coboundary $\in C^1(W,X,G)$. So there are $g_j \in G(\omega_j)$ with $\psi(g_j - g_i) = \psi(\phi(\rho f)) = 0$ in $\omega_i \cap \omega_j$. Hence $h_i = \psi(g_i)$ determines a section $h \in H(X)$ because $\delta h = \psi(\delta g) = 0$. This section h is mapped by $\tilde{\delta}$ onto $\rho f \in C^1(W,X,F)$, which determines the same cohomology class in $H^1(X,F)$ as $f \in C^1(U,X,F)$. Thus this construction of h is the inverse of the map $\tilde{\delta}$. (Note that for a covering U of X which is acyclic for F and G we do not have to pass to a refinement (theorem 7)). Hence the

sequence $H^0(X,H) \xrightarrow{\tilde{\delta}} H^1(X,F) \xrightarrow{\tilde{\phi}} H^1(X,G)$ is exact.

In the next chapter, concerning the Penrose transformation, we will perform the construction given here explicitly using acyclic coverings. It will turn out that in fact formula (4.3) of the preceding chapter represents the map $\tilde{\delta}^{-1}$ and formula (4.2) the map $\tilde{\phi}$.

REFERENCES

HÖRMANDER, L., *An introduction to complex analysis in several variables* (chapter 7.3), North-Holland P.C., Amsterdam, 1973.

SHAPIRA, P., *Théorie des hyperfonctions* (chapitre B), Lecture Notes in Math. 126, Springer, Berlin, etc. 1970.

WELLS, R.O., *Differential analysis on complex manifolds* (chapters II and III), Grad.Texts in Math. 65, Springer, Berlin, 1980.

For a more complete study:

GODEMENT, R., *Topologie algébrique et théorie des Faisceaux*, Hermann & Cie. Paris, 1964.

THE PENROSE TRANSFORMATION

by

J.W. de ROEVER
Department of Mathematics
Twente University of Technology

1. INTRODUCTION

At several places the Penrose transformation has already been mentioned, cf. p.195-196 [proceedings of this seminar 1981/82], chapter "Self-dual Yang-Mills Equations" section 2, chapter "Computation of the Yang-Mills Potentials" section 3.B and chapter "Linear Fields and Yang-Mills Theory" section 4. It expresses holomorphic solutions of the linear massless field equations in terms of certain cohomology classes over a domain in $P_3(\mathbb{C})$. Now one may take advantage of the fact that cohomology groups can be expressed in different ways, cf. the preceding chapter. So Dolbeault-cohomology has been used in chapter "Linear Fields and Yang-Mills Theory" section 4, especially in theorems 7 and 8, while in this chapter we will use Čech-cohomology. The reason that cohomology groups are related to fields is that the field equations occur in a resolution of a certain sheaf on a certain space.

The main purpose of the Penrose transformation is not to produce new solutions, but to serve as a theoretical background pointing out the lines according to which more difficult equations could be solved. Especially, the construction of Ward which indeed produced new solutions of the non-linear self-dual Yang-Mills equations is directly based on the Penrose transformation for the Maxwell field. Actually the Penrose transformation solves the linear field equations in such a way that it is just as easy as the non-linear equations.

Holomorphic solutions in complex Minkowski space $M_{\mathbb{C}}$ can be used to

produce non-holomorphic solutions in real Minkowski space as follows. By choosing an origin in $M_{\mathbb{C}}$ we identify it with $\mathbb{C}^4$ and we consider the map

$$\alpha: \mathbb{C}^4 \times P_1(\mathbb{C}) \longrightarrow P_3(\mathbb{C}) \setminus [Z^1, Z^2, 0, 0]$$

defined by

$$\alpha(z, [\pi]) = [iz^{AA'}\pi_{A'}, \pi_{A'}],$$

where the matrix $z^{AA'} \equiv A_z$ is given in section 1 of chapter "Linear Fields and Yang-Mills Theory" (cf. also [formula (1.12), p.186 proceedings of this seminar 1981/82]) and where the physicists notations $z^{AA'}\pi_{A'} = (z^{00'}\pi_{0'}+z^{01'}\pi_{1'}, z^{10'}\pi_{0'}+z^{11'}\pi_{1'})$ and $\pi_{A'} = (\pi_{0'}, \pi_{1'}) = \pi$ has been used. The space $P_3(\mathbb{C})$ is divided into three parts $P_3^+ \subset P_3(\mathbb{C}) \setminus [Z^1, Z^2, 0, 0]$, P_3^0 and P_3^-, cf. [p.194 proceedings of this seminar 1981/82]. Then

$$M_{\mathbb{C}}^+ = \{z \in M_{\mathbb{C}} \mid \alpha(z, [\pi]) \in P_3^+, \forall [\pi] \in P_1(\mathbb{C})\}$$

is defined; it is a convex open subset of $M_{\mathbb{C}}$ such that the real Minkowski space M lies on its boundary. Thus the limit $x^\mu + iy^\mu \to x^\mu$ can be taken and so a holomorphic function f in $M_{\mathbb{C}}^+$ has a boundary value, often denoted by $f(x^\mu+i\varepsilon)$, which no longer need to be analytic, but might be a distribution (a hyperfunction in general) in M. The same can be done with functions holomorphic in $M_{\mathbb{C}}^-$. Hence it is not unimportant to consider holomorphic solutions in $M_{\mathbb{C}}^+$ of the field equations.

We mention some facts from chapter "Linear Fields and Yang-Mills Theory". Every massless field of helicity $|\frac{n}{2}|$ has a left-handed and a right-handed component (for $n < 0$ or $n > 0$, respectively). These components can be represented by symmetric spinors $\phi_{A\ldots D}$ and $\psi_{A'\ldots D'}$ ($|n|$ indices) satisfying the field equations

$$\nabla^{AA'}\phi_{A\ldots D} = 0, \quad \nabla^{AA'}\psi_{A'\ldots D'} = 0.$$

The fields ϕ and ψ must be functions (or distributions) of the real variables $x^\mu \in M$. Furthermore, the bundle L^m on $P_3(\mathbb{C})$ has been discussed in section 3 of chapter "Linear Fields and Yang-Mills Theory", cf. also [p.167, proceedings of this seminar 1981/82, where it has been denoted by E(-m)]; another notation sometimes is $\mathcal{O}(-m)$. We consider the set of its holomorphic sections as a sheaf $\mathcal{L}^m$ on $P_3(\mathbb{C})$.

The *Penrose transformation*, given by formula (4.3) and theorem 3 for the ϕ-field $(n<0)$ and by formula (4.2) and theorem 7 for the ψ-field $(n>0)$ in chapter "Linear Fields and Yang-Mills Theory", is a bijective correspondence between

$$H^1(P_3^+,\mathcal{L}^{n+2}) \longrightarrow \{\text{holomorphic massless fields in } M_{\mathbb{C}}^+ \text{ of helicity } \tfrac{n}{2}\}.$$

For $n \leqslant -2$ the ϕ-fields will be derived indirectly by means of potentials, but the ψ-fields will be treated directly. We will also discuss the case $n = 0$ which solves the wave equation. Since the self-dual Yang-Mills field is very similar to the left-handed (anti-self-dual) Maxwell field $(n = -2)$ we will pay special attention to this case continuing the discussion started in section 3.B, chapter "Computation of the Yang-Mills Potentials". Then we will place it in a more abstract setting before treating the other fields. Finally, we will give the relation with the construction of Ward, cf. chapters "Self-dual Yang-Mills Equations" and "Computation of the Yang-Mills Potentials". It will turn out that the left-handed Maxwell field is contained in Ward's Ansatz A_n for $n = 0$ for solving the self-dual Yang-Mills equations.

2. THE LEFT-HANDED MAXWELL FIELD (n = -2)

We will use Čech-cohomology and we will explicitly relate the potentials of the field to holomorphic 1-cocycles with respect to some covering of P_3^+. In the next section we will give a more direct interpretation of the Penrose transform using more theory of the preceding chapter.

In section 3.B of chapter "Computation of the Yang-Mills Potentials" potentials B_{PQ}, $P,Q \in \{1,2\}$ have been studied satisfying the self-duality equations (29) and (30). Because of the Euclidean metric the matrix z^{PQ}, given by formula (1) there, has been used. Here we have the Minkowski metric and therefore, instead of z^{PQ}, we rather work with the matrix $iz^{AA'}$. Moreover, here we will denote the potentials in spinor notation. So we start with holomorphic potentials $\chi_{AB'}$ in $M_{\mathbb{C}}^+$ with $A,B \in \{0,1\}$ satisfying the equations (29) and (30). The anti-self-dual Maxwell field ϕ_{AB} is derived from the potentials according to

$$\phi_{AB} = \tfrac{1}{2}\left(\nabla_{AA'}\chi_B{}^{A'} + \nabla_{BA'}\chi_A{}^{A'}\right), \qquad \nabla_{AA'} = \frac{\partial}{\partial z^{AA'}}.$$

We will first show the surjectivity of the Penrose transformation. In section 3.B, chapter "Computation of the Yang-Mills Potentials", for any point $\tilde{z} \in M_{\mathbb{C}}^+$ there have been found domains

$$\Omega = \Omega(\tilde{z}) = \{(z,\zeta) \mid z \in \omega(\tilde{z}),\ |\zeta| < m\gamma\} \subset M_{\mathbb{C}} \times P_1(\mathbb{C})$$

and $$\hat{\Omega} = \hat{\Omega}(\tilde{z}) = \{(z,\zeta) \mid z \in \omega(\tilde{z}),\ |\zeta| > \tfrac{1}{m}\gamma'\} \subset M_{\mathbb{C}} \times P_1(\mathbb{C})$$

with $$\omega = \omega(\tilde{z}) = \{z \in \mathbb{C}^4 \mid |z^{AA'} - \tilde{z}^{AA'}| < \tfrac{1}{m+1}\delta_{AA'}\}$$

for certain γ and γ' depending on $\delta_{AA'}$, and moreover, holomorphic functions $\mu(z,\zeta)$ and $\hat{\mu}(z,\zeta)$ in Ω or $\hat{\Omega}$, respectively, such that

$$D_A\hat{\mu} = \chi_{A0'} - \zeta\chi_{A1'} = D_A\mu, \qquad D_A = \nabla_{A0'} - \zeta\nabla_{A1'}, \quad A = 0,1$$

for a given potential $\chi_{AA'}$ in $M^+_{\mathbb{C}}$. We take m so large that $\Omega \cap \hat{\Omega} \neq \emptyset$ and $\omega \subset M^+_{\mathbb{C}}$. The function $\hat{\mu} - \mu$ is holomorphic in $\Omega \cap \hat{\Omega}$ and satisfies $D_A(\hat{\mu}-\mu) = 0$, $A = 0,1$. Thus $\hat{\mu} - \mu$ is constant on α-planes (an α-plane in the complex Minkowski space $M_{\mathbb{C}}$ is formed by those $z^{AA'}$ for which $iz^{AA'}\pi_{A'} = \omega^A$ holds for given $(\omega^A, \pi_{A'}) \in \mathbb{C}^4 \setminus \{0\}$, or $\alpha(z, [\pi]) = (\omega^A, \pi_{A'})$; the operators D_A are just two independent tangent vectors to such a plane determined by π). Hence $\hat{\mu} - \mu$ determines a holomorphic function f, homogeneous of degree zero, by

$$f(\alpha(z, [\pi])) = \hat{\mu}(z,\zeta) - \mu(z,\zeta), \qquad \zeta = \pi_{0'}/\pi_{1'},$$

defined in $W \cap \hat{W} \subset P^+_3$ with

$$W = W(\tilde{z}) = \alpha(\Omega(\tilde{z})) = \{[Z] \mid Z = \alpha(z, [\pi]), (z, [\pi]) \in \Omega(\tilde{z})\} \subset P^+_3 \cap \{Z^4 \neq 0\}$$

and

$$\hat{W} = \hat{W}(\tilde{z}) = \alpha(\hat{\Omega}(\tilde{z})) \subset P^+_3 \cap \{Z^3 \neq 0\}.$$

This gives us a 1-cocycle with respect to the covering $\{W, \hat{W}\}$ of $W \cup \hat{W}$, hence an element of $H^1(W \cup \hat{W}, \mathcal{O})$. For a local field in ω this would be sufficient. However, we are dealing with global fields in $M^+_{\mathbb{C}}$, so we need a 1-cocycle with respect to the covering

$$\mathcal{W} = \bigcup_{\tilde{z} \in Q} \{W(\tilde{z}), \hat{W}(\tilde{z})\}$$

of P^+_3 where $Q \subset M^+_{\mathbb{C}}$ is a set of sufficiently many points.

The problem is that above one set $W \in \mathcal{W}$ there are many sets from the covering

$$\mathcal{U} = \bigcup_{\tilde{z} \in Q} \{\Omega(\tilde{z}), \hat{\Omega}(\tilde{z})\}$$

of $M^+_{\mathbb{C}} \times P_1(\mathbb{C})$ which intersect $\alpha^{-1}(W) = \{(z, [\pi]) \in M^+_{\mathbb{C}} \times P_1(\mathbb{C}) \mid \alpha(z, [\pi]) \in W\}$. Thus there is no natural way to pass from a function defined in $\Omega_k \cap \Omega_\ell$ to

one in $\alpha(\Omega_k \cap \Omega_\ell)$ with $\Omega_{k,\ell} \in \mathcal{U}$. In the intersection $\Omega_k \cap \Omega_\ell \cap \alpha^{-1}(W)$ of $\alpha^{-1}(W)$ with any two of these sets we have a holomorphic function $\mu_k - \mu_\ell$ which is constant on α-planes, i.e. we have a 1-cocycle with respect to the covering $\mathcal{U}$ of $\alpha^{-1}(W)$ with values in the sheaf $\alpha^*\mathcal{O}$ (the sheaf of holomorphic functions in $M^+_{\mathbb{C}} \times P_1(\mathbb{C})$ which are constant on α-planes). Let us assume that this 1-cocycle is a 1-coboundary with values in this sheaf. Then there are functions ν_k defined in $\Omega_k \cap \alpha^{-1}(W)$ which are constant on α-planes. If instead of μ_k we take $\mu'_k = \mu_k + \nu_k$, we still have the property $D_A\mu'_k = D_A\mu_k$, but now also $\mu'_k - \mu'_\ell = \mu_k - \mu_\ell - (\nu_\ell - \nu_k) = 0$ in $\Omega_k \cap \Omega_\ell \cap \alpha^{-1}(W)$ if this is not empty. Hence for any $W_j \in \mathcal{W}$ there exists a holomorphic function $\tilde{\mu}_j$ in $\alpha^{-1}(W_j)$ with $D_A\tilde{\mu}_j = \chi_{A0'}\zeta\chi_{A1'}$. By taking differences

$$f_{jk}(Z) = \tilde{\mu}_k(z,\zeta) - \tilde{\mu}_j(z,\zeta), \quad (z,\zeta) \in \alpha^{-1}(W_j \cap W_k), \quad Z = \alpha(z,\zeta) \in W_j \cap W_k$$

we have obtained a holomorphic 1-cocycle with respect to the covering $\mathcal{W}$ of P^+_3.

All the neighbourhoods $\omega(\tilde{z})$ of the points $\tilde{z} \in Q \subset M^+_{\mathbb{C}}$ are so small that $\omega(\tilde{z}) \subset M^+_{\mathbb{C}}$, hence the shape of the sets $W \in \mathcal{W}$ does not depend on the boundary of P^+_3 in $P_3(\mathbb{C})$. As a consequence it follows from the construction that all the sets $W \in \mathcal{W}$ are pseudoconvex in a suitable chart, cf. the example of section 4 in the preceding chapter. Hence for each $W \in \mathcal{W}$ we have $H^1(W,\mathcal{O}) = 0$ and in the appendix it is shown that $H^1(\alpha^{-1}(W), \alpha^*\mathcal{O}) = H^1(W,\mathcal{O})$ so that indeed the above constructed 1-cocycle $\{\mu_k - \mu_\ell\}$ is a 1-coboundary $\{\nu_k\}$. Note that necessarily the covering $\mathcal{W}$ of P^+_3 consists of infinitely many sets.

The only property of the functions μ that matters is that $D_A\mu$ has to be fixed for $A = 0,1$. Thus with other functions μ' with this property or with another choice of the 1-coboundary $\{\nu_k\}$ we would have obtained another 1-cocycle $\{f'_{jk}\}$. The difference $\tilde{\mu}_j - \tilde{\mu}'_j = m_j$ determines a holo-

morphic function in $\alpha^{-1}(W_j)$ which is constant on α-planes, hence a holomorphic function m_j of $[Z] \in W_j$. So in $W_j \cap W_k$ we have

$$f_{jk}(Z) - f'_{jk}(Z) = \tilde{\mu}_k(z,\zeta) - \tilde{\mu}_j(z,\zeta) - \tilde{\mu}'_k(z,\zeta) + \tilde{\mu}'_j(z,\zeta) = m_k(Z) - m_j(Z)$$

if $\alpha(z,\zeta) = Z$. Hence the 1-cocycles differ by a 1-coboundary, so that the potentials $\chi_{AA'}$ determine an element of $H^1(P_3^+,\mathcal{O})$.

We will now consider the converse. Let a 1-cocycle f of holomorphic functions be given with respect to an arbitrary covering $\mathcal{W}$ of P_3^+. This determines a 1-cocycle with holomorphic coefficients in $M_{\mathbb{C}}^+ \times P_1(\mathbb{C})$ with respect to the covering $\mathcal{U} = \bigcup_{W\in\mathcal{W}} \alpha^{-1}(W) \cap M_{\mathbb{C}}^+ \times P_1(\mathbb{C})$. Since $M_{\mathbb{C}}^+$ is convex (hence pseudoconvex) $H^1(M_{\mathbb{C}}^+ \times P_1,\mathcal{O}) = 0$. Thus according to theorem 9 of the preceding chapter the 1-cocycle in $M_{\mathbb{C}}^+ \times P_1(\mathbb{C})$ is a 1-coboundary. In particular there exist holomorphic functions $\tilde{\mu}_j$ in $\Omega_j \in \mathcal{U}$ with

$$f_{jk}(Z) = \tilde{\mu}_k(z,\zeta) - \tilde{\mu}_j(z,\zeta), \quad (z,\zeta) \in \Omega_j \cap \Omega_k, \quad \alpha(z,\zeta) = Z.$$

The possibilities for such functions $\tilde{\mu}_j$ are restricted. From analytic continuation and from Liouville's theorem it follows easily that they are determined up to a gauge function g in $M_{\mathbb{C}}^+$:

$$\tilde{\mu}_j(z,\zeta) + g(z), \quad \alpha(z,\zeta) \in W_j.$$

Since $D_A f_{jk} = 0$ we have $D_A\tilde{\mu}_j = D_A\tilde{\mu}_k$. This implies that the functions $\tilde{\mu}_j$ are linear in ζ (Liouville). So we find that

$$D_A(\tilde{\mu}_j+g)(z,\zeta) = \chi_{A0'}(z) + \nabla_{A0'}g(z) - \zeta(\chi_{A1'}(z)+\nabla_{A1'}g(z)) =$$

$$= D_A(\tilde{\mu}_k+g)(z,\zeta), \quad (z,\zeta) \in \Omega_j \cap \Omega_k, \quad A = 0,1.$$

Hence the functions $\chi_{AA'}$ are globally defined in $M_{\mathbb{C}}^+$ and they determine holomorphic potentials. Had f been a 1-coboundary in P_3^+, then

$$f_{jk}(Z) = \nu_k(Z) - \nu_j(Z)$$

and the potentials $\chi_{AA'}$ would vanish, because $D_A\nu_j = 0$.

Thus we have shown that there is a bijective correspondence

$$H^1(P_3^+,L^0) \cong \frac{\text{1-cocycles } f \text{ in } P_3^+}{\text{1-coboundaries}} \longrightarrow \frac{\text{potentials } \chi_{AA'} \text{ in } M_{\mathbb{C}}^+}{\text{gauge } \nabla_{AA'},g \text{ in } M_{\mathbb{C}}^+}$$

A (local) potential $\chi_{AA'}$ in $\omega(\tilde{z})$ can be obtained explicitly from a holomorphic function $f \in \mathcal{O}(W(\tilde{z}) \cap \hat{W}(\tilde{z}))$ as follows:

$$f(iz^{AA'}\pi'_{A'},\pi'_{A'}) = \hat{\mu}(z,\zeta') - \mu(z,\zeta') =$$

$$= \frac{1}{2\pi i} \oint_{\substack{\zeta' \text{ outside the circle}}} \frac{f(iz^{AA'}\pi_{A'},\pi_{A'})}{\zeta-\zeta'} d\zeta - \frac{1}{2\pi i} \oint_{\substack{\zeta' \text{ inside the circle}}} \frac{f(iz^{AA'}\pi_{A'},\pi_{A'})}{\zeta-\zeta'} d\zeta .$$

Since $\nabla_{AA'}f(iz^{AA'}\pi_{A'},\pi_{A'}) = i\pi_{A'}\frac{\partial}{\partial\omega^A}f(iz^{AA'}\pi_{A'},\pi_{A'})$ we have

$$D_A(\zeta')\mu(z,\zeta') = \frac{1}{2\pi i} \oint \frac{i(\pi_{0'},-\zeta'\pi_{1'})}{\pi_{0'}/\pi_{1'},-\zeta'} \frac{\partial}{\partial\omega^A} f(iz^{AA'}\pi_{A'},\pi_{A'})d\zeta =$$

$$= \frac{i}{2\pi i} \oint \frac{\partial}{\partial\omega^A} f(iz^{AA'}\pi_{A'},\pi_{A'})\pi_{1'}\, d\zeta =$$

$$= \text{independent of } \zeta', \text{ hence } = \chi_{A0'}(z), \quad z \in \omega(\tilde{z}).$$

Apparently, we have the gauge for which $\chi_{A1'} = 0$. In this gauge the field becomes

$$\phi_{AB} = -\tfrac{1}{2}(\nabla_{A1'}\chi_{B0'} + \nabla_{B1'}\chi_{A0'}) =$$

$$= -\tfrac{1}{2}\frac{i}{2\pi i} \oint i\pi_{1'}{}^2 \frac{\partial}{\partial\omega^A}\frac{\partial}{\partial\omega^B} f(iz^{AA'}\pi_{A'},\pi_{A'})\pi_{1'}\, d\left(\frac{\pi_{0'}}{\pi_{1'}}\right) =$$

$$= \frac{1}{2\pi i} \oint \frac{\partial}{\partial\omega^A}\frac{\partial}{\partial\omega^B} f(iz^{AA'}\pi_{A'},\pi_{A'})(\pi_{1'}\, d\pi_{0'} - \pi_{0'}\, d\pi_{1'}),$$

exactly the Penrose transformation (4.3) of chapter "Linear Fields and Yang-Mills Theory".

3. A COMPREHENSIVE TREATMENT OF THE CASE n = -2

Before continuing with the other left-handed fields we will summarize the foregoing treatment of the left-handed Maxwell field using the theory of the preceding chapter.

A holomorphic function f, homogeneous of degree zero, in $\mathbb{C}^4 \setminus \{0\}$ determines a holomorphic function in $P_3(\mathbb{C})$. Under the map

$$\alpha: M^+_{\mathbb{C}} \times P_1(\mathbb{C}) \to P^+_3$$

discussed in the introduction, one can pullback such a function to $M^+_{\mathbb{C}} \times P_1(\mathbb{C})$

$$(\alpha^* f)(z, [\pi]) = f(iz^{AA'}\pi_{A'}, \pi_{A'}).$$

The sheaf $L^0(=\mathcal{O})$ on P^+_3 so determines the sheaf $\alpha^* L^0$ on $M^+_{\mathbb{C}} \times P_1(\mathbb{C})$ which is a subsheaf of the sheaf $\mathcal{O}$ of all holomorphic functions of z and ζ. In the appendix it is shown that

$$H^1(P^+_3, L^0) \cong H^1(M^+_{\mathbb{C}} \times P_1, \alpha^* L^0) .$$

In order to calculate this cohomology group we make a resolution of $\alpha^* L^0$ on $M^+_{\mathbb{C}} \times P_1(\mathbb{C})$. Let $\mathcal{O}(m)$ be the sheaf on $M^+_{\mathbb{C}} \times P_1(\mathbb{C})$ determined by holomorphic functions of z and π, homogeneous of degree m in π, similar to the sheaf L^{-m} on $P^+_3(\mathbb{C})$, cf. theorem 3, chapter "Linear Fields and Yang-Mills Theory". Then we have the following resolution of the sheaf $\alpha^* L^0$

$$0 \to \alpha^* L^0 \xrightarrow{i} \mathcal{O} \xrightarrow{D} \mathcal{O}(1)^2 \xrightarrow{D'} \mathcal{O}(2) \to 0$$

where i is the embedding, D is determined by

$$\mu(z, [\pi]) \xmapsto{D} \tilde{D}_A \mu(z, [\pi]) = \pi_{1'} \nabla_{A0'} \mu - \pi_{0'} \nabla_{A1'} \mu = \pi^{A'} \nabla_{AA'} \mu$$

and D' by

$$(\mu_0,\mu_1) \overset{D'}{\longmapsto} \tilde{D}_0\mu_1 - \tilde{D}_1\mu_0.$$

The operators $\tilde{D}_A = \pi_{1'}D_A = \pi^{A'}\nabla_{AA'}$ are defined everywhere on $M_{\mathbb{C}}^+ \times P_1(\mathbb{C})$ now. The exactness of the sequence can be shown with the aid of power series, just as in section 3.B of chapter "Computation of the Yang-Mills Potentials" or by direct integration, cf. the last formula in the appendix; it is a kind of Poincaré-lemma and the sequence is called a relative de Rham complex.

We consider global sections over $M_{\mathbb{C}}^+ \times P_1(\mathbb{C})$ and according to theorem 5, preceding chapter, we get the exact sequence

$$0 \to \alpha^*L^0(M_{\mathbb{C}}^+\times P_1) \to \mathcal{O}(M_{\mathbb{C}}^+\times P_1) \to \operatorname{Ker} D'(M_{\mathbb{C}}^+\times P_1) \xrightarrow{\delta} H^1(M_{\mathbb{C}}^+\times P_1, \alpha^*L^0) \to 0$$

because $H^1(M_{\mathbb{C}}^+\times P_1, \mathcal{O}) = 0$. The elements of $\mathcal{O}(M_{\mathbb{C}}^+\times P_1)$ are holomorphic functions g of $z \in M_{\mathbb{C}}^+$ only, because of Liouville's theorem. The elements of $\operatorname{Ker} D'(M_{\mathbb{C}}^+\times P_1)$ are as follows: holomorphic functions, homogeneous of degree 1, on $\mathbb{C}^2 \setminus \{(0,0)\}$ are polynomials of degree 1 in $\pi_{0'}$ and $\pi_{1'}$; hence $(h_0,h_1) \in \operatorname{Ker} D'(M_{\mathbb{C}}^+\times P_1)$ means

$$h_A(z,\pi) = \pi_{1'}\chi_{A0'}(z) - \pi_{0'}\chi_{A1'}(z) = \pi_{A'}\chi_A{}^{A'}, \quad z \in M_{\mathbb{C}}^+, \quad \pi \in \mathbb{C}^2 \setminus \{0,0\}, \quad A = 0,1$$

for holomorphic functions $\chi_{AA'}$ in $M_{\mathbb{C}}^+$. The equation $\tilde{D}_1h_0 = \tilde{D}_0h_1$ just consists of the field equations (29) and (30) (chapter "Computation of the Yang-Mills Potentials") for the potentials $\chi_{AA'}$. So the field equations appear in the resolution of α^*L^0 by passing over to global sections over a projective space (Liouville).

In section 2 we have constructed the map δ as in the last part of section 4 of the preceding chapter, i.e. the bijection

$$\text{potentials/gauge} \xrightarrow{\delta} H^1(M_{\mathbb{C}}^+\times P_1, \alpha^*L^0) .$$

Furthermore, we have seen that

$$H^1(M_{\mathbb{C}}^{+}\times P_1, \alpha^* L^0) \xrightarrow{\delta^{-1}} \text{potentials/gauge} \longrightarrow \text{fields}$$

is the Penrose transformation.

4. LEFT-HANDED FIELDS

We consider left-handed fields $\phi_{A\ldots D}$ ($|n|$ indices) of helicity $|\frac{n}{2}|$ with $n < 0$. Similarly to the left-handed Maxwell field it is shown that for $n \leqslant -2$ the field can be derived from potentials $\chi_A{}^{B'\ldots D'}$ (spinor fields which are symmetric in the $|n|-1$ indices $B'\ldots D'$) according to

$$\phi_{A\ldots D} = \frac{1}{|n|!} \sum_{\sigma} \nabla_{\sigma_1 D'} \cdots \nabla_{\sigma_{|n|-1} B'} \chi_{\sigma_{|n|}}{}^{B'\ldots D'}$$

where $\sigma = (\sigma_1 \ldots \sigma_{|n|}) \in \{\text{permutations of the } |n| \text{ indices } A, \ldots, D\}$. In order that the $|n|+1$ field components satisfy the $2|n|$ field equations $\nabla^{AA'}\phi_{A\ldots D} = 0$ the $2|n|$ potentials have to satisfy the $|n|+1$ equations

$$\sum_{\sigma'} \nabla^{A\sigma'_1} \chi_A{}^{\sigma'_2 \ldots \sigma'_{|n|}} = 0$$

where $\sigma' = (\sigma'_1 \ldots \sigma'_{|n|}) \in \{\text{permutations of the } |n| \text{ indices } A', \ldots, D'\}$. So there is a freedom of $|n|-1$ gauge functions.

The resolution

$$0 \to \alpha^* L^{n+2} \to \mathcal{O}(|n|-2) \xrightarrow{D} \mathcal{O}(|n|-1)^2 \xrightarrow{D'} \mathcal{O}(|n|) \to 0$$

of the sheaf $\alpha^* L^{n+2}$ on $M^+_{\mathbb{C}} \times P_1(\mathbb{C})$ induces the exact sequence of sections

$$0 \to \alpha^* L^{n+2}(M^+_{\mathbb{C}} \times P_1) \to \mathcal{O}(|n|-2)(M^+_{\mathbb{C}} \times P_1) \xrightarrow{D} \operatorname{Ker} D'(M^+_{\mathbb{C}} \times P_1) \xrightarrow{\delta}$$
$$\xrightarrow{\delta} H^1(M^+_{\mathbb{C}} \times P_1, \alpha^* L^{n+2}) \to 0$$

by theorem 5, preceding chapter. Here we have used that $H^1(M^+_{\mathbb{C}} \times P_1, \mathcal{O}(-n-2)) = 0$ for $n < 0$ (cf. the end of this section). Sections $\chi \in \operatorname{Ker} D'(M^+_{\mathbb{C}} \times P_1)$ are determined by $|n|-1^{\text{th}}$ degree homogeneous polynomials in $\mathbb{C}^2 \setminus \{(0,0)\}$:

$$\chi_A{}^{B'\ldots D'}(z)\pi_{B'}\ldots\pi_{D'}, \qquad z \in M^+_{\mathbb{C}}$$

cf. theorem 3, chapter "Linear Fields and Yang-Mills theory"; the equation $D'\chi = 0$ can be written as

$$\nabla^{AA'}\chi_A{}^{B'\ldots D'}\pi_{A'}\ldots\pi_{D'} \equiv 0, \qquad \forall\pi = (\pi_{0'},\pi_{1'}) \in \mathbb{C}^2\setminus\{(0,0)\}.$$

These are just the equations the potentials for the ϕ-field have to satisfy.

Again we construct the map δ^{-1}. As in the foregoing sections it is clear that for $z \in \omega = \omega(\tilde{z})$ we may consider $f \in \Gamma(W(\tilde{z}) \cap \hat{W}(\tilde{z}), L^{n+2})$ only. If $n \leqslant -2$ we must have (cf. proof of theorem 3, chapter "Linear Fields and Yang-Mills Theory")

$$f(iz^{AA'}\pi_{A'},\pi_{A'}) = \hat{\mu}(z,\zeta)\pi_{0'}^{|n|-2} - \mu(z,\zeta)\pi_{1'}^{|n|-2}, \qquad z \in \omega$$

where $\hat{\mu}$ is defined in $\hat{\Omega}(\tilde{z})$ and μ in $\Omega(\tilde{z})$, i.e. for $|\zeta| > r_1$ and for $|\zeta| < r_2$, respectively, with $r_1 < r_2$. Hence $D_A\mu(z,\zeta) = \zeta^{|n|-2}D_A\hat{\mu}(z,\zeta)$ and by Liouville's theorem this equals a polynomial of degree $|n|-1$ in ζ. We can write this as

$$\pi_{1'}^{|n|-1}D_A\mu(z,\zeta) = \tilde{D}_A\pi_{1'}^{|n|-2}\mu(z,\zeta) = \chi_A{}^{B'\ldots D'}(z)\pi_{B'}\ldots\pi_{D'}$$

for functions $\chi_A{}^{B'\ldots D'}$ in $\omega \subset M^+_{\mathbb{C}}$. The freedom we have in this process is the following

$$\frac{f(iz^{AA'}\pi_{A'},\pi_{A'})}{\pi_{1'}^{|n|-2}} = \underbrace{\zeta^{|n|-2}\hat{\mu}(z,\zeta)+P_{|n|-2}(z,\zeta)}_{\text{power series } \sum_{k\leqslant|n|-2}\hat{\alpha}_k(z)\zeta^k} - \underbrace{(\mu(z,\zeta)+P_{|n|-2}(z,\zeta))}_{\text{power series } \sum_{k\geqslant 0}\alpha_k(z)\zeta^k}$$

where $P_{|n|-2}(z,\zeta)$ is an arbitrary polynomial of degree $|n|-2$ in ζ; thus

$$\pi_{1'}^{|n|-2}P_{|n|-2}(z,\zeta) = g^{C'\ldots D'}(z)\pi_{C'}\ldots\pi_{D'} = g \in \Gamma(M^+_{\mathbb{C}}\times P_1, \mathcal{O}(|n|-2)).$$

$g^{C'\dots D'}$ is a symmetric spinor field with $|n|-2$ indices that determines the gauge. $D_A P_{|n|-2}(z,\zeta)$ is a polynomial of degree $|n|-1$ and the map D becomes

$$Dg = \pi_{1'}^{|n|-1} D_A P_{|n|-2} = \pi^{B'} \nabla_{AB'} g^{C'\dots D'} \pi_{C'\dots D'} = \chi_A{}^{B'\dots D'} \pi_{B'\dots D'},$$

which yields a pure gauge potential χ.

The bijective map δ^{-1} now is

$$H^1(M_{\mathbb{C}}^{+}\times P_1, \alpha^* L^{n+2}) \xrightarrow{\delta^{-1}} \frac{\text{potentials } \chi_A{}^{B'\dots D'}}{\text{gauge } g^{C'\dots D'}} \longrightarrow \text{fields } \phi_{A\dots D}\,.$$

This map is performed as follows: for $z \in \omega$ we have

$$\mu(z,\zeta') = \frac{1}{2\pi i} \oint \frac{f(iz^{AA'}\pi_{A'},\pi_{A'})}{\pi_{1'}^{|n|-2}(\zeta-\zeta')} d\zeta\,, \quad \zeta' \text{ inside the circle}$$

and

$$\hat{\mu}(z,\zeta') = \frac{1}{2\pi i} \oint \frac{f(iz^{AA'}\pi_{A'},\pi_{A'})}{(\zeta'\pi_{1'})^{|n|-2}(\zeta-\zeta')} d\zeta, \quad \zeta' \text{ outside the circle,}$$

so that

$$D_A(\zeta')\mu(z,\zeta') = \frac{i}{2\pi i} \oint \frac{(\pi_{0'}-\zeta'\pi_{1'})\frac{\partial}{\partial\omega^A} f(iz^{AA'}\pi_{A'},\pi_{A'})}{(\pi_{0'}-\zeta'\pi_{1'})\pi_{1'}^{|n|-3}} d\zeta$$

which is independent of ζ' and hence equals $\chi_A{}^{1'\dots 1'}$. Apparently, the gauge is such that only the potentials $\chi_A{}^{1'\dots 1'}$ for $A = 0,1$ do not vanish. Then in ω we get

$$\phi_{A\dots D} = \frac{1}{|n|!} \sum_{\sigma} \nabla_{\sigma_1 1'} \cdots \nabla_{\sigma_{|n|-1} 1'} \frac{i}{2\pi i} \oint \frac{\frac{\partial}{\partial\omega^{\sigma_{|n|}}} f(iz^{AA'}\pi_{A'},\pi_{A'})}{\pi_{1'}^{|n|-1}} \pi_{1'}^2 d\zeta =$$

$$= \frac{i^{|n|}}{2\pi i} \oint \frac{\partial}{\partial\omega^A} \cdots \frac{\partial}{\partial\omega^D} f(iz^{AA'}\pi_{A'},\pi_{A'})(\pi_{1'} d\pi_{0'} - \pi_{0'} d\pi_{1'}),$$

which is just the Penrose transformation (4.3) of chapter "Linear Fields and Yang-Mills Theory".

For $n = -1$ there are no global sections in $\mathcal{O}(|n|-2)(M^+_{\mathbb{C}}\times P_1)$ except 0; for one would have $\mu/\pi_{1'} = \hat{\mu}/\pi_{0'}$, hence

$$\hat{\mu} = \sum_{k\leq 0} \hat{\alpha}_k \zeta^k = \zeta\mu = \sum_{k\geq 1} \alpha_k \zeta^k$$

which is impossible except for $\hat{\alpha}_k = 0 = \alpha_k$. Therefore, we immediately get a bijection between $H^1(M^+_{\mathbb{C}}\times P_1, \alpha^* L^1)$ and $\operatorname{Ker} D'(M^+_{\mathbb{C}}\times P_1) \cong \chi_A$. The functions χ_A are no potentials in this case, but the field $\phi_A = \chi_A$ itself, because $D'\chi = 0$ is the field equation $\nabla^{AA'}\chi_A = 0$, $A' = 0,1$.

REMARK. We have used that $H^1(M^+_{\mathbb{C}}\times P_1, \mathcal{O}(-n-2)) = 0$ for $n < 0$. For $n \geq 0$ this is no longer true and for right-handed fields we will not need this condition. If for left-handed fields we had taken another domain $V \subset M_{\mathbb{C}}$, then we should have required that $H^1(V\times P_1, \mathcal{O}(-n-2)) = 0$ for $n < 0$. Without giving the details we merely mention that this is always the case if $n = -1$ and for $n < -1$ this follows from the condition $H^1(V,\mathcal{O}) = 0$ on V. That $H^1(M^+_{\mathbb{C}}\times P_1, \mathcal{O}(m)) = 0$ for $m \geq -1$ can be seen as follows: divide $M^+_{\mathbb{C}}\times P_1$ in two pieces $M^+_{\mathbb{C}}\times U_A$, $A = 0,1$, with $U_A = \{[\pi] \mid \pi_{A'} \neq 0\}$. Both pieces are pseudoconvex in suitable charts (because $M^+_{\mathbb{C}}$ is convex). Hence a 1-cocycle $f \in \mathcal{O}(m)(M^+_{\mathbb{C}}\times U_0 \cap U_1)$ determines the cohomology group. We can write f in a Laurent series

$$f(z,\zeta) = \sum_k \alpha_k(z)\zeta^k, \quad \alpha_k(z) = \frac{1}{2\pi i} \oint \frac{f(z,\zeta)}{\zeta^{k+1}} d\zeta$$

and

$$f(z,\zeta) = \zeta^m \sum_{k\leq 0} \alpha_{k+m}(z)\zeta^k + \sum_{k\geq m+1} \alpha_k(z)\zeta^k, \quad \zeta = \pi_{0'}/\pi_{1'}.$$

Hence for $m \geq -1$ both terms determine sections in $M^+_{\mathbb{C}}\times U_0$ and in $M^+_{\mathbb{C}}\times U_1$, respectively, so that f is a 1-coboundary.

5. RIGHT-HANDED FIELDS

For right-handed fields $\psi_{A'\ldots D'}$ of helicity $|\frac{n}{2}|$ with $n \geqslant 0$ we proceed as follows.

From the resolution

$$0 \to \alpha^* L^{n+2} \to \mathcal{O}(-n-2) \xrightarrow{D} \mathcal{O}(-n-1)^2 \xrightarrow{D'} \mathcal{O}(-n) \to 0$$

we get the exact sequence of sections

$$0 \to 0 \to 0 \to 0 \to H^1(M^+_{\mathbb{C}} \times P_1, \alpha^* L^{n+2}) \to H^1(M^+_{\mathbb{C}} \times P_1, \mathcal{O}(-n-2)) \xrightarrow{\tilde{D}}$$

$$\xrightarrow{\tilde{D}} H^1(M^+_{\mathbb{C}} \times P_1, \operatorname{Ker} D')$$

because there are no global sections in $\Gamma(M^+_{\mathbb{C}} \times P_1, \mathcal{O}(m))$ except 0 for $m < 0$. Hence

$$H^1(M^+_{\mathbb{C}} \times P_1, \alpha^* L^{n+2}) \cong \operatorname{Ker}(H^1(M^+_{\mathbb{C}} \times P_1, \mathcal{O}(-n-2)) \xrightarrow{\tilde{D}} H^1(M^+_{\mathbb{C}} \times P_1, \operatorname{Ker} D')).$$

Cover $P_1(\mathbb{C})$ with the sets $U = \{\zeta \in \mathbb{C} \mid |\zeta| < r_2\}$ and $\hat{U} = \{\zeta \in \mathbb{C} \mid |\zeta| > r_1\} \cup \{\infty\}$, $r_1 < r_2$, then the covering $\{M^+_{\mathbb{C}} \times U, M^+_{\mathbb{C}} \times \hat{U}\}$ of $M^+_{\mathbb{C}} \times P_1(\mathbb{C})$ is acyclic for the sheafs $\mathcal{O}(m)$. Hence the elements of $H^1(M^+_{\mathbb{C}} \times P_1, \mathcal{O}(-n-2))$ are determined by holomorphic functions h of z and $\pi \in \mathbb{C}^2 \setminus \{(0,0)\}$ for $z \in M^+_{\mathbb{C}}$ and homogeneous of degree $-n-2$ in π for $r_1 < |\zeta| = |\pi_0, / \pi_1, | < r_2$. The condition $\tilde{D}h = 0$ means that $\pi_1, D_A h(z,\pi)$ is a 1-coboundary with coefficients in $\mathcal{O}(-n-1)$:

$$\pi_1, D_A h(z,\pi) = \frac{\hat{\mu}_A(z,\zeta)}{\pi_{0'}^{n+1}} - \frac{\mu_A(z,\zeta)}{\pi_{1'}^{n+1}}$$

and that moreover $D_1 \hat{\mu}_0 = D_0 \hat{\mu}_1$ and $D_1 \mu_0 = D_0 \mu_1$, for holomorphic functions $\hat{\mu}_A$ in $M^+_{\mathbb{C}} \times \hat{U}$ and μ_A in $M^+_{\mathbb{C}} \times U$. We can rewrite this by defining the function

$$\rho(z,\zeta) = \sum_{k=-\infty}^{\infty} \alpha_k(z) \zeta^k \triangleq \pi_1, \pi_{0'}^{n+1} h(z,\pi)$$

so that

$$D_A\rho(z,\zeta) = \hat{\mu}_A(z,\zeta) - \zeta^{n+1}\mu_A(z,\zeta), \quad z \in M_{\mathbb{C}}^+, \quad r_1 < |\zeta| < r_2.$$

Thus for $n > 0$ in the Laurent series of $D_A\rho$ the terms with $\zeta, \ldots, \zeta^n$ should vanish, and no further conditions are required. For $D_0D_1\rho = D_1D_0\rho$ implies $D_0(\zeta)\hat{\mu}_1 - \zeta^{n+1}D_0(\zeta)\mu_1 = D_1(\zeta)\hat{\mu}_0 - \zeta^{n+1}D_1(\zeta)\mu_0$ so that $D_A\hat{\mu}^A = 0 = D_A\mu^A$ automatically holds for $n > 0$. For $n = 0$ this yields the condition $\Box\alpha_0 = 0$, i.e. α_0 should satisfy the wave equation.

If h is itself a 1-coboundary with coefficients in $\mathcal{O}(-n-2)$, then

$$h(z,\pi) = \frac{\hat{v}(z,\zeta)}{\pi_{0'}^{n+2}} - \frac{v(z,\zeta)}{\pi_{1'}^{n+2}}$$

hence

$$\rho(z,\zeta) = \frac{\hat{v}(z,\zeta)}{\zeta} - \zeta^{n+1}v(z,\zeta),$$

which are all the Laurent series without the terms with $\zeta^0, \zeta^1, \ldots, \zeta^n$.

So we have constructed a bijective map:

$$H^1(M_{\mathbb{C}}^+ \times P_1, \alpha^*\mathcal{L}^{n+2}) \cong \operatorname{Ker}\tilde{D} \longrightarrow \{[\rho] = \sum_{k=0}^{n} \alpha_k(z)\zeta^k \mid$$

$$\alpha_k \text{ holomorphic in } M_{\mathbb{C}}^+;\ \Box\alpha_0 = 0 \text{ if } n = 0, \text{ or in}$$

$$D_A[\rho] \text{ the terms with } \zeta, \ldots, \zeta^n \text{ vanish if } n > 0\}.$$

For $n > 0$ the condition on the cohomology class $[\rho]$ is equivalent to $\nabla_{A0'}\alpha_k - \nabla_{A1'}\alpha_{k-1} = 0$ for $k = 1,\ldots,n$. These are just the field equations for the right-handed field, $\nabla^{AA'}\psi_{A'\ldots D'} = 0$, if we put $\psi_{A'\ldots D'} = \alpha_k$ when k of the indices $A'\ldots D'$ have the value $1'$ while the remaining are $0'$.

REMARK. If instead of $M_{\mathbb{C}}^+$ one is interested in fields defined in another domain $V \subset M_{\mathbb{C}}$ which is not pseudoconvex, an element of $H^1(V\times P_1, \mathcal{O}(-n-2))$

is determined by a 1-cocycle with respect to a larger covering (for example $\mathcal{U} = \bigcup_{\tilde{z}} \{\Omega(\tilde{z}), \hat{\Omega}(\tilde{z})\}$ with $\omega(\tilde{z}) \subset V$). Without giving the details we merely mention that also in this case the above construction yields globally defined polynomials $\sum_{k=0}^{n} \alpha_k(z)\zeta^k$ with $z \in V$ and actually $H^1(V\times P_1, \mathcal{O}(m)) \cong H^1(\{V\times U, V\times\hat{U}\}, V\times P_1, \mathcal{O}(m))$ for $m \leqslant -1$.

For a holomorphic function f in $W(\tilde{z}) \cap \hat{W}(\tilde{z}) \subset P_3$, homogeneous of degree $-n-2$, the field ψ for $z \in \omega(\tilde{z})$ is given by

$$\psi_{A'\ldots D'}(z) = \text{term with } \zeta^k \text{ in the Laurent series of}$$

$$\pi_{1'}\pi_{0'}^{n+1} f(iz^{AA'}\pi_{A'}\pi_{A'}) \quad \text{if } k \text{ of the indices } A'\ldots D' \text{ have}$$

$$\text{the value } 1' = \frac{1}{2\pi i}\oint \frac{\pi_{1'}\pi_{0'}^{n+1} f(iz^{AA'}\pi_{A'}\pi_{A'})}{\zeta^{k+1}} d\zeta =$$

$$= \frac{1}{2\pi i}\oint \pi_{1'}^{k}\pi_{0'}^{n-k} f(iz^{AA'}\pi_{A'}\pi_{A'})\pi_{1'}^{2} d\zeta =$$

$$= \frac{1}{2\pi i}\oint \pi_{A'}\ldots\pi_{D'} f(iz^{AA'}\pi_{A'}\pi_{A'})(\pi_{1'}d\pi_{0'} - \pi_{0'}d\pi_{1'})$$

and this is just the Penrose transformation given by formula (4.2), chapter "Linear Fields and Yang-Mills Theory".

6. EXAMPLE

With the aid of the Penrose transformation we construct a holomorphic solution of the wave equation (thus $n = 0$) starting from the homogeneous function of degree -2

$$f(Z) = \frac{1}{Z^0 Z^1}$$

in $U_0 = \{[Z] \mid Z^0 \neq 0\} \cap U_1 = \{[Z] \mid Z^1 \neq 0\}$. The Penrose transformation is

$$\psi(z^{AA'}) = \frac{1}{2\pi i} \oint \frac{\pi_{1'} d\pi_{0'} - \pi_{0'} d\pi_{1'}}{(iz^{00'}\pi_{0'} + iz^{01'}\pi_{1'})(iz^{10'}\pi_{0'} + iz^{11'}\pi_{1'})}.$$

This formula has been expressed in homogeneous coordinates. We choose a coordinate system by setting $\pi_{0'} = \zeta$ and $\pi_{1'} = 1$ (or as usual $\zeta = \pi_{0'}/\pi_{1'}$). Then we get

$$\psi(z^{AA'}) = \frac{1}{2\pi i} \oint \frac{d\zeta}{(iz^{00'}\zeta + iz^{01'})(iz^{10'}\zeta + iz^{11'})},$$

where the integration is performed around one of the singularities $-z^{01'}/z^{00'}$ or $-z^{11'}/z^{10'}$. We choose the contour around $-z^{11'}/z^{10'}$ and we find

$$\psi(z^{AA'}) = \frac{1}{iz^{10'}\left(\frac{iz^{00'}(-z^{11'})}{z^{10'}} + iz^{01'}\right)} = \frac{1}{z^{00'}z^{11'} - z^{01'}z^{10'}} =$$

$$= \frac{1}{(z^0)^2 - (z^1)^2 - (z^2)^2 - (z^3)^2}.$$

The contour around the other singularity would have given this solution with a minus sign.

A solution for real x^μ can be obtained as a distribution by taking the boundary value

$$\psi(x^{\mu}) = \lim_{y^0 \downarrow 0} \frac{1}{(x^0+iy^0)^2-(x^1)^2-(x^2)^2-(x^3)^2}.$$

In the Euclidean metric $(z^0=x^0,\ z^j=ix^j)$ we get an analytic solution of the Laplace equation in $\mathbb{R}^4 \setminus \{(0,0,0,0)\}$:

$$\frac{1}{(x^0)^2+(x^1)^2+(x^2)^2+(x^3)^2}.$$

7. REMARKS

A. Let $W = W_0 \cup W_1$, then the first Čech-cohomology group with coefficients in a sheaf F is mapped injectively into $H^1(W,F)$

$$H^1(\{W_0, W_1\}, W, F) \hookrightarrow H^1(W,F)$$

(theorem 9 of the preceding chapter). This map is not surjective if the covering is not acyclic for F, i.e. if $H^1(W_0,F)$ or $H^1(W_1,F)$ does not vanish. This indeed happens if

$$W = P_3^+, \quad W_0 = P_3^+ \cap \{Z^3 \neq 0\}, \quad W_1 = P_3^+ \cap \{Z^4 \neq 0\}, \quad F = \mathcal{L}^m.$$

Hence a holomorphic function f in $W_0 \cap W_1$, homogeneous of degree $-n-2$ in homogeneous coordinates, determines a solution of the field equations by means of the Penrose transformation, but to get *all* solutions we have to work with an infinitely large acyclic covering. The global formula of the Penrose transformation becomes complicated then. Another possibility is to choose for every solution an appropriate covering of two sets. The advantage is that in one formula (namely (4.2) or (4.3), chapter "Linear Fields and Yang-Mills Theory") the global fields in $M_{\mathbb{C}}^+$ are obtained, but it is unclear how to choose the covering.

B. The method we have performed here for the left-handed field has been inspired by the construction of Ward for the Yang-Mills field. It can more or less be found in [Wells] and [Eastwood, Penrose, Wells]. However our method for the right-handed field is quite different from the literature (where Serre duality has been used by passing to Dolbeault cohomology groups).

For the meaning of $H^1(P_3^+,\mathcal{L}^m)$ as classification of vector bundles of rank 2 over P_3^+ with certain geometric properties, cf. [Atiyah, chapter VI.1].

C. The method with (convergent) power series uses the covering $\mathcal{U}$ of $M_{\mathbb{C}}^{+} \times P_1(\mathbb{C})$ with the sets $\Omega(\tilde{z})$ and $\hat{\Omega}(\tilde{z})$. Since we want to have functions of z for z in a fixed set $\omega(\tilde{z})$, this covering is more usefull than the covering with the sets $\alpha^{-1}(W(\tilde{z}))$ and $\alpha^{-1}(\hat{W}(\tilde{z}))$. Moreover, the method also suggests a method with formal power series:

Let $\rho(z,\zeta) = \sum_k \alpha_k(z)\zeta^k$ be a formal Laurent series (not necessarily convergent for some ζ) with coefficients α_k which are differentiable with respect to z in some set ω that now may be $M_{\mathbb{C}}^{+}$. Let

$$D_A\rho(z,\zeta) = \sum_k \beta_k^A(z)\zeta^k, \qquad D_A = \nabla_{A0'} - \zeta\nabla_{A1'},$$

be the associated formal Laurent series. Then we have

1) *for right-handed fields* ($n \geq 0$): if $\beta_0^A = \ldots = \beta_n^A = 0$, the field is determined by $(\alpha_0, \ldots \alpha_n)$.
2) *for left-handed fields* ($n<0$): if $\beta_n^A = 0$, $\alpha_{n+1} = \ldots = \alpha_0 = 0$, then $(\beta_{n+1}^A, \ldots, \beta_0^A)$ determine potentials in a certain gauge (which can be fixed by requiring $\alpha_{n+1} = \ldots = \alpha_{-1} = 0$ (if $n \leq -2$) so that in this gauge $\beta_{n+2}^A = \ldots = \beta_0^A = 0$).

In fact we have shown that all holomorphic fields in ω can be obtained in this manner when the coefficients α_k are holomorphic functions for $k = 0, \ldots, n$ or $n, \ldots, 0$. Non-holomorphic fields can be found by allowing non-holomorphic coefficients α_k.

8. RELATION BETWEEN THE CONSTRUCTIONS OF ATIYAH, WARD AND PENROSE

On the one hand the construction of Ward (cf. the chapters "Self-dual Yang-Mills Equations" and "Computation of the Yang-Mills Potentials") produced local Yang-Mills potentials and on the other hand the theorem of Atiyah-Ward (cf. [p.210, proceedings of this seminar 1981/82]) related anti-self-dual Yang-Mills fields with certain vector bundles over $P_3(\mathbb{C})$. Looking for $S\ell(2,\mathbb{C})$-Yang-Mills fields rather than SU(2)-Yang-Mills fields the construction of Ward deals with a holomorphic vector bundle over $P_3(\mathbb{C})$ whose transition functions g satisfy condition (2) of chapter "Computation of the Yang-Mills Potentials". The theorem of Atiyah-Ward deals with a holomorphic vector bundle over $P_3(\mathbb{C})$ which is trivial on the fibres of $P_3(\mathbb{C}) \xrightarrow{\pi} S^4$ (cf. [section 4 of chapter "Twistor theories and Yang-Mills Fields", proceedings of this seminar 1981/82]). Here we will first point out the relation between these two properties of a vector bundle over $P_3(\mathbb{C})$ and secondly, we will compare the construction of Ward with the Penrose transform for the left-handed (anti-self-dual) Maxwell field, $n = -2$ (cf. sections 2 and 3 of this chapter).

A. In both points of view the following flagmanifold F plays a central role:

F is the space of pairs of complex 2-dimensional linear subspaces L_2 of $\mathbb{C}^4$ and the 1-dimensional subspaces L_1 they contain, thus

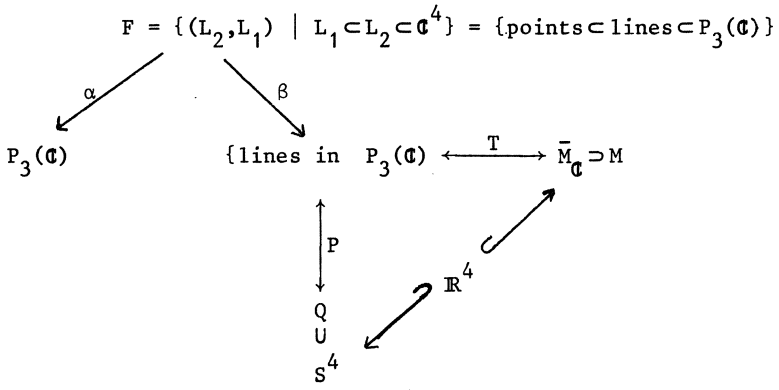
$F = \{(L_2,L_1) \mid L_1 \subset L_2 \subset \mathbb{C}^4\}$.

We also consider the projections $\alpha\colon F \to P_3(\mathbb{C})$ with $\alpha(L_2,L_1) = L_1$ and $\beta\colon F \to \{\text{space of complex lines in } P_3(\mathbb{C})\}$ with $\beta(L_2,L_1) = L_2$.

Furthermore, we remind the twistor correspondence T between the complex compactified Minkowski space $\bar{M}_{\mathbb{C}}$ and the space of complex lines in $P_3(\mathbb{C})$ and also the correspondence P by means of Plücker coordinates between this last space and a certain complex quadric $Q \subset P_5(\mathbb{C})$ containing the real

quadric S^4, cf. [chapter "Twistor theories and Yang-Mills Fields", proceedings of this seminar 1981/82], and also section 2, chapter "Self-dual Yang-Mills Equations".

We have the following diagram



Since we want to describe the fibration $P_3(\mathbb{C}) \to S^4$ we rather use the matrix $z^{PQ} = \sqrt{2}\, iA_{(iz^0,z^1,z^2,z^3)}$ cf. [formula (1.12), p.186 proceedings of this seminar 1981/82]) in the twistor correspondence T instead of the matrix $iz^{AA'} = iA_z$ of the original twistor transformation.
A point $(x^0,x^1,x^2,x^3) \in \mathbb{R}^4$ is mapped onto the point $(ix^0,x^1,x^2,x^3) \in M_{\mathbb{C}}$ and the triangular diagram is commutative if $\mathbb{R}^4$ is mapped by the inverse of stereographic projection into S^4 (from the pole $Y_2 = 1$ in the notation of [p.200, proceedings of this seminar 1981/82]).

We remove the point at infinity from $\bar{M}_{\mathbb{C}}$ and then we have in local coordinates:

$$(z,[\pi]) \in F$$

$$\alpha \swarrow \qquad \searrow \beta$$

$$[z^{1P}\pi_P, z^{2P}\pi_P, \pi_1, \pi_2] \in P_3(\mathbb{C}) \setminus [Z^1,Z^2,0,0] \qquad \{[Z] \mid Z^P = z^{P1}Z^3 + z^{P2}Z^4,\ P = 1,2\}$$

$$\xleftrightarrow[z^{PQ}]{} (iz^0,z^1,z^2,z^3) \in M_{\mathbb{C}}.$$

Thus $z \in \mathbb{C}^4$ determines the plane in $\mathbb{C}^4$ and $[\pi] \in P_1(\mathbb{C})$ the lines in this plane. The map $\alpha : \mathbb{C}^4 \times P_1(\mathbb{C}) \longrightarrow P_3(\mathbb{C}) \setminus [Z^1, Z^2, 0, 0]$ is holomorphic.

We consider real $z = x \in \mathbb{R}^4$, so that $z^{PQ} = \tilde{x}$ is given by formula (1), chapter "Computation of the Yang-Mills Potentials". Then α is analytic in x and holomorphic in $[\pi]$. Moreover, in that case α is injective because $\alpha(x,[\pi]) = \alpha(y,[\pi])$ implies that x and y are null-separated and in the Euclidean metric this means that $x = y$. So the inverse $\alpha^{-1} : P_3(\mathbb{C}) \setminus [Z^1, Z^2, 0, 0] \longrightarrow \mathbb{R}^4 \times P_1(\mathbb{C})$ is well-defined. However, it is not an analytic map (if $\alpha(x,[\pi]) = [\omega^1, \omega^2, \pi_1, \pi_2]$ we have

$$x^{12} = -\overline{x^{21}} = \frac{\bar{\pi}_2\omega^1 - \pi_1\overline{\omega^2}}{|\pi_1|^2+|\pi_2|^2} \quad \text{and} \quad x^{11} = \overline{x^{22}} = \frac{\bar{\pi}_1\omega^1 + \pi_2\overline{\omega^2}}{|\pi_1|^2+|\pi_2|^2}\Big).$$

Now the fibration $P_3(\mathbb{C}) \setminus [Z^1, Z^2, 0, 0] \xrightarrow{\pi} \mathbb{R}^4 \hookrightarrow S^4$ is described in this picture by $\pi = \beta\alpha^{-1}$. The fibre of π above a point x is $\alpha\beta^{-1}\{x\}$. Thus the map β describes the trivial bundle over $\mathbb{R}^4$ with a fibre $P_1(\mathbb{C})$ of π as typical fibre. Hence a holomorphic bundle over $P_3(\mathbb{C}) \setminus [Z^1, Z^2, 0, 0]$ which is trivial on the fibres of π is trivial when lifted by α, i.e. it is trivial when (part of) the holomorphic structure is forgotten. In coordinates this condition yields for a holomorphic transition function g of $[Z]$ the property

$$g([\tilde{x}\pi, \pi]) = \hat{h}(x,[\pi])\, h(x,[\pi])^{-1},$$

where $\hat{h}$ is holomorphic for $[\pi]$ near $[1, 0]$ and h for $[\pi]$ near $[0, 1]$ and where both are analytic in x. This is exactly condition (2) of chapter "Computation of the Yang-Mills Potentials" of the construction of Ward.

B. In Wards construction of local Yang-Mills potentials the choice of the metric, Euclidean or Minkowski, is not important. So in order to compare the construction of Ward to the Penrose transformation we will from now on

work with the Minkowski metric. For that purpose instead of $\tilde{z} = z^{PQ} = \sqrt{2}\,iA_{(iz^0,z^1,z^2,z^3)}$ we rather use the matrix $iz^{AA'} = iA_z$ in the definition of the projections α and β. The above mentioned condition for a holomorphic function g which is homogeneous of degree zero (serving as a transition function in a holomorphic vector bundle over $P_3(\mathbb{C}) \setminus [z^1,z^2,0,0]$) now becomes

$$(*) \qquad g(iA_z\pi,\pi) = \hat{h}(z,[\pi])\,h(z,[\pi])^{-1}\,.$$

It has been pointed out in the chapter "Computation of the Yang-Mills Potentials" that holomorphic Yang-Mills potentials can be constructed from such a function g provided that holomorphic solutions of the occurring differential equations (32) and (35), chapter "Computation of the Yang-Mills Potentials", are taken.

In both, Wards construction and the Penrose transform, we use the holomorphic map α with α-planes as fibres. Both fit in the following general scheme:

Let G be a Lie group with associated Lie algebra g and let ρ be a representation. Furthermore, let F be a sheaf of rings over X, then we denote by $F_{\rho(G)}$ the sheaf of (non-Abelian, multiplicative groups of) matrices with F-valued entries of the representation ρ of G and by $F_{\rho(g)}$ the sheaf of (Abelian, additive groups of) matrices with F-valued entries of the associated representation of g. In particular, denote by L^0 the sheaf of holomorphic functions on $P_3(\mathbb{C}) \setminus [z^1,z^2,0,0]$; this has also been denoted by $\mathcal{O}$, but here the notation $\mathcal{O}$ will be reserved for the sheaf of holomorphic functions on $\mathbb{C}^4 \times P_1(\mathbb{C})$. We consider the following complex of maps between sheafs on $\mathbb{C}^4 \times P_1(\mathbb{C})$

$$\mathbb{1} \to \alpha^* L^0_{\rho(G)} \xrightarrow{i} \mathcal{O}_{\rho(G)} \xrightarrow{D} \mathcal{O}(1)^1_{\rho(g)} \xrightarrow{D'} \mathcal{O}(2)^2_{\rho(g)},$$

where $\mathbb{1}$ is the identity in the representation ρ of G, i is the in-

jection map, F^j is the sheaf of j-forms above the α-planes with coefficients in F and where the map D is defined by

$$D(h(z,[\pi]) = h(z,[\pi])^{-1} D_A h(z,[\pi]) dw^A$$

with $D_A = \pi^{A'} \nabla_{AA'}$ as independent differentiations in the α-planes for $A = 0,1$, and D' by

$$D'(h_A(z,[\pi])dw^A) = (D_1 h_0(z,[\pi]) - D_0 h_1(z,[\pi]) + [h_0, h_1])dw^0 \cap dw^1 .$$

(Here w^0 and w^1 may be considered as coordinates in an α-plane). Unless G is Abelian, the maps D and D' are no group homomorphisms. The above given complex is an exact sequence (which can even be terminated as an exact sequence by $\mathcal{O}(2)^2_{\rho(g)} \to 0$, but we will not need this). The exactness at the place $\xrightarrow{D} \mathcal{O}(1)^1_{\rho(g)} \xrightarrow{D'}$ expresses the fact that if the curvature vanishes the connection is completely integrable (or in more physical terms that a vanishing field comes from a potential which is pure gauge). The important thing is that in the above sequence the equation

$$D'(h_A\, dw^A) \equiv (-\pi^{A'} \nabla_{AA'} h^A + h_A h^A) dw^0 \cap dw^1 = 0$$

for sections over $\omega \times P_1(\mathbb{C})$ is the self-duality equation for the potentials in $\omega \subset \mathbb{C}^4$ we want to solve (cf. Liouville's theorem and proof (ii) of Wards theorem, section 3.2 of chapter "Self-dual Yang-Mills Equations").

As in the Abelian case we want to have the associated exact sequence of sections. For that purpose we need non-commutative cohomology. One can define this by means of Čech cohomology: a 1-cochain $\{g_{ij}\}$ of non-Abelian groups is a 1-cocycle if $g_{ij} g_{jk} = g_{ik}$ and two 1-cocycles $\{g_{ij}\}$ and $\{\tilde{g}_{ij}\}$ belong to the same cohomology class if $\tilde{g}_{ij} = h_i g_{ij} h_j^{-1}$ for a 0-cochain $\{h_i\}$. As in theorem 8 of the preceding chapter the direct limit defines the first cohomology set $H^1(X,F)$, which is not a group unless F is a sheaf of Abelian groups. The set $H^1(X,F_{\rho(G)})$ characterizes the

(classes of isomorphic) F-vector bundles associated to principle G-bundles over X.

If in theorem 5 of the preceding chapter F is a sheaf of subgroups of the sheaf G of non-Abelian groups and $H = G/F$ is the associated sheaf of homogeneous spaces then this theorem can be generalized only for the first five terms in the exact sequence of maps between sections, cf. [Frenkel, théorème I.1]. However, this is just far enough for us here. So from the above written exact sequence we derive the exact sequence of maps between sections over $\omega \times P_1(\mathbb{C})$ with $\omega \subset \mathbb{C}^4$:

$$1 \to \Gamma(\omega\times P_1, \alpha^* L^0_{\rho(G)}) \xrightarrow{i} \Gamma(\omega\times P_1, \mathcal{O}_{\rho(G)}) \xrightarrow{D} \Gamma(\omega\times P_1, \operatorname{Ker} D') \xrightarrow{\delta}$$

$$\xrightarrow{\delta} H^1(\omega\times P_1, \alpha^* L^0_{\rho(G)}) \xrightarrow{i^*} H^1(\omega\times P_1, \mathcal{O}_{\rho(G)}) .$$

Furthermore, two sections $h_A\, dw^A$ and $\tilde{h}_A\, dw^A$ in $\Gamma(\omega\times P_1, \operatorname{Ker} D')$ have the same image under the map δ if and only if they are gauge equivalent:

$$\tilde{h}_A = \Lambda^{-1} h_A \Lambda + \Lambda^{-1} D_A \Lambda, \qquad A = 0,1$$

for a section $\Lambda \in \Gamma(\omega\times P_1, \mathcal{O}_{\rho(G)})$, cf. [Frenkel, complément du théorème I.1].

The image of the map δ is the set of classes of those sections g in $(\alpha^* L^0_{\rho(G)})(\omega \times U\cap\hat{U})$ which are mapped by i^* onto the trivial class in $H^1(\omega\times P_1, \mathcal{O}_{\rho(G)})$, i.e. such a class is determined by a holomorphic function g homogeneous of degree zero in $W = \alpha(\omega \times U\cap\hat{U}) \subset P_3(\mathbb{C}) \setminus [Z^1, Z^2, 0, 0]$ such that

$$g(iA_z\pi, \pi) = \hat{h}(z, [\pi])\, h(z, [\pi])^{-1}$$

with $\hat{h} \in \Gamma(\omega\times\hat{U}, \mathcal{O}_{\rho(G)})$ and $h \in \Gamma(\omega\times U, \mathcal{O}_{\rho(G)})$. This is exactly condition (*).

Just as in the Penrose transformation, modulo gauge Yang-Mills potentials $\chi_{AA'} \in \mathcal{O}_{\rho(G)}(\omega)$ satisfying the self-duality equations

$D'(\pi_{A'}\chi_A^{A'}dw^A) = 0$ are obtained from such a g by δ^{-1} if we take $G = SL(2,\mathbb{C})$ represented by itself. This is the construction of Ward.

The left-handed Maxwell field, $n = -2$, has the (additive) Abelian gauge group $G = \mathbb{C}$. By exponentiation we get the 1-dimensional representation $S\ell(1,\mathbb{C}) \cong \mathbb{C}^*$ and so the above description yields the Penrose transformation. In this form all holomorphic solutions correspond to those holomorphic line bundles over $W \subset P_3(\mathbb{C}) \setminus [Z^1,Z^2,0,0]$ which are trivial when lifted by α, i.e. whose transition functions g satisfy

$$\frac{1}{2\pi i} \oint_{|\zeta|=1} \left(\frac{d}{d\zeta} g(iA_z\pi,\pi)\right)/g(iA_z\pi,\pi)d\zeta = 0,$$

provided that $H^1(\omega\times P_1, \mathcal{O}) = 0$ (just as in the remark in section 4). For, in that case, one can take logarithms and so one passes from the multiplicative to the additive description and $H^1(\omega\times P_1, \mathcal{O}) = 0$ means that the additive analogue of (*) is satisfied, cf. [p.167 in the proceedings of this seminar, 1981/82].

Another form of Penrose's method is obtained if we take the 2-dimensional (multiplicative) representation $\left\{\begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \mid u \in \mathbb{C}\right\}$ of the additive gauge group $\mathbb{C}$ with the associated Lie algebra represented by the additive group of the matrices $\left\{\begin{pmatrix} 0 & u \\ 0 & 0 \end{pmatrix} \mid u \in \mathbb{C}\right\}$. The condition $H^1(\omega\times P_1, \mathcal{O}) = 0$ is equivalent to $H^1(\omega\times P_1, \mathcal{O}_{\rho(G)}) = 0$ in this representation, so that no condition (*) is needed (it is automatically satisfied). Now it is clear that Ward's Ansatz A_n for $n = 0$ and $f \equiv 0$ (in formula (5), chapter "Computation of the Yang-Mills Potentials") yields the Penrose transform for the left-handed Maxwell field!

APPENDIX

THEOREM. *Then* $W \subset P_3(\mathbb{C}) \setminus \{[\omega^A, 0]\}$ *and* $\Omega \subset M_{\mathbb{C}} \times P_1(\mathbb{C})$ *be domains with* $\alpha\Omega = W$ *such that for every* $[Z] \in W$ *the set* $\alpha^{-1}[Z] \cap \Omega$ *is connected and simply connected. Then*

$$H^1(W, L^{-m}) \cong H^1(\Omega, \alpha^* L^{-m}).$$

REMARK. We have used this theorem with $W \subset P_3^+$ and $\Omega = \alpha^{-1}W \cap M_{\mathbb{C}}^+ \times P_1(\mathbb{C})$ so that every $\alpha^{-1}[Z] \cap \Omega$ is convex.

PROOF.

Denote by E^{-m} the sheaf of C^∞-functions on $P_3(\mathbb{C})$ with values in the line bundle determined by homogeneous functions of degree m (similar to L^{-m} on $P_3(\mathbb{C})$). For convenience we denote the similar sheaf on $M_{\mathbb{C}} \times P_1(\mathbb{C})$ by $E(m)$ (just as $\mathcal{O}(m)$ on $M_{\mathbb{C}} \times P_1(\mathbb{C})$). On W we consider the Dolbeault resolution of L^{-m} and we obtain the derived sequence of sections

$$0 \to \Gamma(W, L^{-m}) \to \Gamma(W, E^{-m}) \xrightarrow{\bar\partial} \Gamma(W, E_{0,1}^{-m}) \xrightarrow{\bar\partial} \Gamma(W, E_{0,2}^{-m}) \to \dots$$

where $E_{0,p}^{-m}$ denotes the sheaf of p-forms of type $(0,p)$ with coefficients in E^{-m}. Since the fibres of $\alpha\colon \Omega \to W$ are connected, this complex is isomorphic to

$$0 \to \Gamma(\Omega, \alpha^* L^{-m}) \to \Gamma(\Omega, \alpha^* E^{-m}) \to \Gamma(\Omega, \alpha^* E_{0,1}^{-m}) \to \dots .$$

Hence $H^1(W, L^{-m})$ is isomorphic to the first cohomology group of the last complex.

However, the sheafs $\alpha^* E^{-m}$, $\alpha^* E_{0,1}^{-m}$ are not soft on Ω, so we cannot apply definition 5 (preceding chapter) to decide that this cohomology group is also isomorphic to $H^1(\Omega, \alpha^* L^{-m})$. We have to show that $H^1(\Omega, \alpha^* E^{-m}) = 0$ in order to draw this conclusion (cf. the discussion following theorem 5, preceding chapter):

$$H^1(W, L^{-m}) \cong \frac{\text{Ker} : E^{-m}_{0,1}(W) \to E^{-m}_{0,2}(W)}{\text{Im} : E^{-m}(W) \to E^{-m}_{0,1}(W)} \cong$$

$$\cong \frac{\text{Ker} : \Gamma(\Omega, \alpha^* E^{-m}_{0,1}) \to \Gamma(\Omega, \alpha^* E^{-m}_{0,2})}{\text{Im} : \Gamma(\Omega, \alpha^* E^{-m}) \to \Gamma(\Omega, \alpha^* E^{-m}_{0,1})}.$$

We consider the soft resolution of $\alpha^* E^{-m}$ on $M_{\mathbb{C}} \times P_1(\mathbb{C})$ similar to the resolution of $\alpha^* L^{-m}$ given in sections 3, 4 and 5. Then if the sequence of sections

$$0 \to \Gamma(\Omega, \alpha^* E^{-m}) \to \Gamma(\Omega, E(m)) \xrightarrow{D} \Gamma(\Omega, E(m+1))^2 \xrightarrow{D'} \Gamma(\Omega, E(m+2)) \to 0$$

is exact at the third place, we have $H^1(\Omega, \alpha^* E^{-m}) = 0$, cf. definition 5, preceding chapter.

Let $\{W_j\}$ be a locally finite covering of W, let $\{\chi_j\}$ a partition of unity with respect to this covering and let $\psi_A = (\psi_0, \psi_1) \in \text{Ker}\, D' \subset \subset \Gamma(\Omega, E(m+1))^2$. We will show that there are $\phi_j \in \Gamma(\alpha^{-1} W_j \cap \Omega, E(m))$ with $\tilde{D}_A \phi_j = \psi_A$ in $\alpha^{-1} W_j \cap \Omega$. Then we define $\phi = \sum_j (\alpha^* \chi_j) \phi_j \in \Gamma(\Omega, E(m))$, so that

$$\tilde{D}_A \phi = \sum_j \tilde{D}_A (\alpha^* \chi_j) \phi_j = \sum_j (\alpha^* \chi_j) \tilde{D}_A \phi_j = \sum_j (\alpha^* \chi_j) \psi_A = \psi_A.$$

Hence the sequence is exact at the right place to conclude that $H^1(\Omega, \alpha^* E^{-m}) = 0$

For every $[Z] \in W$ the set $\alpha^{-1}[Z] \cap \Omega$ is a simply connected domain of an α-plane. For $Z^4 \neq 0$ we can parametrize $\alpha^{-1}[Z]$ by

$$(w^A, -w^A \zeta + \omega^A; \zeta) = \left(\begin{pmatrix} w^0 & -w^0 \zeta + \omega^0 \\ w^1 & -w^1 \zeta + \omega^1 \end{pmatrix}; \zeta \right), w^A = (w^0, w^1) \in \mathbb{C}^2$$

if $[Z] = [\omega^0, \omega^1, \zeta, 1]$. Then the operators D_A are the exterior differential operators with respect to w^0 and w^1 and the complex can be consid-

ered as a de Rham-complex (actually, it is called a relative de Rham-complex).

First we choose W_j such that the bundle over $\alpha^{-1}W_j \cap \Omega$ is trivial, thus $W_j \subset P_3(\mathbb{C}) \setminus \{[\omega^A,\pi_{0'},0]\}$ or else $W_j \subset P_3(\mathbb{C}) \setminus \{[\omega^A,0,\pi_{1'},]\}$. We can find ϕ_j by integrating $\psi_A dw^A$ from a fixed point w_j^A in an α-plane to an arbitrary point in the same α-plane. This is independent of the path of integration within Ω because the domain cut by Ω in the α-plane is simply connected (and because $D_A\psi^A = 0$), and every point is reached because the domain is connected. Therefore, secondly we choose W_j so small that for fixed $w_j^A = (w_j^0,w_j^1)$, such that for a fixed point $[Z_j] = [\omega_j^A,\zeta_j,1] \in W_j$ the point $(w_j^A, -w_j^A\zeta_j + \omega_j^A; \zeta_j)$ belongs to $\alpha^{-1}[Z_j] \cap \Omega$, also

$$(w_j^A, -w_j^A\zeta + \omega^A; \zeta) \in \alpha^{-1}W_j \cap \Omega$$

if $[\omega^A, \zeta, 1]$ varies in W_j. The function ϕ_j can be found then by integration of $\psi_A dw^A$ from the point $(w_j^A, -w_j^A\zeta + \omega^A; \zeta)$ to an arbitrary point $(z,\zeta) \in \alpha^{-1}W_j \cap \Omega$ with $\alpha(z,\zeta) = [\omega^A,\zeta,1]$.

For example, if $\psi_A \Big|_{\alpha^{-1}W_j \cap \Omega}$ is represented by the function $\psi_A(z;\zeta)$, then the function ϕ_j is determined by

$$\phi_j(z;\zeta) = \int_{w_j^1}^{z^{10'}} \psi_1(w_j^0,w^1, (z^{00'}-w_j^0)\zeta + z^{01'}, (z^{10'}-w^1)\zeta + z^{11'}; \zeta)dw^1 +$$

$$+ \int_{w_j^0}^{z^{00'}} \psi_0(w^0,z^{10'}, (z^{00'}-w^0)\zeta + z^{01'}, z^{11'}; \zeta)dw^0,$$

$$(z,\zeta) \in \alpha^{-1}W_j \cap \Omega .$$

provided that the paths of integration are contained in Ω.

A similar formula is valid in the chard $Z^3 \neq 0$. ∎

REMARK. Instead of using power series this last formula can also be used to show that the resolution of $\alpha^* L^{-m}$ on $M_{\mathbb{C}} \times P_1(\mathbb{C})$, given in sections 3, 4 and 5, is exact.

REFERENCES

ATIYAH, M.F., *Geometry of Yang-Mills fields*, Acad.Naz.Lincei.Sci.Norm. Sup., Pisa (1979).

EASTWOOD, M.G. - PENROSE, R. - WELLS, R.O., *Cohomology and massless fields*, Comm.Math.Phys. 78(1981), p.305-351.

FRENKEL, J., *Cohomologie non Abélienne et espaces fibrés*, Bull.Soc.math. France, 85(1957), p.135-220.

JAGER, E.M. de - PIJLS, H.G.J. (eds), Proceedings seminar 1981-1982, *Mathematical structures in field theories*, CWI Syllabus 2, Mathematisch Centrum, Amsterdam (1984).

WELLS, R.O., *Complex geometry in mathematical physics*, Université de Montréal (1982).

MC SYLLABI

1.1 F. Göbel, J. van de Lune. *Leergang besliskunde, deel 1: wiskundige basiskennis.* 1965.

1.2 J. Hemelrijk, J. Kriens. *Leergang besliskunde, deel 2: kansberekening.* 1965.

1.3 J. Hemelrijk, J. Kriens. *Leergang besliskunde, deel 3: statistiek.* 1966.

1.4 G. de Leve, W. Molenaar. *Leergang besliskunde, deel 4: Markovketens en wachttijden.* 1966.

1.5 J. Kriens, G. de Leve. *Leergang besliskunde, deel 5: inleiding tot de mathematische besliskunde.* 1966.

1.6a B. Dorhout, J. Kriens. *Leergang besliskunde, deel 6a: wiskundige programmering 1.* 1968.

1.6b B. Dorhout, J. Kriens, J.Th. van Lieshout. *Leergang besliskunde, deel 6b: wiskundige programmering 2.* 1977.

1.7a G. de Leve. *Leergang besliskunde, deel 7a: dynamische programmering 1.* 1968.

1.7b G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7b: dynamische programmering 2.* 1970.

1.7c G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7c: dynamische programmering 3.* 1971.

1.8 J. Kriens, F. Göbel, W. Molenaar. *Leergang besliskunde, deel 8: minimaxmethode, netwerkplanning, simulatie.* 1968.

2.1 G.J.R. Förch, P.J. van der Houwen, R.P. van de Riet. *Colloquium stabiliteit van differentieschema's, deel 1.* 1967.

2.2 L. Dekker, T.J. Dekker, P.J. van der Houwen, M.N. Spijker. *Colloquium stabiliteit van differentieschema's, deel 2.* 1968.

3.1 H.A. Lauwerier. *Randwaardeproblemen, deel 1.* 1967.

3.2 H.A. Lauwerier. *Randwaardeproblemen, deel 2.* 1968.

3.3 H.A. Lauwerier. *Randwaardeproblemen, deel 3.* 1968.

4 H.A. Lauwerier. *Representaties van groepen.* 1968.

5 J.H. van Lint, J.J. Seidel, P.C. Baayen. *Colloquium discrete wiskunde.* 1968.

6 K.K. Koksma. *Cursus ALGOL 60.* 1969.

7.1 *Colloquium moderne rekenmachines, deel 1.* 1969.

7.2 *Colloquium moderne rekenmachines, deel 2.* 1969.

8 H. Bavinck, J. Grasman. *Relaxatietrillingen.* 1969.

9.1 T.M.T. Coolen, G.J.R. Förch, E.M. de Jager, H.G.J. Pijls. *Colloquium elliptische differentiaalvergelijkingen, deel 1.* 1970.

9.2 W.P. van den Brink, T.M.T. Coolen, B. Dijkhuis, P.P.N. de Groen, P.J. van der Houwen, E.M. de Jager, N.M. Temme, R.J. de Vogelaere. *Colloquium elliptische differentiaalvergelijkingen, deel 2.* 1970.

10 J. Fabius, W.R. van Zwet. *Grondbegrippen van de waarschijnlijkheidsrekening.* 1970.

11 H. Bart, M.A. Kaashoek. H.G.J. Pijls, W.J. de Schipper, J. de Vries. *Colloquium halfalgebra's en positieve operatoren.* 1971.

12 T.J. Dekker. *Numerieke algebra.* 1971.

13 F.E.J. Kruseman Aretz. *Programmeren voor rekenautomaten; de MC ALGOL 60 vertaler voor de EL X8.* 1971.

14 H. Bavinck, W. Gautschi, G.M. Willems. *Colloquium approximatietheorie.* 1971.

15.1 T.J. Dekker, P.W. Hemker, P.J. van der Houwen. *Colloquium stijve differentiaalvergelijkingen, deel 1.* 1972.

15.2 P.A. Beentjes, K. Dekker, H.C. Hemker, S.P.N. van Kampen, G.M. Willems. *Colloquium stijve differentiaalvergelijkingen, deel 2.* 1973.

15.3 P.A. Beentjes, K. Dekker, P.W. Hemker, M. van Veldhuizen. *Colloquium stijve differentiaalvergelijkingen, deel 3.* 1975.

16.1 L. Geurts. *Cursus programmeren, deel 1: de elementen van het programmeren.* 1973.

16.2 L. Geurts. *Cursus programmeren, deel 2: de programmeertaal ALGOL 60.* 1973.

17.1 P.S. Stobbe. *Lineaire algebra, deel 1.* 1973.

17.2 P.S. Stobbe. *Lineaire algebra, deel 2.* 1973.

17.3 N.M. Temme. *Lineaire algebra, deel 3.* 1976.

18 F. van der Blij, H. Freudenthal, J.J. de Iongh, J.J. Seidel, A. van Wijngaarden. *Een kwart eeuw wiskunde 1946-1971, syllabus van de vakantiecursus 1971.* 1973.

19 A. Hordijk, R. Potharst, J.Th. Runnenburg. *Optimaal stoppen van Markovketens.* 1973.

20 T.M.T. Coolen, P.W. Hemker, P.J. van der Houwen, E. Slagt. *ALGOL 60 procedures voor begin- en randwaardeproblemen.* 1976.

21 J.W. de Bakker (red.). *Colloquium programmacorrectheid.* 1975.

22 R. Helmers, J. Oosterhoff, F.H. Ruymgaart, M.C.A. van Zuylen. *Asymptotische methoden in de toetsingstheorie; toepassingen van naburigheid.* 1976.

23.1 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 1.* 1976.

23.2 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 2.* 1977.

24.1 P.J. van der Houwen. *Numerieke integratie van differentiaalvergelijkingen, deel 1: eenstapsmethoden.* 1974.

25 *Colloquium structuur van programmeertalen.* 1976.

26.1 N.M. Temme (ed.). *Nonlinear analysis, volume 1.* 1976.

26.2 N.M. Temme (ed.). *Nonlinear analysis, volume 2.* 1976.

27 M. Bakker, P.W. Hemker, P.J. van der Houwen, S.J. Polak, M. van Veldhuizen. *Colloquium discretiseringsmethoden.* 1976.

28 O. Diekmann, N.M. Temme (eds.). *Nonlinear diffusion problems.* 1976.

29.1 J.C.P. Bus (red.). *Colloquium numerieke programmatuur, deel 1A, deel 1B.* 1976.

29.2 H.J.J. te Riele (red.). *Colloquium numerieke programmatuur, deel 2.* 1977.

30 J. Heering, P. Klint (red.). *Colloquium programmeeromgevingen.* 1983.

31 J.H. van Lint (red.). *Inleiding in de coderingstheorie.* 1976.

32 L. Geurts (red.). *Colloquium bedrijfsystemen.* 1976.

33 P.J. van der Houwen. *Berekening van waterstanden in zeeën en rivieren.* 1977.

34 J. Hemelrijk. *Oriënterende cursus mathematische statistiek.* 1977.

35 P.J.W. ten Hagen (red.). *Colloquium computer graphics.* 1978.

36 J.M. Aarts, J. de Vries. *Colloquium topologische dynamische systemen.* 1977.

37 J.C. van Vliet (red.). *Colloquium capita datastructuren.* 1978.

38.1 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part I.* 1979.

38.2 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part II.* 1979.

39 O.J. Vrieze, G.L. Wanrooy. *Colloquium stochastische spelen.* 1978.

40 J. van Tiel. *Convexe analyse.* 1979.

41 H.J.J. te Riele (ed.). *Colloquium numerical treatment of integral equations.* 1979.

42 J.C. van Vliet (red.). *Colloquium capita implementatie van programmeertalen.* 1980.

43 A.M. Cohen, H.A. Wilbrink. *Eindige groepen (een inleidende cursus).* 1980.

44 J.G. Verwer (ed.). *Colloquium numerical solution of partial differential equations.* 1980.

45 P. Klint (red.). *Colloquium hogere programmeertalen en computerarchitectuur.* 1980.

46.1 P.M.G. Apers (red.). *Colloquium databankorganisatie, deel 1.* 1981.

46.2 P.G.M. Apers (red.). *Colloquium databankorganisatie, deel 2.* 1981.

47.1 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60: general information and indices.* 1981.

47.2 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 1: elementary procedures; vol. 2: algebraic evaluations.* 1981.

47.3 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3A: linear algebra, part I.* 1981.

47.4 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3B: linear algebra, part II.* 1981.

47.5 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 4: analytical evaluations; vol. 5A: analytical problems, part I.* 1981.

47.6 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 5B: analytical problems, part II.* 1981.

47.7 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 6: special functions and constants; vol. 7: interpolation and approximation.* 1981.

48.1 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 1.* 1982.

48.2 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 2.* 1982.

49 T.H. Koornwinder (ed.). *The structure of real semisimple Lie groups.* 1982.

50 H. Nijmeijer. *Inleiding systeemtheorie.* 1982.

51 P.J. Hoogendoorn (red.). *Cursus cryptografie.* 1983.

CWI SYLLABI

1 Vacantiecursus 1984 *Hewet - plus wiskunde.* 1984.

2 E.M. de Jager, H.G.J. Pijls (eds.). *Proceedings Seminar 1981-1982 Mathematical structures in field theories.* 1984.

3 W.C.M. Kallenberg et al. *Testing statistical hypotheses: worked solutions.* 1984.

4 J.G. Verwer (ed.). *Colloquium topics in applied numerical analysis, volume 1.* 1984.

5 J.G. Verwer (ed.). *Colloquium topics in applied numerical analysis, volume 2.* 1984.

6 P.J.M. Bongaarts, J.N. Buur, E.A. de Kerf, R. Martini, H.G.J. Pijls, J.W. de Roever. *Proceedings Seminar 1982-1983 Mathematical structures in field theories.* 1985.